Overviewing the emerging methods for predicting urban Sprawl features

Ange-Gabriel Bélinga[1,⋆] and *Mohamed* El Haziti[1,2]

[1] *LRIT, Faculty of Sciences in Rabat, Mohammed V University in Rabat, Morocco.*
[2] *Mohammed V University, High School of Technology, Sale Morocco*

Abstract. Urban sprawl, a common phenomenon characterized by uncontrolled urban growth, has far-reaching socio-economic and environmental implications. It's a complex phenomenon, and finding a better way to tackle it is essential. Accurate simulation and prediction of urban sprawl features would facilitate decision-making in urban planning and the formulation of city growth policies. This article provides an overview of the techniques used to this end. Initially, it highlights the use of a certain category of so-called traditional methods, such as statistical models or classical machine learning methods. It then focuses particularly on the intersection of deep learning and urban sprawl modelling, examining how deep learning methods are being exploited to simulate and predict urban sprawl. I finally studies hybrid approaches that combine deep learning with agent-based models, cellular automata, or other techniques offer a synergistic way to leverage the strengths of different methodologies for urban sprawl modelling.

1 Introduction

Urbanization is an ongoing global phenomenon that presents both opportunities and challenges for sustainable development and urban planning [1, 2]. As cities expand, the phenomenon of urban sprawl, characterized by the rapid and sometime uncontrolled outward growth of urban areas, has gained prominence as a critical issue affecting land use patterns, infrastructure development, and environmental sustainability [3]. This raises the question to know what exactly is urban sprawl? In other words, what are its characteristics, so that we can describe, model, predict and monitor them, in order to find better compromises to guide city growth. The issue of urban sprawl has been nkown as complex and difficult to be generalized from on author to another [4, 5]. Despite the multitude of works in the literature addressing this issue, the exact definition of urban sprawl and the exhaustiveness of its features are even not fixed. They vary from one author or context to another. For example, some authors equate this phenomenon with problems of density (population, buildings, etc.), while others approach it in terms of land use and/or cover [6], or layout/lanscape [7]. Others, on the other hand, focus on environmental issues, infrastructure and socio-economic indicators to characterize urban sprawl [8].

Generally, addressing a such complex dynamics of urban sprawl requires advanced tools and methodologies capable of predicting and simulating the spatial evolution of urban areas. Thus, several works in the literature have attempted to propose both conceptual and empirical models to try and characterize this phenomenon, which is a headache for many cities. Among the wide variety of methods used to characterize urban sprawl indicators, machine learning methods, and in particular those based on a deep learning (DL) approaches, have been the most promising in recent years [5]. However, little work has been done to establish a state-of-the-art conceptual framework to guide future work, which promises to be very numerous according to the trend observed. An important aspect of such a study would be to draw up a comparative analysis of the emerging models, their strengths and weaknesses, and to discuss their adaptation to fit up one or more urban sprawls' indexes.

Indeed, in the remainder of this paper section 2 will overview the concept of urban sprawl, its mean characteristics and challenges while section 3 will deal with emerging urban sprawl's predictive methods. Finally, section 4 will summerize and conclude our study.

2 About urban Sprawl and its Challenges

2.1 Definition and key characteristics

Urban sprawl refers to the unrestricted and often unplanned expansion of urban areas into surrounding rural or undeveloped land [2]. It is characterized by the outward growth of a city or metropolitan area, resulting in low-density development, inefficient land use, and a dispersion of population and infrastructure [9]. Urban sprawl can have far-reaching impacts on land use patterns, infrastructure, transportation, the environment, and the overall quality of life within a region [10, 11].

As shown in Figure 1, key characteristics of urban sprawl include: Low Population Density [9, 11, 12], Car-Centric Transportation [13], Fragmented and Discontin-

⋆e-mail: agbelinga@gmail.com

https://doi.org/10.1051/e3sconf/202341803008

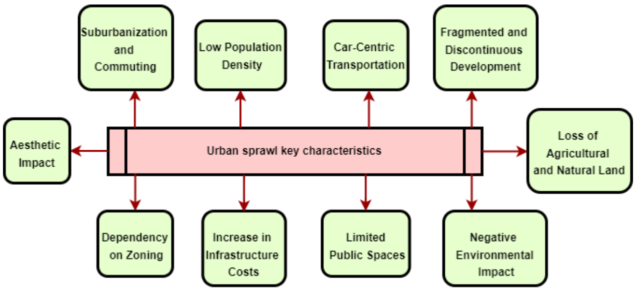

Figure 1. Key characteristics of urban sprawl.

uous Development [12], Loss of Agricultural and Natural Land [2, 6, 14], Increase in Infrastructure Costs [2, 15], Limited Public Spaces [16], Negative Environmental Impact [1–3, 14], Suburbanization and Commuting [2, 3, 13, 17], Dependency on Zoning [6], Aesthetic Impact [18].

2.2 Challenges from urban sprawl

Urban sprawl as described in the previous sections would pose a number of challenges on several fronts. These challenges can be grouped into three main categories: environmental, socio-economic and infrastructural as hown in figure 2. Each of these categories contains specific challenges that the city will have to overcome as a result of the emergence of urban sprawl.

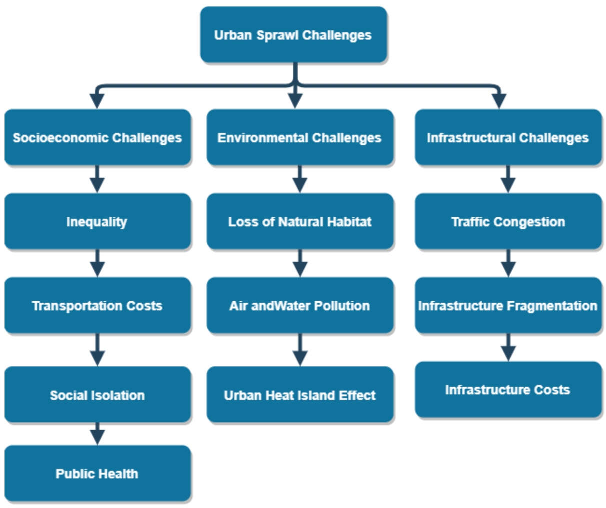

Figure 2. Challenges of urban sprawl.

For example, the socio-economic challenges will raise issues such as growing social and/or economic inequalities, susceptibility to racial segregation caused by low densities, social isolation, high transportation costs, public health problems, etc. In the category of infrastructural challenges, the sprawling city will have to cope with high infrastructure construction costs, traffic congestion, infrastructure fragmentation, etc. And finally, in terms of environmental challenges, the sprawling city will have to overcome climatic challenges (heat islands, high temperatures, etc.), water and air pollution, loss of natural habitats, destruction of green spaces, and so on.

Addressing these challenges requires a shift towards more sustainable urban planning approaches. Compact and mixed-use development, improved public transportation systems, green infrastructure, and the promotion of walkable neighborhoods are key strategies to mitigate the negative impacts of urban sprawl. Such strategies can enhance socioeconomic equity, reduce environmental degradation, and improve the overall quality of life in urban areas.

2.3 Needing accurate and predictive urban sprawl modelling

The need for accurate and predictive urban sprawl modelling techniques arises from the profound impact of urbanization on the environment, society, and economy. As cities continue to expand, urban sprawl has become a pressing challenge with far-reaching implications. Accurate modelling is essential for understanding, anticipating, and managing these implications effectively. Here are key reasons highlighting the importance of precise urban sprawl modeling: Informed Decision-Making [1, 15, 16], Resource Management [3, 19], Infrastructure Planning [2, 3, 6], Environmental Conservation [1], Mitigating Environmental Impact, [2],Climate Change Mitigation [15], Economic Planning [6, 13, 15], Policy Evaluation [2, 6, 15], Sustainable Development [2, 6, 15, 20].

Predictive urban sprawl modelling techniques can serve as essential tools for guiding the trajectory of urban development in a sustainable, informed, and proactive manner. By facilitating effective decision-making, resource management, and environmental stewardship, these models play a pivotal role in shaping the future of our cities.

3 Urban sprawl prediction methods

Predicting urban sprawl is a complex task that involves analyzing various factors and trends related to urban development. It is essential to consider a wide range of indicators and factors that can influence the extent and patterns of urban expansion. These indicators are typically derived from various data sources, including demographic, economic, environmental, and spatial data. Tekouabou S.C.K. [4] has highlighted a generic framework for urban form indicators modelling that is also valuable for urban sprawl features prediction. Urban form's features are key indicators and factors to take into consideration. These indicators can be population growth ; economic indicators such as job growth, income levels; Land Use and Land Cover (LULC) such as urban, agricultural, forested, and vacant; Transportation Infrastructure like road networks, highways; Zoning and Land Use Regulations; Proximity to essential services like schools, hospitals, grocery stores; proximity to natural features like rivers, lakes, parks.
Urban data are collected from relevant data sources, including historical urban development data, land use data,

socioeconomic data, transportation data, and environmental data. These datasets may be obtained from government agencies, satellite imagery, surveys, or open data platforms. Predicting urban sprawl involves integrating these indicators and using them as features in predictive models. The choice of indicators may vary depending on the specific region, context, and goals of the prediction task.

3.1 Traditional urban sprawl modelling techniques

Traditional urban sprawl modelling techniques encompass a range of approaches that have been used to simulate and predict the expansion of urban areas. These methods have evolved over time and have contributed to our understanding of urban growth dynamics [4]. Two notable categories of traditional urban sprawl modelling techniques are statistical models and cellular automata.

Statistical Models approach urban sprawl modeling from a quantitative perspective [21–23], leveraging historical data and various statistical techniques to understand the factors influencing urban expansion. These models often focus on relationships between socio-economic variables [24, 25], land use patterns [26], and demographic trends [27].

Cellular Automata (CA) are spatially explicit models that divide the study area into discrete cells or pixels, each of which can have multiple states (e.g., urban, agricultural, vacant) [2]. These models simulate changes in land use based on predefined rules that determine how cells transition from one state to another [28]. Cellular automata offer a dynamic representation of urban sprawl and can capture interactions between neighbouring cells.

Both statistical models and cellular automata have contributed valuable insights into the drivers and dynamics of urban sprawl. However, they also come with limitations, such as oversimplification of complex processes or assumptions that might not hold in all scenarios. Modern urban sprawl modelling often integrates these traditional approaches with newer techniques, such as machine learning [1] and remote sensing [21], to enhance accuracy and provide a more comprehensive understanding of urban growth patterns.

3.2 Deep learning methods for urban sprawl modelling

Deep learning (DL) methods have been adapted and integrated into urban sprawl modelling to enhance the accuracy, complexity, and predictive capabilities of these models. Leveraging the power of deep neural networks, these approaches can capture intricate spatial relationships, learn from large datasets, and make more informed predictions about urban expansion. [4] presented all ML and DL techniques that can be used for predicting urban sprawl. The ways in which DL has been utilized in urban sprawl modeling involves :

1) High-Resolution Satellite Imagery Analysis: Deep learning methods excel at processing and analyzing high-resolution satellite imagery [3, 29]. Convolutional Neural Networks (CNNs) have been applied to classify land cover types [30], detect urban areas [31], and identify changes in land use over time [31]. This enables accurate mapping of urban expansion and the identification of fine-grained patterns.

2) Feature Extraction: Deep learning models can automatically learn relevant features from raw data, eliminating the need for manual feature engineering. This is particularly useful when dealing with complex, multi-dimensional data such as satellite images [3, 29], where identifying relevant features can be challenging.

3) Spatial and Temporal Dynamics: Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks can capture spatial and temporal dependencies in urban sprawl [30]. By processing sequential data over time, these networks can learn how urban areas evolve and expand, accounting for trends and patterns.

4) Multi-Scale Analysis: Deep learning models can operate at multiple scales, from local neighbourhoods to entire metropolitan areas. This ability to process data at different resolutions allows for a comprehensive analysis of urban sprawl dynamics [32].

5) Data Fusion: Deep learning techniques enable the fusion of diverse data sources, such as satellite imagery [3, 29], socioeconomic data and environmental variables [24]. This integrated data can provide a more holistic understanding of the factors driving urban expansion.

6) Transfer Learning: Pre-trained deep learning models, such as CNNs trained on large image datasets, can be fine-tuned for specific urban sprawl modelling tasks. This approach leverages the knowledge captured by the pre-trained model and adapts it to the urban context.

7) Generative Models for Scenario Planning: Generative models like Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) can generate synthetic urban growth scenarios based on historical data [33]. These scenarios aid in exploring potential future urban expansion patterns and evaluating the impacts of different policy interventions.

8) Hybrid Models: Deep learning methods can be integrated with traditional urban sprawl modelling techniques, such as cellular automata [10] or regression models [34]. This hybrid approach leverages the strengths of both methodologies, resulting in more accurate and comprehensive predictions.

3.3 Hybrid Approaches

Integrating deep learning techniques with traditional simulation methods can offer a powerful approach that leverages the strengths of both methodologies. This hybrid approach capitalizes on the accuracy and complexity of deep learning while incorporating the interpretability and domain-specific knowledge of traditional simulation methods. Here's how this integration can be beneficial for urban sprawl modeling:

1) Enhanced Accuracy and Complexity knowing that DL-based models could capture intricate spatial patterns, complex relationships, and high-dimensional data representations. By integrating deep learning with traditional methods, the accuracy of predictions can be improved,

especially in capturing non-linear interactions and fine-grained details [35].
2) Interpretability and transparency because the traditional methods are often more interpretable, allowing stakeholders to understand and validate the underlying logic of the model. By combining deep learning with traditional methods, the hybrid model can provide both accurate predictions and explanations for decision-making, fostering better understanding and trust.
3) Data Efficiency indeed, DL methods typically require a large amount of labeled data for training. Integrating deep learning with traditional methods can help mitigate data scarcity by using expert-defined rules as a basis, reducing the demand for extensive training data.
4) Robustness and Generalization as the traditional methods can be designed to capture fundamental trends and principles, making them more robust against noisy or limited data. By combining these robust principles with the data-driven capabilities of deep learning, the hybrid model can achieve better generalization and adaptability.
5) Customization for Specific Goals knowing that the Hybrid models can be tailored to specific urban planning goals and objectives. By combining the predictive capabilities of deep learning with the targeted focus of traditional models, the hybrid approach can be fine-tuned to address specific challenges in urban sprawl modeling.

Hybrid approaches that combine deep learning with agent-based models, cellular automata, or other techniques offer a synergistic way to leverage the strengths of different methodologies for urban sprawl modelling. These hybrid models can provide accurate predictions while incorporating complex interactions and domain-specific rules. Here are some ways in which deep learning can be integrated with these traditional techniques.

- Deep Learning-Enhanced Agent-Based Models : In an agent-based model, individual agents (representing entities like households, businesses, or people) interact with each other and their environment based on rules and behaviours. Deep learning can enhance agent-based models by providing accurate predictions for agent behaviour, such as movement patterns or decision-making. For example, deep learning can predict commute patterns using historical data, informing agent mobility decisions in an urban simulation.

- Deep Learning-Driven Cellular Automata : Cellular automata models divide a study area into cells, each of which can transition between different states based on predefined rules. DL could enhance cellular automata models by predicting transition rules based on spatial and temporal patterns learned from data. For instance, DL can identify land use change patterns and relationships, improving the accuracy of state transitions in the cellular automata model.

- Data-Driven Model Initialization : Deep learning can provide initial conditions or parameters for traditional models, ensuring that they start with accurate baseline data. For example, deep learning models can estimate population distributions, which can be used as input for agent-based models or other simulation methods.

- Data-Driven Model Calibration : Deep learning can assist in calibrating traditional models by learning relationships between model parameters and observed data. This can lead to more accurate parameter estimation, improving the performance of traditional models.

- Fusion of Data Sources : Deep learning can fuse diverse data sources, such as satellite imagery and demographic data, enhancing the input data for traditional models. This enriched data can lead to better-informed simulations in cellular automata or agent-based models.

So, combining DL-based approaches with agent-based models, cellular automata, or other traditional techniques is efficient for urban sprawl modelling for accurate predictions, nuanced relationships, domain-specific rules, and the ability to capture complex interactions. These hybrid approaches enable a more comprehensive understanding of urban growth dynamics and contribute to more effective urban planning and policy decisions.

4 Conclusion

Having reached the end of this work, it has been a question for us of overviewing the methods for predicting urban sprawl, with the emphasis on deep learning methods. However, given the complexity of this phenomenon, we were first called upon to define it, recall its main characteristics and its challenges before tackling the methods used to predict it. As a result, three main categories of methods can be identified in the literature for this purpose. The first category includes so-called traditional methods, more specifically statistical-based methods [26, 27] and cellular automata approaches [2, 28]. The second category concerns emerging deep learning methods, such as the use of convolutional neural networks, transfer learning, recurrent networks such as long short term momory (LSTM) and recurrent gatered units (GRU) [30], etc. While the first category of methods has the advantage of being easy to explain, the second seems more effective for modelling complex behaviours such as urban sprawl indicators. Hence the third category, which consists of hybrid methods that enable a compromise to be reached between the two previous categories of method [35]. In our future work, we will carry out more in-depth studies exploring the adaptations between mixing methods and targeted indicators.

References

[1] E. Purswani, S. Verma, S. Jayakumar, M. Khan, B. Pathak, Journal of Urban Management **11**, 82 (2022)
[2] A. Sarkar, P. Chouhan, Modeling Earth Systems and Environment **5**, 1723 (2019)
[3] A. Hua, P. Gani, Applied Ecology & Environmental Research **21** (2023)
[4] S.C.K. Tekouabou, E.B. Diop, R. Azmi, R. Jaligot, J. Chenal, Journal of King Saud University-Computer and Information Sciences **34**, 5943 (2022)

[5] S.C. Koumetio Tekouabou, E.B. Diop, R. Azmi, J. Chenal, Archives of Computational Methods in Engineering **30**, 1421 (2023)

[6] T. Osman, D. Shaw, E. Kenawy, Journal of Land Use Science **13**, 565 (2018)

[7] A. Akın, M.A. Erdoğan, Modeling Earth Systems and Environment **6**, 1331 (2020)

[8] M. Seevarethnam, N. Rusli, G.H.T. Ling, ISPRS International Journal of Geo-Information **11**, 442 (2022)

[9] M.F. Baqa, F. Chen, L. Lu, S. Qureshi, A. Tariq, S. Wang, L. Jing, S. Hamza, Q. Li, Land **10**, 700 (2021)

[10] G. Prayitno, N. Sari, A.W. Hasyim, N.W. SW, GEOMATE Journal **18**, 64 (2020)

[11] B.H. Aithal, S. Vinay, T. Ramachandra, Modeling Earth Systems and Environment **4**, 1297 (2018)

[12] H. Dadashpoor, F. Salarian, Environment, Development and Sustainability **22**, 593 (2020)

[13] D. Abudu, R.A. Echima, G. Andogah, The Egyptian Journal of Remote Sensing and Space Science **22**, 315 (2019)

[14] K. Kundu, P. Halder, J. Mandal, Journal of the Indian Society of Remote Sensing **48**, 1535 (2020)

[15] S. Dinda, K. Das, N.D. Chatterjee, S. Ghosh, Modeling Earth Systems and Environment **5**, 331 (2019)

[16] J.J. Arsanjani, M. Helbich, W. Kainz, A.D. Boloorani, International Journal of Applied Earth Observation and Geoinformation **21**, 265 (2013)

[17] P. Pokojska (2019)

[18] M. Jafari, H. Majedi, S.M. Monavari, A.A. Alesheikh, M. Kheirkhah Zarkesh, Sustainability **8**, 810 (2016)

[19] M. Jat, P.K. Garg, D. Khare, International Journal of Remote Sensing **29**, 543 (2008)

[20] P. Jayasinghe, V. Raghavan, G. Yonezawa, Spatial Information Research pp. 1–14 (2021)

[21] R. Padmanaban, A.K. Bhowmik, P. Cabral, A. Zamyatin, O. Almegdadi, S. Wang, Entropy **19**, 163 (2017)

[22] S.C.K. Tékouabou, I. Chabbar, H. Toulni, W. Cherif, H. Silkan et al., Expert Systems with Applications **189**, 115975 (2022)

[23] S.C. Tékouabou, S.C. Gherghina, H. Toulni, P.N. Mata, J.M. Martins, Mathematics **10**, 2379 (2022)

[24] A. Khan, M. Sudheer, The Egyptian Journal of Remote Sensing and Space Science **25**, 541 (2022)

[25] S.C.K. Tekouabou, W. Cherif, H. Silkan, *A data modeling approach for classification problems: application to bank telemarketing prediction*, in *Proceedings of the 2nd International Conference on Networking, Information Systems & Security* (2019), pp. 1–7

[26] D.G. Brown, B.C. Pijanowski, J.D. Duh, Journal of Environmental Management **59**, 247 (2000)

[27] L. Congedo, S. Macchi, Current Opinion in Environmental Sustainability **13**, 1 (2015)

[28] Y. Huang, J.L. Huang, T.J. Liao, X. Liang, H. Tian, Science of the total environment **643**, 1553 (2018)

[29] M. Adel Saharkhiz, B. Pradhan, H. Mojaddadi Rizeei, H.S. Jung, Korean Journal of Remote Sensing (2020)

[30] W. Boulila, H. Ghandorh, M.A. Khan, F. Ahmed, J. Ahmad, Ecological Informatics **64**, 101325 (2021)

[31] A. Ahmadi, S. Khalesi, A. Golroo, International Journal of Pavement Engineering **23**, 3536 (2022)

[32] H. Chen, C. Wu, B. Du, L. Zhang, arXiv preprint arXiv:1906.11479 (2019)

[33] K. Man, J. Chahl, Journal of Imaging **8**, 310 (2022)

[34] S.S. Somvanshi, O. Bhalla, P. Kunwar, M. Singh, P. Singh, Environment, Development and Sustainability **22**, 1073 (2020)

[35] M. Li, S. Gao, P. Qiu, W. Tu, F. Lu, T. Zhao, Q. Li, Transportation Research Part C: Emerging Technologies **144**, 103908 (2022)