A GEOSPATIAL MODELLING FRAMEWORK TO SIMULATE FUTURE SCENARIOS OF URBAN FORM - A CASE STUDY IN PHNOM PENH, CAMBODIA

Richard Burke

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the School of Engineering and and the Faculty of Science, Agriculture and Engineering

School of Engineering
Newcastle University
Newcastle Upon Tyne
United Kingdom

14 April 2025

ACKNOWLEDGEMENTS

This PhD was funded by the United Kingdom's Engineering and Physical Sciences Research Council as part of the Centre for Doctoral Training in Geospatial Systems, under grant number EP/S023577/1.

I thank my supervisors, Dr Alistair Ford and Professor Richard Dawson, for their invaluable support and guidance throughout my PhD research. Their insightful critiques and direction during supervision meetings have been instrumental in shaping this project. I also thank my former supervisor, Dr Jin Xing, for guiding the project in its early stages.

I was honoured to receive a PhD studentship with the Centre for Doctoral Training in Geospatial Systems in 2020. The past five years have been a once-in-a-lifetime opportunity and I am deeply grateful for the experience. Being part of the Centre for Doctoral Training has been invaluable and I would like to thank the the centre's manager, Jamie Stogden, and director, Professor Jon Mills, for organising such an excellent and well-structured programme.

I am also grateful to my fellow PhD students for their support. In particular, I would like to thank David Gregg and Kaitlyn Ries for their expertise in assisting me in writing mathematical equations and using the Linux command line for AWS instance transfers. A special thanks to Adam Booth for allowing me to run deep learning models on his machine, helping me reduce AWS costs and ensure modelling efficiency.

Finally, I extend my heartfelt thanks to my family for their support and love throughout my PhD journey.

ABSTRACT

Urbanisation is proliferating globally, with the majority of the world's population now living in urban areas. This rapid urban growth places significant pressure on cities to expand, often leading to physical expansion rates that outpace population growth. This growth is particularly pronounced in developing regions, where it poses numerous challenges. Urban growth modelling offers decision-makers a valuable tool to understand the complexities of urban dynamics, anticipate future transformations, and design sustainable and liveable cities. However, existing modelling approaches have largely focussed on large-scale, rasterised, and gridded projections of urban expansion that lack relevance and specificity to the regions being modelled. Furthermore, the models lack a systematic assessment procedure and there has been insufficient attention to generating detailed urban form layouts. To address these gaps, it is essential to use vector data at finer spatial scales to enhance contextual relevance and apply robust assessment methodologies that improve model credibility. Additionally, generating future urban form layouts in developing regions is critical for informed urban planning. Therefore, this PhD thesis aims to develop a geospatial modelling framework to simulate, explore and characterise future urban form scenarios.

Phnom Penh, Cambodia, was selected as a case study due to its rapid urban expansion, data-scarce environment, and the relatively small proportion of built-up area. The framework employs two state-of-the-art techniques: agent-based modelling (ABM) and generative adversarial networks (GAN). The ABM replicates the urban development process based on academic literature and government reports, representing key actors, such as urban planners, property developers, and landowners, involved in this process as agents to simulate future urban growth. This ABM undergoes a thorough and systematic evaluation process, including stabilisation analysis, calibration and validation, and sensitivity analysis, to ensure its robustness. Dasymetric mapping is performed to convert the ABM's development probability outputs into a raster of future development demand density. The GAN then uses this future development density to perform image-to-image translation, converting a density raster into future road network layouts and subsequent building footprints based on the previously generated road network. The generated urban form outputs are then assessed against ground truth urban form using a range of pixel and vector-based accuracy metrics to identify the best-performing GAN experiment for future urban form generation.

The ABM simulated five urban growth scenarios in Phnom Penh by 2040 which range from low to high growth under regulated and unregulated conditions. These scenarios are based on likely socioeconomic forecasts from grey literature and are variations of the baseline simulation. The evaluation protocol's stabilisation analysis revealed that the ABM's stochastic behaviour was effectively captured after 240 model runs. The ABM achieved good performance in calibration, achieving area-weighted fuzzy kappa simulation values of 0.241 in calibration and 0.509 in validation showing sufficient performance for future simulation. The sensitivity analysis explored 8 model variables and determined their importance, interactions, and polarity with the fuzzy kappa simulation output, ensuring model credibility. The best performing GAN for road network generation in the study area achieved a mean absolute error of 40.96%. Whilst the best performing GANs for building footprint generation in the study area achieved mean absolute errors of 54.29% and 29.48%. The baseline or business as usual future simulation from the ABM projects 18.04 km^2 of new urban area by 2040 in the study area. For urban form, this results in 227.68 km of new roads and 26,521 new building footprints, this represents a 71.52% increase in road networks and 106.63% in building footprints from the 2023 urban form. These projections indicate large increases in new urban area and urban form over the following decades.

The ABM scenarios achieve varied and distinct patterns of growth, with differing locations and probabilities of growth, demonstrating the ABM is responsive to variations in the input variables. The evaluation protocol highlighted issues with separating the calibration and validation period in equal temporal periods and weaknesses of the fuzzy kappa simulation as a result of this. The protocol could be developed further by employing explicit spatiotemporal sensitivity analyses to enhance model understanding in the study area. GANs are demonstrated to be a feasible and practical technique in creating realistic and convincing future urban form that integrate seamlessly with existing layouts. However, GAN training performance was sensitive to input image resolution and content and they do generally underestimate urban form features. Overall, this PhD thesis advances urban growth modelling approaches by presenting a novel geospatial modelling framework as a new standard in the field, which can be easily re-applied in other data-scarce areas experiencing rapid growth. It has strong implications for academic researchers, policy makers, and practitioners. It offers an innovative approach to generating future urban form with multidisciplinary applications in natural disaster preparedness, urban planning, and transportation modelling.

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ABM	Agent-Based Model
AI	Aritificial Intelligence
AWS	Amazon Web Services
BAU	Business as Usual
CA	Cellular Automata
CBD	Central Business District
CNN	Convolutional Neural Network
CV	Coefficient of Variation
DGM	Deep Generative Model
EB	Economic Boom
EO4SD	Earth Observation for Sustainable Development
ER	Economic Recession
ESRI	Environmental Systems Research Institute
FID	Fréchet Inception Distance
FKS	Fuzzy Kappa Simulation
FoM	Figure of Merit
GA	Genetic Algorithm
GAMA	GIS Agent-Based Modelling Architecture
GAN	Generative Adversarial Network
GDP	Gross Domestic Product
GGGI	Global Green Growth Institute

GIS	Geographical Information Science
GSA	Global sensitivity analysis
GWR	Geographically Weighted Regression
IP	Intelligent Planning
IPCC	Intergovernmental Panel on Climate Change
KS	Kolmogorov-Smirnov
LB	Land Bubble
LHS	Latin Hypercube Sampling
LR	Logistic Regression
LSA	Local Sensitivity Analysis
MAE	Mean Absolute Error
mIoU	Mean Intersection Over Union
ND	Node Degree
OAT	One-at-a-time
ODD	Overview, Design concepts, and Details
OFAT	One-factor-at-a-time
OSM	Open Street Map
PCM	Percent Correct Match
PSO	Particle Swarm Optimisation
RF	Random Forests
S_i	First-Order Sensitivity Index
SA	Sensitivity Analysis
SLEUTH	Slope Landuse Exclusion Urban Transportation and Hillshade
SOLV	Stochastic Urban Growth Model Evaluation
SVM	Support Vector Machines
TS_i	Total-Order Sensitivity Index
UN	United Nations

Chapter 1

INTRODUCTION AND BACKGROUND

1.1 Background

Urbanisation around the world is one of the most visible and transformative drivers of human activity. How cities grow and change impacts not only the spatial structure of urban environments, but also broader issues such as sustainability, economic development, climate resilience, and quality of life. However, predicting the form and direction of urban growth remains a complex challenge due to the dynamic interplay of socioeconomic factors, limited data availability, and inherent uncertainty in human decision-making. This thesis seeks to address this gap by developing a geospatial modelling framework to simulate and generate future urban form. This study focusses on Phnom Penh, Cambodia, which is an example of a rapidly developing and data-scarce city. Different scenarios are simulated and analysed to explore how different socioeconomic trajectories may influence urban expansion. This research contributes to a broader understanding of how urban areas are likely to evolve and how socioeconomic conditions and urban planning policies can shape more sustainable and liveable cities.

1.1.1 Urbanisation and its impacts

Global urbanisation is increasing. In 1950, 29% (0.75 billion people) of the global population lived in urban areas, while 71% lived in rural areas and in 2018, 55% (4 billion) lived in urban areas. This is projected to increase to 68% (6 billion) by 2050 (Kundu et al., 2020). Rapid urbanisation since the mid-20th century has significantly expanded the size and spatial extent of cities. Cities are rapidly expanding, covering approximately 0.5% of the world's land area (Angel et al., 2011). This expansion has become twice as fast as the urban population growth rates, resulting in less dense urban land which is unsustainable and inefficient in the long term. Angel et al. (2011) identified that of 200 cities, only 23% of the population increase was accommodated by densification within existing urban footprints and 77% in new expansion areas. This highlights that despite efforts to promote densification, urban expansion remains the dominant mode of accommodating urban population growth.

Currently, urbanisation and growth are shifting from developed to developing regions. In 2018, Asia represented half of the urban population in the world, while Europe and North America comprised only 20% (Kundu et al., 2020). By 2050, 83.17% of the global urban population will be in less developed regions, mainly Africa and Asia (Kundu et al., 2020). It is also projected that by 2050, urban land cover in developed and developing countries will increase by 31.10% and 60.77%, respectively (Angel et al., 2011).

The most significant driver of urban growth is the migration of people from rural to urban areas within a country (Bhatta, 2010). There are numerous socioeconomic causes for this migration that ultimately leads to urbanisation and expansion of cities. For example, as countries undergo economic development, economic opportunities shift from the rural to the industrial and service sectors. The roles emerging in urban areas provide higher incomes and attract populations to cities (Zhang, 2016). As income increases, people pursue improved lifestyles, preferring larger homes and convenient access to amenities and recreation. This drives further urban development. Advancements in infrastructure and transportation significantly drive urban growth (Zhang, 2016). The construction of major highways and railways improves a city's capacity for outward expansion. Improved access to the city centre encourages residents to live outside the main urban area (Bhatta, 2010). These advancements lead to increased housing affordability and accessibility, attracting residents to urban areas. Government decisions also influence infrastructure development, housing policies, and zoning regulations, demonstrating its overarching role in the direction and pace of urban expansion. For example, restrictive land use policies in one area can redirect growth to regions with fewer controls, whereas decisions about the placement of highway or metro lines can concentrate growth around new infrastructure (Bhatta, 2010).

Urban sprawl, characterised by low density, fragmented and leap-frogging expansion, is a common symptom of rapid urban growth in developing regions (Yasin et al., 2020). Developing areas are vulnerable to urban sprawl due to improper land use regulations, economic pressures, and rapid population growth. Sprawl has a range of impacts on the urban and natural environment. Expansive urban areas result in more urban activities that increase greenhouse gas emissions and local temperatures, while the direct loss of vegetation and green space reduces the capacity to absorb carbon dioxide (Zhang, 2016). Cities account for approximately 78% of global energy consumption and generate 60-70% of energy-related greenhouse gas emissions, which exacerbates climate change (Kundu et al., 2020). Another impact of sprawl is intensification of the urban heat island effect, where cities experience higher temperatures compared to surrounding rural areas due to concentrated buildings, impervious surfaces, reduced greenery, and increased human activity (Frumkin, 2002). There is a higher risk of surface wa-

ter flooding due to the expansion of impermeable surfaces, which significantly reduces natural infiltration compared to permeable ground (Li et al., 2020b). Water quality also suffers from increased effluent pollution, with higher concentrations of sewage, oils, and greases entering rivers (Frumkin, 2002). Personal health and well-being are negatively affected. Lack of greenspace and natural environments causes mental health declines (Bhatta, 2010). Increased pollution and temperature increase health risks, particularly for vulnerable populations (Zhang, 2016). Sprawl also impacts rural livelihoods by encroaching on agricultural land. Consequently, this leads to the loss of arable land, a decrease in agricultural productivity, and a reduction in income for those dependent on agriculture (Kundu et al., 2020).

1.1.2 Urban form

Urban form refers to the organisation and physical structure of the city, encompassing various characteristics that define its spatial and functional identity. The specific components of urban form vary depending on the theoretical framework or the analytical perspective. Scheer (2001) categorised urban form into site, superstructure, infill, buildings, and objects, while Dempsey et al. (2010) focused on density, land use, transport infrastructure, layout, and buildings. More recently, the literature has condensed these elements into four core and concise elements (Kropf, 2017; Miao et al., 2017; Lin et al., 2024):

- Road networks - systems of interconnected streets and roads that facilitate movement and accessibility within a city. Road networks often follow a hierarchical structure based on function and connectivity.
- Blocks - areas bounded by roads and composed of multiple plots and blocks serve as fundamental organisational elements within urban spaces.
- Plots - (also referred to as parcels or lots) these are subdivisions of blocks and act as the fundamental units of land ownership within a city.
- Buildings - an integral component of urban form, buildings characterise urban form significantly by their placement, layout, and function.

These characteristics of urban form are strongly interrelated. The structure of road networks defines the size and arrangement of plots, which are further aggregated into blocks. Blocks, in turn, determine the density and organisation of buildings, shaping the overall character and functionality of the urban environment. Urban form significantly impacts various dimensions of sustainability, including social, economic, environmental, and energy-related aspects.

The configuration of these elements is important because it has a direct impact on factors such as accessibility, affordability, social equity, and the environmental performance of cities (Lin et al., 2024). For example, while compact urban form with higher densities can promote access to amenities and public transport, it can also lead to challenges related to social interaction, affordability, and availability of green space (Dempsey et al., 2010). The layout of road networks and the density of the buildings can significantly affect the efficiency of evacuation in natural disasters, such as floods and earthquakes, making urban form an important part of the mitigation of such events. Cities with well-planned grids and multiple exit options facilitate faster evacuations during emergencies. In contrast, areas with narrow streets and limited exits can hinder movement, posing risks during crises (Srinurak et al., 2016). The spatial arrangement of a city significantly influences energy consumption and greenhouse gas emissions. Compact, mixed-use developments reduce the need for long commutes by promoting public transport, walking, and cycling. They also offer lower carbon footprints and air temperatures due to smaller living spaces that require less heating and cooling, as well as better thermal insulation compared to single-family homes (Wiedenhofer et al., 2018). In contrast, sprawling suburban areas with larger buildings tend to have higher energy demands and local temperatures and increase dependence on vehicles, leading to higher carbon emissions (Wiedenhofer et al., 2018). Overall, the geometric and spatial configuration of urban form must be efficiently understood, planned, and designed to accommodate complex urban structures and ensure cities are liveable and sustainable (Fang et al., 2022).

1.1.3 Urban models

Models are simplifications of reality that are designed to capture the essential features of a system that are critical to a theory and its applications (Wang et al., 2024). They act as a tool for exploring and experimenting with theoretical ideas, offering predictive insights and fostering understanding. Urban models are computer simulations of how cities function, translating theory into a testable and applicable form without experimenting on the real thing (Batty, 2009). These models represent the spatial structure of the urban area in terms of land use, population, employment, and transportation. They are typically embodied in computational environments that allow theories to be tested against data and projections of future locational patterns to be generated (Batty, 2009).

Wang et al. (2024) classified three types of urban models according to their application (Figure 1.1):

- Economic and spatial interaction - These models apply macroeconomic and microeconomic principles to analyse the spatial distribution of socioeconomic and transportation activities in response to urban policies.

- Urban dynamic models - Based on the system dynamics methodology, these models represent urban socioeconomic processes as interconnected elements, capturing non-linear growth and chain reactions in urban systems.
- Behavioural models of individual agents - These simulate individual agents' behaviours and interactions within spatial environments, though quantifying and justifying these behaviours can be challenging.

There are two types of modelling approaches across these classifications: top-down and bottom-up. Top-down models use spatial interaction principles, such as the gravity model and entropy maximisation, to assume that individual behaviours are shaped by overarching macro trends, grouping individuals into homogeneous categories to simplify analysis and reduce the need for detailed behavioural information (Wang et al., 2024). However, bottom-up models focus on individual-level behaviours and interactions over space and time, using defined rules to simulate complex systems without requiring full understanding of the entire system (Wang et al., 2024).

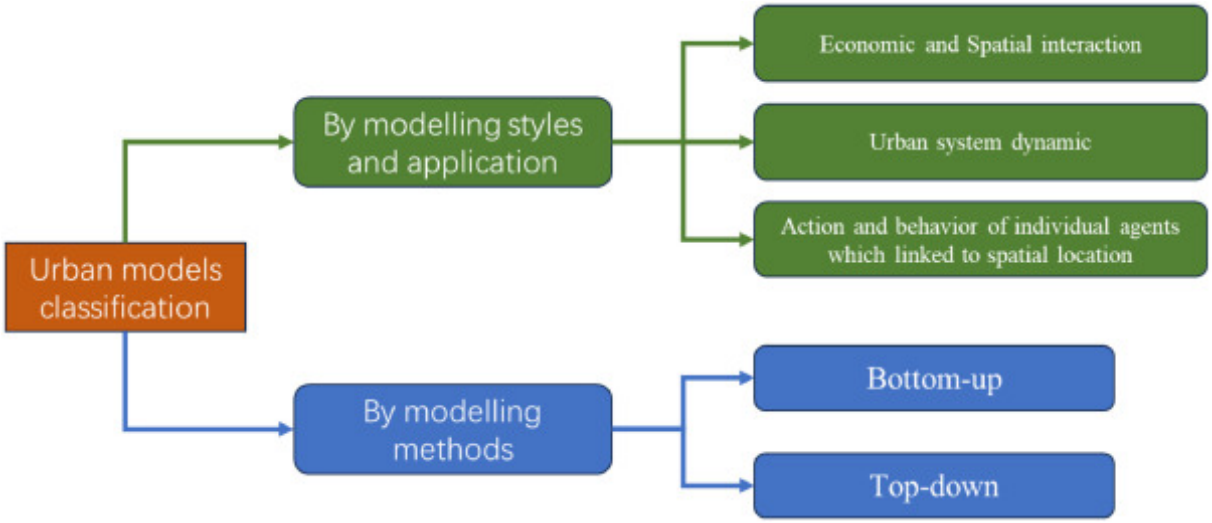

Figure 1.1: Urban model classification by application and method (Wang et al., 2024).

Among these urban models, the behavioural models are the most relevant and widely used in research for this thesis as their bottom-up approach is better suited to describe complex systems (Wang et al., 2024). These models typically involve simulating the behaviour of individual entities (e.g. people, households, or agents) using bottom-up techniques to observe emergent outcomes at a macro level (Li and O'Donoghue, 2013). For example, behavioural techniques could model traffic flows, land use change, or population dynamics. These models often rely on detailed data, explicit rules, and stochastic processes to replicate individual behaviours and interactions. Two of the most popular behavioural approaches are agent-based models (ABM) and cellular automata (CA), which have a strong focus on applications in urban growth modelling (Batty, 2009). These two techniques within the context of urban growth modelling are further discussed and explained in the literature review (Chapter 2). Urban models are useful be-

cause they contribute to theoretical development and more practical applications, such as decision making, physical planning, projecting future patterns, and understanding urban dynamics (Batty, 2009).

1.2 Research questions, aims and objectives

1.2.1 Questions

This background has highlighted the importance of addressing proliferating urban areas over the next century. It also underscores the critical role of urban form in ensuring cities remain liveable under such pressures. Spatial urban growth models have also been identified as relevant and viable for simulating and understanding the evolution of cities within a computational environment. This has raised the following research questions:

- What are the current shortcomings of current urban growth models and how can they be addressed?
- What is the best approach to generating future urban form and how can the necessary elements be effectively generated within the context of existing urban form?
- What potential future socioeconomic scenarios can be modelled and how will these scenarios be realised and quantified within the framework?

1.2.2 Aim and objectives

The aim of this thesis is to develop a geospatial modelling framework capable of simulating, exploring and characterising future scenarios of urban form by utilising state-of-the-art techniques. The following objectives have been formulated to achieve this aim:

1. To simulate future urban growth in an area experiencing rapid development using a spatial behavioural model.
2. To systematically assess the uncertainty, performance, and sensitivity of the spatial urban growth model.
3. To generate plausible future urban form based on the outputs of the urban growth model while considering existing urban form patterns.
4. To develop and analyse multiple scenarios of future urban form based on likely socioeconomic forecasts and planning policy changes relevant to the study area.

1.3 Thesis structure

This thesis is organised into eight chapters. Chapter 2 reviews the evolution of the previous literature related to behavioural urban modelling and urban form generation. This chapter explores urban growth models, justifying their importance and common techniques used, such as CA and ABMs, and then identifies limitations of current approaches. Urban form modelling techniques are then explored ranging from conventional techniques to deep generative models. Behavioural model evaluation procedures, such as calibration, validation, and sensitivity analysis, and urban form accuracy metrics are reviewed.

Chapter 3 introduces and justifies the case study for the modelling framework - Phnom Penh, Cambodia. It details the city's current growth trends, provides an explanation of its key features in and locates a specific study area in Phnom Penh to test the modelling framework. Chapter 4 then provides a detailed description of data preparation and processing required for the framework. It also outlines and justifies the framework's design, stages, and key components.

Chapter 5 addresses objectives 1 and 2 of the research by describing the urban growth model developed and the protocol used to evaluate it. This is the first half of the framework and the results of the model are reported here, as they are a direct input into the framework's second half. Chapter 6 describes the second half of the framework, which is the approach used to generate the future urban form for the study area using the urban growth model outputs and addresses objectives 3. Objective 4 is met in both chapters 5 and 6 because the creation of the scenarios requires alterations to specific variables and inputs in both parts of the framework

.

Chapter 7 discusses the results from the previous two chapters. It examines the performance, effectiveness, and contributions of the geospatial modelling framework by comparing it with existing literature. It also explains the benefits of future urban form modelling by highlighting various potential multidisciplinary applications to existing models, such as those for flooding and transport. Chapter 8 reviews the fulfilment of the aim and objectives, summarises the key findings, highlights research implications, and recommends future work to develop this PhD research.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

This chapter critically reviews the current literature on urban growth and urban form modelling. It begins by reviewing CA and ABMs of urban growth, highlighting their importance in understanding urban dynamics. Next, it explores techniques for urban form generation, ranging from conventional approaches such as procedural modelling to contemporary methods like DGMs, including GANs. The third section addresses rule-based model evaluation procedures, discussing stochasticity analysis, calibration and validation, sensitivity analysis, and accuracy metrics to establish robust and best practice approaches for model assessment. The fourth section delves into urban form accuracy metrics, comparing image-based and vector-based methods to identify optimal metrics. Finally, the chapter highlights and summarises key literature gaps and limitations that are addressed in the thesis. This review and synthesis provides the necessary context for the research presented.

2.2 Urban growth modelling

2.2.1 Urban growth modelling approaches

Many spatial urban growth models exist as they provide useful insights into future urban transformations, detect likely future trends, enable investigation into urban expansion driving factors and enable decision makers to plan urban environments more sustainably (Mozaffaree Pour and Oja, 2021). There are numerous techniques to simulate future urban growth, such as regression models, Markov chains, and deep learning, and integrated approaches. This section will describe these techniques and identify their respective advantages and limitations.

Regression models, such as logistic regression, are widely used in urban expansion and land use change studies to estimate the likelihood of land transitioning to urban

use based on various explanatory variables, such as proximity to roads and population density (Salem et al., 2021). Logistic regression offers clear interpretability of driving forces and flexibility in handling different data types, making it valuable for identifying key factors influencing urbanisation and generating probability maps for spatial simulations (Nong and Du, 2011). However, its limitations include assuming a linear relationship between driving factors and urban change and treating spatial units independently in the model, which oversimplifies complex urban systems (Achmad et al., 2015). Due to its static nature it is difficult to use this technique to capture non-linear dynamics (Salem et al., 2021).

Markov chain models are stochastic tools used to project land use change by estimating the probability of land transitioning between categories, such as agricultural to urban, based on past patterns (Agyemang and Sahana, 2025). They rely on a transition probability matrix, calculated from historical land use maps, which assumes that future changes depend only on the current state. The primary advantages of Markov chain models is are they are simple, data-efficient, and effective for quantifying overall land use transitions and making short-term projections (Badshah et al., 2024). However, their limitations include ignoring spatial patterns and the underlying drivers of change and assuming constant transition probabilities over time. As a result, while useful for estimating how much change might occur, they are insufficient at predicting where it will happen due to their non-spatial nature (Dinda et al., 2019).

Artificial intelligence (AI) models are increasingly used to analyse and predict urban growth and land use change. These rely on data-driven learning to identify patterns and relationships in large datasets, without requiring explicit transition rules. Common AI techniques include artificial neural networks, random forests (RF), support vector machines (SVM), and convolutional neural networks (CNN) (Badshah et al., 2024; Guo and Liu, 2024). Primary advantages of AI techniques are they can capture complex and non-linear dynamics making them well-suited for modelling urban systems and they outperform the previously mentioned modelling techniques, especially when given large and rich datasets (Wang et al., 2019; Mithun et al., 2023). However, they are a black box approach as it is often difficult to understand how predictions are made which makes them problematic in public planning contexts where accountability is critical (Agyemang and Sahana, 2025). AI techniques also require large volumes of high-quality, representative data making them challenging to implement in data-scarce regions. Models can inherit biases present in training data, potentially skewing simulation output (Tsagkis et al., 2023; Mithun et al., 2023).

As identified by Wang et al. (2024) in Chapter 1, CA and ABMs are current state-of-the-art behavioural, bottom-up techniques most relevant to this thesis. Recent reviews

by Li and Gong (2016), Wang et al. (2024), Santé et al. (2010), Tong and Feng (2020), Gaur and Singh (2023), Musa et al. (2017), Sharma (2019) and Agyemang and Sahana (2025) consistently highlight the applicability and popularity of CA and ABM methods for urban growth modelling.

2.2.2 Cellular automata

CA was conceptualised in the latter half of the 20th Century and is now an established paradigm in the field of urban growth modelling (Figure 2.1). CA is a versatile and intuitive rule-based method that uses local neighbourhood interactions to effectively capture complex spatial and temporal systems (Chakraborty et al., 2022). In CA, land cover is represented as a lattice of cells, with cell values typically being binary: values 0 and 1 indicating non-urban and urban land cover, respectively. To simulate urban expansion, individual cells change their state to urban based on user defined transition rules and the state of their neighbouring cells (Li et al., 2017). CA is a popular approach to simulating large-scale urban expansion, but this approach does have limitations. For example, CA cannot capture multidimensional processes of urban change, such as urban densification, infill development, and urban regeneration (Tong and Feng, 2020). Additionally, CA does not incorporate human decision behaviours into the modelling process and are difficult to implement top-down drivers of growth, such as accessibility to employment, to the bottom-up rules which also drive CA processes (Liu et al., 2021).

Figure 2.1 shows the intellectual developments of CA over the last 50 years. Li and Gong (2016) categorised CA model into two branches: probability-based and rule-based models. Probability-based CA models are characterised by cell conversion being primarily determined by computing a synthesised probability. Rule-based CA are models where urban growth is represented with explicit deterministic rules (Wang et al., 2024; Silva and Clarke, 2002).

In probability-based CA, constrained CA is representative of this branch where constraints on urban expansion can vary in form, including growth types, development intensity, or dynamic constraints in both space and time (Li and Gong, 2016). Integrated-CA models involve coupling CA with an exogenous model to quantify urban demands at different times. These coupled models are expected to provide a top-down constraint on urban demand, directly influencing the number of cells converted within a specific period. For example, CA has previously been coupled with socioeconomic, environmental, population growth, dynamic system, and Markov models (e.g. Li et al., 2021; Yaagoubi et al., 2024).

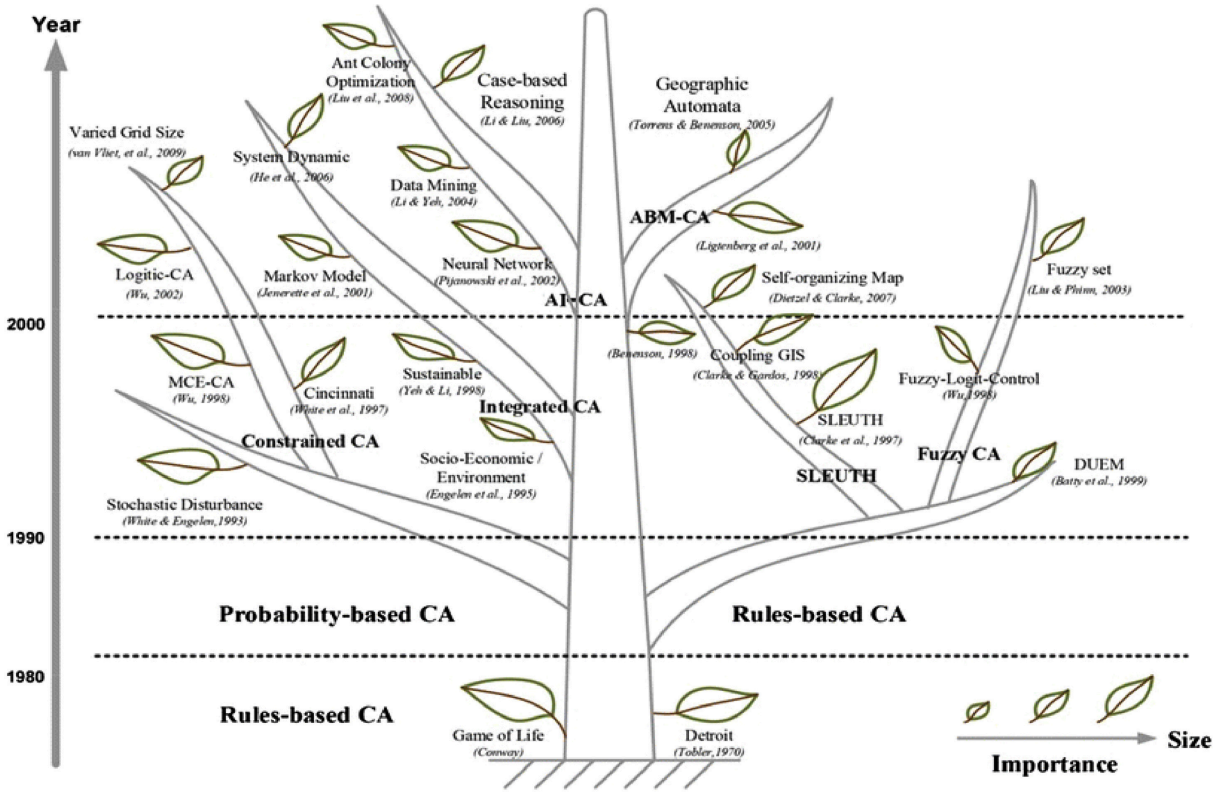

Figure 2.1: Evolution of urban growth models over the past 50 years (Li and Gong, 2016).

CA models have also used AI to establish transition rules based on techniques such as SVM, RF, and neural networks. These can result in better model performance than conventional statistical methods (Li and Gong, 2016). Kamusoko and Gamba (2015) used a RF model to generate a map of the urban change transition potential for their CA model. The good model performance was attributed to how RF could establish non-linear relationships between land cover change and the driving factors used which resulted in accurate transition rules. Xing et al. (2020) developed a CA model of land use change by integrating CNNs and long short-term memory (LSTM). CNNs were used to capture latent spatial features, providing a complete representation of neighbourhood effects and spatial heterogeneity. LSTM was employed to extract temporal dependencies from historical land use maps, effectively modelling the long-term nature of land use change, which traditional CA-Markov models failed to capture. LSTM generated transition probabilities for each cell by incorporating its temporal history, spatial context, and driving factors (Xing et al., 2020). However, Li and Gong (2016) noted these AI approaches can function as a black box and do not provide clear causal relationships.

One of representative example of rule-based CA urban growth models is the Slope, Land use, Excluded areas, Urban extent, Transportation, and Hillshade (SLEUTH) model developed by Clarke et al. (1997) (Figure 2.1). SLEUTH uses raster-based GIS data to predict how cities expand and how land use evolves under different conditions. The model operates using five key growth coefficients: diffusion, breed, spread, slope, and road gravity. These control how urbanisation spreads based on factors such as

terrain and infrastructure (Clarke et al., 1997). SLEUTH enables researchers and policymakers to test different urban planning scenarios, assess environmental impacts, and make informed decisions about future land use. Its flexibility allows it to be applied across diverse geographic regions, making it a valuable tool in sustainable development and policy making (Silva and Clarke, 2002).

Also under rule-based CA are Fuzzy CA models which view urban expansion as a continuous process in both space and time. The fuzzy inference process can be simplified through conditional functions and thresholds, and fuzzification has been applied to different components within CA models, such as input variables or rule definitions (Cattaneo et al., 1997). Li and Gong (2016) also identified that CA has been integrated into ABMs, where agents' behaviours and decision-making processes are governed by explicit and adjustable rules. The integration of ABM and CA has been widely explored in land use change studies recently, where CA-driven environmental dynamics, such as local neighbourhood effects, serve as constraints that shape agent behaviours and decision-making processes (e.g. Mozaffaree Pour and Oja, 2021; Mustafa et al., 2017).

2.2.3 Agent-based models

ABMs are a sophisticated technique for simulating dynamic spatial processes by modelling self-driven activities and interactions among rule-based agents, each responsible for changes within a system (Xu et al., 2020). This approach enables the modelling of land market processes, such as land competition and the preferences of actors such as households and real estate developers (Agyemang et al., 2023). ABMs simulate urban development by representing the behaviours and choices of autonomous agents, such as residents, planners, landowners, and governments. Agent behaviours are determined by predefined rules and interactions with other agents and the environment. For example, a model might include rules about a developer's willingness to purchase land or a resident's price threshold for living in a specific location (Koch et al., 2019). By explicitly modelling the feedback between people and their environment, ABMs provide a powerful tool to simulate realistic urban change (Korah et al., 2024).

ABMs have several advantages over CA. ABMs are more dynamic and flexible models that provide numerous options for developing sophisticated and realistic simulations of urban growth (Xu et al., 2020). For example, ABMs can capture land value and pricing dynamics from the perspectives of different agents, something that CA cannot do (Tang and Yang, 2020). They can incorporate individual-level decision making, enable agent interaction, have heterogeneous representation, capture non-linear and emergent behaviour, incorporate adaptive behaviour in complex systems. Previous urban growth modelling studies have developed multi-scale ABMs that have simulated urban land cover change by modelling the interaction of macro and microlevel agents (e.g.

Saeedi, 2018; Zhang, 2016). Macro agents would represent the government, whereas micro agents would represent resident behaviour. Hybrid CA-ABM models have also been developed as previously mentioned in Section 2.2.2 (Figure 2.1). Although there are deterministic CA and ABMs of urban growth, stochastic modelling approaches are often considered superior to capture the heterogeneous and unpredictable nature of urban development (Guzman et al., 2020; Brown et al., 2005).

As described in Section 2.2.2, there has been recent advances in the use of AI in defining CA transition rules. These techniques have also been applied to ABMs to model agent dynamics more effectively, with the volume of research combining ABMs with AI haven grown rapidly in recent years (Guo and Liu, 2024). AI techniques have proven to be a useful technique for addressing ABM's rigid rule-based approach by applying them to define agent behaviour and learning. For example, Bayesian learning has been used in ABMs to simulate adaptive decision-making processes, where agents update their strategies based on feedback from other agents, enabling more realistic and efficient modelling of urban negotiation and planning dynamics (Pooyandeh and Marceau, 2014). ANNs, decision trees, and RF, can learn from historical data to derive complex behaviour and define choices for agents. This includes applications in urban land use planning and residential land growth simulation, where agents' actions, attitudes, or probabilities of action are informed by historical behaviours or changes in the simulation (Xu et al., 2020; Mithun et al., 2023). Large Language Models are a recent cutting-edge advancement in urban modelling which allows agents to exhibit highly human-like social behaviours such as thinking, conversing, forming opinions, remembering, and reflecting, which has led to the emergence of sociality within urban growth simulations (Gao et al., 2024; Guo and Liu, 2024).

2.2.4 Current state-of-the-art urban growth models

Agyemang et al. (2023) developed a stochastic CA-ABM model of urban residential growth in Ghana that included household, developer, and government agents. The model was calibrated and validated between 2000-2010 and simulated future growth up to 2030. The paper used location choice factors, such as proximity to Central Business Districts (CBDs) and geographical characteristics of land parcels. It also assessed the sensitivity of a development control parameter which determined the influence of zoning laws on the government agent's decision-making process. This parameter was varied three times to see how weak, average and strong development control impacted model outputs.

Alternatively, Lin et al. (2023) developed a stochastic CA model of urban land use change in Guangzhou, China. The study performed calibration from 2005-2010, where the agreement between cells and landscape similarity was maximised as a measure

of model performance followed by validation from 2010-2015. The paper used driving factors, such as distance from roads, train stations, CBDs, and percentage of urban area in a Moore neighbourhood. It used the figure of merit (FoM) to assess accuracy. Lin et al. (2023) noted that stochastic processes in models introduce significant biases and uncertainties in the output.

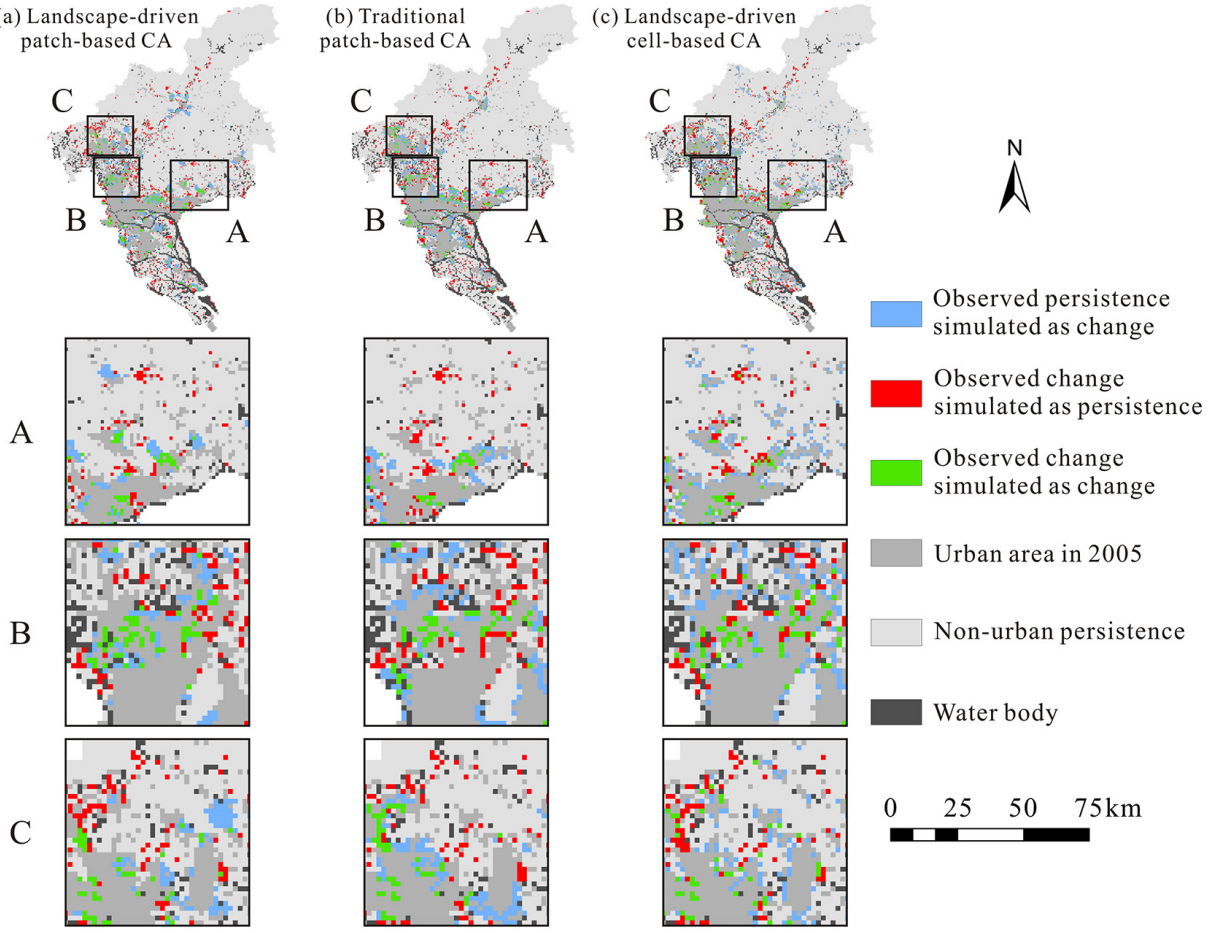

Figure 2.2: Calibration (2005-2010) outputs for three CA models assessed by Lin et al. (2023).

Li et al. (2020a) created a learning-based ABM in Shenzhen, China, with a resident and policymaker agent and performed calibration, validation, and sensitivity analysis between 2005-2015. The sensitivity of model parameters in the learning submodel was tested with 25 different parameter sets, run 10 times to account for stochasticity and acquire an average total accuracy. The simulation accuracy did not show notable variations when the values of the learning models' parameters were altered. The weights used in the calculation of the utility function of the resident agent were also tested with 9 different sets and the average simulation accuracy showed a variation less than 0.5%. This demonstrated that the model developed by Li et al. (2020a) was insensitive to variation in weighting values. The model simulated future growth for 2025 and 2035 (Figure 2.2).

Similarly, a multilevel CA-based model was developed by (Meentemeyer et al., 2013) to simulate the urban landscape structure in Charlotte, North Carolina. The study analysed the average output from 50 stochastic outputs and iteratively adjusted a patch compactness parameter until the patch shape distribution accuracy was maximised. It also included a joint calibration and validation analysis from 1996-2006. The model used per capita demand to estimate the rate of land conversion based on historical changes. Driving factors also included proximity to forests and farmlands, open spaces, water bodies, roads, and CBDs.

Koch et al. (2019) expanded Meentemeyer et al. (2013) CA model into a hybrid CA-ABM model which examined the impacts of landowner and developer interactions on development patterns. The study assessed the sensitivity of the peer influence parameter on the model by experimenting with five different values. Peer influence refers to the impact that neighbouring agent's decisions and characteristics have on a focal agent's willingness to sell their land. The study found that higher levels of peer influence lead to a lower overall willingness to sell among landowners, resulting in fewer parcels available for development (Meentemeyer et al., 2013).

Zhai et al. (2020) developed a vector-based CNN CA model to enhance the realism of simulating urban land use change in Shenzhen, China. The study addressed the limitations of previous vector-based CA models in considering neighbourhood effects of driving factors on transition suitability. The paper used used 14 driving factors, such as distance from transport, density of points of interest (e.g. hospitals and train stations) and topography. It used the FoM to assess model accuracy, calibrating the model between 2009-2014 and simulating future urban growth from 2014-2025.

Dahal and Chow (2014) presented an agent-integrated irregular automata stochastic model designed to simulate urban land use dynamics in San Marcos, Texas, using land parcels as its spatial unit. Agents in the model included a city planner agent, land developer agents for the supply side of development, and land consumer agents for the demand side of development such as residential agents and retailer agents. The study found that incorporating commercial, industrial, and institutional agents, alongside residential and planning agents, and using social theories captured urban growth processes well, as reflected by the superior model performance compared to the baseline model. It employed 16 driving factors, such as distance to rivers, population density, and the density of existing developments within a parcel's neighbourhood. The model was validated between 2000-2010 and then two future simulation scenarios were generated between 2010-2020, a baseline scenario and a compact development scenario which focussed growth around the city centre.

Yao et al. (2024) investigated the applicability and spatial scale sensitivity of the Vector-based CA model for land cover change simulation, comparing its performance with two raster-based models. The study was conducted in Shenzhen, China. It used 14 driving factors, such as distance to road networks and amenities. Future simulation was not performed and calibration and validation occurred between 2008-2018, using FoM and the Kappa coefficient to assess model performance. The land parcel model demonstrated excellent simulation results in rapidly expanding newly developed urban areas and was superior in simulating actual land cover changes compared to the two raster models, achieving the highest FoM.

Lu et al. (2020) similarly compared the performance of a vector-based land parcel CA model with a gridded raster CA model to determine whether vector CA models are more accurate for simulating land use changes in Ipswich, Australia. Driving factors included parcel area, population pressure, and distance to commercial services. Model accuracy was evaluated over the period 1999–2016, with the vector CA model demonstrating superior performance based on the kappa coefficient accuracy metric. Both CA models were run 30 times to account for their stochastic nature. This number of runs was determined by monitoring when the cumulative mean and standard deviation of the accuracy metric had stabilised.

2.2.5 Modelling human behaviour

In ABMs, a key challenge is accurately simulating human behaviour within the agents. One approach is to use surveys to define agent behaviour. For example, Koch et al. (2019) parameterised land owner agents based on a survey conducted among wood owners in North Carolina, identifying subcategories such as economic maximisers, whose primary motivation is selling their land. Similarly, Ku (2024) developed an ABM land use model to capture stakeholder interactions during urban development in Taipei City, Taiwan. Household surveys and city planner interviews were used to define behavioural rules for households and planner agents, providing information on the various attitudes of different types of households, such as single-person households and families. Additionally, interview surveys helped reveal how local authorities influence land use change processes and the criteria they consider in decision-making (Ku, 2024). However, these survey-based approaches can be laborious, time-consuming, and can still result in incomplete knowledge. For example, the survey in Koch et al. (2019) focused solely on woodland landowners but applied its findings to owners for all land cover types. Furthermore, due to the lack of empirical data on the spatial distribution of landowner types in the study area, they were randomly assigned throughout the landscape.

Another more practical and commonly used approach in ABMs is to use behavioural theories to capture decision-making (Groeneveld et al., 2017; Crooks et al., 2018). Bid-rent theory, game theory, theory of planned behaviour rational choice and bounded rationality have all been used in modelling studies. The bid-rent theory was explains how land values vary based on distance from CBDs. It posits that agents compete for locations near the city centre, where proximity to markets, employment opportunities, services, and transportation provides the greatest benefits. The bid-rent theory assumes that cities have a single central point of economic activity, that transportation costs increase with distance, and that land users make rational decisions to maximise utility while minimising costs (Alonso, 1964). Game theory provides a mathematical framework for analysing strategic interactions between rational decision makers. Game theory assumes that individuals or groups, referred to as players make decisions with the goal of maximising their own payoffs while anticipating the actions of others. The theory classifies games into different types, including zero-sum games, where one player's gain is another's loss, and non-zero-sum games, where cooperation can lead to mutually beneficial results (von Neumann and Morgenstern, 1944).

The theory of planned behaviour explains human behaviour through the role of intentions, which are influenced by attitudes, subjective norms, and perceived behavioural control. An individual's intention to engage in a behaviour is shaped by their personal attitude toward that behaviour, the perceived social pressure from others, and their belief in their own ability to perform the behaviour. The stronger these three factors, the more likely an individual is to follow through with a given action (Ajzen, 1991). Rational choice theory suggests that people make decisions systematically weighing the costs and benefits of each available option to maximise their personal advantage. It assumes that people act rationally, have stable preferences, and have complete information about their choices (Monroe, 2001). Bounded rationality is a critique of rational choice theory. It argues that people make rational decisions, but their cognitive limitations, time constraints, and imperfect access to information prevent them from achieving true optimisation. Instead of seeking the best possible outcome, people engage in satisficing where they settle for a solution that is sufficient rather than optimal. It provides a more realistic perspective on human behaviour compared to rational choice theory, highlighting the importance of cognitive processes in decision-making (Simon, 1956).

2.2.6 Urban growth modelling limitations

Although the highlighted urban modelling literature does employ some form of model evaluation, there are large discrepancies in this process among authors and it is typically executed in an unsystematic way. For example, the stochastic models that performed numerous runs to account for their stochastic process did not determine whether the output had stabilised or if further runs were needed. Instead, an arbitrary num-

ber, such as 100 Monte Carlo runs in Tang and Yang (2020), was defined from the modeller's intuition or previous studies and a stabilisation point was not identified (Wu, 2002). Pontius Jr et al. (2004) identified that land use models require a clear distinction between calibration and validation procedures to avoid confusion and minimise model overfitting. Although some urban growth models now adopt this approach of separating calibration and validation either temporally or spatially, there is still a presence of failure to separate both procedures or adopt only one procedure (Wang et al., 2019; Yang et al., 2023a). van Vliet et al. (2016) reviewed models between 2010-2014 and identified that 18% did not report calibration and 31% did not report validation out of the the 307 publications reviewed. In addition, sensitivity analysis is rudimentary, with studies often varying only one parameter to evaluate its impact on outputs, not considering interaction between model parameters and not using statistical techniques to better understand the model. van Vliet et al. (2016) identified that only 12% of models performed a sensitivity analysis and recent publications have also omitted this entirely from their analysis (e.g. Kafy et al., 2020; Mustafa et al., 2017; Kisamba and Li, 2023). Sensitivity analysis can improve model understanding, robustness, and credibility, which makes it a useful procedure to perform (van Vliet et al., 2016). This review highlights the need for urban growth models to use more comprehensive, systematic, and robust approaches to assess the model and its outputs.

Despite the fact that CA and ABMs have strong scenario potential, few recent studies have fully implemented them. Only a minority of urban growth ABMs simulate multiple urban development scenarios. In particular, no scenarios were simulated by Xu et al. (2020); Mozaffaree Pour and Oja (2021); Lin et al. (2023); Li et al. (2020a), despite these studies claiming that their models could effectively investigate future urban planning scenarios. The papers that created different future scenarios conflated them with basic sensitivity analyses. For example, Tang and Yang (2020) varied the annual demand by 20%, 40%, 140%, and 160% of the baseline value, where the first two percentages represent slow land use development and the last two percentages indicate rapid development. Similarly, Koch et al. (2019) adjusted peer influence at five levels, from zero to total. However, these what-if scenarios were comparable to colloquial sensitivity analyses rather than true scenario-based modelling. There remains a significant opportunity to develop more sophisticated scenarios that go beyond isolated parameter adjustments, capturing comprehensive and realistic urban planning situations. Van Vuuren et al. (2014) put forward the idea of a scenario matrix to include as many possible alternatives for future conditions via intersecting conditions of the different factors in order to derive multifaceted future estimations. An example of a scenario matrix is shown in Figure 2.3. The x-axis is the socioeconomic pathway that varies conditions, such as population growth and governance efficiency. On the y-axis are climatic (forcing) trends such as greenhouse gas emissions and temperature changes.

To generate more comprehensive forecasts, it is essential to adopt a scenario matrix approach that accounts for various socioeconomic realities rather than relying on adjustments to a single model parameter. By considering what axes are feasible based on the model developed, researchers can build a more nuanced and realistic representation of future urban dynamics (Van Vuuren et al., 2014).

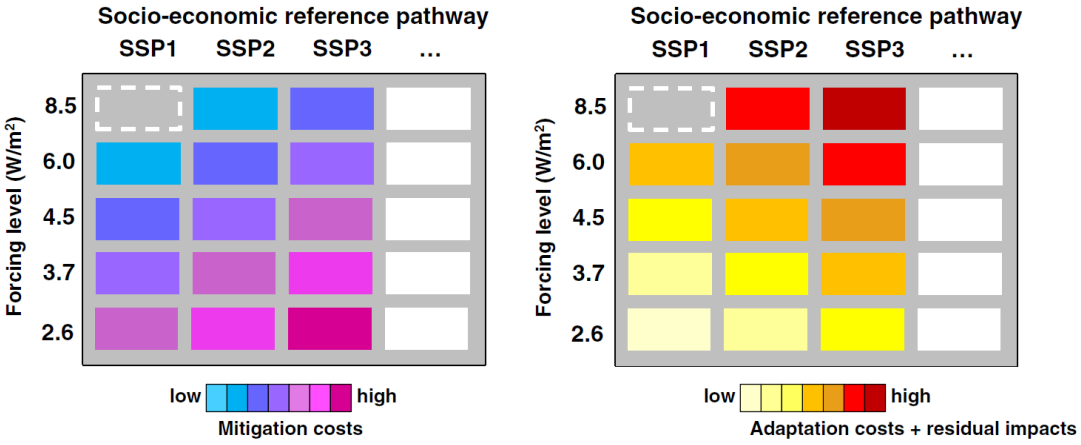

Figure 2.3: Scenario matrix developed by Van Vuuren et al. (2014).

Furthermore, the aforementioned urban growth models have mainly focused on raster-based and regional simulations (Figure 2.2) of urban growth (e.g. Li et al., 2020a; Agyemang et al., 2023). Raster datasets used in these urban modelling studies have gridded formats, lack granularity, and contain coarse pixel resolutions ranging from 30 m^2 to 1 ha in size (e.g. Kaviari et al., 2019). This results in raster urban growth models being inadequate and uninterpretable at a more detailed spatial scale when localised urban growth patterns need to be examined. Vector data is superior at capturing the shape of irregular land parcel data at a high resolution, especially when considering that land parcels are the unit of urban planning and development, which ensures greater applicability of models in real-world settings (Chakraborty et al., 2022). Dahal and Chow (2014) also argued that they eliminate cell size sensitivity.

Although there is a presence of vector-based urban growth models ((e.g. Zhai et al., 2020; Lu et al., 2020; Dahal and Chow, 2014; Yao et al., 2024)), there is a paucity of applying these models to rapidly urbanising regions. Their applications face significant challenges, as delineated land parcel data and temporal information on the state of the parcel are usually unrecorded and unavailable, especially in developing regions (Lin et al., 2023). Mozaffaree Pour and Oja (2021) noted a lack of readily available spatiotemporal land parcel data for urban growth models, which means that modellers have to work with gridded cells instead, limiting the reality of the model. The parcel-based models reviewed in Section 2.2.4 mainly focused on developed countries, such as Australia and the USA, with a strong emphasis on using CA, with only one identified paper employing an ABM approach (Dahal and Chow, 2014).

2.3 Urban form modelling

Urban form modelling refers to the computational generation of city structures, including road networks and building footprints, using algorithmic approaches. Traditional procedural techniques are based on predefined rules, simulations, and optimisation methods. Deep generative models (DGM) are a class of machine learning algorithms that use deep neural networks to model complex, high-dimensional probability distributions from training data. These models learn the underlying distribution of the data and can generate new similar data samples. By understanding and capturing patterns within the data, DGMs can synthesise realistic content, making them useful for generating urban form layouts(Lin et al., 2024). They include variational autoencoders (VAEs), generative adversarial networks (GANs), and diffusion models, which offer a more automated and data-driven approach to urban form generation, requiring less manual input and domain expertise.

2.3.1 Procedural generation

Procedural generation refers to the creation of data using algorithms and computational methods rather than manual design (Beneš et al., 2014). It is a popular technique due to its efficiency and flexibility in generating diverse configuration of synthetic urban form. Procedural road network and building footprint generation can be achieved using several methods: pattern-based approaches that use predefined patterns to create road layouts, L-systems that use a set of rules to generate complex structures, agent simulation that involves simulating the actions of individual agents to generate road networks, and simulated annealing that uses a probabilistic optimisation technique inspired by the annealing process in metallurgy (Galin et al., 2010).

Song and Whitehead (2019) developed TownSim, a novel agent-based city simulator that generates road networks. The model simulated city evolution conceptually, driven by agents engaging in a simplified consume-work-trade-rest cycle. The generated outputs had applications in urban planning, game development and autonomous vehicle testing. Raimbault et al. (2016) presented a hybrid CA-ABM model to simulate urban morphogenesis. The model used a lattice of cells to represent buildings with their status being empty or occupied and new buildings were built overtime in empty cells in based on the land value (Figure 2.4). The model also had a dynamic network topology to represent roads over the lattice where vertices represent city centres, and edges represent roads. The model incorporates heterogeneous agent urban activities and uses various metrics—including density, distance to roads, accessibility to activities, and a Schelling segregation model—to analyse and optimise generated urban patterns.

Brasebin et al. (2018) used a simulated annealing algorithm to generate building footprint configurations within parcels that comply to regulations. It assesses the impact of local urban building regulations on 3D building constructibility. The resulting open-source SimPLU3D project offers a tool to assist both urban planners in designing regulations and citizens in understanding their implications. The paper demonstrates the method's effectiveness through various simulations and explores its potential for assisting in urban planning and design. Whereas Taillandier et al. (2016) developed a vector-based ABM to simulate building footprint densification within urban blocks, defined by roads and rivers. New building footprints were placed within these blocks, based on the shapes of existing buildings, available area of the block, and constructability scores based on density, proximity to transportation, and access to services.

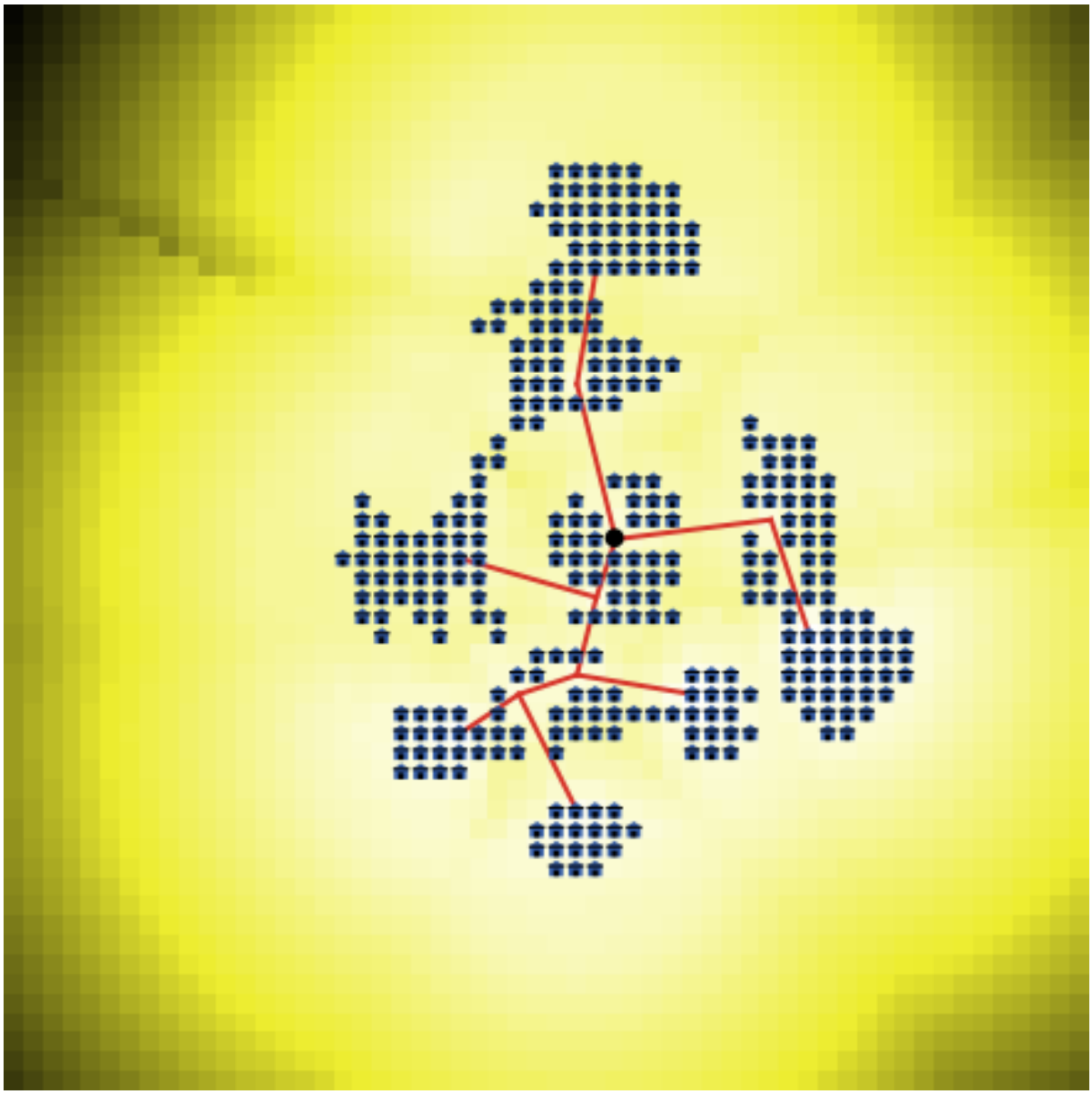

Figure 2.4: Urban form generated by Raimbault et al. (2016). Houses (blue squares) and roads (red edges). Brighter yellow cells represent closer distances to built cells.

In these procedural examples, realism is limited with generated networks that do not always conform to real-world road design, and some unusual real-world patterns may be outside of the capabilities of rule-based generation. Road network generation in the examples was purely conceptual, without considering the expansion of the road network from existing road configurations. Building footprint approaches had limitations such as rule formulation and they were unable to capture all nuances of regulations. In addition, regulations are often time-consuming to find and apply to a computational environment with lots of interaction with relevant government bodies being required. Finally, the simulation results are sensitive to the chosen parameters and the model's validation is limited to a small number of case studies. Procedural models could be more automated. DGMs for urban form generation offer this automation as they require less user input and domain expertise compared to these procedural techniques.

2.3.2 Deep learning generative models

2.3.2.1 Variational autoencoders

VAEs are a DGM introduced by Kingma (2013) used for unsupervised learning of complex data distributions. They consist of an encoder and decoder architecture. The encoder compresses the data into a lower-dimensional latent representation, which captures the essential features of the input data. The decoder takes this latent representation and reconstructs the input data. The VAE training process aims to minimise the reconstruction loss between the reconstructed and original data while ensuring that the latent representations are diverse and distributed approximately according to a normal distribution. A regulariser also encourages the encoder to create latent representations that follow a standard normal distribution. This ensures that the latent representations are diverse and that new data can be generated by sampling from this distribution. A trained VAE can be used to generate new data similar to the training data by sampling from the learnt latent distribution and passing it through the decoder (Kingma, 2013).

Kempinska and Murcio (2019) trained a VAE on 12,479 street network images from OpenStreetMap data. The images were centred within a 3 km^2 of world cities, had a pixel resolution of 64x64 and were binary with roads as 1 and no roads as 0. During VAE training, the images were augmented to increase the size and diversity of training data by randomly cropping and flipping images. When the VAE was applied to the test data, the reconstruction quality was unsatisfactory, the VAE being incapable of generating a dense road network. New and synthetic road network images were then generated by randomly sampling the learnt latent distribution and passing this through the decoder. Synthetic road networks generated from the learnt latent distribution worked well for sparse network images, but produced blurred results that lacked detail for dense road network images (Figure 2.5). Kempinska and Murcio (2019) attributed this to the low

image resolution and acknowledged the need for quantitative methods to assess the generated network quality.

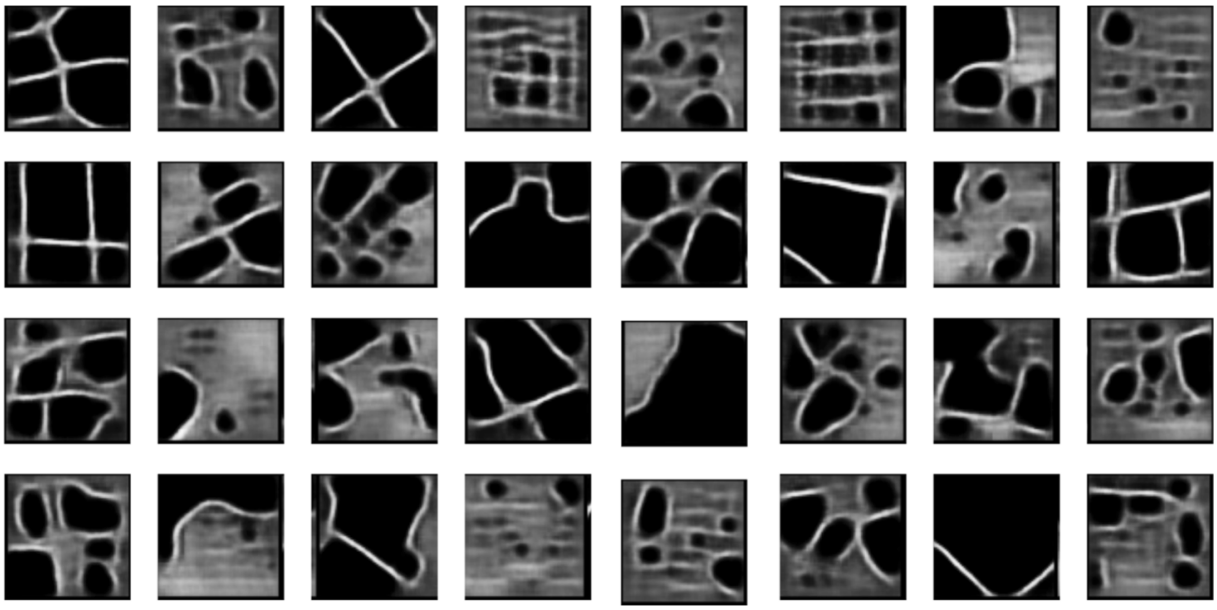

Figure 2.5: VAE generated synthetic road networks by passing a randomly sampled learned latent distribution through the decoder (Kempinska and Murcio, 2019).

Birsak et al. (2022) designed a VAE for generating road networks guided by conditioning data, such as road density and land-water boundaries. The input to the VAE was a road network image which was processed and compressed into the encoder which learns latent representations from it. These conditioning maps are used to guide the generation of the street network in the decoder. The decoder output was post-processed by thinning and tracing the rasterised road networks and converting the images to vector, ready for analysis and use in urban planning contexts. It was noted that future work needs to be done to develop more robust and informative evaluation frameworks, potentially involving studies of human perception to validate generated road networks (Birsak et al., 2022). Additionally, VAEs needs to have an explicit road network input to generate networks and using the latent distribution produces unsatisfactory results.

2.3.2.2 Generative adversarial networks

GANs were introduced by Goodfellow et al. (2014) as a framework for data generation. They function as a two-player minimax game with two neural networks: a generator and a discriminator, trained simultaneously in an adversarial setup. The generator learns to produce plausible data from random noise sampled from a simple distribution, which serves as fake examples for the discriminator. The discriminator learns to distinguish between real data from the training set and fake data from the generator. During training, the generator aims to maximise the probability that the discriminator misclassifies its fake data as real. In contrast, the discriminator aims to minimise this probability, improving its ability to differentiate real and fake data. This adversarial process drives

both models to improve iteratively, resulting in generated data that become increasingly indistinguishable from real data. GANs present several advantages over VAEs, including gradient-based optimisation through backpropagation, greater modelling flexibility, and the ability to represent sharp distributions. Since their inception, GANs have rapidly advanced and shown robust performance in generating realistic images across diverse datasets. Conditional GANs generate images conditioned on inputs, enabling context-specific data generation. This makes conditional GANs ideal for tasks like image-to-image translation, such as converting daytime images to nighttime or satellite imagery to basemaps (Figure 2.6).

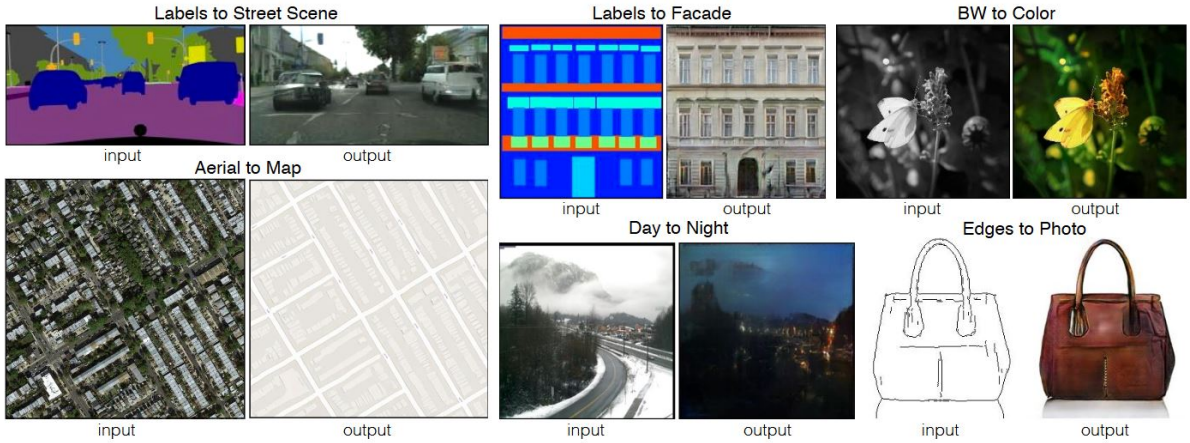

Figure 2.6: Examples of paired images applicable to image-to-image GANs (Isola et al., 2017).

Hartmann et al. (2017) developed StreetGAN to generate synthetic road networks. The StreetGAN workflow involved converting OpenStreetMap (OSM) road data into a binary image where white pixels indicate roads and black pixels indicate the absence of roads. StreetGAN is trained on these binary images, and once trained, uniformly sampled spatial noise seeds the generator to create novel and synthetic road networks. This use of spatial noise ensures the generation of diverse road networks, even from a limited dataset of real-world examples. StreetGAN enables the generation of arbitrary-sized road networks that replicate the style and structural properties of the training data (Hartmann et al., 2017). However, a key limitation of this work is that the networks are generated independently and cannot automatically extend or adapt to existing street networks.

Fang et al. (2020) and Fang et al. (2022) developed a conditional GAN based approach to interpolate and generate road networks in real cities to extend the work by Hartmann et al. (2017). DeepStreet used image inpainting, CNNs, and GANs to interpolate road patterns in a defined region based on surrounding road network patterns and elevation (Fang et al., 2020). DeepStreet effectively replicated various complex patterns, demonstrating strong potential as a novel urban design tool. Fang et al. (2022) enhanced their DeepStreet model by incorporating more topography data (e.g. slope and aspect) and

road network data (e.g. intersection nodes) to generate road networks (Figure 2.7). It introduced a geo-extractor module which dynamically weighted the topographic information based on their importance for specific locations. It also used a geo-merging bypass module to align the network with the terrain. This resulted in more realistic and plausible road network generation. However, both road network generation models were better suited for interpolation rather than road expansion. Their predictive performance decreased as the size of the masked region increased, relying heavily on a dense surrounding road network for accurate generation (Fang et al., 2020, 2022).

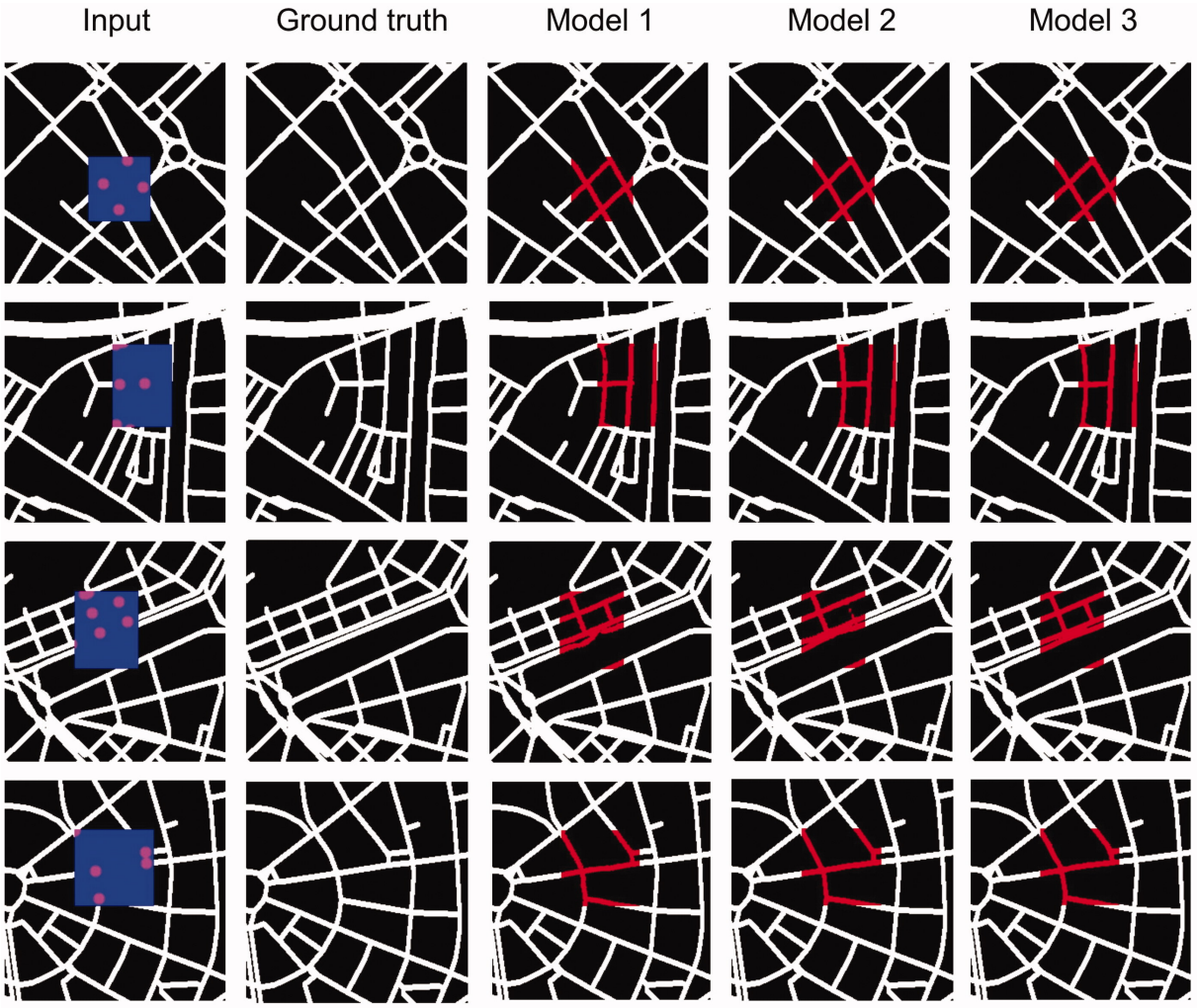

Figure 2.7: GAN road network interpolation and generation in an urban area (Fang et al., 2022).

GANs have also been used to generate building footprints based on road layouts, with two of the most well-known examples being GANmapper and InstantCITY (Wu and Biljecki, 2022, 2023). Both took an input image of the road layout and had a corresponding target image of the road layout with building footprints. GANmapper operated with images at 256x256 pixel resolution and performed numerous experiments, such as assessing the impact of GAN performance on spatial scales (Figure 2.8) (Wu and Biljecki, 2022). For the experiments, the image datasets were divided into 80% for training and 20% for testing. The quantity of training images at the zoom levels

in Singapore ranged from 150 to 2825 images, depending on the zoom level. Wu and Biljecki (2022) identified that denser road networks returned better results due to more contextual information. However, the low resolution in GANmapper led to irregular footprint shapes and individual buildings in close proximity being merged, reducing accuracy and detail. Additionally, only raster data was produced without converting it back to georeferenced vector data for further analysis. Wu and Biljecki (2023) developed InstantCITY to generate high-resolution vector building data from road layouts. It introduced pre-processing which converted a target city into image tiles at 1024x1024 pixel resolution and post-processing which stitched generated outputs together and converted and tidied the footprints to vector data. It conducted experiments in 8 cities at 500 m^2 and 1000 m^2 to assess GAN performance. The training images ranged from 91 to 781 images, depending on the city and scale. InstantCITY generated synthetic building footprints of much higher quality than GANmapper. Wu and Biljecki (2023) did note that the training data should be compromised of mainly complete tiles and that the under-mapped tiles should be removed.

Figure 2.8: GANmapper building footprint outputs at four different zoom levels (Wu and Biljecki, 2022).

2.3.2.3 Diffusion models

Diffusion models provide a competitive alternative to GANs for generating high-quality data. They offer a probabilistic modelling approach that balances flexibility and tractability, as introduced by Sohl-Dickstein et al. (2015). Diffusion models use a Markov chain to transform a complex data distribution into a simple, tractable distribution, such as a

Gaussian, through a forward diffusion process. This process introduces small amounts of noise iteratively until the original data distribution becomes indistinguishable from random noise. The generative model then learns a reverse diffusion process (denoising), starting from the simple distribution and progressively removing the noise to generate data resembling the original distribution. This reverse process is modelled by estimating the small perturbations (noise) added in each forward step, which makes the problem computationally tractable compared to directly modelling the full data distribution. A significant strength of diffusion models lies in their generality: any smooth target distribution can be represented using an appropriate diffusion process. This capability enables them to capture complex, multimodal data distributions, making them highly versatile. Diffusion models allow for precise sampling from the learnt model, providing high-fidelity outputs. They support operations like combining with other distributions, which is valuable for tasks such as denoising, inpainting, and conditional generation. The model's logarithmic likelihood and the probabilities of intermediate states are efficiently to compute, facilitating rigorous evaluation of the model.

Gu et al. (2024) used a conditional diffusion model to generate road networks using land use, elevation, slope, and road intersection rasters. OSM road data were converted to 1 km^2 patches at 128x128 resolution and the conditional factors were encoded as separate image channels. Once trained, the model can produce new road networks for a given area guided by the provided geospatial raster layers. The generated road network image is binarised, then road pixels are thinned to eliminate artefacts, and the image is converted to vector to form the final road network. One major limitation of this research was that image generation was restricted to a low pixel resolution due to the U-net architecture's numerous self-attention layers which make the model computationally demanding (Gu et al., 2024). A multimodal diffusion model, called ControlCity, was designed by Zhou et al. (2024) to generate building footprints based on road networks, land use, and text prompts. It builds on GANmapper and InstantCity by utilising text captions from Wikipedia's GeoSearch using GPT-4o mini to describe building types, such as residential or commercial. ControlCity was tested on data from 22 cities and processed city data into 1.2 km^2 cells with 1024x1024 resolution stored in a WMTS directory. This resulted in 3,140 image pairs for training and testing. Stable Diffusion XL by Podell et al. (2023) was used as a basis for ControlCity as the model extended Stable Diffusion's text-to-image capabilities and demonstrated state-of-the-art performance as a result of these enhancements.

Diffusion models outperform GANs and VAEs in generating high-quality, detailed samples with enhanced stability and fewer artefacts. Compared to VAEs, which can produce blurred outputs due to their reliance on approximate posterior distributions, diffusion models operate in latent space with fine-grained control over the generation

process (Sohl-Dickstein et al., 2015). Their scalability and adaptability make diffusion models particularly well-suited for high-resolution and diverse data generation, affirming their status as a leading choice in contemporary generative modelling. In contrast to GANs, which undergo adversarial training prone to mode collapse and unstable convergence, diffusion models use a denoising process to iteratively generate samples from Gaussian noise, enhancing robustness. This denoising process enables diffusion models to effectively capture complex data distributions, yielding higher fidelity and realism in outputs (Zhou et al., 2024).

2.3.3 Urban form generation limitations

There are many promising applications of DGMs for creating synthetic future urban form, with these models being a viable option for the aim of this research. However, the following limitations are present in existing literature when trying to generate urban form:

- Road network generation is restricted to small areas, with a maximum application seen by Gu et al. (2024) at a scale of 1.2 km^2.
- All road network generation attempts have required explicit information on the existing network, such as the actual network in Birsak et al. (2022) VAE or road intersections in Fang et al. (2020, 2022); Gu et al. (2024).
- For building footprint generation, the pixel resolution is often low, but some high-resolution examples at 1024x1024 are present (Wu and Biljecki, 2023).

To generate future road networks across a large scale, more ambiguous contextual information is necessary to guide the DGM. The main barrier to generating future urban form is the generation of a high-resolution and large-scale road network. Following this, building footprint generation should be more straightforward.

Of the DGMs explored for urban form generation, there is a strong focus and wide application of using GANs, showing they are a trusted and known entity for this task. Diffusion models may be a more advanced technique, but their record of urban form generation is less present. Due to the complexity in creating DGMs from scratch, this research requires a readily available and standardised solution that can be used for the task of urban form generation. GANs indicate a promising DGM for generating urban form for this research.

2.4 Model evaluation and assessment procedures

Beyond urban growth modelling, various protocols and frameworks have been developed to evaluate and explore ABMs, enhancing model understanding and confidence.

2.4.1 Variance stability and uncertainty analysis

Lee et al. (2015) noted that the number of model runs in ABMs is too low, high, or selectively chosen. Complex ABMs with extended runtimes require a minimum number of runs to adequately capture stochasticity, while simpler ABMs with short runtimes may lead to excessive runs and heightened sensitivity to statistical tests. An excessive number of runs can make trivial outcome differences statistically significant, potentially lacking practical significance (Lee et al., 2015). Variance stability quantifies the variability of a stochastic model using a metric to measure the uncertainty surrounding the variance. Law and Kelton (2007) based the number of runs when variability stays within a specified proportion of the confidence interval around the mean. However, this approach assumes normality of the model runs and that the mean has a Gaussian distribution. Lorscheid et al. (2012) avoided these assumptions and took a different approach to identify the appropriate number of Monte Carlo runs, using the coefficient of variation (CV) which is a statistical measure of the relative variability of the data and a fixed epsilon as a metric to measure the uncertainty surrounding the model output variance. The approach involved comparing the CV of increasing runs (e.g. 10, then 100, 500, 1000 runs etc.) and the run at which the difference between consecutive CVs falls below the epsilon value and remains so was considered the minimum number of runs until stabilised. Furthermore, analysis of ABMs often entails an exploration of the parameter space to understand the dependencies between key input parameters and their outcomes, including their variability. Given the complexity of ABMs, the variance of the outcomes may or may not be constant in the parameter space (Lee et al., 2015).

Uncertainty analysis explores the parameter space and involves Monte Carlo simulations, where the model is run multiple times with different sets of randomly sampled input parameters drawn from a uniform distribution. This allows exploration and understanding of a wide range of potential outcomes that a model may produce due to variability in its inputs (Kang et al., 2022). Previously, Tang and Jia (2014) defined the total number of simulations required based on the number of parameters being explored and Monte Carlo simulations, typically between 100-500 for sophisticated spatial ABMs. Parameter exploration is performed systematically, by sampling across the entire space using a variety of techniques. Random sampling is an efficient way to explore the parameter space, as it uniformly draws values within the parameter range. However, it can result in uneven coverage, leaving gaps or clusters in certain areas. Factorial sampling explores all parameter value configurations sequentially within the space and provides complete coverage, but can become computationally intensive with numerous model parameters. Latin hypercube sampling (LHS) divides each parameter into equal intervals and selects one sample per interval. This helps achieve better coverage of the space compared to random sampling and is more efficient than factorial sampling when dealing with high-dimensional spaces (Lee et al., 2015).

2.4.2 Calibration and validation

There are numerous benefits in separating calibration and validation; it is useful to clarify what these terms mean and what they involve. Calibration is the process of adjusting model parameters to improve the model's relevance or fit to a specific case (van Vliet et al., 2016). Validation is a demonstration that the model has a satisfactory and credible accuracy consistent with the intended application and purpose (Houet et al., 2016).

2.4.2.1 Calibration techniques

There are numerous techniques to calibrate a model and find the optimal model parameters which result in the best performance. Some of the most prominent of these techniques include exhaustive exploration, genetic algorithm (GA), particle swarm optimisation (PSO), RF, and regression (Agyemang and Sahana, 2025).

Exhaustive exploration is a brute-force calibration strategy that systematically evaluates every possible combination of model parameters within a defined search space. The SLEUTH model performs exhaustive exploration when performing sequential multistage calibration, where all permutations of growth coefficients across multiple resolution stages are tested to identify the optimal configuration (Silva and Clarke, 2002). Exhaustive exploration guarantees the discovery of a global optimum within the parameter space defined by the modeller. However, it suffers from significant computational inefficiency, particularly as the dimensions of the parameter space increase. The approach becomes computationally unfeasible for complex models or when the number of parameters and their value ranges are large, rendering it impractical for many real-world applications, especially under constrained resources (Liu et al., 2019).

In contrast, GA is a popular machine learning technique for parameter optimisation. It employs an evolutionary approach to find the optimal solution to a search problem based on natural selection. Starting with an initial population of solutions, it iteratively applies evolution operators (crossover and mutation) until it identifies the optimal solution using a fitness function (García et al., 2013). The crossover operator exchanges genes between two parent solutions to create two offsprings. The offspring inherit traits from both parents, and the cutting point for gene exchange is randomly determined. The mutation operator alters one or more genes of a single parent by randomly flipping bits (0 to 1 or 1 to 0). Mutation prevents the algorithm from getting stuck at local optima or minima, ensuring continued evolution (Li and Yeh, 2005). In urban growth modelling, Mustafa et al. (2017) and Lin et al. (2023) used GA to identify optimal model parameters that would maximise landscape similarity and cell-by-cell agreement. GA is a robust heuristic method for calibrating urban expansion models, demonstrating greater computational efficiency in high-dimensional parameter spaces and improved accuracy in

complex and non-linear environment problems such as urban growth capturing urban growth compared to exhaustive exploration (Saxena and Jat, 2020). However, as it is a heuristic method it does not guarantee the identification of the absolute global optimum and it can be highly dependent on the tuning of its internal parameters which, if poorly tuned, may converge slowly or stagnate (García et al., 2013).

A similar technique to GA for parameter calibration is PSO - a population-based meta-heuristic algorithm inspired by the social dynamics of collective behaviour in biological systems. In PSO each particle represents a candidate set of parameters for the model (Freitas et al., 2020). It operates by iteratively updating the positions of particles based on individual and collective experiences within a parameter space. Each particle adjusts its search behaviour based on three parameters: inertia (past momentum), cognitive learning (its own best solution), and social learning (the best-known solution found by the swarm or a neighbourhood) (Freitas et al., 2020). PSO has proven effective for calibrating driving factor weights and parameter values in urban growth models (Lu and Laffan, 2018; Naghibi and Delavar, 2016). Like GA, it can efficiently explore complex, non-linear and high-dimensional parameter spaces. In addition, when using small swarm sizes it can often yield good solutions rapidly, reducing computational cost. However it suffers from similar limitations to GA, such as premature convergence and stagnation due to improper parameter values (Freitas et al., 2020).

RF is another widely used machine learning technique that supports urban model calibration through indirect parameter optimisation (Belgiu and Drăguț, 2016). It constructs multiple decision trees on bootstrapped subsets of the training data and aggregates their outputs to improve prediction accuracy and reduce overfitting. By randomly selecting features for node splits, RF introduces diversity among trees, enhancing robustness and reducing variance (Belgiu and Drăguț, 2016). Although RF is not a direct optimiser like GA or PSO, it is useful at identifying the most influential driving factors in a model through its feature importance measures which enables these factors to be prioritised, reducing the dimensionality of the parameter space and enabling more efficient optimisation (Rienow et al., 2021). RF's non-parametric nature, high predictive power, and ability to handle high-dimensional data make it a powerful tool for guiding parameter selection in urban growth modelling, especially in data-constrained setting. However, it can be computationally intense and requires significant processing times. In addition, it can only be implicitly used in parameter estimation in terms of selection, whereas GA and PSO, offer explicit optimisation of actual parameter values. RF is more commonly associated with defining non-parametric models elements such as the computation of transition potential maps (Kamusoko and Gamba, 2015).

Calibration using regression involves setting weights of driving factors, such as distance from roads, using techniques such as logistic regression (LR) or geographically weighted regression (GWR). Mustafa et al. (2017) mapped historic change as binary where cells with observed change were 1 and no-change were 0. Their driving factors were the predictor variables and 1000 cells without change and 1000 cells with change were randomly sampled. These sampled values were input into a logistic regression model and the output coefficient for each predictor variable was treated as a weight for each driving factor. Burke et al. (2024) advanced this approach by using GWR coefficients to define driving factor weights. GWR is a more nuanced method than LR because it captures local spatial patterns and variations. It generates local coefficients by fitting a separate regression equation for each location. This provides information on how the influence of the predictor variables varies geographically, providing a better fit to spatial data (Brunsdon et al., 1996). However, a limitation of statistical models for calibration is their applications to model driving factors and not parameters. If a driving factor is dynamic and has model feedback, such as distance from simulated development, then this cannot be calibrated with GWR or LR, as these techniques use only historical change and simulated development is a model-only output.

2.4.2.2 Calibration accuracy metrics

Urban growth models use various metrics to evaluate their performance. Cell-based metrics involve classifying cells into four types: true positive - the model correctly predicts an urban cell, true negative - the model correctly predicts a non-urban cell, false positive - the model incorrectly predicts an urban cell where it should be non-urban and false negative - the model incorrectly predicts a non-urban cell where it should be urban (Table 2.1).

Table 2.1: A confusion matrix used to evaluate the cell-based performance of urban growth models, comparing simulated cells against observed cells.

	Simulated urban	Simulated non-urban
Observed urban	True Positive (TP)	False Negative (FN)
Observed non-urban	False Positive (FP)	True Negative (TN)

Using the confusion matrix from Table 2.1, the percent correct match (PCM) and the Figure of Merit (FoM) can be calculated using the following equations, with higher values indicating better agreement between the simulated and observed results from these equations:

$$\text{PCM} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.1}$$

$$\text{FoM} = \frac{TP}{TP + TN + FP + FN} \tag{2.2}$$

PCM measures the proportion of correctly simulated change and persistence relative to the total number of cells. In contrast, FoM focusses more on change, evaluating the precision of the correctly simulated change against all cells where change or persistence was simulated or observed (Kaviari et al., 2019). However, PCM values may be inflated in areas with minimal change and significant non-developed cells, as a large number of true negative cells can artificially increase the score. Furthermore, FoM values, which emphasise simulated change, tend to be low in regions dominated by persistence (Tang and Jia, 2014; Lin et al., 2023).

A very popular technique used in urban growth models is Cohen's Kappa coefficient. This statistic is a more sophisticated approach than the previously mentioned cell-based metrics to determine accuracy by observed agreement and also to measure the agreement that would be expected by chance (Cohen, 1960). The kappa coefficient (κ) is given as follows:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{2.3}$$

P_o represents the observed accuracy, defined as the proportion of instances where observed and simulated classifications align, reflecting the model's classification accuracy, as calculated in equation 2.1. In contrast, P_e represents the expected precision, indicating the proportion of instances that are expected to be correctly classified by chance. Calculated by taking into account the probability that each class occurs in both observed and simulated classifications. The expected accuracy (P_e) is given by (Chicco et al., 2021), where N is the total cells:

$$P_e = \left(\frac{TP + FP}{N} \times \frac{TP + FN}{N}\right) + \left(\frac{TN + FP}{N} \times \frac{TN + FN}{N}\right) \tag{2.4}$$

Hagen (2003) developed Fuzzy Kappa, enhancing traditional Kappa by integrating fuzzy set theory to address spatial discrepancies. Kappa addresses exact pixel matches, while Fuzzy Kappa assigns partial agreement for matching pixels within a defined proximity. This statistic employs a distance decay function to define agreement levels based on spatial distance, providing a refined evaluation of spatial alignment in binary outputs. The Kappa Simulation was introduced by van Vliet et al. (2011) to address limitations in the traditional Kappa model of expected agreement. Unlike Kappa, which assumes uniform probability across pixels based on frequency, land use models are highly dependent on the initial spatial distribution of land uses. Kappa Simulation improves upon this by incorporating conditional probabilities, assessing the likelihood of occurrence at specific locations based on the initial land use map. This adaptation makes it better suited for evaluating land use model outputs. The Fuzzy Kappa Simulation was cre-

ated by van Vliet et al. (2013) to merge the Fuzzy Kappa and Kappa Simulation into a unified statistic. It defines agreement between observed and simulated cells, correcting for chance agreement based on the initial map, and employs fuzzy interpretation for transitions. Fuzzy Kappa Simulation (FKS) combines the spatial tolerance of Fuzzy Kappa with the conditional probability approach of Kappa Simulation, enhancing the assessment of land use transitions with nuanced spatial alignment (van Vliet et al., 2013).

2.4.3 Sensitivity analysis

Sensitivity analysis (SA) in ABMs is vital to assessing how variables influence specific outcomes. SA identifies which inputs affect output variability and to what extent. Its primary purposes are to evaluate the robustness of findings and to determine the relative sensitivity of inputs, helping to understand influenced processes (Saltelli et al., 2004). By analysing which inputs change outputs, SA serves as a design of experiments, enhancing comprehension of system dynamics. This involves identifying driving inputs and quantifying model non-linearity (Kang et al., 2022). SA also tests ABM credibility as it helps identify model limits and improve model transparency (Grimm et al., 2014; Saltelli et al., 2004). SA has been categorised into local and global methods.

2.4.3.1 Local sensitivity analysis

Local sensitivity analysis (LSA) explores the immediate space of sample input values. It assesses how a relatively small change in an input affects the output, while holding all other inputs constant. Essentially, LSA investigates the impact of minor variations in individual inputs on model outcomes. Some of the local methods used in ABMs include (Ligmann-Zielinska et al., 2020):

- One-at-a-time (OAT): This is the simplest form of SA to understand and implement. OAT evaluates one input at a time, assuming that if the results change greatly with slight variations in an input, more effort should be invested in obtaining more accurate estimates of that input. OAT is suitable for exploring the mechanistic explanations of single input effects on outcomes.
- One-factor-at-a-time (OFAT): This method is an extension of OAT that uses stochastic modifications to inputs, rather than small variations, to produce a graphical result depicting output against inputs. The procedure involves uniform sampling n times throughout an input's distribution and executing the model n times while holding all other inputs constant.

The primary limitation of OAT and OFAT is their assessment of the effects of individual input variables on model behaviour. They neglect scenarios where interactions among

multiple inputs lead to significant changes in outputs. These methods only cover a small portion of the input space and should not be used to understand the comprehensive dynamics of a target system. For example, when dealing with ten input variables, OFAT only evaluates 0.0025 percent of the input space (Du and Wang, 2011). As a result, SA results using single input approaches, like OAT and OFAT, may generate erroneous inferences about the relative importance of model sensitivity to particular inputs

2.4.3.2 Global sensitivity analysis

Global sensitivity analysis (GSA) examines the complete input space and the variability of outcomes stemming from both individual inputs and their interactions. Unlike local sensitivity analysis (LSA) which emphasises minor input changes, GSA seeks to elucidate the full spectrum of model behaviour across all input values. Some of the GSA methods used in ABMs include (Ligmann-Zielinska et al., 2020):

- Screening SA: aim to efficiently identify the most influential inputs with minimal computational cost, making them well-suited for high-dimensional models. They allow input sensitivity to be assessed with a relatively small number of model runs, providing a ranking of variables based on their importance. However, the resulting sensitivity measures are primarily qualitative and are not sufficient for accurately comparing the relative magnitude of input effects (Ayllón et al., 2016).
- Variance-based SA: decomposes the total variance of the model output into contributions from individual inputs and their interactions. These methods are model-independent and capable of handling non-linear relationships as well as complex input interactions. They provide a comprehensive and quantitative assessment of input influence and their interactions. However, they are more computationally demanding, assume input independence, and may be influenced by the assumption that the output distribution is approximately normal (Ayllón et al., 2016).

These GSA methods provide unique ways to analyse how inputs affect model outputs and are often used together to balance computational demands with analytical depth (Ligmann-Zielinska, 2013; Borgonovo et al., 2022). GSA does not assume linearity and can handle the non-linear relationships between inputs and outputs. This is a significant advantage in ABMs, where non-linear dynamics are common (Borgonovo et al., 2022). LSA assume that the model is additive and can become impractical when dealing with a large number of inputs (Kang et al., 2022). LSA methods, such as OAT and OFAT, are easy to understand and implement, their limitations, such as the inability to capture input interactions and the restriction to a small portion of the input space, make them less suitable for complex ABMs. GSA provides a more robust and comprehensive approach by exploring the entire input space, identifying input interactions, and handling non-linearities, making it a superior choice for gaining a deeper understanding of ABM behaviour (Ligmann-Zielinska et al., 2020).

2.4.4 Limitations of existing procedures

Uncertainty analysis, calibration, and sensitivity analysis have been applied in existing evaluation protocols to address the credibility of ABMs. For example, Borgonovo et al. (2022) offered a six-step approach for performing ABM sensitivity analyses that involved identifying SA goals, specifying parameters to vary, choosing SA tests and visualising the results. Their protocol was applied to the conceptual Garbage Can ABM, and emphasised the importance of interactions between model parameters in ABMs. Kang et al. (2022) developed a global sensitivity analysis and calibration framework named GSA-CAL for spatially explicit ABMs. The framework reduced the computational intensity of ABMs when working with numerous model parameters by excluding less influential parameters prior to calibration. It was applied to an ABM of influenza transmission in Miami, Florida and the approach successfully reduced computational intensity and improved efficiency by focussing on important model parameters. Patel et al. (2018) analysed ABM outputs in a framework using exploratory data analysis, sensitivity analysis, and data mining techniques. The paper provided a better understanding between model inputs and outputs and identified the most important model parameters on the model output. The conceptual DITCH model was used as a demonstration, which simulates inter-ethnic partnerships in the UK. Ligmann-Zielinska et al. (2020) provided a comprehensive roadmap for selecting appropriate SA methods based on the modelling purpose. The authors argued that different ABMs require different SA approaches based on their purpose and the paper developed a roadmap with three different SA pathways based on the application and complexity of the ABM. It noted the importance of using a combination of SA methods, to capture input-output variability and interactions and presents emerging methods that handle multidimensional aspects of ABMs (Ligmann-Zielinska et al., 2020).

There is an opportunity to bridge the gap between sensitivity and exploration procedures performed in more general ABMs and to apply this specifically to urban growth models. These previous modelling procedures provide useful examples to build a new protocol that is specific to the urban modelling domain and also introduce variance stability and the importance of the distinction between calibration and validation to allow future models to follow this example. Among the existing protocols, the majority have been applied to conceptual ABMs. However, to ensure their practical application, protocols should instead be demonstrated on non-conceptual ABMs that model real-world phenomena (Lorscheid et al., 2012). Furthermore, Ligmann-Zielinska et al. (2020) argues that while there are compulsory protocols for experiments in physical laboratories, similar protocols should also be mandatory for evaluating ABMs. Designing a protocol with these added features and ensuring its automation to any type of urban model will enhance a model's credibility.

2.5 Urban form modelling accuracy metrics

2.5.1 Pixel-based metrics

The direct image outputs of urban form GANs are commonly assessed against ground truth images using the Fréchet Inception Distance (FID) and the Mean Intersection over Union (mIoU) (e.g. Wu and Biljecki, 2022, 2023; Gu et al., 2024). The FID score compares the distribution of the generated images to the distribution of real images by computing the distance between the feature vectors of real and generated images, which are extracted using an Inception-v3 image classification model (Salimans et al., 2016). It assesses the similarity between the two sets of images in terms of their statistical properties (mean and covariance) in a feature space. A lower FID score indicates that the generated images are more similar to the real images, suggesting that the generated images are of higher quality and more morphologically similar to the target images. Eq. 2.5 expresses the FID score mathematically:

$$FID = \|m - m_w\|_2^2 + \mathsf{Tr}\left(C + C_w - 2\left(CC_w\right)^{1/2}\right) \tag{2.5}$$

Here m and m_w are the mean vectors of the generated and real image distributions, respectively. C and C_w are the covariance matrices of the generated and real image distributions, respectively. $\|\cdot\|_2$ represents the Euclidean norm. Tr denotes the trace of a matrix. $(CC_w)^{1/2}$ represents the matrix square root of the product of the two covariance matrices.

The mIoU is a performance metric commonly used in semantic segmentation and has also been applied in GANs. It measures how well a model predicts pixel-level categories in an image. The mIoU is calculated between input and generated images to assess the overlap with ground truth. Mathematically, mIoU is the ratio of common pixels between input and generated images to the total number of pixels in both data sets. A high mIoU indicates considerable overlap with the input, reflecting effective data replication. It is calculated in the following equations, where c is the total number of classes (Sulistiyo et al., 2018):

$$\mathsf{IoU}_c = \frac{TP_c}{TP_c + FP_c + FN_c} \tag{2.6}$$

$$\mathsf{mIoU} = \frac{1}{C}\sum_{c=1}^{C} \mathsf{IoU}_c \tag{2.7}$$

2.5.2 Vector-based metrics

In urban form GANs that convert images to vector road networks, numerous metrics have been used to evaluate the accuracy of generated networks against ground truth networks. Birsak et al. (2022) used segment circuity, transport ratio, and node degree to assess the road network against the ground truth. Segment circuity quantifies the curvature of road segments. The transport ratio evaluates movement efficiency, defined as the ratio of the shortest path distance to the straight-line distance between points. A higher transport ratio reflects a less direct route, influenced by curvature or obstacles. Node degree quantifies the connections at intersections, providing insight into network complexity. A higher average node degree signifies greater interconnectivity. Hartmann et al. (2017) assessed StreetGAN based on block area generated by roads, compactness, and aspect ratio. Compactness is the ratio of block area to its minimal bounding box, while aspect ratio compares the lengths of the bounding box sides. Yang et al. (2023b) evaluated generated road networks using connectivity index, intersection density, total road length, and convenience of transportation. The connectivity index represents the average node degree between generated and real networks; intersection density counts intersections (vertices with degrees greater than 2) in a layout, reflecting connectivity and accessibility. Transportation convenience assesses travel ease between points using Euclidean distance and shortest graph distance, averaged over 100 sampled source-destination sets (minimum distance of 0.25 km). These metrics collectively provide information on connectivity and accessibility.

For vector building footprint accuracy metrics, Wu and Biljecki (2023) measured the median percentage difference between the generated and ground truth building areas and perimeters in InstantCITY. Zhou et al. (2024) and Wu and Biljecki (2023) also calculated the percentage ratio of generated to real footprints to assess the proportion of generated footprints relative to the actual number. In addition, both studies measured the site cover ratio, building area to image tile area, to evaluate overall coverage.

2.5.3 Accuracy metric limitations

The limitations of pixel-based and vector-based metrics are important to consider when evaluating road networks and building footprints created from DGMs. FID and mIoU are primarily designed to assess visual similarity between natural images, such as landscapes, by comparing distributions of high-level features like textures and object shapes. However, these metrics do not directly account for the completeness, topology, and connectivity of road networks or the completeness, layout, and size of the building footprints. As a result, critical elements may be missing or incorrectly generated without significantly affecting the FID score, leading to a misleading assessment of the generated road and building realism. To ensure a more accurate evaluation, vector-

based analysis is necessary to verify structural integrity and spatial relationships within the generated data.

Vector-based road network and building footprint metrics are robust, but have certain limitations. For road networks, metrics such as average node degree, connectivity index, and intersection density (e.g. Birsak et al., 2022; Yang et al., 2023b) provide global insights, but do not assess the performance of GAN in capturing each node degree, which could reveal local variations. Additionally, transportation convenience depends on sampled source-destination pairs that may not fully represent real-world travel behaviours. This metric is also computationally intensive, requiring significant processing time for Euclidean and graph distance calculations (Yang et al., 2023b). For building footprints, the site cover ratio measures overall coverage but ignores spatial distribution, meaning that two layouts with similar coverage can have vastly different densities and spatial patterns. Similarly, the total building count may be misleading if the layout and size of the footprints are not considered.

Furthermore, there has been limited research on how sensitive DGMs are to image scale and resolution, especially for road network generation. One way to quantify this sensitivity would be to examine the impact on pixel and vector metric values under various spatial scales and pixel resolutions. Only Wu and Biljecki (2022, 2023) examined the FID of the building footprint and the accuracy of the vector metrics on different spatial scales. Furthermore, while previous GAN studies used several vector-based metrics to assess performance, they did not calculate a cumulative error across all metrics (Yang et al., 2023b; Wu and Biljecki, 2023). Averaging performance across multiple metrics could provide a more comprehensive evaluation of GAN accuracy, helping to mitigate the limitations of relying on individual metrics in isolation.

2.6 Summary

This review has explored the current state-of-the-art in urban growth modelling, urban form modelling, model evaluation procedures and urban form accuracy metrics. From this comprehensive review and critique of the literature, the following limitations have been identified:

- Urban growth modelling - current models primarily focus on pixelated or regional scales, which are challenging for urban planners and decision makers to interpret effectively at their operational scale. To address this, there is a need to shift modelling efforts toward simulating urban dynamics at the land-parcel level in undeveloped and rapidly urbanising regions, ensuring high spatial granularity and usability at finer spatial scales with vector data. In addition, rule-based urban

growth models often lack robust evaluations of their outputs, limiting the understanding of their underlying processes and behaviours. Although these models are important in projecting simulating future urban development.

- Urban form modelling - there is a need not only to simulate future urban dynamics but also to generate synthetic representations of urban form. Current approaches of urban form generation have struggled across large areas or focused on purely conceptual examples. Such approaches would provide unprecedented information on future potential configurations and a deeper understanding of the impacts arising from these scenarios.

- Model evaluation procedures - these exhibit an overreliance on conceptual examples, raising concerns about their practical applications despite numerous ABM protocols available for model assessment. Furthermore, these protocols fail to differentiate calibration from validation and lack variance stability analysis for stochastic models, and have yet to be adapted for urban modelling.

Currently, no single modelling approach can comprehensively address these limitations or achieve the aim of this thesis of simulating future urban form. There is no overarching technique to fill this urban modelling gap. Instead, a framework that integrates multiple methodologies is essential to achieve this vision. This thesis proposes such a framework, designed for easily reproducible application to cities worldwide, positioning it as a new standard for urban growth modelling. This framework addresses the aforementioned limitations and gaps, both enhancing urban growth model credibility and detail of outputs. Before presenting the proposed modelling framework in Chapters 4, 5, and 6, a suitable area must be selected that is experiencing rapid urban growth and provide the appropriate context for the framework preparation steps. The study area selection is outlined in the next chapter.

Chapter 3

CASE STUDY - PHNOM PENH, CAMBODIA

3.1 Introduction

This chapter explains why Phnom Penh in Cambodia was selected as a case study for this research, justifying its suitability by exploring its historical population, economic, and physical growth within the wider context of Southeast Asian growth. It describes key characteristics of the city, including its climate, administrative structure, and geography. Future scenarios of Phnom Penh and Cambodia are also examined. An overview of Cambodia's spatial planning system is then provided to establish the relevant context and justify the development of the modelling framework. Finally, a study area is selected on the urban periphery of Phnom Penh to demonstrate and test the validity of the geospatial modelling framework.

3.2 Phnom Penh

3.2.1 Rationale

Cities in Asia are expected to increase significantly in size and population over the 21st Century. Southeast Asia is a hotspot for this growth, as its cities host 52.2% of the region's population of 695 million (UN, 2024). Cambodia is experiencing rapid economic growth, with an increase in Gross Domestic Product (GDP) of more than 1000% , from $2.5 billion in 1990 to $30 billion in 2022 (UN, 2024). Phnom Penh, the country's capital, is the largest and fastest-growing city, with its share of the national population nearly doubling over the past 34 years. In 1990, Phnom Penh's population was 615,000 people (6.9% of the national population), increasing by 282% to 2.35 million by 2024 (13.33% of the national population) (UN, 2024).

Although the 2.35 million population of Phnom Penh is dwarfed by neighbouring capital cities such as Hanoi in Vietnam at 5.43 million and Bangkok in Thailand at 11.23 million, the academic literature identified that Phnom Penh had one of the most intense urban expansion patterns within Southeast Asia over the past several decades (Cao et al., 2019; Mialhe et al., 2019). The increased urbanisation in Phnom Penh has led to a significant expansion of the city's urban footprint. Thanh Son et al. (2022) mapped an increase in urban area of 9.2% between 1990-2005 and a 14.9% increase between 2005-2020.

However, this rapid growth, driven by large-scale urban development projects from foreign direct investment, has caused urban sprawl, resulting in a fragmented urban landscape that struggles to accommodate its diverse population (Mialhe et al., 2019). Key issues include inadequate infrastructure, loss of green spaces, and traffic congestion (Mialhe et al., 2019). Phnom Penh's rapid growth makes it a suitable location for testing a framework to model future urban scenarios because the city's abundant history of past expansion provides sufficient data to thoroughly calibrate and validate the performance of the framework. Although Cambodia, like many developing countries, lacks detailed topographic geospatial data with extensive temporal records, a framework applicable to Phnom Penh will prove its robustness and should be easily adaptable and repeatable to other regions with limited geospatial data availability.

3.2.2 Description

Phnom Penh is located in the south-central part of Cambodia at the confluence of the Mekong, Tonlé Sap, and Bassac rivers (Figure 3.1). Covering an area of approximately 684 km^2, the city serves as the political, economic, and cultural hub of Cambodia (Mialhe et al., 2019). Its strategic location along major waterways has historically facilitated trade and connectivity, making it a key urban centre in Southeast Asia. The city is low-lying and flat, with an average elevation of 11 metres above sea level. This makes Phnom Penh very vulnerable to flooding, especially during the heavy rains of the monsoon season (Thanh Son et al., 2022).

Phnom Penh has a population density of 3436 people per km^2. The city is characterised by a young and rapidly growing population. The majority of the population is Khmer, but there are significant communities of Chinese and Vietnamese (World Bank, 2024).

The climate of Phnom Penh is classified as atropical wet and dry climate under the Köppen climate classification system. The city has two distinct seasons: the wet season from May to October and the dry season from November to April. During the wet season, monsoon rains contribute to high humidity and frequent flooding in low-lying areas. The dry season is characterised by warm temperatures and lower humidity,

with average annual temperatures ranging from 22-35°C. The hottest months are typically March and April, while the coolest period occurs between December and January (Thanh Son et al., 2022).

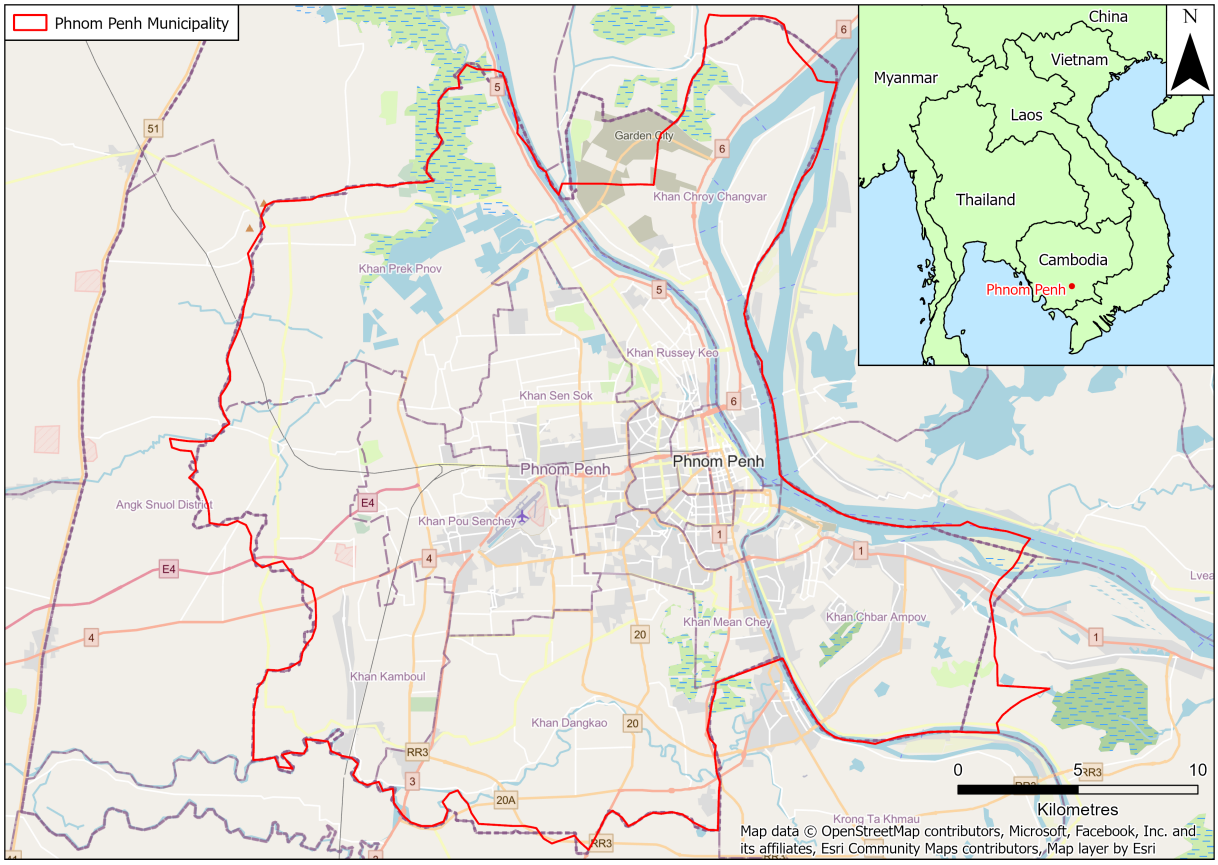

Figure 3.1: The location of Phnom Penh within the wider context of Cambodia and Southeast Asia.

Phnom Penh is the economic centre of Cambodia and contributes significantly to national GDP through various sectors. The city's economy is driven by manufacturing, trade, real estate, and tourism. The garment industry is one of the largest employers, attracting foreign investment from China and South Korea. The city is home to a growing financial sector, with banks, microfinance institutions, and insurance companies playing an increasing role in economic development. Tourism is another key contributor, with historical sites and cultural landmarks that attract visitors from all over the world (World Bank, 2024).

The city is administered as an autonomous municipality, separate from the provinces of Cambodia. The city is governed by the Phnom Penh Municipality, which is headed by a governor appointed by the central government. The municipality is divided into 12 districts, which are subdivided into 96 communes (Figure 3.2) (Runfola et al., 2020). Local governance focusses on urban planning, public services, and infrastructure development. The administration also works in collaboration with national authorities to address issues such as traffic management and public safety (Diepart et al., 2016*a*).

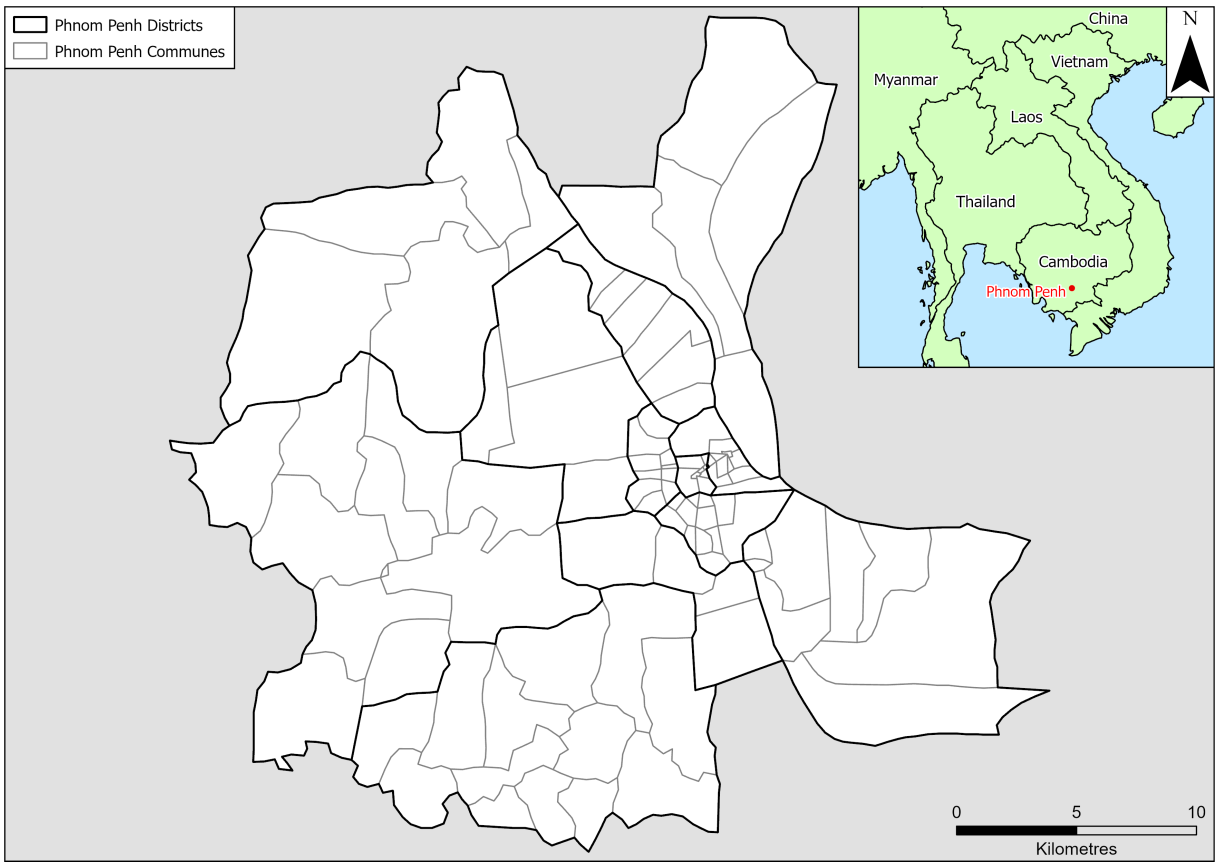

Figure 3.2: Phnom Penh administrative levels (Runfola et al., 2020).

Phnom Penh has undergone significant modernisation in recent decades. High-rise buildings, commercial centres, and residential developments are reshaping the city's skyline. The government is investing in infrastructure projects such as road expansions, bridges, and drainage systems to modernise the city. Public transportation remains limited, but recent initiatives, such as bus services and planned mass transit systems, aim to improve mobility in the city.

Overall, the city is an evolving metropolis that is experiencing intense economic, population, and urban growth. As the political, economic, and cultural centre of Cambodia, it continues to play a crucial role in the development and regional integration of the country. With ongoing investments in infrastructure and sustainability, the city is poised to maintain its importance in the urban landscape of Southeast Asia.

3.2.3 Future socioeconomic scenarios

Think tanks, consultancies, nongovernmental organisations, and intergovernmental organisations have developed scenarios and strategies for Cambodia's future. For example, Future Forum, an independent think tank founded in 2015, has conducted rigorous research to formulate numerous scenarios up to the year 2040 using a measured, analytical, and evidence-based approach to identify potential trends (Udom et al., 2020). Similarly, the Global Green Growth Institute collaborated with the Royal Government

of Cambodia and the Phnom Penh Capital Administration to develop sustainable city plans to 2030 that promote green growth and urban resilience (GGGI, 2018). The World Bank is also actively involved in Cambodia's urban development, particularly in Phnom Penh. Its reports provide critical information on urbanisation trends, institutional and financial frameworks, urban planning, and service delivery. The World Bank identifies key challenges to efficient urbanisation and outlines strategic priorities to improve competitiveness, sustainability, and social inclusion in Phnom Penh (World Bank, 2024). The most important, relevant, and likely scenarios have been extracted from these works and are reported below to shed light on potential socioeconomic pathways.

Both Future Forum and the World Bank propose an economic boom scenario. By 2040, Cambodia's economy is expected to transition from an agricultural-based system to an industry-driven and skills-based economy. This shift is envisioned to bring diversification, higher living standards, improved healthcare, education, and inclusive social security. The global economic transformation, fuelled by advancements in communication, production, and energy, presents significant opportunities for Cambodia's industrialisation. Enhancements in agriculture, labour market efficiency, and industry are expected to drive production growth(Udom et al., 2020). Phnom Penh, in particular, has experienced a significant economic boom, driven by strong national growth and substantial investments in the construction sector. Between 1994 and 2015, it was the sixth fastest growing economy globally, surpassing Vietnam and most other countries in the region. The gross national income per capita more than tripled from $300 in 1994 to an estimated $1,070 in 2015, elevating Cambodia to a lower middle income status (World Bank, 2024). This growth has been fuelled by garment exports, agriculture, tourism, and, more recently, construction and real estate. It is likely to become an upper middle income economy by 2030 and an upper income economy by 2050.

Future Forum has identified a potential economic recession in Cambodia due to its economic fragility. The absence of an inclusive social protection programme may hinder the country's ability to break the poverty cycle amid economic shocks. A predicted slowdown in the Chinese economy poses a significant risk, as China is a major investor in Cambodia. Declines in Chinese capital flows could severely impact Cambodia's economic stability. Experts emphasise that Cambodia's dependence on garment exports, which comprise 77% of total exports and primarily target eight partner countries, constitutes a weak risk management strategy, making the economy vulnerable to recessions (Udom et al., 2020). This lack of diversification exposes Cambodia to global market fluctuations and increased competition from other low-wage countries. Nations with larger and more affordable labour forces may surpass Cambodia's garment industry over time. Moreover, reliance on China may restrict Cambodia's policy flexibility in international relations and economic development. A sudden reduction or reversal of

Chinese foreign direct investment would significantly impede private sector growth. The COVID-19 pandemic led to the first economic contraction in 25 years, with a 3.1% decline in 2020. Post-COVID-19, the economy has gradually recovered, with growth rates of 3.0% and 5.2% in 2021 and 2022, respectively, reflecting the country's dependency on tourism, a volatile industry (World Bank, 2024).

The Global Green Growth Institute (GGGI) envisions Phnom Penh as a green city with a clean environment, aiming to minimise environmental impact while improving its role as a competitive economic, business, and cultural hub. This vision is outlined in the Phnom Penh Master Plan for Land Use 2035, which emphasises sustainable and equitable development. By 2030, Phnom Penh will become a clean, green, and competitive city, providing a safe and high-quality lifestyle for its residents (GGGI, 2018). To achieve this, the city has set the following key objectives:

- Decouple economic growth from environmental impact, ensuring sustainability.
- Promote social inclusion, reduce poverty, and improve urban welfare.
- Improve urban resilience to natural disasters, climate change, and other risks.
- Boost urban competitiveness and business attractiveness, making Phnom Penh a desirable place for investment

Future Forum has identified a potential real estate bubble in Phnom Penh. From 2004 to 2008, land prices increased from approximately $250 to more than $2,000 per square metre (Udom et al., 2020). In 2021, Chinese fixed asset investment in Cambodia reached $2.32 billion, an increase of 67% from the previous year (World Bank, 2024). Chinese investment significantly inflates real estate prices, necessitating proactive measures to prevent unsustainable growth in the sector. High property prices and demand can cause a bottleneck in the market due to demand outpacing land supply. Future Forum identified government intervention as crucial to mitigate the risk of a real estate bubble. Property tax reform could address the economic instability caused by vacant land and speculative price inflation. Taxing vacant land would discourage hoarding, promote development, and increase housing supply and economic activity. A flat tax on commercial and rental properties, gradually increasing over two decades, could prevent speculation and stabilise the market. Without such measures, unchecked speculation by local and foreign developers could distort the market and exacerbate economic downturns. Implementing these reforms foster a sustainable economic environment, reducing the likelihood of real estate bubbles (Udom et al., 2020).

3.2.4 Earth Observation for Sustainable Development

The Earth Observation for Sustainable Development (EO4SD) initiative, led by the European Space Agency, aims to integrate satellite-derived environmental data into

the operations of International Financial Institutions and their partner countries to support sustainable development efforts. In Phnom Penh, Cambodia, the EO4SD Urban Project focused on using Earth observation data to support urban development initiatives. The project produced detailed geospatial datasets for the city, including 2017 land use, urban extent, transport infrastructure, and flood hazard assessments. These data sets provide information at the block level, classifying land blocks into categories such as low, medium, and high density urban fabric (Figure 3.3). By providing these comprehensive datasets, the EO4SD Urban Project supports informed decision-making and strategic planning for sustainable urban development in Phnom Penh (ESA, 2019). This dataset provides invaluable geospatial information to inform the scenarios generated. More information on how the 2017 EO4SD land use dataset was used is described in Chapters 5 and 6.

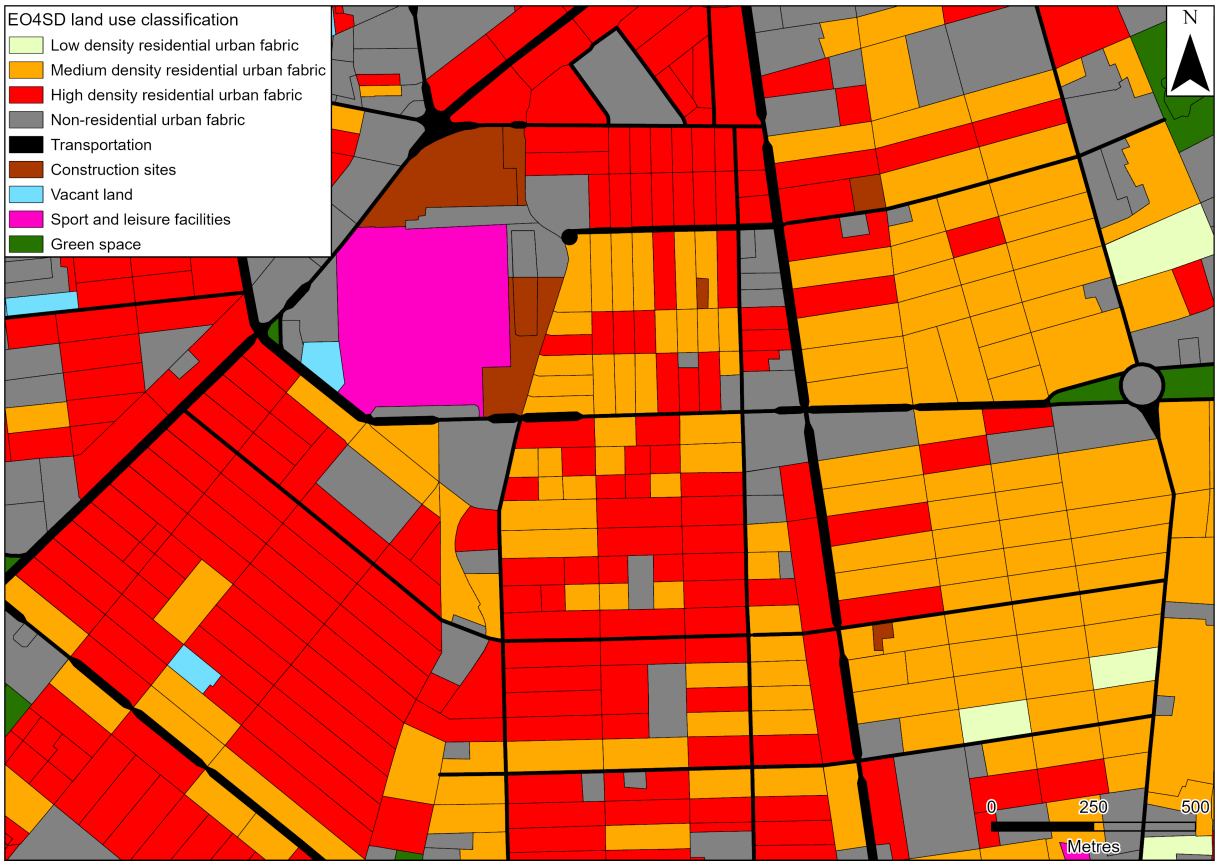

Figure 3.3: Example of EO4SD block-level land use information in central Phnom Penh (ESA, 2019).

3.3 The spatial planning system

The Ministry of Land Management, Urban Planning, and Construction (MLMUPC), in collaboration with the German development agency GIZ, published a series of spatial planning handbooks in Cambodia. These handbooks provide an overview of the spatial planning system and serve as guides for formulating spatial plans at subnational

levels to improve their quality and strengthen the capacity of spatial planners. Designed as practical resources, they outline a step-by-step approach to urban planning at different administrative levels. The handbooks offer invaluable information that is fundamental for the proposed geospatial modelling framework. The information from these resources was used for this section (Diepart et al., 2016*a*,*b*,*c*).

3.3.1 History of Cambodian spatial planning

The MLMUPC, founded in 1999, is a central government body in Cambodia that oversees land management and urban planning. The MLMUPC serves as the secretariat of the National Committee for Land Management and Urban Planning, which is the main agency for the implementation of spatial planning in Cambodia. The ministry addresses competing land uses and aims to balance them for sustainable development. The MLMUPC is involved in the development of planning procedures for subnational spatial plans and is the key institution in Cambodia that drives spatial planning policy and practice, ensuring that land use and development align with the goals of sustainability, economic growth, and social equity.

Cambodia's spatial planning system is relatively new, having evolved over the past 30 years through various legal and policy documents. These establish the legal framework and institutional structures for spatial planning at national and subnational levels. The purpose of the system is to enhance administrative decentralisation. The Law on Land Management and Urban Planning, introduced in 1994, is the key legislation for spatial planning, allowing provinces, cities, and municipalities to create their own development master plans for territorial organisation and development. The Law on Administrative Management of the Capital, Provinces, Municipalities, Districts, and Khans, enacted in 2008, transformed the existing planning system by giving each subnational administration the mandate to produce a 5-year development plan.

The 2011 National Policy on Spatial Planning outlined the vision, goals, and strategic objectives for spatial planning in Cambodia. It specified the mechanisms, preparation, and hierarchy of spatial plans at the national and subnational levels. The 2015 National Land Policy Paper is another key policy document for the sector. It requires that relevant authorities initiate, coordinate, and approve spatial planning across the four administrative levels: National, Provincial, District, and Commune. Sub-Decree No. 42 on Urbanisation of the Capitals, Municipalities and Urban Areas, introduced in 2015, defined the scope, content, and procedures for developing three urban physical plans: Land Use Master Plan, Land Use Plan, and Urban Detailed Plan.

3.3.2 Hierarchy of spatial plans

The 2011 National Policy on Spatial Planning established a hierarchy of spatial plans at various administrative levels, each with specific scopes. Cambodia's spatial planning system encompasses provincial, district, and commune spatial plans. The plans are designed to work together to create a cohesive and integrated spatial planning system. Lower-level plans (e.g. Commune) are intended to support the implementation of higher-level plans (e.g. Province).

The provincial plan represents the highest level of the subnational system hierarchy, offering a strategic vision for land use and physical development within the province. It evaluates the characteristics of the province, identifying existing spatial structures and development patterns. The plan establishes an ideal spatial framework for settlements, infrastructure, and open spaces based on local needs and potential while adhering to national directives. In addition, it allocates functions to various urban and rural areas, fostering sustainable development. The Provincial Spatial Plan is developed by the Provincial Committee for Land Management and Urban Planning, with the a time horizon of 20 years.

The district plan provides more detailed directions for spatial development. It is meant to determine the specific spatial structure and main functions of the district or municipality, considering its integration into the province. It defines spatial development strategies to guide the physical and functional development of the district, addressing future transport and communication networks, infrastructure, public services, and the management of agriculture, forests, and water. The District Plan is intended to determine general land use types, differentiating between buildable and control categories. The District Land Management and Urban Planning Committee is responsible for developing the master plan, under the supervision of the District Council.

The commune plan represents the most localised spatial plan which considers land and water resources, potential land uses, and social and economic characteristics to ensure appropriate land use allocation and zoning. This includes identifying areas for agriculture, residential, commercial, or other uses, as well as areas that need protection or conservation. The plan considers various interests and ensures that land use decisions are made in a transparent and participatory manner through stakeholder engagement workshops, public meetings, and community consultations. It is a five-year development plan and is the responsibility of the Commune council to create and implement. The commune plan is a legally binding document that precisely allocates future land use units and includes detailed ordinances regulating zoning, construction, land development, and conservation.

All of these plans are interrelated and provide invaluable information for the proposed framework. For example, it has clarified that communes set five-year plans to meet land use demand over this time period, local governments are also a key actor in the development process by setting the master plans and base their decisions off physical factors, such as water resources and existing land use. The spatial planning process is complex and lengthy, which means that simplifications were necessary when replicating it in a computational environment. Elements of all these plans can be integrated and utilised in the modelling framework, especially the commune spatial plan.

3.4 Study area selection

As identified in Chapter 2, urban growth models should focus on smaller areas and use actual land parcel information to provide more contextual future growth simulations. Creating a parcel-scale model for all of Phnom Penh would be too large and computationally demanding when focussing on small-scale processes as there are potentially up to a million land parcels within the municipality. Therefore, a smaller area was selected as the study area in the southwest of Phnom Penh to demonstrate this modelling framework (Figure 3.4). It compromises three communes: Pong Tuek, Prey Sa, and Prey Veaeng, which are merged and treated as a single area for modelling purposes.

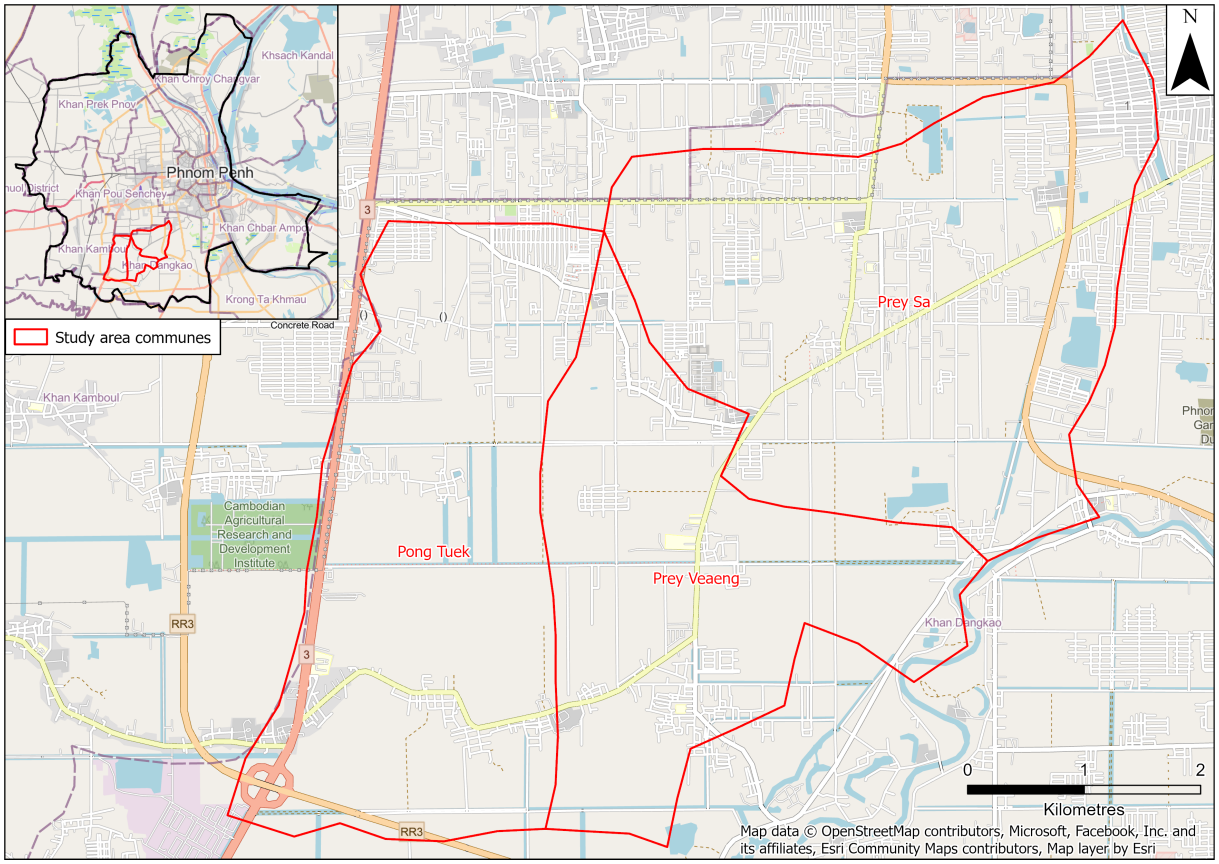

Figure 3.4: The three communes used as the study area used for the geospatial modelling framework.

This study area spans 33.30 km^2 and has a collective population of 40,915 as of the 2019 census. The 1998 census reported a population of 12,795 showing an increase of almost 220%. These communes have also experienced rapid growth, in 2014 urban area compromised 3.41 km^2 and by 2023 this increased by 205%, reaching 10.41 km^2. However, in 2023, the non-urban area compromised 22.63 km^2 and still dominates over 50% of the land in the study area, indicating that it has sufficient space to accommodate future development. In addition, combining three communes creates an area large enough to capture key characteristics of the wider region, such as surrounding roads and urban areas. The study area's population and urban growth, predominant non-urban character and size make this location suitable for applying and demonstrating the proposed modelling framework.

3.5 Summary

This chapter has identified Phnom Penh as a suitable case study within a region that is expected to undergo significant urban growth throughout the 21^{st} century. The city has a history of rapid and intense urban development, which has led to urban sprawl. In addition, future scenarios that the country may experience have been identified and explored. These scenarios include an economic boom, economic recession, sustainable planning, and real estate bubble, all of which are feasible and realistic pathways the city may experience supported by research and evidence found in grey literature.

Cambodia's spatial planning system has also been outlined, which is relatively new, with its foundations laid in the 1990s. The system follows a bottom-up hierarchy and is interrelated between the administrative levels, beginning at the commune level and extending to the provincial level. An explanation of the planning system provides a solid and useful understanding of spatial planning processes necessary for the research in this thesis.

Given Phnom Penh's size, a smaller study area was selected consisting of three communes on its urban periphery due to recent urban expansion and sufficient non-urban space to allow meaningful future simulations. Overall, this chapter has laid the foundation for testing the proposed framework by justifying Phnom Penh as a study area based on its intense urban growth, planning system, and future scenarios.

Chapter 4

FRAMEWORK DESIGN AND PREPARATION

4.1 Introduction

This chapter presents the geospatial modelling framework and outlines the process for urban form generation. Each stage of the framework is justified based on the literature review and the specific techniques are explained. Data preparation and acquisition for the framework is also described.

4.2 Framework structure

As concluded in Chapter 2, a variety of techniques need to be linked to achieve the aim of this research to simulate future urban form. A methodological framework has been developed to achieve this goal, first by modelling potential urban growth locations and then generating urban form within those areas projected to change. The framework has been divided into four stages to address the each research objective described in Chapter 1. Each stage leads into the other with the outputs generated being necessary for the next stage. Figure 4.1 summarises the modelling framework, with each stage and the respective modelling technique used. The first half of the framework focusses on simulating the location of future urban growth and the second half focusses on the future urban form in the areas expected to transition to urban.

4.2.1 Agent-based model

An ABM was selected to simulate future urban growth as it allows greater modelling flexibility and heterogeneous agent decision-making. CA is limited to extrapolating past trends and only operates on fixed transition rules that lead to predictable outcomes (Mozaffaree Pour and Oja, 2021). With an ABM, there are more options for generating scenarios compared to CA, which makes this technique particularly suitable for the purpose of this research. Additionally, ABMs can incorporate CA elements, such as local neighbourhood rules, allowing them to combine the strengths of both approaches.

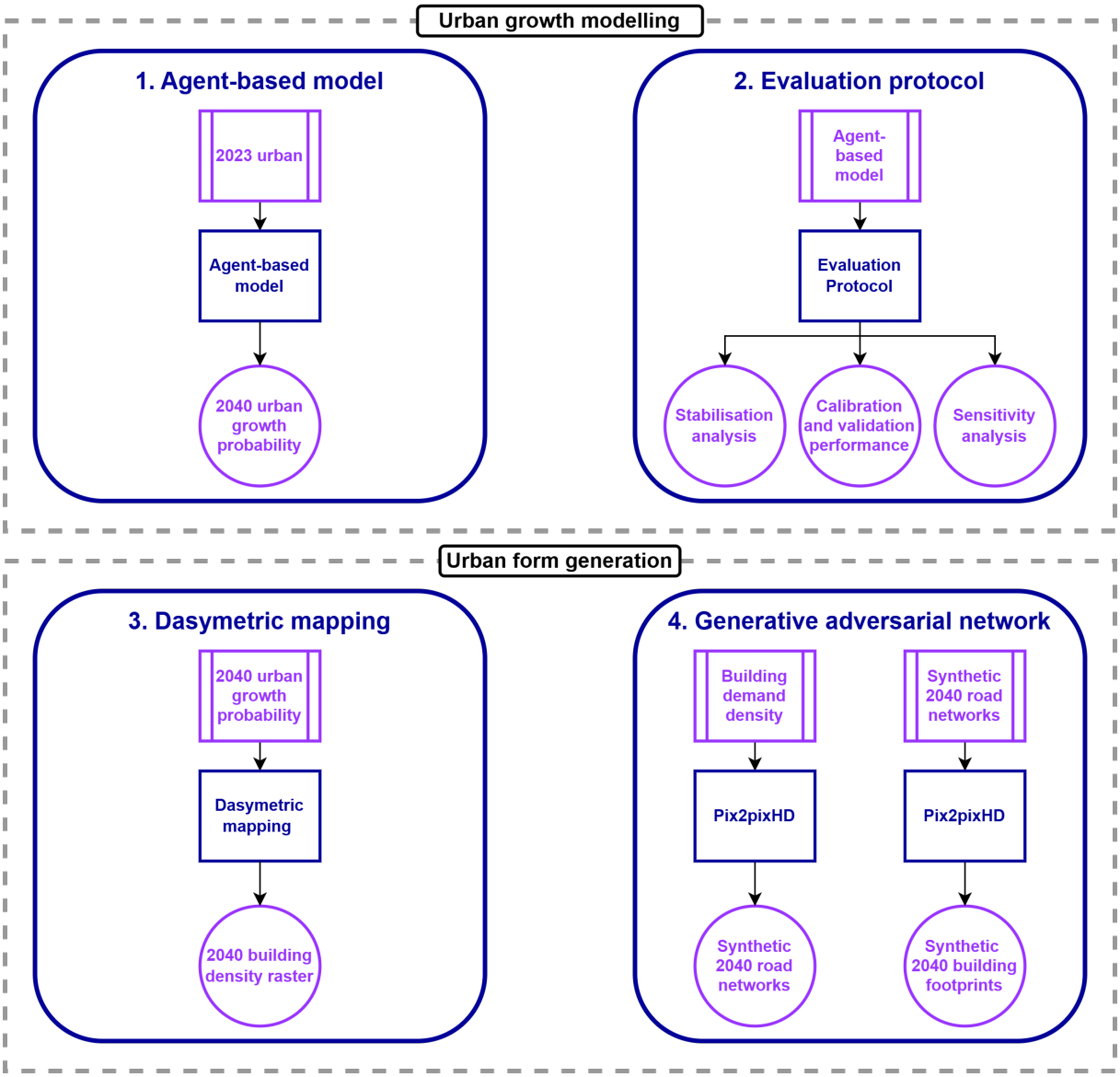

Figure 4.1: Overview of the proposed geospatial modelling framework to simulate future urban form. Stages 1 and 2 are described in Chapter 5. Stages 3 and 4 are described in Chapter 6.

An ABM is used to simulate future urban growth extent up to the year 2040. Stochastic processes have a strong presence in the ABM to reflect the heterogeneous and unpredictable nature of urban development (Guzman et al., 2020). The structure and processes of the model were based on the urban planning system at the commune level outlined in Section 3.3. It incorporates key actors, such as urban planners, developers, and landowners, into the simulation and use driving factors based on available geospatial data and previous urban growth modelling literature. The ABM is described in full in Section 5.2 of Chapter 5.

The open source GIS Agent Based Modelling Architecture (GAMA) platform was selected for this project as it is a modern software developed to overcome the limitations of previous ABM toolkits, such as NetLogo and MASON (Taillandier et al., 2019). GAMA utilises its own straightforward and agent-orientated modelling language called GAML,

has an intuitive graphical user interface, and has built-in exploration options, such as GA and LHS. It is a computationally powerful software that has strong geospatial data support and was specifically designed to support sophisticated spatially explicit ABMs that can support millions of geospatial agents (Taillandier et al., 2019).

4.2.2 Evaluation protocol

The ABM applied to an evaluation protocol created in this research to assess and understand its behaviour and dynamics. Initial stabilisation analysis is performed to identify the number of runs required until the ABM has converged, incorporating both variance stability and uncertainty analysis techniques. Calibration and validation evaluates the performance of the ABM against historical growth and ensure that it is suitable for future simulation. A sensitivity analysis then identifies model parameter importance, direction of influence and interactions quantification. This enables a full understanding of the ABM and ensure it's credibility in future urban growth projections. The protocol is described in full in Section 5.3 of Chapter 5.

This protocol addresses limitations of previous approaches by incorporating stabilisation analysis to determine model uncertainty and variance, clarifying the importance of separating calibration and validation, and develop more sophisticated and empirically-founded scenarios. It provides a practical example of how such a protocol can be applied to a real-world ABM and not purely a conceptual example. In addition, it is an urban simulation-specific protocol to guide best practice in systematically evaluating future urban growth models. The evaluation protocol was performed primarily on the GAMA platform.

4.2.3 Dasymetric mapping

The ABM's output development probabilities provide a proxy for building demand. These probabilities are converted into a building density raster using dasymetric mapping, which is used as an input image for road network generation. The dasymetric mapping approaches are described in full in Section 6.2 of Chapter 5.

Dasymetric mapping is a technique that subdivides a source zone into smaller areas to better represent spatial variations in a geographic feature based on ancillary data (Qiu et al., 2022). In this case, dasymetric mapping was used to refine the conversion by redistributing spatially aggregated development probabilities using additional data, such as existing building footprints. By subdividing parcels into smaller polygons based on existing curtilage area and probability distributions, this approach ensures a more precise and nuanced spatial representation of development demand. The resulting density raster enables a deep generative model to feasibly generate road networks.

4.2.4 Generative adversarial network

A GAN was selected for the urban form generation from the DGMs listed in the literature review. Although diffusion models identified in the literature review (Chapter 2 do provide an advancement over GANs, few have been developed for use at high resolutions, and current research has been focused on perfecting these models at lower resolutions (e.g. 256x256 pixels) making them inadequate for generating road networks across large-scale areas such as the study area selected for this framework. GANs also have a stronger track record of applications in the built environment, because research on using diffusion models in the built environment has not yet been widely adopted (Wu et al., 2022).

The first logical step to begin the generation of urban form in the study area is to train a GAN to generate road networks based on the density output of the study area from the dasymetric mapping stage. Once a synthetic network has been generated, the building footprints can be generated based on that network as previously done by Wu and Biljecki (2023). Therefore, the urban form generation involves two stages: GAN road network generation from ABM density outputs and GAN building footprint generation from future road layouts. The road network and building footprint generation are described in full in Sections 6.4 and 6.5, respectively, of Chapter 6. Before GAN training and testing begins, a suitable GAN must be selected which is capable of translating an input image to a target image, for example, an image of road density to road networks.

The pix2pixHD GAN was selected as an appropriate deep generative model to use to create a synthetic future urban form due to its proven track record in the urban modelling domain (e.g. Wu and Biljecki, 2022, 2023). Pix2pixHD is an open source state-of-the-art GAN developed to a high standard by NVIDIA (Wang et al., 2018). Its image-to-image capabilities mean that it has a strong potential to convert density rasters to road networks and generate building footprints based on road layouts. Due to the size of the study area (w: 8 km, h: 7.1 km), the entire area must be expressed in a single high-resolution image of at least 1024x1024 pixels to capture the intricate details of road networks as shown in Figure 4.2. Therefore, since the pix2pixHD GAN has been specifically designed for high-resolution images, it provides the ideal DGM to generate urban form.

Various accuracy metrics have been used in the literature to assess the performance of generated road networks and building footprints, providing a comprehensive understanding of how well the generated urban form reflect reality. These metrics have looked at completeness of the data, such as total road length or building count, or the general morphology, such as node degree and building perimeter. A variety of these metrics will be used to evaluate the cumulative accuracy of the generated urban form.

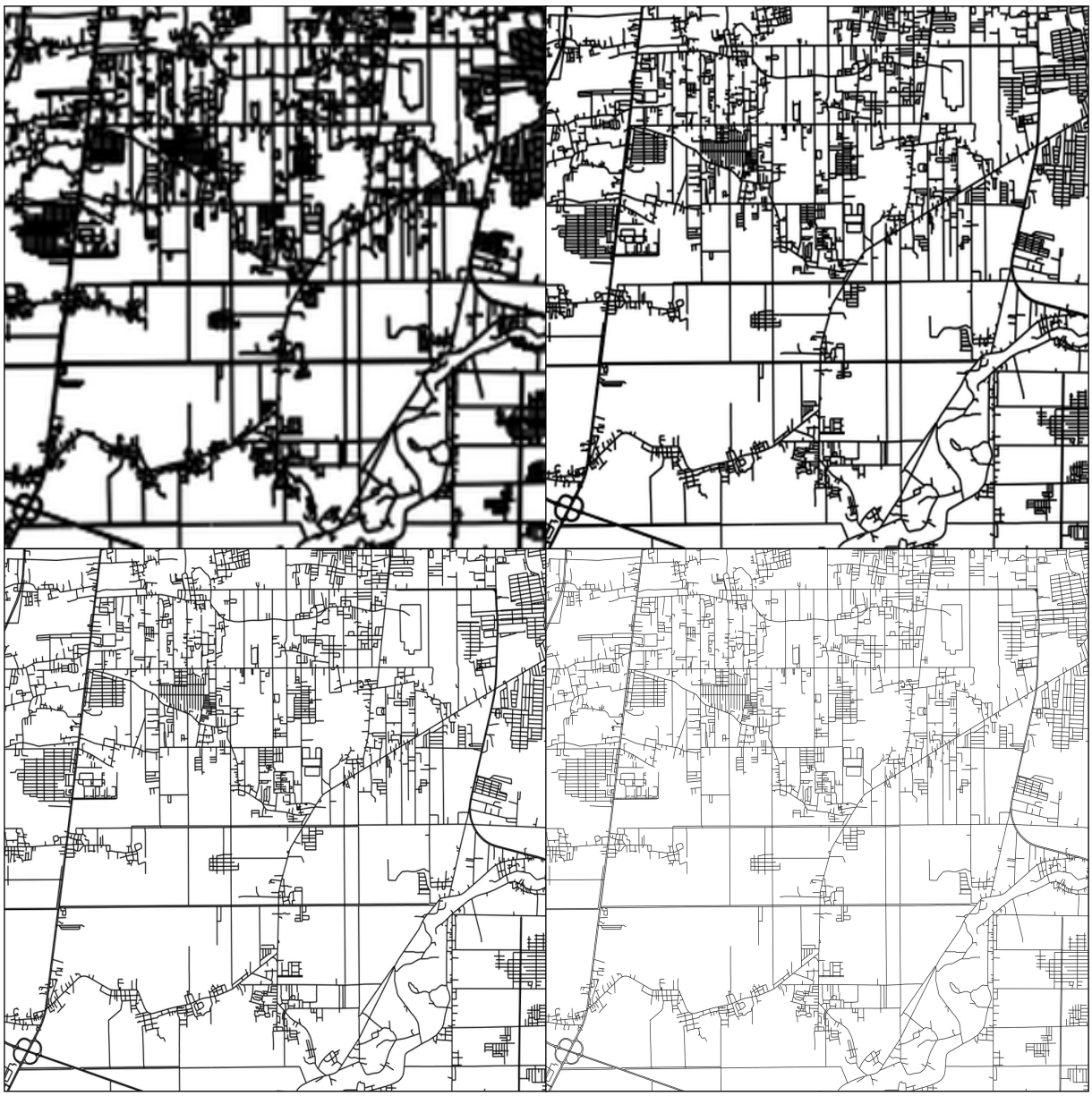

Figure 4.2: Different image resolutions of road networks in the study area. Top row: 256x256 pixels (left) and 512x512 pixels (right). Bottom row: 1024x1024 pixels (left) and 2048x2048 pixels (right).

4.2.5 Scenario development

Various urban growth scenarios are developed to explore potential forecasts for the study area by 2040. As noted in Section 2.2.6 of the literature review, a suitable approach involves generating what-if scenarios using a scenario matrix. These alternative scenarios are based on those described in Section 3.2.3, although some trade-offs were necessary, as not all nuances and details of the four scenarios can be fully captured in a computational environment. Adjustments to the urban growth ABM variables (framework stage 1) and variations in dasymetric mapping (framework stage 3) are made to reflect these four scenarios (Figure 4.1).

4.3 Urban growth modelling data preparation

4.3.1 Parcel extraction

One of the key pieces of data that needed to be acquired and digitised for this work were the land parcel boundaries in the study area. These are the spatial unit of the ABM. Land parcels provide a detailed representation of the urban landscape which represents the actual ownership boundaries directly applicable to urban planning and policy making. These are building blocks of urban planning and their irregular nature better represents the reality of urban development than a regular grid (Chakraborty et al., 2022). Historic satellite imagery of Phnom Penh reveals that land parcels often serve as a basis for urban development transactions. Imagery from 2014 was used to extract parcel boundaries, as they remained consistent over time and were frequently referenced in land negotiations during urban conversion (Figure 4.3).

Figure 4.3: Construction of residential development between 2019 (left) and 2023 (right), showing development occurs based on land parcel boundaries in red.

Data extraction and acquisition for the ABM were conducted in ArcGIS Pro, utilising Esri's pre-trained deep learning models and high-resolution satellite imagery (30-50cm). Wayback Imagery offers a digital archive of satellite imagery since 2014 from providers such as Maxar, Microsoft, and DigitalGlobe (Esri, 2024). 1000 parcel boundaries were manually digitised in the study area using high-resolution 2014 imagery as training data. An existing Esri deep learning model for parcel extraction was optimised, with a pre-trained accuracy of 73.2%. It is a Bi-Directional Cascade Network model for pixel classification and was initially trained on similar resolution imagery (40-50 cm), making it suitable for delineating land parcels (Esri, 2021). The model was optimised for 200 epochs, with a batch size of 16 and 10% of training samples were used for vali-

dation. Early stopping was implemented and the model reached a maximum accuracy of 85.47%, stopping at the 89th epoch indicating good performance.

This deep learning model delineated parcel boundaries within the study area and a 1 km buffer around it as raster pixels which were then converted to polygon vectors and tidied. Post-processing involved simplifying polygons by removing extraneous vertices, merging small polygons (less than 400 m^2) with neighbouring ones that shared the longest border, and subdividing large polygons (greater than 15,000 m^2) into smaller ones based on the average polygon size (approximately 2,465 m^2) in the study area. Figure 4.4 shows the tidied land parcels used in the ABM.

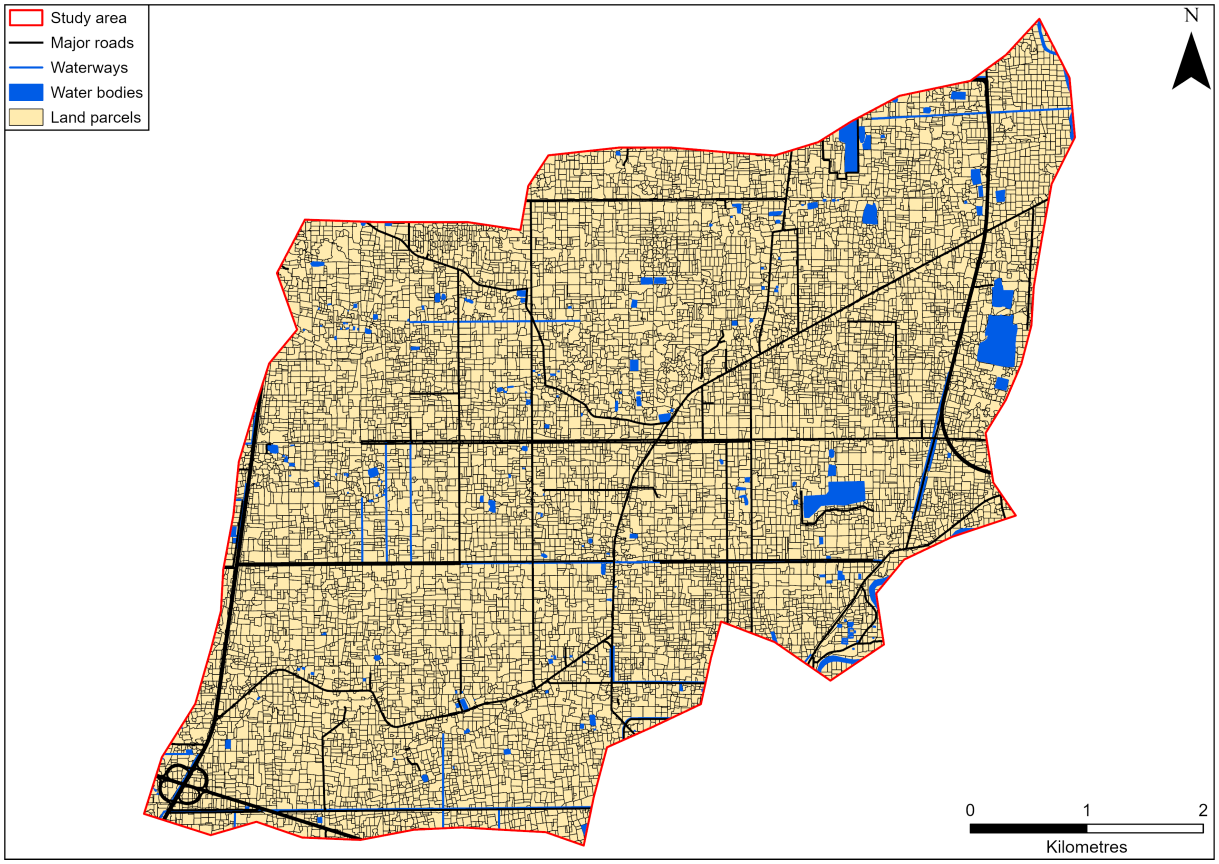

Figure 4.4: Tidied 2014 land parcels in the study area.

4.3.2 Building footprint extraction

To capture urban growth at the land parcel scale between 2014 and 2023 for model calibration and validation, the building footprints were delineated for 2014 and 2019. For 2023, the Google Buildings v3 dataset, inferred in May 2023, was used (Sirko et al., 2021). The building footprints were extracted using a deep learning model across an area of 58 million km^2 that included Africa, Southeast Asia, Latin America, and the Caribbean (Sirko et al., 2021). Although there is an OSM building footprint dataset for Cambodia, the coverage of these is largely incomplete (Figure 4.5). In Cambodia, the OSM and Google Buildings have a count of 364,292 and 13,558,239, respectively. The

Google building footprints have also been generated to an extremely high standard in Cambodia, with footprints in Cambodia having an average softmax confidence score of 73.58%, with confidence scores ranging from 60% to 90.49% (Sirko et al., 2021). Therefore, it was more appropriate to use the building footprints provided by Google. Detecting building footprints in 2014 and 2019 ensured a calibration period for the ABM from 2014 to 2019 and a validation period from 2019 to 2023.

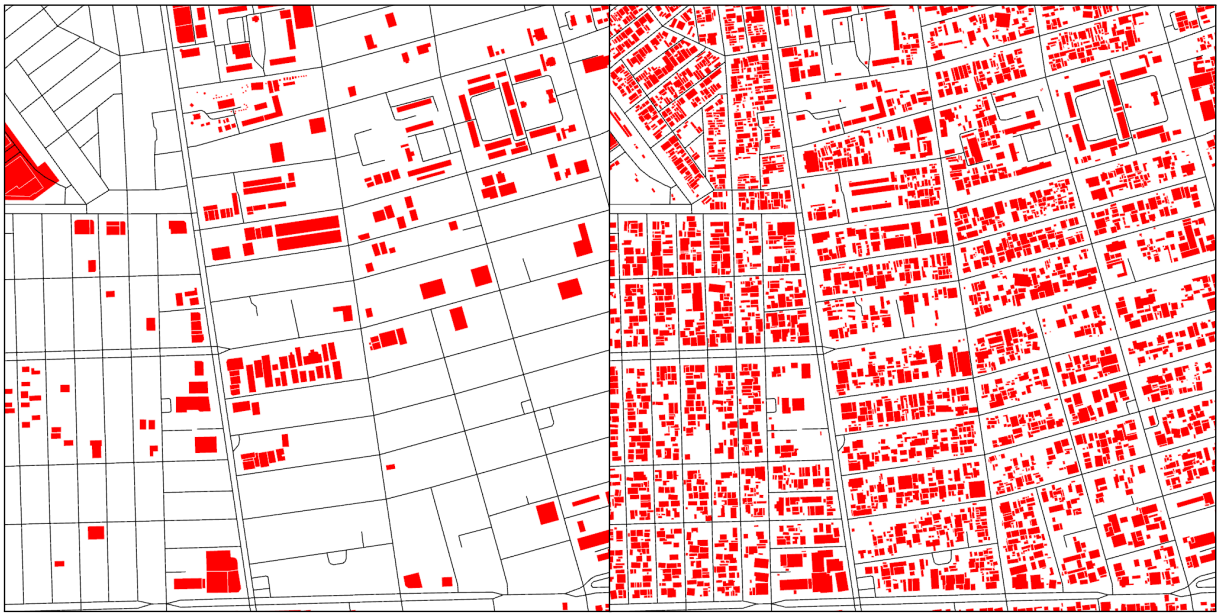

Figure 4.5: OSM building coverage (left) and Google Buildings coverage (right) in Phnom Penh.

Similarly to land parcel extraction, a pre-trained model was optimised for both 2014 and 2019 imagery. Using Google buildings data as training data, 1000 samples were taken for 2014 and 2019 for each period and used to optimise the model. These 1000 footprints were manually verified against buildings visible in the imagery from those years. A building footprint extraction model from Esri was used with a pre-trained precision of 0.718. This model uses a Mask-R-CNN architecture for object detection, enabling precise instance segmentation in images. It was previously trained on high-resolution imagery (40-50 cm). The model was optimised for 200 epochs, with a batch size of 16, using 10% of training samples for validation. Training for 2014 ended at the 49th epoch, achieving a precision of 0.746, while 2019 reached 95 epochs with a precision of 0.812. Once both models were trained, they were used to detect building footprints in 2014 and 2019 within the study area and a 1 km buffer around it. Since these building footprints were used to determine urban areas at the parcel level, precise tidying of their exact size and shape was not necessary. Minimal adjustments were made, with only spurious areas removed. Specifically, only footprints with an area greater than 10 m^3 and less than 5,000 m^2 were retained.

4.3.3 Parcel processing

Parcels covering water bodies were erased because urban development could not occur here. Now that 2014 parcels and building footprints in 2014, 2019 and 2023 had been extracted or acquired, the urban footprint of the study area could be quantified. For each year, the parcels were classified as urban if their building density was equal to or greater than 20% or had a building count of 6 or more based on visual assessment. Otherwise, they were classified as non-urban. The values for classifying a parcel as urban or non-urban were selected based on visual inspection of random parcels to see if they had sufficient building presence. Any non-urban parcels outside of the study area were removed. Figures 4.6 show the urban footprint of the modelling area in 2023. Parcels that turned urban between 2014-2019 and 2019-2023 were also classified for ABM calibration and validation. This spatial data acquisition was fundamental for the evaluating ABM performance.

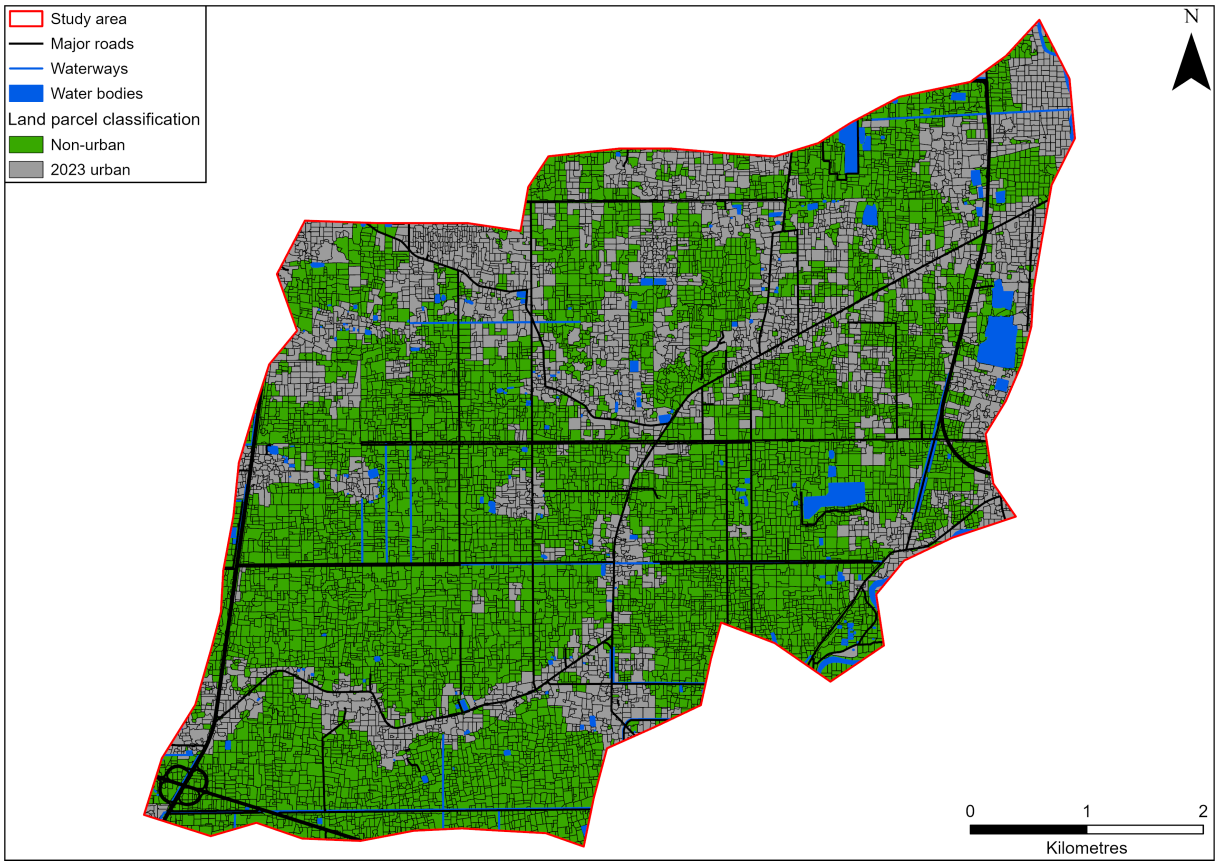

Figure 4.6: 2023 urban footprint of the study area.

For the non-urban parcels, driving factors (Table 5.1) used in the ABM were also pre-calculated, such as parcel centroid distance from Phnom Penh centre. The density of driving factors were also calculated within an 800 m radius of each non-urban parcel. The previously mentioned 1 km buffer for parcel delineation and building detection was implemented to avoid edge effects at the boundary of the study area. This was especially important when calculating the density of driving factors.

4.4 Urban form generation data preparation

4.4.1 Road network pre-processing

To train pix2pixHD to generate future road networks based on a density raster, a portion of residential roads were removed while retaining some for context. Since the study area contains both major and residential roads (Figure 4.2), preserving major roads ensures the model is properly conditioned to generate realistic networks. If trained solely on density images without existing road context, the GAN would struggle to produce meaningful outputs. By including partial road information, the model can better integrate newly generated roads into the existing network. This approach helps the GAN learn how to expand the residential road network effectively.

A fishnet was generated over Cambodia, Thailand, Vietnam, Laos, and Myanmar, as these countries share similar road network patterns. Given the size of the study area (57 km^2), a large geographic region was necessary to obtain a sufficient quantity of cells. Each fishnet cell had the same dimensions as the study area (width: 8 km, height: 7.1 km). OSM roads were used because Cambodia lacks national government-created GIS data and OSM roads are a viable alternative that have good coverage in Southeast Asia. To identify sets of residential road neighbourhoods from OSM data, the following steps were taken. This resulted in a single attribute being linked to a set of minor roads whilst the original OSM road data has an attribute per road line:

1. Filter OSM road type to non-residential and identify and remove small topologically isolated networks (fewer than 10 connected features). This removes minor disconnected segments that do not contribute significantly to the main network.
2. Merge these isolated sets with OSM residential roads and classify them as type 2 roads. This results in a coherent type 2 (minor roads) network.
3. Refine the type 2 network by identifying and removing small topologically isolated networks (fewer than 10 connected features). This ensures that only substantial and contiguous minor road network sets remain in the type 2 network.
4. With the type 2 network established, go back to the original road dataset and classify roads that do not share a segment with type 2 roads as type 1 (major roads with small sets of minor roads).
5. Finally, merge type 1 and type 2 into a complete, classified road network for further processing (Figure 4.7).

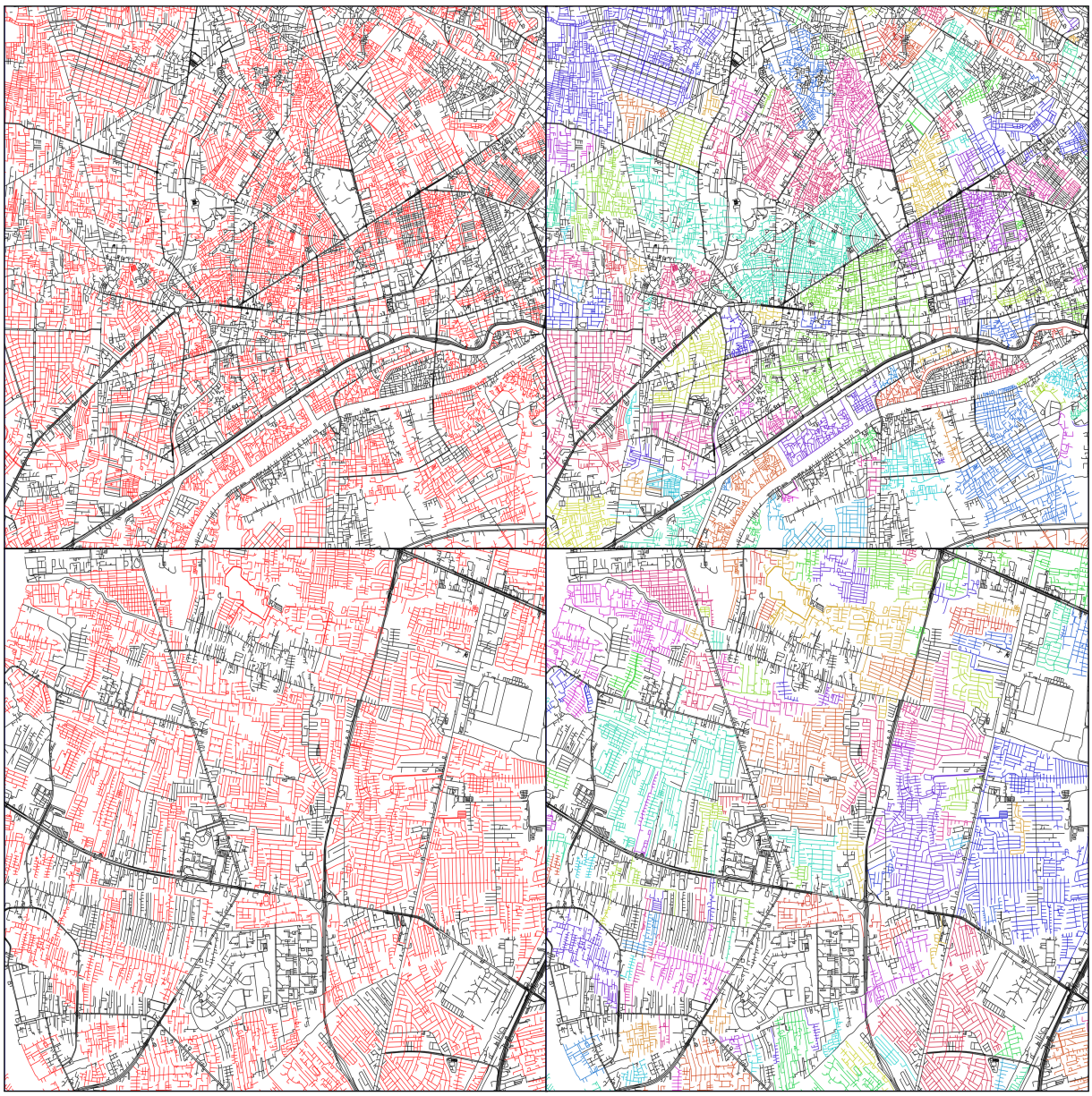

Figure 4.7: Type 1 roads in black represent major roads with small sets of minor roads and Type 2 in red represent minor residential roads (left). The different coloured roads (right) represent the individually connected sets or clusters of continuous type 2 roads.

The newly classified road network was intersected with the fishnet and type 2 density was calculated for each cell. The 2,400 cells with the highest density were selected for use in the GAN (Figure 4.8). Of these, 400 were randomly chosen for testing, while the remaining 2,000 were used for training, following the standard 80/20 train-test split. Wu and Biljecki (2023) used a maximum training size of 781 images for instantCITY, while GANmapper used up to 2,825 images, both following an 80/20 train-test split (Wu and Biljecki, 2022). Both studies found that increasing the number of training images improved performance. Given the spatial scale of images across Southeast Asia, selecting 2,400 cells for training and testing was a reasonable choice. A larger dataset could have led to less dense networks per cell, reducing the amount of meaningful information available to the GAN to learn and increasing the risk of GAN failure. Further

details on how these cells are converted to images for the various road network GAN experiments is provided in Section 6.4.

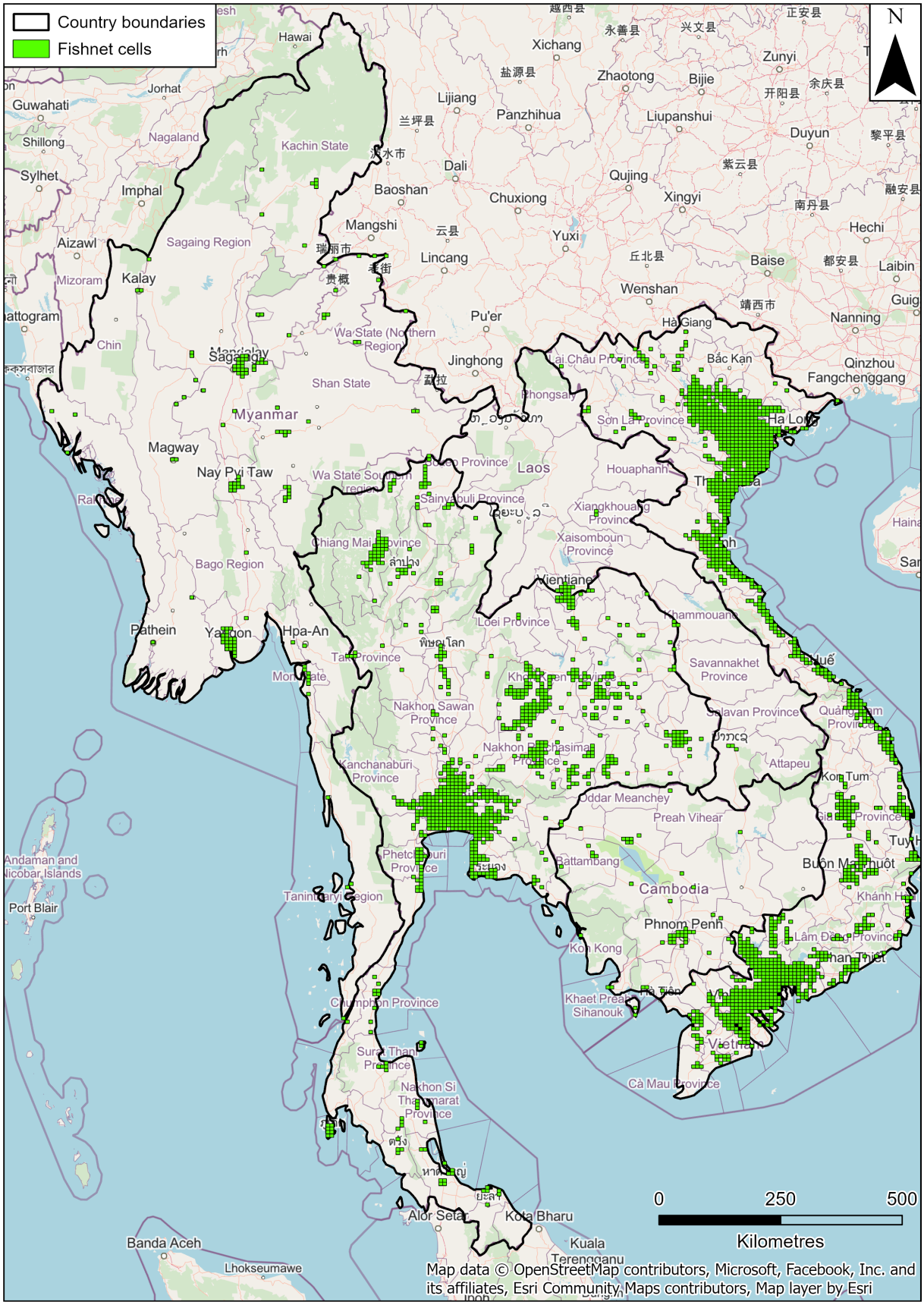

Figure 4.8: The 2,400 fishnet cells with the highest type 2 density in Southeast Asia for the road network GAN.

4.4.2 Building footprint pre-processing

Google buildings data were used as input for this part of the urban form generation. As the building footprint generation uses the road network as input, the fishnet cells for image generation could be a much smaller spatial area due to the more explicit information of the road network when compared with a density raster. Therefore, like Wu and Biljecki (2023), two fishnets with cell sizes of 500 m^2 and 1000 m^2 were created in Cambodia only because both provided a sufficient number of cells to generate training and testing polygons. Any building footprints that erroneously intersected roads were removed and the top 2400 fishnet cells for the 500 m^2 and 1000 m^2 with the highest building density were filtered for use in the GAN experiments (Figure 4.9). 400 of the cells were randomly selected for testing and the other 2000 were used for training. Further details on how these cells are converted to images for the various building footprint GAN experiments is provided in Section 6.5.

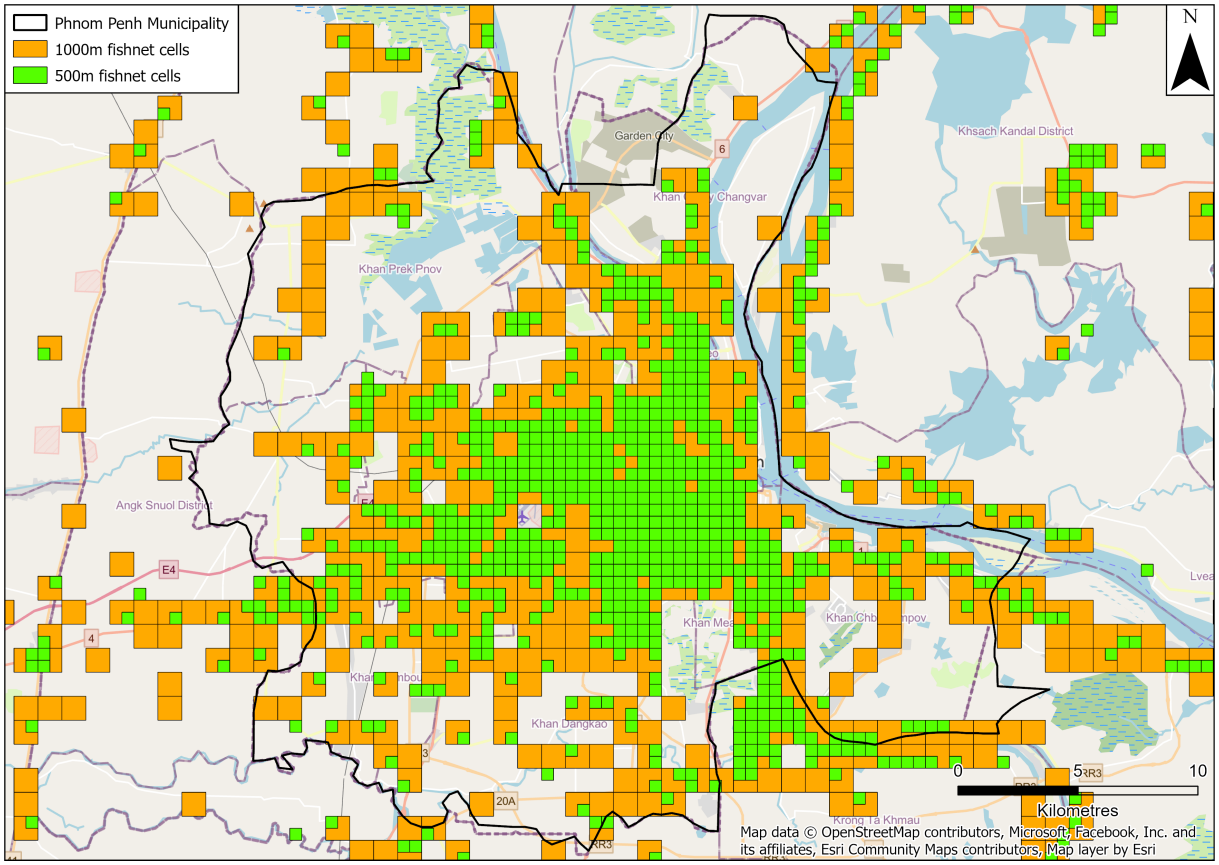

Figure 4.9: Example of the top 2400 500 m^2 and 1000 m^2 building footprints cells with the highest building density in Cambodia, zoomed into Phnom Penh.

4.5 Summary

This chapter has outlined the structure of the geospatial modelling framework, providing a guide through its four stages, beginning with the ABM and finishing with the GAN.

Justification has been given for each step. Additionally, it has outlined the data preparation required for the ABM and GAN.

Data preparation for the framework's first half, focussed on the urban growth modelling (Chapter 5), involved delineating and tidying parcel boundaries. The deep learning parcel extraction and building footprint delineation models all achieved sufficient accuracy scores and were based on pre-trained models for the same feature tasks proving their credibility. Urban extent at the parcel level in 2014, 2019 and 2023 and urban change from 2014-2019 and 2019-2023 has also been captured to assess model performance through calibration and validation and to simulate future change from 2023 onwards.

Data preparation for the framework's second half, focussed on the urban form generation (Chapter 6), involved acquiring fishnet cells with road networks and building footprints. For the road network GAN experiments, 2,400 cells have been selected based on minor road density across countries in Southeast Asia. For the building footprint GAN experiments, 2,400 cells at 500 m^2 and 1000 m^2 scale were selected based on building density. 2,000 of these cells were used to train the pix2pixHD GAN and 400 to test the pix2pixHD GAN in various experiments. All of this data preparation is required to perform the framework. Chapter 5 addresses the first two stages of the framework related to urban growth modelling and Chapter 6 addresses the subsequent two stages on urban form modelling.

Chapter 5

URBAN GROWTH MODELLING

5.1 Introduction

This chapter features the first and second stage of the modelling framework which includes developing the ABM to simulate future urban extent and then applying the ABM to an assessment protocol to ensure results are robust and credible.

First, an appropriate ABM software is selected and justified. Next, the Overview, Design concepts, and Details (ODD) protocol developed by Grimm et al. (2020) is used to design the ABM, based on behavioural theories, academic literature, and Cambodian planning documents for objective 1. Using this protocol provides a thorough outline of the purpose and key actors of the ABM, the modelling concepts of the ABM and details of the submodels and initialisation of the ABM.

The evaluation protocol for objective 2 is then presented which outlines a way to assess stochastic urban growth models. This was created to address the research gap of the lack of systematic processes to evaluate stochastic urban growth models in the literature. The protocol adopts a systematic workflow that involves preparation, stabilisation analysis, calibration and validation, and sensitivity analysis. It is applied to the ABM to demonstrate the benefits of using the protocol to extract enhanced insights and to demonstrate the credibility of the ABM for future urban growth simulation.

Finally, four alternate scenarios are generated that deviate from the baseline simulation to explore the ABM. All five scenarios are visualised and contrasted which provide potential dynamic changes in the ABM and show potential future urban pathways.

5.2 Agent-based model description

The Overview, Design concepts, and Details (ODD) protocol by Grimm et al. (2020) is a widely accepted guideline to describe ABMs. It provides a standardised, logical, and

readable framework for documenting the structure and dynamics of ABMs, making it commonly used in the ABM literature (e.g., Koch et al., 2019; Tang and Yang, 2020). The ODD protocol is divided into three categories, each serving a distinct purpose and comprising specific elements (Figure 5.1):

- Overview – provides a general overview of the model by outlining its purpose, key entities, spatio-temporal scale, and sequence of main processes.
- Design Concepts – explains the methodological concepts and characteristics relevant to the model, situating the model within a common conceptual framework. The design concepts document important ABM characteristics that are not addressed by traditional modelling techniques. The model presented in this thesis incorporates 8 of the 11 standard design concepts outlined in the protocol.
- Details - describes complete details of the model, including initial conditions, input data, and the specific processes governing agent behaviour. It provides a comprehensive breakdown of the submodels that make up the model, ensuring that the model and its results can be reproduced.

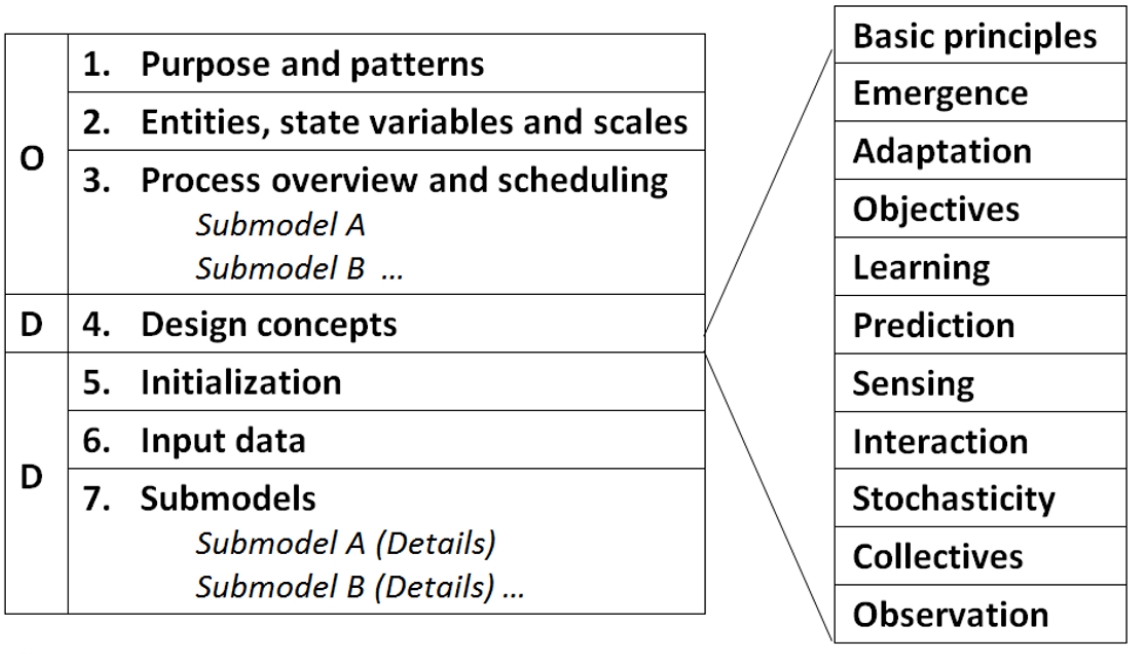

Figure 5.1: Structure of the ODD protocol (Grimm et al., 2020).

5.2.1 Overview

5.2.1.1 Purpose and patterns

The ABM's purpose is to simulate residential urban growth by replicating the Cambodian urban development process outlined in Chapter 3. It focusses on supply-side

dynamics of urban expansion by incorporating key interactions, goals, and decision-making processes of major actors that influence residential urban expansion in Phnom Penh. The model intends to capture how stakeholders, represented as agents, value land for development and interact with one another to influence urban growth.

The model's success in simulating residential growth patterns within the study area is based on its ability to capture historic growth rates and extents and the general location and direction of urban expansion. This gives the model a solid foundation and ensure its predictive power in projecting future growth. Replicating exact growth locations and patterns is not anticipated as historic growth in Phnom Penh has been sporadic and unregulated and it is unclear whether historic growth has followed the government planning processes. The utility of the model is to explore what if scenarios that explore how land development could play out if certain policies are followed. This is why there is a strong focus on Cambodian spatial planning policies in the ABM.

5.2.1.2 Entities, state variables and scales

5.2.1.2.1 Entities

The model's spatial unit are land parcels because these represent actual ownership boundaries and serve as a basis for land use transactions (Figure 4.3). The agents in this model are an urban planner, developer, and landowner. These agents were selected as integral to replicate the urban development process based on agents utilised in previous urban growth models from the literature review and the spatial planning reports. The three agents interact with each other and the land parcels. These interactions and decision-making processes are ultimately determined by the utility functions they calculate for each land parcel based on the state variables outlined in Table 5.1.

The urban planner agent represents the government, which is responsible for developing master plans every five years and ensuring that developers follow the zoning regulations established. This agent captures the institutional macro-processes shaping urban development and the role regulatory bodies have in defining top-down constraints on land use change (Parker et al., 2003). Government agents have been a popular choice in other modelling studies, which highlights their necessity in an urban growth ABM (e.g. Agyemang et al., 2023; Li et al., 2020a). Its characteristics include distance-based state variables which are standard proxies for planning suitability and reflect urban economic theory (Alonso, 1964; Diepart et al., 2016*a*). Government agents typically operate using regional knowledge and prioritise physical factors when zoning (Xu et al., 2020). The agent's decision-making for defining master plans uses a random weighted selection based on their utility function from Long and Zhang (2015).

The developer agent represents the private sector central to Phnom Penh's market-led urban expansion and is a standard inclusion in urban growth ABMs due to their role in identifying feasible development locations and initiating new urban development (Agyemang et al., 2023; Dahal and Chow, 2014). Their profit-driven behaviour aligns with rational choice theory and reflects dominant economic incentives in land conversion (Sent, 2018). The developer's state variable characteristics are same as the urban planner but only operate within master plan areas. The agent's decision-making is profit-driven with them performing utility maximising, determined by their utility function, in the zoning areas set by the urban planner. Developers are reactive to regulatory constraints but proactive in development selection. This profit-driven decision making is supported by Dahal and Chow (2014) where developers select the most profitable land based on their utility function.

The landowner agent was selected to reflect the role of individual property rights in determining development feasibility and to capture land supply constraint dynamics (Parker et al., 2003). Koch et al. (2019) also identified that many land use change ABMs focused primarily on the assumption of profit maximisation by agents. However, individual decisions regarding land ownership and management are not solely driven by profit (Groeneveld et al., 2017). Landowners are therefore included to represent different priorities to the other two agents. Landowners characteristics are defined by density-based state variables calculated at the neighbourhood scale. Their decision-making combines bounded rationality with stochastic elements. When approached by a developer, they assess utility within a radius and decide whether to sell based on willingness to sell and local peer influence. This logic is based on Tang and Yang (2020) who argues that it reflects the complexity and unpredictability of real-world landowner behaviour, which ensures heterogeneous landowner agents.

Collectively, the urban planner, property developer, and landowner agents were selected to capture supply-side dynamics of urban growth. The ABM's purpose is to simulate key decision-making processes behind urban land conversion, rather than to completely replicate a housing market. Demand-side consumer agents, such as residents and households, were excluded due to the ABM's focus on supply-side processes and limited data to define these agents (Burke et al., 2024; Dahal and Chow, 2014). The urban planner acts to regulate the supply of developable land through zoning. The developer initiates supply of urban land by selecting profitable parcels within zoning areas. The landowner acts as a gatekeeper of supply through their decision to sell to the developer which enables the land to be converted to urban. This sequence demonstrates that the agents are linked decision-making entities in the development process. The combined characteristics, dynamics, and decision-making processes of these agents are fundamental to replicate urban growth in the study area.

5.2.1.2.2 State variables

State variables are attributes that characterise the current state of entities in an ABM. Table 5.1 displays the state variables for each of the entities mentioned previously. Section 5.2.3.3 describes where the state variables are used throughout the model.

Table 5.1: State variables of entities in the ABM.

Land parcels				
State variable	**Dynamic**	**Type**	**Value(s)**	**Description**
area	No	Int	100-16,945	Area in m^2.
urban	No	Bool	true/false	If a parcel is existing urban.
new_urban	Yes	Bool	true/false	If a parcel is simulated urban.
build_year	Yes	Int	2023-2040	Year simulated parcel was built.
Urban planner agent				
State variable	**Dynamic**	**Type**	**Value(s)**	**Description**
pl_eurb_ds	No	Float	0-1	Distance from existing urban.
pl_rd_ds	No	Float	0-1	Distance from major roads.
pl_cbd_ds	No	Float	0-1	Distance from Phnom Penh.
pl_wtrcrs_ds	No	Float	0-1	Distance from waterways.
pl_wtrbdy_ds	No	Float	0-1	Distance from water bodies.
pl_nurb_ds	Yes	Float	0-1	Distance from simulated urban.
pl_uf	Yes	Float	0-1	Planner utility function.
Property developer agent				
State variable	**Dynamic**	**Type**	**Value(s)**	**Description**
dv_eurb_ds	No	Float	0-1	Distance from existing urban.
dv_rd_ds	No	Float	0-1	Distance from major roads.
dv_cbd_ds	No	Float	0-1	Distance from Phnom Penh.
dv_wtrcrs_ds	No	Float	0-1	Distance from waterways.
dv_wtrbdy_ds	No	Float	0-1	Distance from water bodies.
dv_nurb_ds	Yes	Float	0-1	Distance from simulated urban.
dev_uf	Yes	Float	0-1	Developer utility function.
Landowner agent				
State variable	**Dynamic**	**Type**	**Value(s)**	**Description**
ow_eurb_dens	No	Float	0-1	Density of existing urban.
ow_rd_dens	No	Float	0-1	Density of major roads.
ow_wtrcrs_dens	No	Float	0-1	Density of waterways.
ow_wtrbdy_dens	No	Float	0-1	Density of water bodies.
ow_nurb_dens	Yes	Float	0-1	Density of simulated urban.
ow_uf	Yes	Float	0-1	Urban planner utility function.

5.2.1.2.3 Scales

As mentioned in Chapter 4, edge effects were mitigated when pre-calculating the state variables (e.g. distance from major roads, density of existing urban parcels) for each agent by applying a 1 km buffered bounding box around the immediate study area, resulting in a total modelling area of 90.95 km^2. The urban simulation was only in the core area of 33.3 km^2 core area of the adjoining communes of Pong Tuek, Prey Sa, and Prey Veaeng in Phnom Penh (Figure 3.4). Each land parcel in the study area has an average size of 2465 m^2. One time step in the ABM represents one year, with the simulation starting in 2023 and ending in 2040, running for 18 years or 18 simulation cycles.

5.2.1.3 Process overview and scheduling

Figure 5.2 shows a model schedule flow chart that specifies the chronological order of each submodel per time step. The schedule has been broken down into four submodels. A summary of the agent's actions that occur in a time step and the frequency of submodel execution is provided below:

1. Urban planner utility function calculation submodel - At the beginning of the first time step, the urban planner agent calculates their utility function using their state variables defined in Table 5.1.

2. Parcel zoning submodel - the urban planner randomly selects (weighted by their utility function) initial seed parcels for the master plan. The neighbouring parcels of the seed parcels are continuously added to the plan until land area demand for the next five years is met. These first two submodels are only run every five simulation time steps.

3. Developer and owner utility function calculation submodel - the developer and owner agents then calculate their utility functions for each parcel from their state variables in Table 5.1.

4. Land transactions submodel - Developer agents select parcels with the highest utility values within the master plan until total area of the selected parcels meets the annual land demand. The developer agent interacts with the owner agent of these parcels and a successful land transaction occurs and turns the parcel to urban if the owner's selected utility function value is greater than a random probability.

The first and second submodels are only executed every 5^{th} time step with the ABM starting a new time step from submodel three if 5 time steps have not elapsed. The rationale and complete details of each submodel is provided in Section 5.2.3.3.

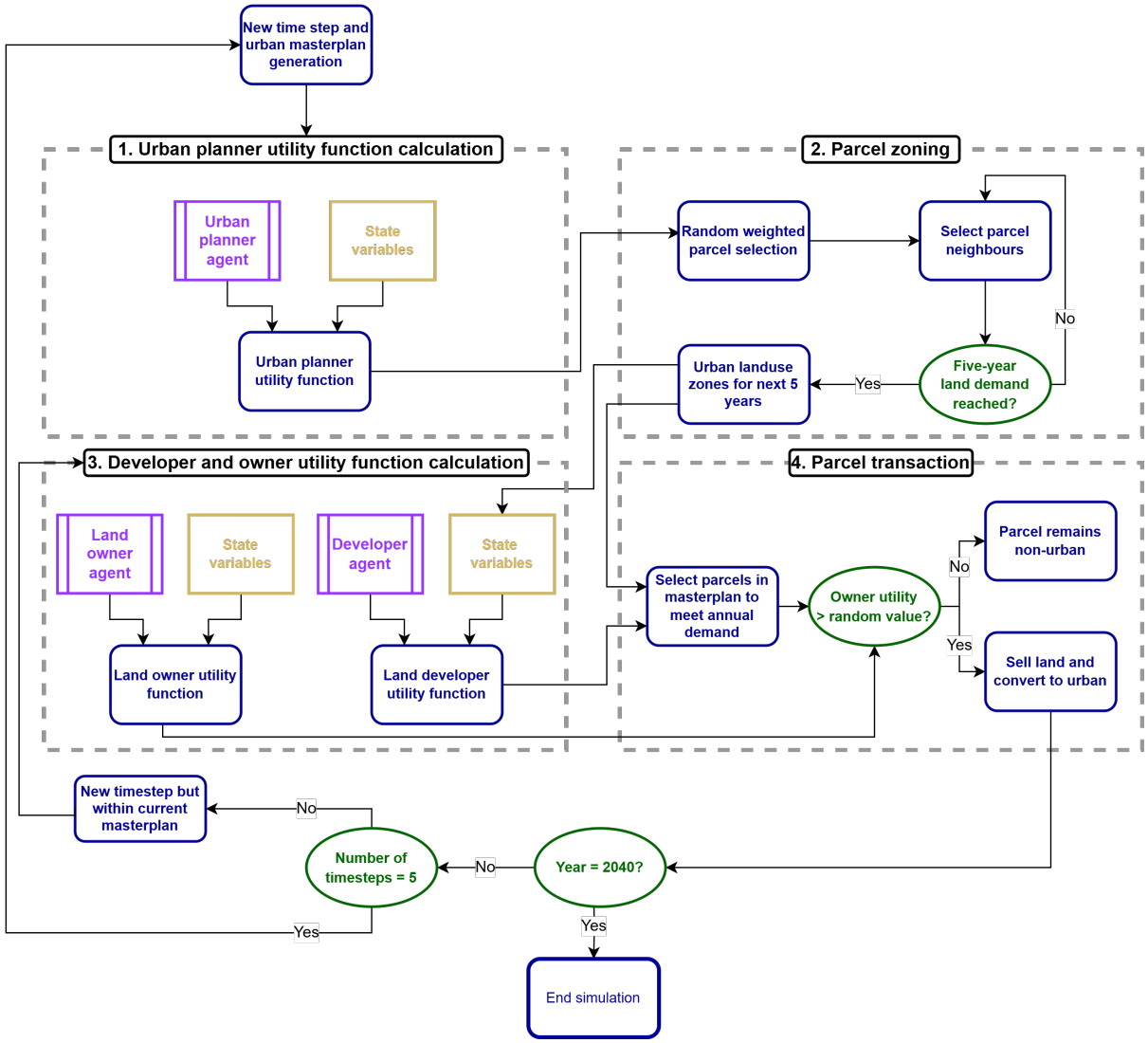

Figure 5.2: Overview of the ABM design showing the submodel structure and schedule.

5.2.2 Design concepts

5.2.2.1 Basic principles

This ABM is based on relevant theories and techniques used in similar ABM studies, providing a solid foundation for simulating urban growth in Phnom Penh. The key principle of this ABM is a bottom-up approach that replicates the urban development process in Phnom Penh. Given the absence of readily available empirical surveys and detailed census data, the ABM uses relevant theories outlined in Section 2.2.5 to determine agent behaviour and decision-making.

At the global level, the model employs urban economic theory generally to explain the varied valuation of land parcels based on proximity to city centres, infrastructure, and urban areas (Fujita, 1989). In particular, it uses bid-rent theory which posits that land users compete for locations closer to the city centre, where access to markets, jobs, services, and transport is most advantageous (Alonso, 1964). The model simulates this

competition by using utility functions that reflect the desirability of different land parcels based on their characteristics and location, mirroring the real-life economic trade-offs.

At the agent level, the model uses rational choice and bounded rationality theory (Sent, 2018). Rational choice theory assumes that individuals make decisions by maximising their utility using all available information and preferences. However, bounded rationality suggests that people make decisions with limited information, cognitive constraints, and time restrictions (Sent, 2018). Urban planners use rational choice and bounded rationality theory together when developing master plans for development. In contrast, developers use rational choice theory to maximise profits when selecting parcels. Landowners revert back to bounded rationality by deciding whether to accept a developer's offer by taking into account their neighbour's preferences and the associated risks of selling. These theories are modelled using maximising and satisficing techniques of the utility functions to encapsulate the agents' preferences and the constraints they face, allowing the model to capture the complex interplay between individual decision-making and broader urban growth dynamics.

5.2.2.2 Emergence

The emergent outcomes of the model include the spatial patterns and locations of urban land use that evolve in the simulation. These patterns are not explicitly programmed or predetermined, but rather arise from the adaptation and interactions of the urban planner, developer, and landowner agents. This results in unpredictable, dynamic, and organic urban growth patterns. This emergent system reflects the complex realities of urban growth, where diverse stakeholders with varying objectives and information shape the urban landscape's evolution.

5.2.2.3 Adaptation

The urban planner, developer, and landowner agents demonstrate adaptive behaviour by adjusting their decisions based on changes in their environment. All agents adapt their decisions based on simulated urban developments, which are incorporated into their utility functions. This dynamic adjustment reflects realistic decision-making processes within the evolving urban landscape. If certain areas experience significant development, planners can adapt by zoning adjacent parcels for future growth, ensuring continued urban expansion. In contrast, areas where development fails to occur may be deprioritised in future zoning decisions by planners. Developers adapt by developing only within the allowed zoning areas set by urban planners and use a ranking system to select the best urban parcels within a zoning area. Landowners are influenced by their neighbours' utility, adjusting their decisions to sell based on local development.

5.2.2.4 Objectives

In this model, each agent has specific objectives that guide their decision-making processes. These objectives are quantified using utility functions that serve as the primary measure for evaluating decision options. These objectives ensure that the agents' actions are aligned with their interests, driving the overall dynamics of urban development within the model.

- Urban planner - their objective is to optimise land use by identifying parcels that maximise overall urban growth potential. They perform utility satisficing and maximising when selecting parcels for master plan creation to achieve this objective.
- Developer - their objective is to maximise profit by selecting parcels that offer the highest potential return on investment. They are driven by utility maximising by acquiring parcels with the highest utility values in the master plan.
- Landowners - their objective is to make rational, profitable decisions when selling their parcels to developers. Their choices depend on local factors, neighbouring parcel values, and a bounded rationality approach.

5.2.2.5 Sensing

Each agent has varying levels of information about their environment due to their sensing abilities. The variation in sensing capabilities between agents introduces realism, as it is unlikely that all agents have complete knowledge of the study area (Koch et al., 2019). Urban planners have global knowledge, so their state variables refer to distance from driving factors (Xu et al., 2020). This complete spatial awareness ensures that their zoning decisions are well-informed. However, developers are restricted to zoning areas, which means that they have less autonomy than planners to select the best parcels for development, but are driven by the same state variables (Table 5.1). Landowners have a more localised sensing capability. They primarily sense the density of state variables within an 800 m neighbourhood radius of their parcel to calculate their utility function (Koch et al., 2019). In addition, landowners are influenced by neighbour utility values that they sense and incorporate into their decision-making processes.

5.2.2.6 Interaction

This ABM features numerous interactions between agents to represent how decision-making processes impact urban growth. One primary interaction occurs between landowners and developers. Developers contact landowners to purchase their parcel and landowners decide whether to sell it based on a negotiation process that determines whether a parcel transitions from non-urban to urban use. Another interaction is the zoning enforcement by urban planners on developers so they can only develop in parcels within

the approved master plan. This interaction ensures that development occurs in areas that align with the urban planner's strategic goals, thereby directly influencing the decisions that developers can make about land acquisition. Owner agents also interact with each other by exchanging utility function information when deciding whether to sell their land parcel.

5.2.2.7 Stochasticity

The stochastic components of urban growth models play two roles: to reflect the uncertainty and heterogeneity of real-world processes and to ensure that every model run can result in a different output (van Vliet et al., 2016; Grimm et al., 2020). Urban planners use a random weighted selection method to choose a parcel to develop the master plan. This reflects the unpredictable nature of zoning decisions due to factors such as political pressure, unforeseen environmental changes, or stakeholder influences. Landowners use a random probability approach, when they are approached by developers. This means that there is still a chance that a landowner may choose not to sell, even if their utility function is high, due to reasons such as land attachment. This stochasticity aims to acknowledge that human behaviour is inherently unpredictable and is influenced by a wide range of factors (Tang and Yang, 2020).

5.2.2.8 Observation

A graphical user interface provided in the GAMA platform which displays a map of the study area is used to observe simulations (Figure 5.3). Outputs generated during the model evaluation were saved as shapefiles or text files for further processing, visualisation, and analysis.

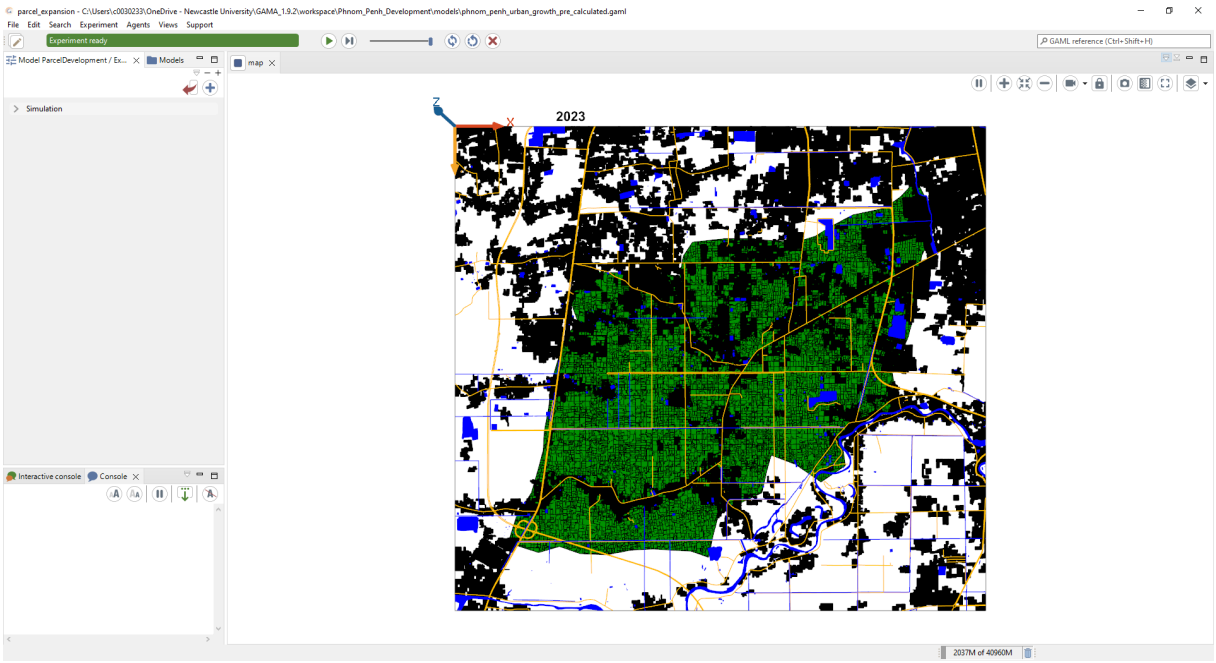

Figure 5.3: Graphical user interface of GAMA Platform.

5.2.3 Details

5.2.3.1 Initialisation

The ABM environment was initialised by loading the land parcels GIS shapefile as an entity. This defined the urban and non-urban footprint of the study area and served as the initial state of the simulation. To provide context, additional GIS datasets, including roads, water bodies, and waterways, are also loaded (OpenStreetMap contributors, 2024). During initialisation, none of the static agent state variables (e.g. distance from major roads, density of existing urban areas) are calculated within the ABM, as these have been pre-processed externally to enhance model efficiency. However, normalisation and utility function calculations occur within the ABM in submodels 1 and 3 (Figure 5.2). The normalisation values of the state variables depend on the urban planner's selection for the master plan, which varies in each simulation run. The utility function calculations must account for newly simulated urban development as it unfolds, meaning it must be dynamically updated through the simulation.

5.2.3.2 Input data

During the simulation, time series input was not required once the model has been initialised. However, the new urban dynamic state variables for the three agents was modified during the simulation due to changes in the study area environment resulting from the ABMs emergent behaviour.

5.2.3.3 Submodels

A full description and rationale for the four submodels are provided here. Table 5.2 outlines the global model parameters that are used throughout the ABM.

Table 5.2: Global model parameters. Uncertain parameters in the value column are defined in calibration.

Model parameter	Type	Value	Description
ow_nbhd	Int	800 m	Landowner neighbourhood area.
mstr_plns	Int	Uncertain	Number of master plan seeds.
one_yr_lnd_dmd	Int	Calculated from five_yr_lnd_dmd	Yearly land area demand.
five_yr_lnd_dmd	Float	2.57/4.17 km^2	Five-year land area demand.
wts	Float	Uncertain	Landowner willingness to sell.

5.2.3.3.1 Urban planner utility function calculation submodel

This submodel begins at the first time step and repeats every five time steps (Figure 5.2). The urban planner state variables (Table 5.1) were normalised on a scale of 0-1 for all non-urban parcels using the following equation:

$$sn = \frac{s_{\text{max}} - s}{s_{\text{max}} - s_{\text{min}}}. \tag{5.1}$$

Where s is the original state variable value for a parcel, sn is the normalised value, and s_{max} and s_{min} are the maximum and minimum values of that state variable, respectively. This normalisation approach assumes that lower values of the state variable are more desirable, which is appropriate for the planner's distance-based state variables. These normalised state variables were then used to calculate the planner's utility function for all non-urban parcels. The urban planner's utility function ($planner_{uf}$) for a given parcel is defined as the weighted average of its normalised state variables with the following equation:

$$planner_{uf} = \frac{\sum_{i=1}^{N} w_i \cdot sn_i}{\sum_{i=1}^{N} w_i}. \tag{5.2}$$

Where sn_i is the normalised value of state variable i for the parcel, w_i is the corresponding weight for state variable i and N is the total number of the planner's state variables. This generated a suitability layer for development from the urban planner's perspective and defined their decision-making in the next submodel. As the urban planner represents the regional authority, they are considered to have complete knowledge of all urban development driving factors in the study area. Therefore, their state variables are selected to represent a global understanding of the whole study area and they assess all parcels in the region to calculate their utility function. During the development of the Commune Plan, officials are tasked with performing situation analysis such as acquiring elevation data, population, current land use and land tenure data, topographic data, population growth, and population density data (Diepart et al., 2016*c*). This information provides the basis for the state variables used in the utility function.

5.2.3.3.2 Parcel zoning submodel

The time step dynamics of the land use zoning model are the same as the urban planner utility function calculation submodel. Figure 5.4 shows that colours were used to represent the status of parcels in the ABM, with green and black parcels representing available non-urban and urban, respectively. As the urban planner performs utility satisficing based on bounded rationality, based on Long and Zhang (2015) the agent performs a random weighted sample using their utility function to select non-urban parcels (Figure 5.4, step 2). The selected non-urban parcels turn purple and are the seed parcel in the master plan. Its neighbouring parcels continue to be added to the plan until

the five-year land demand is met (Figure 5.4, step 3). The 1 and 2 numbers in Figure 5.4 represent the iteration number of the neighbours to the stating parcel. All of these purple parcels are master plan potential parcels and the planner selects the parcels with the highest utility function (utility maximising) until the five-year land demand is met to avoid overestimation of the master plan's area. (Figure 5.4, step 4). These parcels make up the final master plan and be converted to blue. The mstr_plans and five_yr_lnd_dmd parameters from Table 5.2 operate in this submodel, which define how many master plan seeds are laid out by the urban planner and the total land area they occupy, respectively. The mstr_plans are defined during model evaluation. Whereas, the five_yr_lnd_dmd parameter is set at 2.57 km^2 for calibration and 4.17 km^2 for validation as these were the observed urban footprint increases for both time periods. For future simulations from 2023 onwards, it is kept at 4.17 km^2.

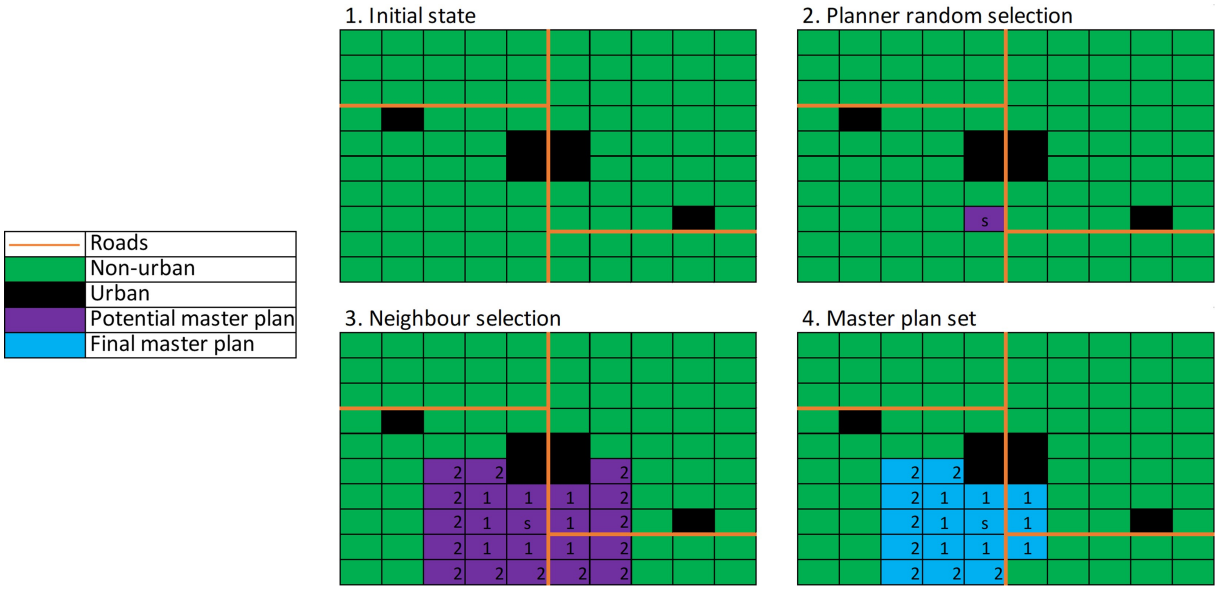

Figure 5.4: Visual description of parcel zoning submodel shown in a conceptual study area.

In Cambodia, land use zoning is defined by regional authorities, and developers must comply with the regulations established. Land is divided into two types, areas where development is permitted and controlled areas. Spatial plans are legally binding through the Law on Land Management and Urban Planning, these regulations on land use plans are part of the legal framework, which directly affects the eligibility of any development within the scope of the plans (Diepart et al., 2016*a*). This submodel therefore focusses on developing land use zones for the study area to develop over the next five years. This five-year land demand rule is based on the Commune's spatial plan and their responsibility to develop the five-year development and land use plan (Diepart et al., 2016*c*). It is logical for the ABM to begin with the urban planner setting the urban land use zones, as this regulates where development can occur in the land transaction submodel. The National Committee for Land Management and Urban Planning in Cambodia has developed planning procedures that include the overall scope and the desired content of each subnational spatial plan. The development of master plans is complex because it

involves environmental analysis, stakeholder consultation, scenario analysis, technical reports, and approval by the national government (Diepart et al., 2016*c*). This complex planning process is difficult to capture in a computational model, so a stochastic process was used to simplify the creation of the master plan. Grimm et al. (2020) argues that stochastic processes are appropriate to use when it is not possible to model a processes complex mechanisms.

5.2.3.3.3 Developer and landowner utility function calculation submodel

This submodel is run from the second time step onwards once the urban plan has been defined for the next five years. The developer's state variables were normalised onto a scale of 0-1 for the master plan parcels using Eq. 5.1 like the planner's normalisation as the developer has distance-based state variables where lower values would be more favourable. Whereas, the landowner's state variables were normalised on a scale of 0-1 (Table 5.1) for the master plan parcels using the following equation:

$$sn = \frac{s - s_{\text{min}}}{s_{\text{max}} - s_{\text{min}}}. \tag{5.3}$$

Where s is the original state variable value for a parcel, sn is the normalised value, and s_{max} and s_{min} are the maximum and minimum values of that state variable, respectively. This normalisation approach assumes that lower values of the state variable are more desirable, which is appropriate for the owner's density-based state variables. The developer's and owner's normalised state variables were then used to calculate the utility function for the master plan parcels. The developer's ($developer_{uf}$) and owner's ($owner_{uf}$) utility function for a given parcel is defined as the average of its normalised state variables using the following equation:

$$developer_{uf}, owner_{uf} = \frac{1}{N}\sum_{i=1}^{N} sn_i. \tag{5.4}$$

Where sn_i is the normalised value of state variable i for the parcel and N is the total number of either agent's state variables. The developer uses the same state variables as the urban planner as they are considered to also have a global knowledge of real estate conditions in the study area due to having a greater expertise and knowledge of development potential likelihood. However, the owner agent only has a limited neighbourhood perception of the study area as the landowners are most interested in what their neighbours are doing. The owner agent uses separate state variables based on density within a defined neighbourhood of each master plan parcel in the study area. Here, the ow_nbhd parameter is used from Table 5.2 to define the landowner's search radius. The owner's state variables were pre-calculated within an 800 m radius before initialisation of the model. This represents the 15 minute neighbourhood area and the limited sensing range of owner agents (Feng et al., 2020; Shan and Zhu, 2011).

5.2.3.3.4 Parcel transaction submodel

This submodel begins with the developer agent selecting the master plan parcels with their highest utility value until the annual land demand is reached (Figure 5.5, step 1) to represent utility maximising. The developer then approaches the landowner of the parcels they want to acquire. The landowner selects a random owner utility function value from all the parcels selected by the developer and compare it to a random probability within the same range of the owner utility function pool (utility satisficing). If the utility value is greater than the random probability, the owner agrees to sell the land and the sold parcel is turned urban. The decision pool is introduced to capture the variability and diversity in decision-making among landowner agents (Tang and Yang, 2020). Landowners make different decisions even under similar circumstances. The year before the next master plan generation, parcels that were not converted to urban revert back to non-urban parcels. The random probability is weighted by the wts parameter (Table 5.2) with a higher value making the transaction less likely (Koch et al., 2019).

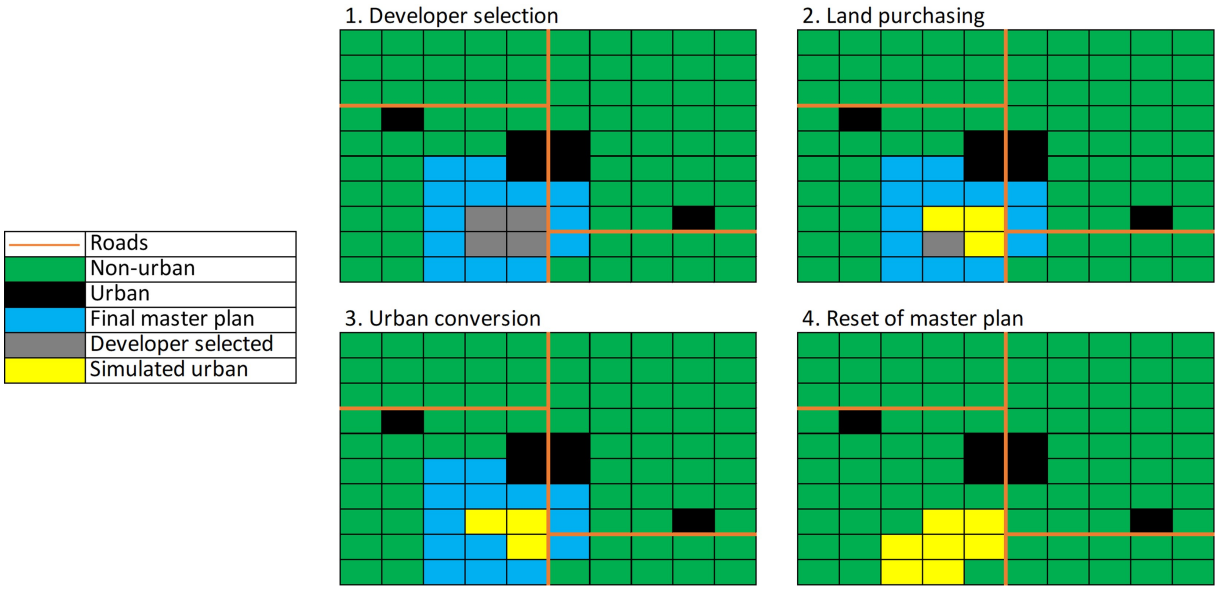

Figure 5.5: Visual description of land transaction submodel shown in a conceptual study area.

This submodel is intended to capture the heterogeneous decision-making process between a macro and micro agent that function at two different scales to understand how urban expansion is affected by their interactions. Koch et al. (2019) assessed that human decision-making can greatly influence where urban development can occur due to heterogeneous priorities. However, most academic research has focused solely on decision-making from the perspective of developer preference's for suitable locations (e.g. Alghais and Pullar, 2018; Xu et al., 2020). It should not be assumed that parcels popular among developers automatically convert to urban as it also depends upon the landowner's preferences and perspective of land value and risk. It is also important to consider the landowner's decision making process and improve the understanding of the bottom-up processes that generate urban growth patterns (Koch et al., 2019).

5.3 The stochastic urban growth model evaluation protocol

As identified in the review of the literature, many urban growth models lack clarity and detail about how the model was assessed. Therefore, a thorough and robust model evaluation procedure has been developed to enable a good understanding of the ABM developed and ensure applicability to a real-world issue. This procedure proposed is titled the **S**tochastic urban gr**O**wth mode**L** e**V**aluation (SOLV) protocol which features a four-stage process which is demonstrated on this ABM: preparation, stabilisation analysis, calibration and validation, and sensitivity analysis.

5.3.1 Preparation

The following criteria must be defined in the SOLV protocol:

- Output of interest - this can be a model performance metric, used to assess predictive accuracy. This protocol recommends the use of the FKS coefficient (Section 2.4.2.2) because it provides a nuanced approach that quantifies the agreement between simulated and observed outputs (van Vliet et al., 2013). However, it can be an output unrelated to model performance, such as quantity of urban change or number of unsold parcels, based on the model's purpose.
- Evaluation period - the time frame for the evaluation period to explore the model. This should be a historical period so that the model can be benchmarked against observed data.
- Calibration and validation - The evaluation period should be equally divided, either temporally or spatially, into calibration and validation phases. This prevents the model from overfitting the observed data and ensures that its predictive performance is not compromised.
- Variable selection and ranges - as the protocol explores model parameters and agent driving factor weights (collectively referred to as model variables), it is necessary to define a range of possible values for the model variables being explored. The increments should be equal and have a sensible gap between them. For example, if historical yearly growth is 2 km^2, then two values on either side of this value should be inputted into the range, such as a 10% buffer on either side of this value.

As the success of the ABM was defined in Section 5.2.1.1 as simulating historic growth rates, extents, location, and direction of urban expansion rather than exact spatial matches, the FKS coefficient was selected as the output of interest. FKS enables a

more nuanced and lenient evaluation by allowing fuzzy classifications between urban and non-urban areas, making it a more flexible and realistic measure of agreement. A strict cell-by-cell comparison would be inappropriate, as it demands exact spatial matches between simulated and observed data, which is not the primary objective of the model. When defining the fuzzy distance, it is typically set at a 4-pixel radius for gridded models (Guzman et al., 2020; Visser and De Nijs, 2006). Since the land parcels are irregular in this ABM, the average parcel area of 2,465 m^2 was used. Assuming the parcels are square, the average parcel length would be approximately 49.65 m. The fuzzy distance was therefore set at 200 m to replicate the radius of approximately four parcels. In addition, the FKS calculation was weighted by parcel area, so if larger parcels matched between the simulated and observed data, this was classified as a greater agreement compared to smaller parcels. This adequately handled the irregular nature of the land parcels.

As framework data preparation involved extracting historical changes of the study area between 2014-2023, this was set as the evaluation period, with calibration from 2014-2019 and validation from 2019-2023. For variable exploration, the two uncertain model parameter values in Table 5.2 were explored to optimise the quantity of urban growth and the state variables of the urban planner for the calculation of utility functions were also introduced for weighting in terms of relative importance to help optimise the simulated urban location. Table 5.3 presents the range of these variable values used to explore the model. For example, mstr_plns was set to a value between 25 and 100, with steps of 25. It could therefore be 25, 50, 75, or 100 during model evaluation (Table 5.3).

Table 5.3: Model variables exploration ranges used in stabilisation and calibration.

Model variable	Description	Min/Max	Step
mstr_plns	Number of master plan seeds.	25 / 100	25
wts	Landowner willingness to sell	0.2 / 1.0	0.2
pl_eurb_ds_w	Planner distance from existing urban weight	0.2 / 1.0	0.2
pl_rd_ds_w	Planner distance from major roads weight	0.2 / 1.0	0.2
pl_cbd_ds_w	Planner distance from Phnom Penh weight	0.2 / 1.0	0.2
pl_wtrcrs_ds_w	Planner distance from watercourses weight	0.2 / 1.0	0.2
pl_wtrbdy_ds_w	Planner distance from water bodies weight	0.2 / 1.0	0.2
pl_nurb_ds_w	Planner distance from new urban weight	0.2 / 1.0	0.2

5.3.2 Stabilisation analysis

The first stage of the SOLV protocol integrates the variance stability and uncertainty analysis procedures (Section 2.4.1) into a stabilisation analysis. This stage aims to

identify when the model's stochasticity has stabilised based on an appropriate number of runs and variable space exploration so that random variations are statistically insignificant. The following equation by Tang and Jia (2014) was used to determine the initial number of runs to use in the stabilisation analysis:

$$N_{\text{run}} = (2k + 2) \cdot n_{\text{mc}} \tag{5.5}$$

Here k is the number of variables and n_{mc} is the number of Monte Carlo runs. Tang and Jia (2014) suggests that the number of model runs for spatial ABMs is set between 100-500. Based on the number of runs, LHS divides each variable's range into equal intervals in Table 5.3. In each run, the procedure randomly selects one interval per variable, draws a variable value within that interval's boundaries, runs the model over the evaluation period, and then excludes that slice from further selection. This process continues until all intervals have been used. LHS ensures a thorough and well-distributed exploration of the variable space while maintaining randomness. This approach prevents any particular variable configuration from dominating the results, reducing bias and improving coverage compared to simple random sampling (Lee et al., 2015). Furthermore, since the evaluation period (2014-2023) is being used for this part, the five-year land demand was set at 3.37 km^2 as the average land demand value over calibration and validation.

After generating a number of runs output of interest values, their order was shuffled and reordered the same amount of times to eliminate any potential order bias. For each shuffle, the CV (Eq. 5.6) was calculated in groups of 30 runs up to the final run. The CV, which shows the extent of variability in relation to the mean of the population, helps to assess the consistency of the stochastic simulation results (García et al., 2011; Lee et al., 2015).

$$\text{CV} = \left(\frac{\sigma}{\mu}\right) \times 100\%. \tag{5.6}$$

In this equation, σ is the standard deviation and μ is the mean of the data set. The differences between successive CV values were calculated for each shuffle, starting with the CV difference between the first 30 and 60 runs. After completing the required number of shuffles, the median CV difference across all shuffles was computed for each sample size pair. For example, the median difference between 30 and 60 model runs based on all shuffles. To determine the stabilisation point, the Knee Locator algorithm was applied to the median CV differences. This algorithm identifies the point where the rate of change in CV differences decreases significantly, indicating the sample size at which additional runs provide diminishing reductions in variability. This stabilisation point represents the minimum number of model runs required to achieve reliable estimates with minimal uncertainty.

As an additional check, two groups of FKS values are created, each group's quantity of values was equal to the number of runs needed until stabilisation. The data distribution of both groups is verified and a Student's t-test or Mann-Whitney U test would be applied to the groups to confirm there is no statistically significant difference between the means of both groups. The hypotheses are as follows:

- Null hypothesis (H_0): The mean of group 1 is equal to the mean of group 2.
- Alternative hypothesis (H_1): The mean of group 1 is not equal to the mean of group 2.

If the p-value is less than 0.05, the null hypothesis is rejected and there is a statistically significant difference between the means of the two groups. If the p-value is greater than 0.05, the null hypothesis is not rejected, and there is not enough evidence to suggest a significant difference between the means of the two groups.

5.3.3 Calibration and validation

In calibration, the model variable values being explored (Table 5.3) are fine-tuned to determine the optimal configuration of variable values that maximise the FKS (output of interest). The steps of the calibration and validation process are described below:

1. In calibration, the GA (Section 2.4.2.1) explored different configurations of model variable values and determined the optimal configuration that maximises the average FKS. Each configuration was run multiple times, based on the stabilisation point defined in the previous step, and the average FKS was used across these runs to assess the performance of that configuration.
2. Once the best configuration of model variable values was identified, the model was rerun to get spatial outputs with this configuration for both the calibration and validation period. These outputs were merged and the probability of a parcel's conversion to urban was calculated (e.g. number of times a parcel is turned to new urban across the model runs).
3. A suitable threshold value was selected to create a binary urban vs non-urban map which ensures the highest average FKS value in the calibration and validation period. For example, if the ABM was run 20 times and parcels that are converted at least 8 times (a 40% probability of parcel conversion) yield the highest FKS value, then this threshold was used to define simulated urban areas.
4. The FKS for calibration and validation was reported to demonstrate that the ABM is not overfitted and performs well over both periods.

As the ABM was developed in the GAMA platform, its built-in implementation of a simple GA was used for model calibration (Taillandier et al., 2019). The initial population was generated randomly within defined parameter bounds. In GAMA's GA implementation this population can be constructed by generating multiple random preliminary populations of candidate solutions. However, In this study, the number of random populations used to build the initial population was set to 1. From this population, the top 20 solutions (defined by the population size parameter) with the highest fitness scores were selected to form the initial population, introducing a fitness-based bias before the evolutionary process began.

Roulette-wheel selection was used as the selection operator during evolution. This mechanism probabilistically selects solutions based on their relative fitness, favouring better solutions while maintaining diversity. Crossover and mutation operators were applied to explore the solution space. No additional heuristic seeding was used beyond GAMA's default initialisation process. The GA parameters were set to the following values:

- number of random populations used to build the initial population - 1 (GAMA default value)
- number of generations - 20 (GAMA default value)
- population size - 20
- crossover probability - 0.9
- mutation probability - 0.01

These parameter values were based on the following literature recommendations. The population size of GA should be set between 20-200 for robust results, higher values require longer runtimes and higher computational demand (Li and Yeh, 2005). The crossover operator should be set to a high value, such as 0.9, to promote the exchange and recombination of genetic information between solutions, while the mutation operator should be set to a lower value, such as 0.01, to introduce rare random changes. This balance between crossover and mutation allows the GA to explore the solution space efficiently while maintaining diversity in the population (Lin et al., 2023).

The convergence criteria is the number of generations parameter, the evolutionary process terminates after completing these generations, without additional stopping criteria based on fitness improvements, a lack of improvements over generations, or convergence thresholds. Each solution's fitness is evaluated using the FKS coefficient which acts as the only cost function to be maximised during the calibration process.

5.3.4 Sensitivity analysis

The purpose of the sensitivity analysis in this protocol is to ensure that the model can be trusted and confirm its ability as a tool for urban development simulation with confidence (Ligmann-Zielinska et al., 2020). The ultimate goals of the sensitivity analysis in this protocol are to determine variable importance and confidence levels; variable direction of influence; and the significance of variable interactions. To achieve these insights, two GSAs have been selected - the Morris and Sobol SA (Morris, 1991; Sobol, 2001). Both sensitivity tests are performed over the evaluation period (2014-2023), with five-year land demand set to 3.37 km^2.

5.3.4.1 Morris

The Morris method is a computationally inexpensive screening method that probes the entire input space. It involves performing OAT sensitivity tests, where only one input factor is varied at a time while the others remain fixed. It belongs to GSA because at each step the baseline input values change and sensitivity is assessed across the entire input space (Ligmann-Zielinska et al., 2020). This method produces two key sensitivity measures (Tian, 2013):

- μ - This reflects the average change in the output due to variations in the input, indicating whether the input has a positive or negative influence on the output.
- μ^* - This measure takes the absolute value of the elementary effects before averaging, which means it captures the overall influence of the input, regardless of the direction of influence

Morris is used primarily to identify influential inputs rather than provide a complete quantitative evaluation of the significance of each input relative to others (Ligmann-Zielinska et al., 2020). In addition, higher-order interactions between input variables are lumped into σ, making it difficult to isolate and understand the nature of these interactions (Campolongo et al., n.d.). Sobol addresses this shortfall of Morris.

5.3.4.2 Sobol

The Sobol sensitivity analysis is a global variance-based method (Section 2.4.3.2) that decomposes the total variance of the model output into contributions from individual input variables and their interactions. Unlike the Morris method, Sobol provides a more detailed understanding of how inputs and their interactions contribute to the variability of the output. This method calculates several sensitivity indices that quantify both the direct and interaction effects between input variables (Saltelli, 2002):

- First-Order Sensitivity Index (S_i): Measures the proportion of output variance that is attributable to each input variable individually. It reflects the direct effect of a single input on the output.
- Total-Order Sensitivity Index (TS_i): Measures the proportion of output variance due to each input variable, including both its individual effect and its interactions with other input variables. It captures the overall influence of each input on the output.

In addition to these two indices, Sobol produces confidence values that reflect the degree of uncertainty associated with the indices. Sobol sensitivity analysis is often more computationally intensive than Morris, but provides a deeper understanding of input-output relationships, especially in complex models where interaction effects are significant (Kang et al., 2022).

5.3.4.3 Sensitivity test usage

The Morris and Sobol outputs are both used to meet the objectives of variable importance and confidence levels, variable direction of influence, and variable interaction quantification. The μ^*, S_i and TS_i outputs are indicators of variable importance. The confidence values of S_i and TS_i are then used to assess the uncertainty of the variable's importance. For variable direction of influence, μ^* reveals the relationship of the variable with the output of interest. For interaction quantification, the sum of S_i can gauge how much of the total output variance is explained by the individual effects of the input variables alone (Kang et al., 2022):

$$1 - \sum S_i \tag{5.7}$$

The difference between TS_i and S_i reveals the extent to which a variable interacts with other variables (Lee et al., 2015). Furthermore, fitting a linear regression model to the pairwise interactions in the raw Sobol output data can determine which pairwise interactions are statistically significant at the 1% threshold (Borgonovo et al., 2022).

5.4 Scenario development

The 2040 future simulation of the calibrated and validated ABM variables is considered a Business As Usual (BAU) scenario where the historic trends are considered to continue into the future without significant changes following their behaviour since 2014. The four scenarios described in Chapter 3 offer an empirical basis and use the BAU scenario as a baseline to develop a scenario matrix.

Firstly, the axes of the matrix need to be defined: based on contrasting forecasts of low versus high growth in Section 3.2.3 from the economic recession and economic boom identified by Udom et al. (2020), this was set as the x-axis (Figure 5.6). This choice of axis is supported by Tang and Yang (2020) who also explored variations from their baseline value in annual demand ranging from slow land use development to rapid land use development scenarios. For the y-axis, the presence of planning was identified as an important spectrum (Figure 5.6). The ABM developed assumes the urban planner's zoning areas to be followed by law. However, World Bank (2024) argued that despite the various laws on land management and urban planning, their implementation has been hindered. There is weak regulatory control of development from the government, which is due to the lack of technical capacity and institutional fragmentation. The choice of this axis is also supported my modelling literature as Agyemang et al. (2023) who explored scenarios of weak, average, and strong development control in their ABM of residential growth.

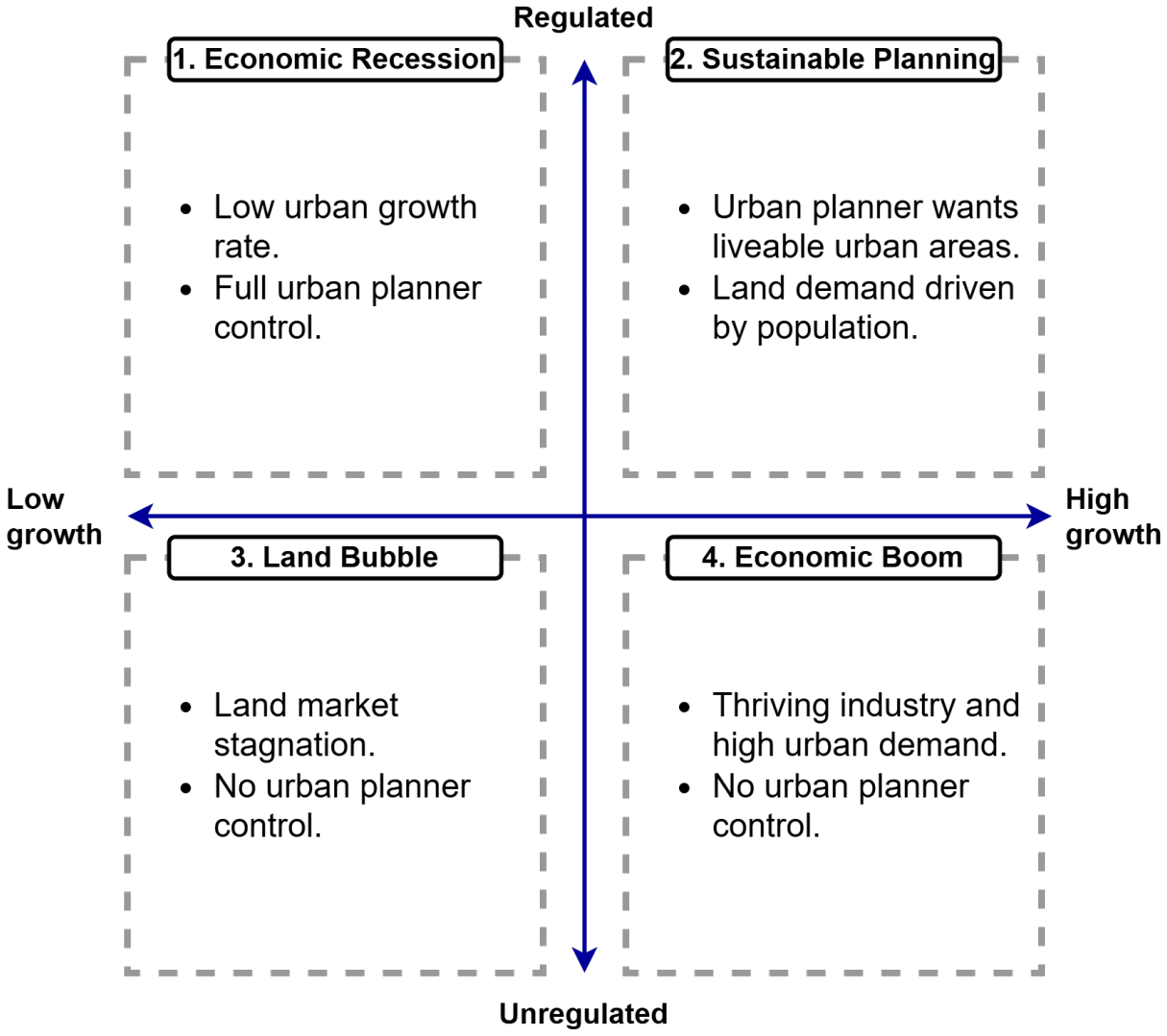

Figure 5.6: Outline of the four socioeconomic scenarios to be simulated in the ABM up to 2040. Changes to model variable values is provided in Table 5.4.

Therefore, two main spectrums in the ABM were explored: the rate of growth and the presence of planning which are ultimately a product of the scenario projections for Phnom Penh in Chapter 3, previous urban growth modelling studies, and modelling feasibility in the ABM developed. Based on the four quadrants, four forecasts were developed: economic recession (ER), sustainable planning (SP), land bubble (LB) and economic boom (EB) (Figure 5.6).These four scenarios represent the different projections outlined in Chapter 3. Since the scenarios in the grey literature are complex, all their intricacies and details can not be captured in the ABM developed, so simplifications and assumptions were made for them. The ABM was adjusted in parts to reflect the different conditions of the likely future socioeconomic pathways. For example, to represent high or low growth conditions, the variables mstr_plns, five_yr_dmd, and wts were adjusted accordingly. To reflect regulated versus unregulated conditions, the urban planner agent is either included or omitted in the scenario (Figure 5.6). More detail is given below about the four scenarios and how model variables were changed from BAU to reflect different socioeconomic conditions and agent behaviour. Each scenario is essentially a deviation from BAU, with model variable values contrasted around this scenario.

Table 5.4 displays the changes in model variable values for the four scenarios. Logic and justification for these variations in variable values is given in the subsequent four sections. In addition, the dasymetric mapping approach is slightly different for each scenario, which is described in the next chapter.

Table 5.4: Model variable values in each scenario. N/A variables appears in the LB and EB scenarios because these variables are only relevant to the planner agent, which is omitted in both cases.

Variables	ER	SP	LB	EB
mstr_plns	50	50	N/A	N/A
five_yr_dmd	2.57	3.28	4.17	6.74
wts	1.0	0.5	1.5	1.0
pl_eurb_ds_w	1.0	0.0	N/A	N/A
pl_rd_ds_w	0.2	0.0	N/A	N/A
pl_cbd_ds_w	0.8	1.0	N/A	N/A
pl_wtrcrs_ds_w	0.4	0.0	N/A	N/A
pl_wtrbdy_ds_w	0.2	0.0	N/A	N/A
pl_nurb_ds_w	0.8	0.0	N/A	N/A

5.4.1 Economic recession scenario

This scenario involves a severe decline in economic activity which has resulted in low growth rates and urban development. Strict regulation development is still maintained

as the government does not want the recession to be exacerbated by privatised development. In addition, the recession has led to pessimistic selling attitudes from landowners who take a rational and cautious approach to selling their land. The urban planner is still active in this scenario and regulates building development.

In the ABM, the urban planner agent still generates master plans which the developer must build in to reflect that strong planning is still maintained. However, the 5-year land demand and the number of master plans are set at 2.57 km^2 and 50, respectively, to reflect lower growth rates and less development (Figure 5.6). Willingness to sell is set at 0.8 to show that owners are more reluctant to sell, meaning transactions are significantly less likely to take place as a result. The urban planner assumes the same weights as the BAU scenario as they want to maintain the status quo of development and not exacerbate the current economic conditions.

5.4.2 Sustainable planning scenario

The urban planner wants to develop sustainable and liveable urban areas that effectively accommodate future population growth. The urban planner bases the land demand for the study area on a 2040 population projection. This scenario is the most prudent and sustainable forecast that aims to simulate development based on the planner's perspective to provide sensible urban development potential. This scenario is characterised by a strong planning influence and high growth rate.

For the population projection, census information from 1998 and 2019 was used to understand the total growth rate of the three communes in the study area. The land area demand over the simulation was calculated by projecting future population using Cambodian census data from 1998 and 2019. The percentage difference in population between 2019 and 1998 was calculated for the communes (5.8) where P_{final} is the 2019 population and P_{initial} is the 1998 population:

$$population\ difference\ (\%) = \frac{P_{\text{final}} - P_{\text{initial}}}{P_{\text{initial}}} \times 100. \tag{5.8}$$

The population difference was divided by the number of years (21) to obtain the annual growth rate as 10.47%. This percentage rate was used to increase the population each year from 2019 adding the percentage rate to each new yearly population value to get an exponential population growth projection as shown in Figure 5.7. This results in a projected population of 330,852 people in the study area by 2040.

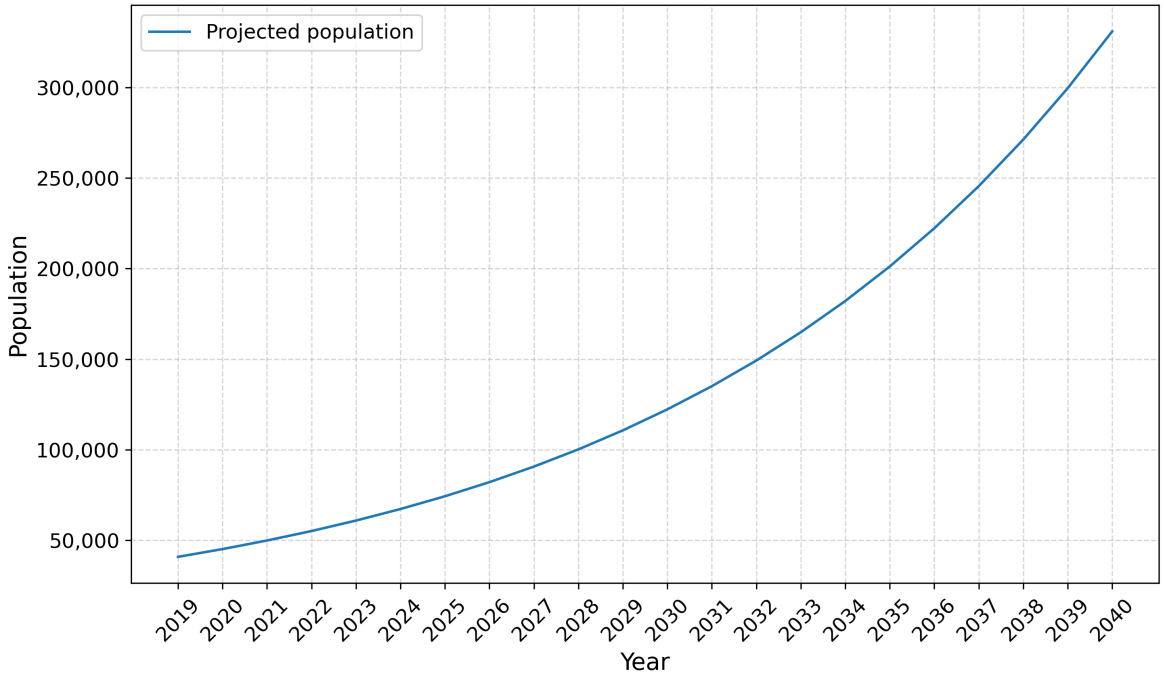

Figure 5.7: Projected population by 2040

Converting the 2040 population to land demand involved the following steps:

1. Using EO4SD parcel classifications (low, medium, and high density) described in Section 3.2.4 and Google building footprints that had their centroid within these parcels. The average building area with curtilage was calculated for each parcel classification.

2. Using the average 2019 household size for the three communes (4.8 people per household), the number of houses was calculated as 68,928 required to meet this population demand by 2040.

3. Using the low, medium, and high density urban parcels from EO4SD, the proportion of each parcel in Phnom Penh was calculated. Then the total number of houses for each parcel classification was calculated to quantify how many houses could realistically fit in each category of EO4SD parcel.

4. The number of houses per parcel type was then multiplied by their average area (step 1) and summed together to get total land demand by 2040.

5. Finally, this was then divided by the number of five year intervals that occur in the ABM (4) to get five year land demand.

In the ABM, mstr_plns is set to 50 to reflect more cohesive development and fewer chances of sprawl. The wts is set to 0.5 and the five-year land demand is set at 3.28 km^2 based on the 2040 population projection previously described. As this is a Phonm Penh-centric scenario, only the distance from Phnom Penh is considered by the urban planner to try and minimise sprawl.

5.4.3 Land bubble scenario

The housing market has begun to stagnate due to a property bubble - a significant rise in property prices due to demand outpacing supply of land. In response to the stagnation caused by the bubble, the government has enabled a free market to allow a greater opportunity for supply. This scenario features a decline through supply and demand dynamics and more liberal planning compared to the ER scenario (Figure 5.6).

In the ABM, the model variables have the same values as the BAU scenario to represent the same demand, apart from the wts which is increased to 1.5 to reflect high land values and limited supply (Table 5.4). This captures the high demand and low level of supply typically experienced in real-estate bubble. Developers can ultimately choose wherever they want, as the urban planner is not present in this scenario. Developers and owners weights are kept equal.

5.4.4 Economic boom scenario

This scenario envisions a period of rapid and sustained economic growth in Phnom Penh driven by foreign investment and new government policies that promote a free market. The government adopts a laissez-faire approach, refraining from regulating urban development so that favourable economic conditions are not disrupted. The city experiences rapid growth, necessitating extensive infrastructure development. Property developers have full autonomy to choose and develop sites based on their own strategic interests. The urban development process is entirely unregulated, which means that land transactions are completely private. This scenario experiences large-scale urban development.

In the ABM, five-year land demand is the sum of overall growth in the evaluation period (6.74 km^2). Developer and owner weights are kept equal. As in the LB scenario, the urban planner is removed from this scenario. The owner wts is increased to 0.8, as they want to maximise profits and ensure that their land is not sold at an undervalued price. Tables 5.4 summarise the model variable values in the five scenarios.

5.5 The stochastic urban growth model evaluation protocol results

5.5.1 Stabilisation analysis outputs

The ABM was run 1800 times using Equation 5.5, with the n_{mc} set at 100 due to the long ABM runtime and the median stabilisation point across the 1800 shuffles was 240 runs,

with a knee-defined epsilon of approximately 1.15×10^{-3}. Figure 5.8 shows that from the 30th run to the 240th model run the median CV difference drops steeply by 5.33×10^{-3} from 6.48×10^{-3} to 1.15×10^{-3}, indicating a rate of decline of 2.54×10^{-5} per run. From this run onwards to 1800 the rate of decline is much lower, with a difference of 9.95×10^{-4} from 1.15×10^{-3} to 1.55×10^{-4}, indicating a much lower rate of decline to just 6.38×10^{-7} per run. This highlights the diminishing returns, supporting the 240th run as a natural stabilisation point beyond which further iterations yield negligible benefits.

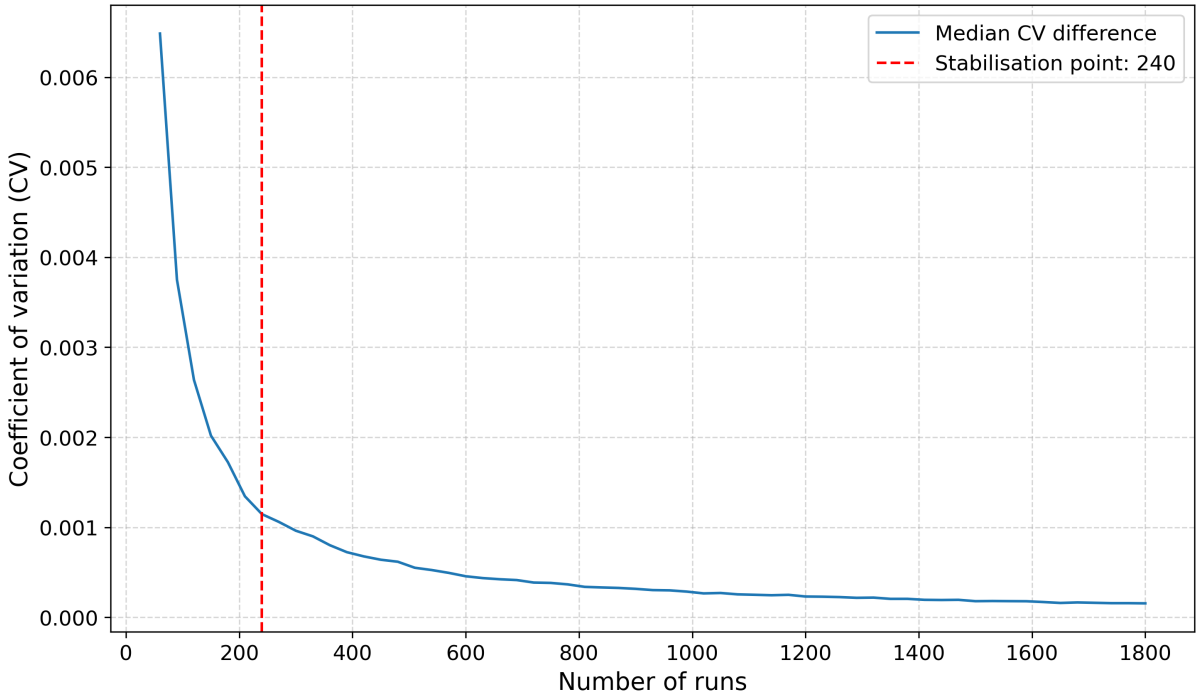

Figure 5.8: CV difference between every 30 model runs.

480 FKS values were randomly selected across the 1800 runs and split into two groups. Both groups were normally distributed, so a student's t-test was performed which outputted a p-value of 0.330. Therefore, this p-value is much higher than 0.05 and the null hypothesis in Section 5.3.2 is not rejected and there is no significant difference between the means of the two groups.

5.5.2 Calibration and validation

The GA ran through 194 variable configurations and these variable values resulted in the highest accuracy score with an average FKS value of 0.325:

- mstr_pln: 100
- wts: 0.2
- pl_eurb_ds: 1.0
- pl_rd_ds: 0.2

- pl_cbd_ds: 0.8
- pl_wtrcrs_ds: 0.4
- pl_wtrbdy_ds: 0.2
- pl_nurb_ds: 0.8

The model was run 240 times for the calibration and validation period with these optimal variable configuration to get spatial outputs for each run. After merging the runs and getting urban probability values for the parcels, visual inspection of the calibration and validation outputs showed the FKS was sensitive to overestimating area to get a better score when there was little change in the study area, particularly in calibration. For example, in Table 5.5 the 10% threshold results in a FKS of 0.365 for calibration, but the area of simulated growth is extremely vast at 7.68 km^2 a 200% overestimation of the 2.57 km^2 that occurred in reality.

Table 5.5: FKS, area and weighted area FKS accuracy scores (Eq. 5.10) in calibration and validation.

	Calibration			Validation			
Threshold	**FKS**	**Area**	**Weighted**	**FKS**	**Area**	**Weighted**	**Average**
10%	0.365	2.00	0.122	0.453	1.93	0.155	0.138
20%	0.366	1.17	0.169	0.556	0.76	0.316	0.242
30%	0.340	0.41	0.241	0.550	-0.08	0.509	0.375
40%	0.306	-0.13	0.271	0.416	-0.62	0.257	0.264
50%	0.234	-0.54	0.152	0.188	-0.89	0.099	0.126

Due to this unforeseen issue, the area difference (area column, Table 5.5) was calculated for the probability thresholds in calibration and validation using the following equation:

$$\Delta\ Area = \left(\frac{A_{\text{Sim}}}{A_{\text{Obs}}} - 1\right) \tag{5.9}$$

Where A_{Sim} is the area of the simulated parcels and A_{Obs} is the area of the observed parcels. FKS values in calibration and validation were then weighted (weighted column, Table 5.5) by $\Delta\ Area$ with the following equation:

$$Area - weighted\ FKS = \frac{FKS}{(1 + |\Delta\ Area|)} \tag{5.10}$$

The threshold that produced the highest average area-weighted FKS value (average column, Table 5.5) was used for calibration and validation, rather than relying on the traditional FKS value as originally proposed. The probability threshold for classifying a parcel as a simulated new urban in the 240 runs was set at 30% or a higher conversion

rate, as this resulted in the highest area-weighted FKS value of 0.375 across calibration and validation (Table 5.5). Furthermore, Table 5.5 shows that as the threshold increases, the weighted FKS value first increases and then decreases, indicating that the lower thresholds are too lenient, while the higher thresholds become too restrictive to accurately capture historical growth in calibration and validation. In calibration, the simulated urban growth overestimates observed growth by 41%, with the simulated growth at 3.62 km^2 and observed growth at 2.57 km^2 (Figures 5.9). The calibration area-weighted FKS is 0.241.

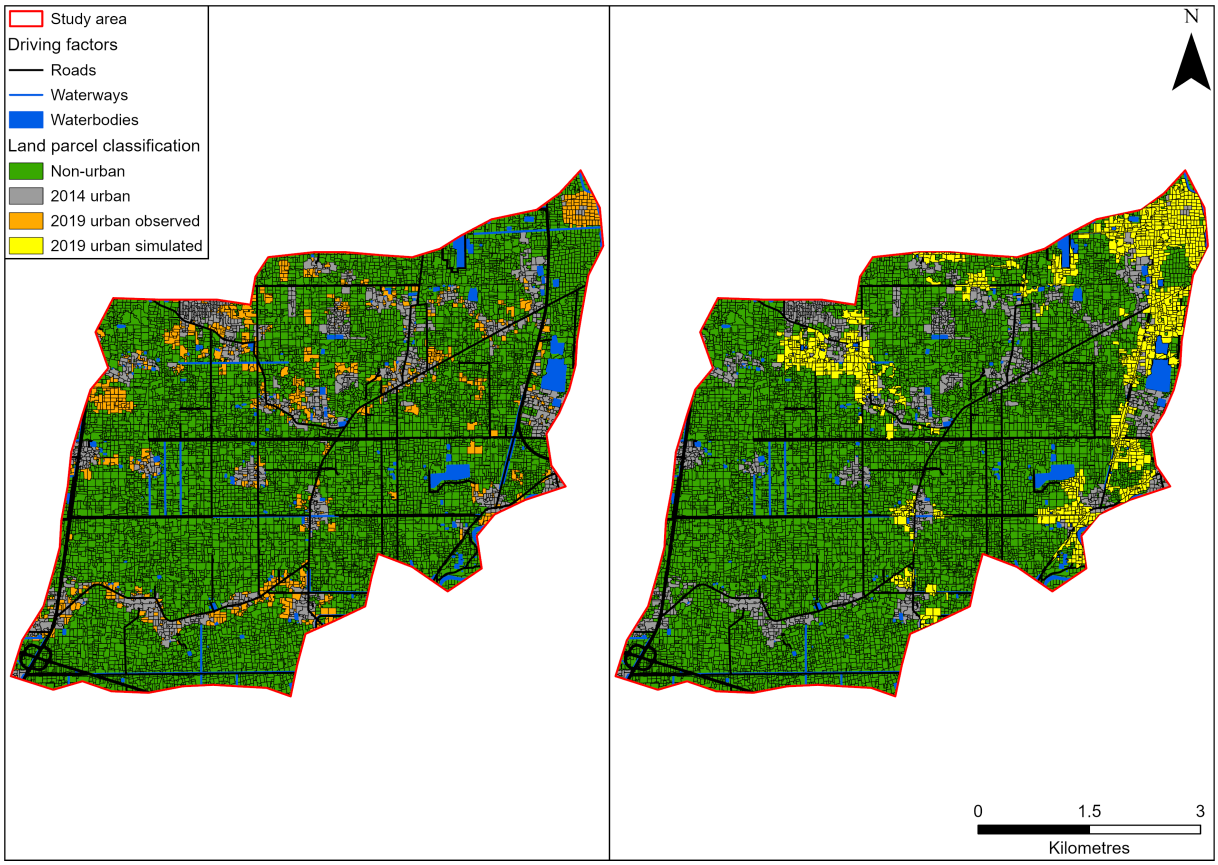

Figure 5.9: Visual comparison of observed urban growth (left) and simulated urban growth (right) in calibration period (2014-2019).

Figure 5.10 In the validation phase the weighted FKS is 0.509, the ABM now slightly underestimates the area of growth by 8%, where the observed growth was 4.17 km^2, but simulated growth reached 3.84 km^2. As both calibration and validation FKS values are between 0.241-0.509, the strength of agreement can be interpreted as fair for calibration and moderate for validation(Henry et al., 2017). The superior performance in validation is a positive sign that shows that the model is capturing more recent trends better, which is important for future simulation. In addition, Figures 5.9 and 5.10 are visually convincing, indicating that the model is replicating historic patterns to a good degree and is suitable for future projections.

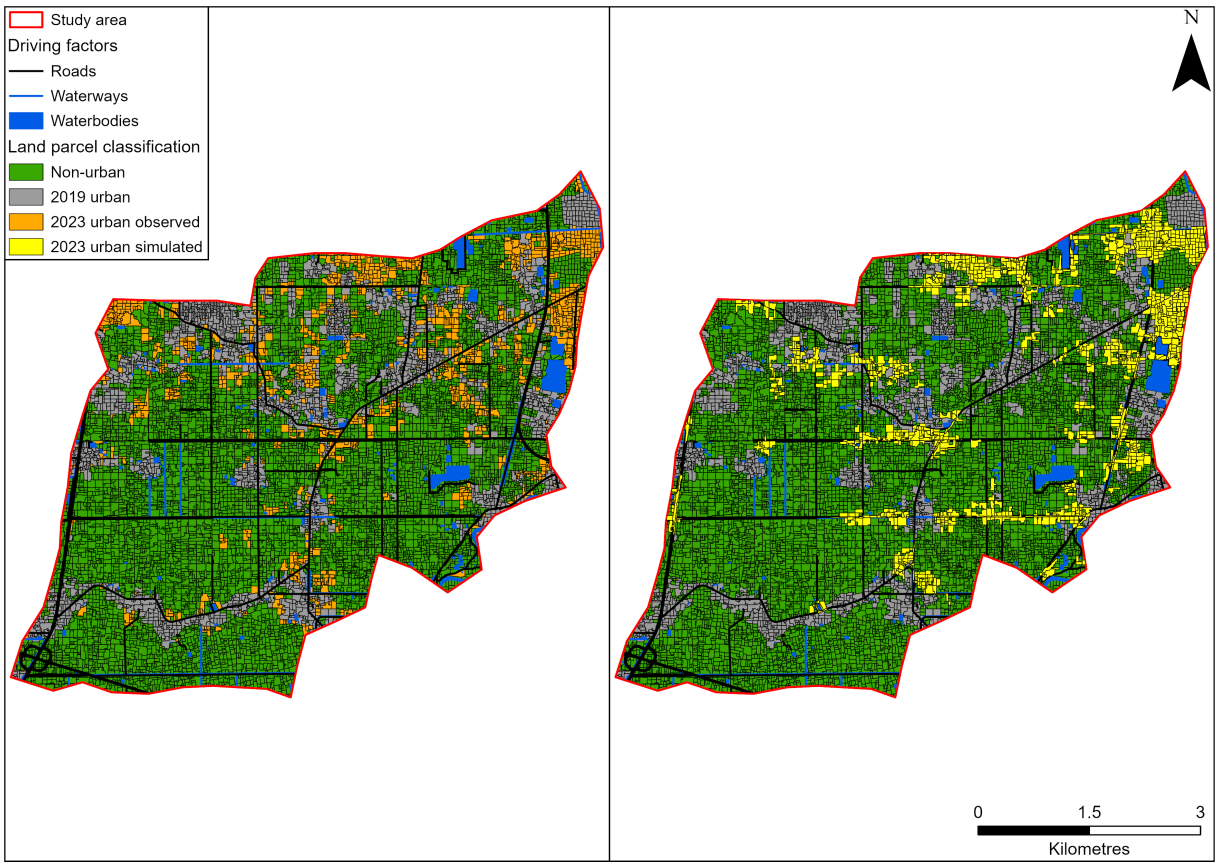

Figure 5.10: Visual comparison of observed urban growth (left) and simulated urban growth (right) in validation period (2019-2023).

5.5.3 Sensitivity analysis outputs

5.5.3.1 Variable importance and confidence

Figure 5.11 displays the importance of the model variables based on the μ^* from Morris and S_i and TS_i from Sobol. The variables were sorted by TS_i as it is a comprehensive indicator of importance that accounts for both the factor's direct effect and its interaction effects with other variables. The S_i and TS_i values sum to less than and at least 1, respectively, which is an expected result of the Sobol analysis. Table 5.3 gives the description of each variable name. The mstr_plns TS_i is over 0.8, meaning this variable interacts heavily with other variables and has a large overall influence on the output. In addition, it also has the highest S_i at 0.23. The lowest S_i are pl_eurb_ds_w at 0.0112 and 0.01324, respectively, suggesting a relatively small direct effect on the output of the model. μ^* values appear very similar, suggesting that all variables contribute similarly to the output variability when considering their direct effects across the input space.

The TS_i have an average confidence level of 0.118 and a standard deviation of 0.014, whereas the S_i confidence have an average of 0.154 and standard deviation of 0.014. This suggests that the results from S_i can be relied upon more with slightly higher confidence. Among the variables, mstr_plns has the highest confidence level for its S_i

sensitivity at 0.1599, followed closely by pl_cbd_ds_w with a confidence of 0.1709. These variables have relatively high confidence in their S_i sensitivity values, indicating a more consistent effect between different settings. On the other hand, pl_wtrbdy_ds_w has the lowest S_i confidence at 0.1477, but still maintains a fairly strong level of reliability. For TS_i, the confidence levels tend to be lower overall. The highest confidence for TS_i is observed in pl_cbd_ds_w, with a confidence of 0.1231, while mstr_plns shows a lower confidence of 0.1173, suggesting more variability in its TS_i. Variables like pl_wtrcrs_ds_w and pl_eurb_ds_w (Eurb Demand) have slightly higher TS_i confidence levels (0.1194 and 0.1162, respectively), indicating a greater uncertainty in their total sensitivity contributions compared to the S_i effects.

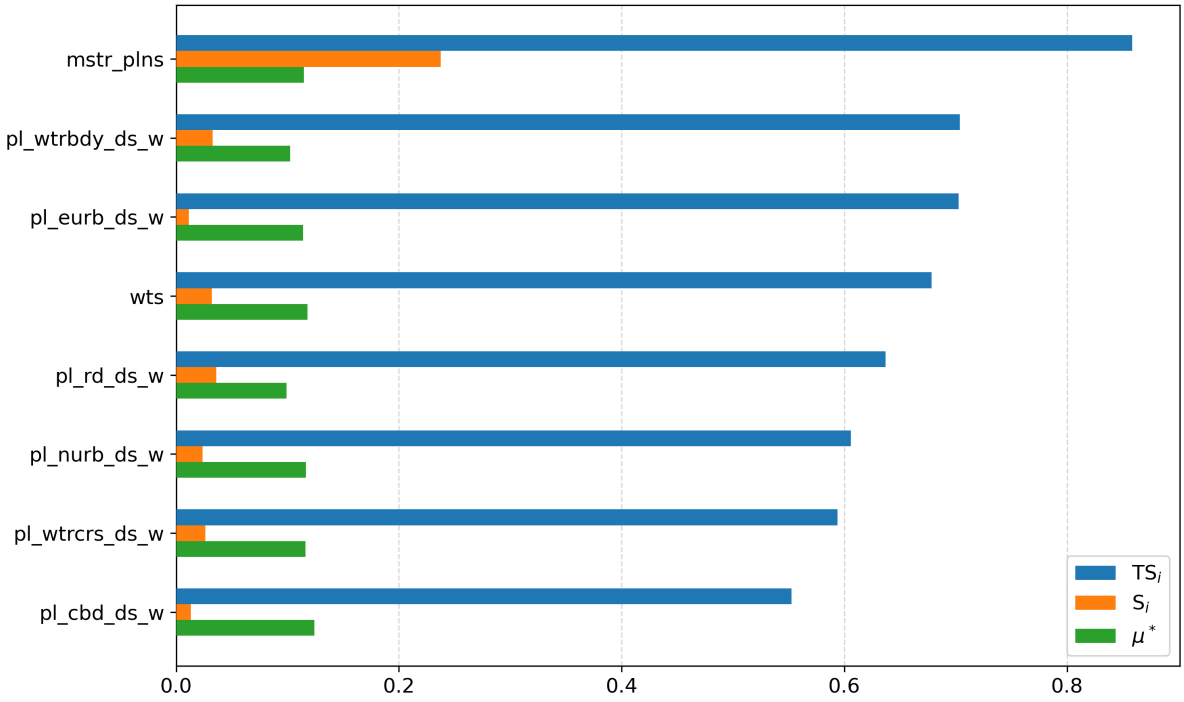

Figure 5.11: Ranking of model variables, sorted by TS_i which measure both the direct effect of each input and its interactions with other inputs. S_i and μ^* focus only on the individual impact of each input, ignoring interactions.

5.5.3.2 Variable direction of influence

Figure 5.12 shows the relationship of the variables with the FKS using μ from Morris. Five of the eight variables have a negative value, indicating that increasing their value decreases the FKS output. For wts, this is expected as increasing this value makes parcel conversion to urban less likely, impacting the FKS by lowering it as it is looking for accuracy of simulated change. In addition, this variable has the strongest negative relationship with FKS showing it has a large influence on FKS. Whereas, mstr_plns has the strongest out of all positive and negative relationships with a value of 0.039 showing a large influence of improving the FKS value.

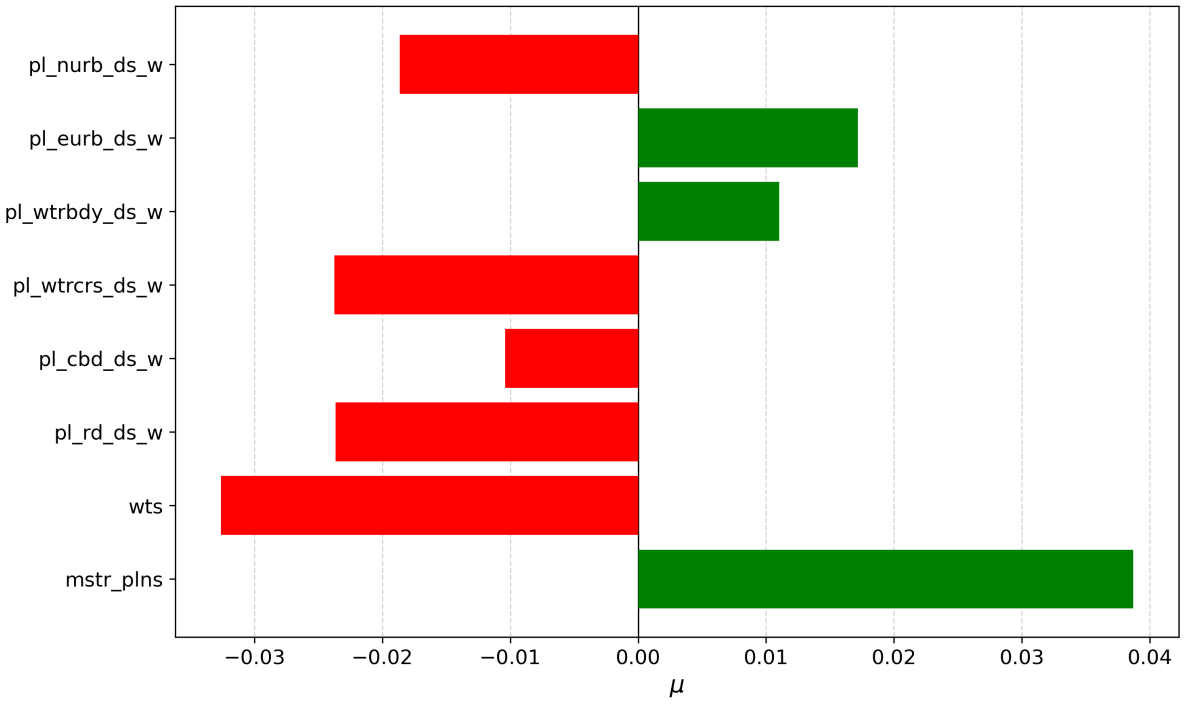

Figure 5.12: Mean effect of the model variables on the FKS.

The direction of influence also shows that only higher weightings for pl_eurb_ds_w and pl_wtrbdy_ds_w directly improve the FKS score in isolation. Whereas the other weightings decrease the FKS score.

5.5.3.3 Variable interaction quantification

The interaction effect (Eq. 5.7)) has a value of 0.5870. This means that approximately 58.7% of the total variance in the model output is attributed to interactions between the variables, indicating that a substantial portion of the variability in the model output is due to the combined effects of multiple variables working together rather than the independent effects of each variable alone. Table 5.6 displays the difference between the TS_i and S_i. A larger difference suggests that a variable's influence is largely due to interactions, whereas a smaller difference implies a more direct effect on the model output. pl_eurb_ds_w and pl_wtrbdy_ds_w have the highest interaction effects, meaning their influence is largely due to indirect effects through other variables. Since all differences (TS_i - S_i) are over 0.5, this indicates that every variable has a strong influence primarily due to interactions with other variables. This suggests that the model output is highly dependent on higher-order interactions rather than individual variable effects alone. The high total order and low first order suggest that these inputs owe a significant part of their importance to their interactions with the remaining variables.

Table 5.6: Difference between TS_i and S_i of model variables.

Model variable	TS_i	S_i	Difference
pl_cbd_ds_w	0.55243	0.01324	0.53919
pl_wtrbdy_ds_w	0.70380	0.03274	0.67106
pl_rd_ds_w	0.63703	0.03611	0.60092
wts	0.67832	0.03189	0.64643
mstr_plns	0.85836	0.23772	0.62064
pl_wtrcrs_ds_w	0.59388	0.02630	0.56758
pl_eurb_ds_w	0.70259	0.01121	0.69138
pl_nurb_ds_w	0.60572	0.02375	0.58197

The linear regression results indicate that 23 out of the 28 pairwise interactions were statistically significant at the 1% threshold. These interactions exhibit a range of positive and negative effects on the model output, suggesting that certain configurations of variables have a substantial impact on the model's behaviour.

5.6 Urban growth scenario results

5.6.1 Business as usual urban growth

Figure 5.13 shows the BAU output for the study area. The five classes of urban probability in Figure 5.13 were split based on natural breaks in the data and this classification was kept consistent across all scenarios to enable comparison. Higher probabilities of urban expansion (76-88% and 89-100%) are concentrated in the northeastern and central parts of the study area. The 89-100% 2040 urban is clustered around existing urban zones and major road networks, indicating that future development is likely to follow current urban centres and infrastructure corridors. The 61-75% typically occurs on the edge of the high 2040 urban probabilities, although in some areas, this probability class occurs independently. Lower 2040 urban probability classes (30-45% and 46-60%) are common in the southwestern part of the study area, where there are driving factor features.

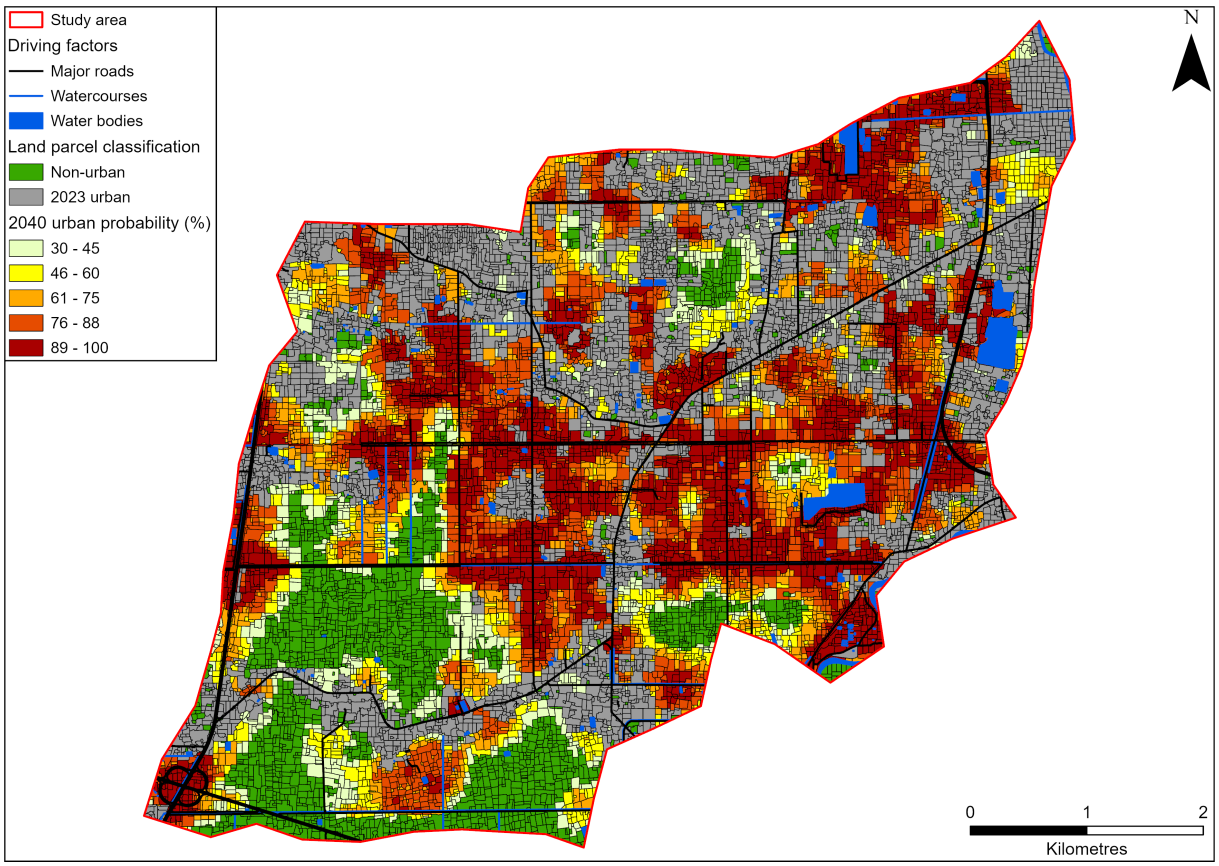

Figure 5.13: BAU 2023-2040 urban growth scenario, displaying ABM-simulated probability of non-urban parcels being converted to urban.

5.6.2 Economic recession urban growth

Figure 5.14 displays the ER 2040 scenario map. Both the 89-100% and 76-88% probability classes appear only on a few parcels and have a highly restrained distribution and reduced extent compared to the BAU scenario. The three lower probability classes (30-75%) appear to dominate across the study area, which contrasts with the BAU scenario where the high and very high probability classes are more frequent. The locations of the simulated parcels in this scenario frequently match with the BAU 76-88% and 89-100% simulated parcels. Overall, the extent of development is much less in the ER scenario, although the core locations of where development occurs remain consistent with the BAU scenario.

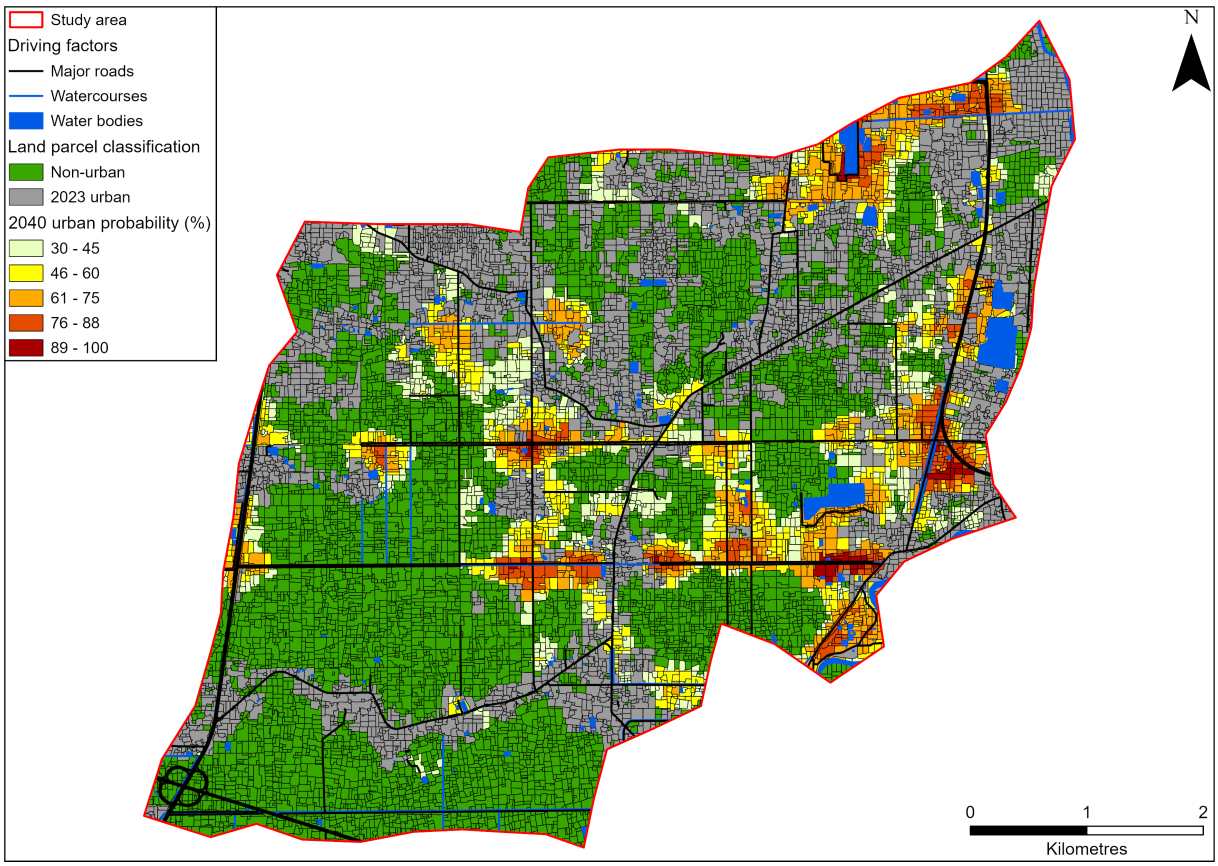

Figure 5.14: ER 2023-2040 urban growth scenario, displaying ABM-simulated probability of non-urban parcels being converted to urban.

5.6.3 Sustainable planning urban growth

Figure 5.15 shows the SP 2040 scenario. This scenario shows very probability patterns to the BAU scenario. but their is a much stronger pull of development occurring towards the north-east corner of the map, as the planner is minimising sprawl and prioritising development close to Phnom Penh (Figure 5.13). There is no development in the south-east region of the map. The occurrence of development around water bodies in the north-eastern is also very similar to the BAU, but parcels around these water bodies seem to have slightly lower probability than the BAU, likely due to the low weighting given to them in this scenario. Having the urban planner utility function only weighted by the distance from Phnom Penh appears to have a strong impact on the location of where development occurs and only slight change to the probability classes of simulated parcels.

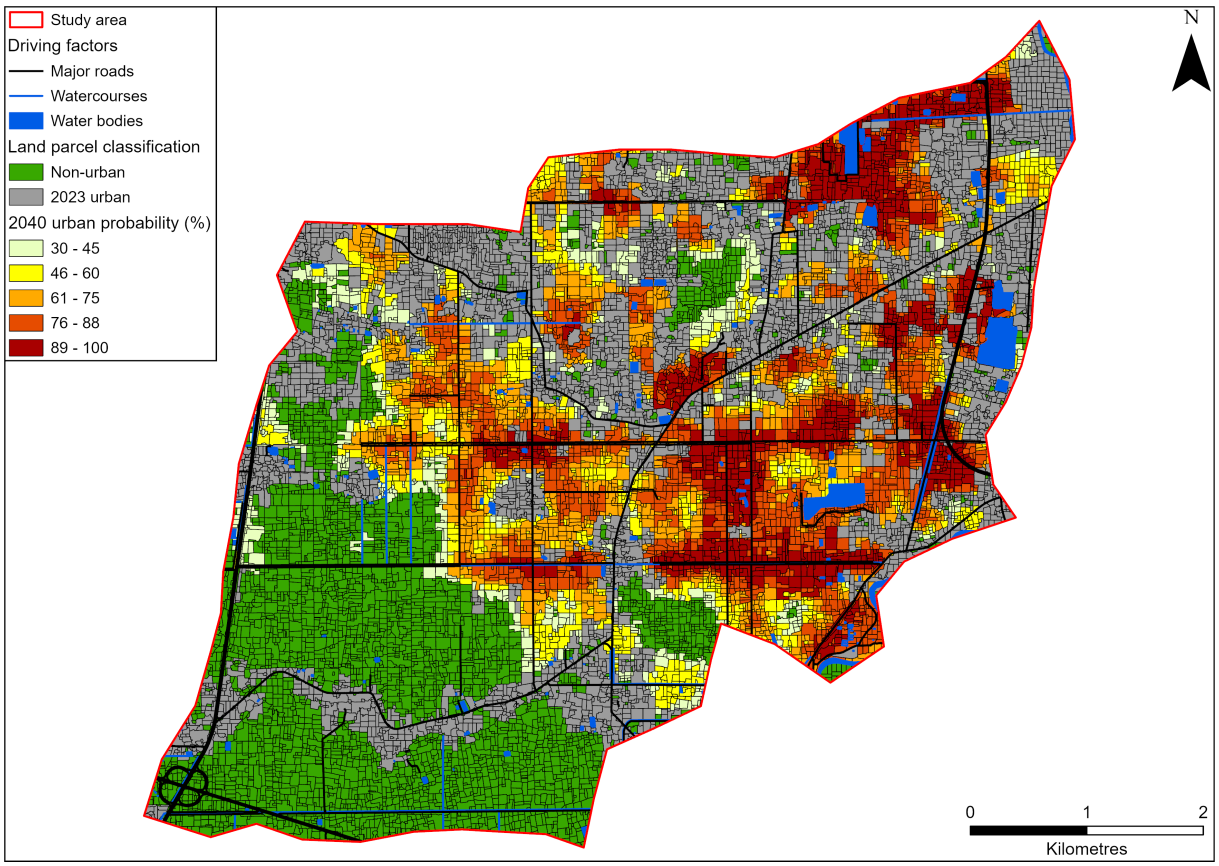

Figure 5.15: SP 2023-2040 urban growth scenario, displaying ABM-simulated probability of non-urban parcels being converted to urban..

5.6.4 Land bubble urban growth

The LB scenario (Figure 5.16) shows large amounts of cohesive clusters with very high probability development (89- 100%) and lower probability classes on the periphery of these areas. Most of the development occurs on the eastern side of the study area, some in the centre. Unlike other scenarios that show a more contiguous pattern of growth, this scenario displays some isolated pockets of high probability, likely reflecting speculative land development. In general, the quantity of development is low and not as widespread as in the other scenarios, with some areas, such as the ring road in the southwestern corner of the study area not experiencing development by 2040. In the BAU scenario, the parcels were converted to urban on the ring road.

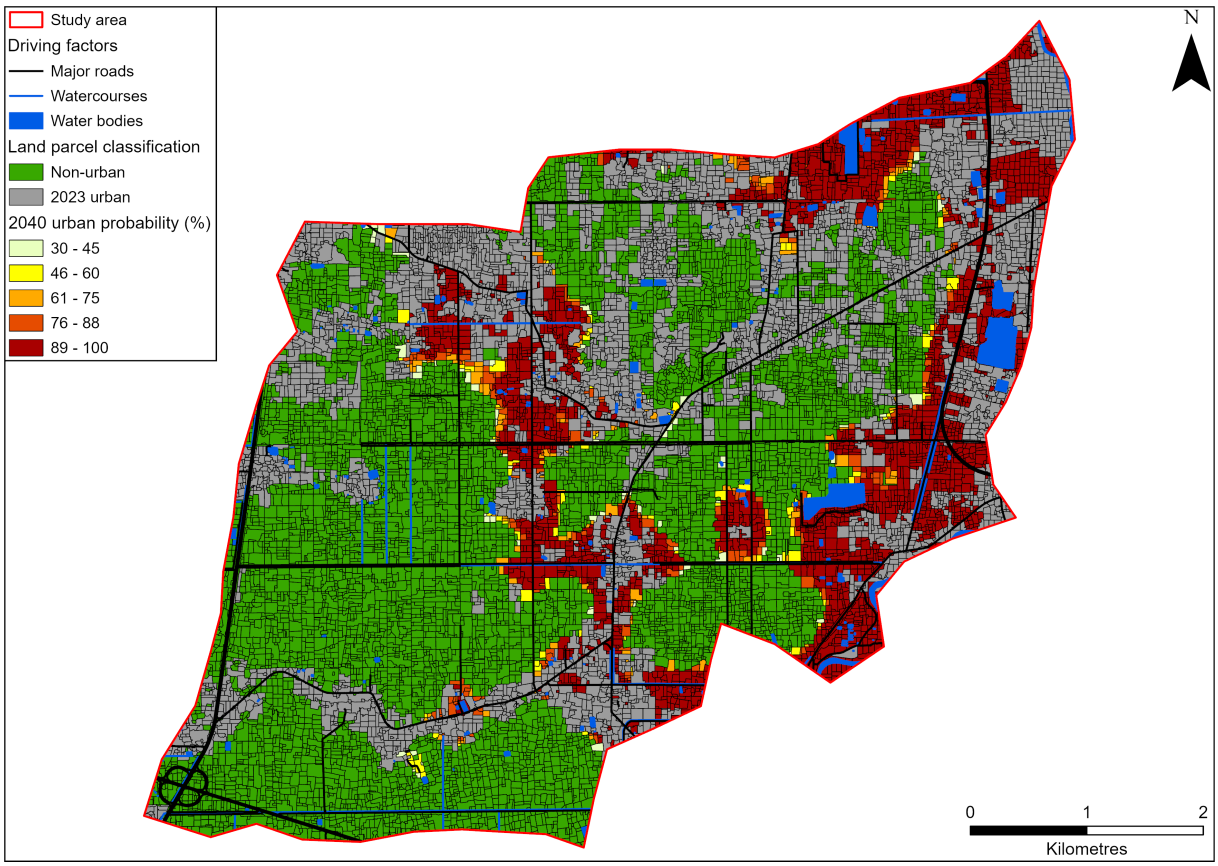

Figure 5.16: LB 2023-2040 urban growth scenario, displaying ABM-simulated probability of non-urban parcels being converted to urban.

5.6.5 Economic boom urban growth

The EB scenario displays outputs similar to the LB scenario, but to a much larger extent (Figure 5.17). There is extensive and widespread urban expansion all within the 80-100% probability band and growth occurs in areas that were not developed in any of the other scenarios. In general, this map highlights a pattern of expansive urban growth, with high urban probabilities in the 89-110% class that covers most of the study area. This scenario suggests that an economic increase would lead to rapid development across nearly all available land, significantly reducing non-urban areas and potentially leading to extensive urban sprawl.

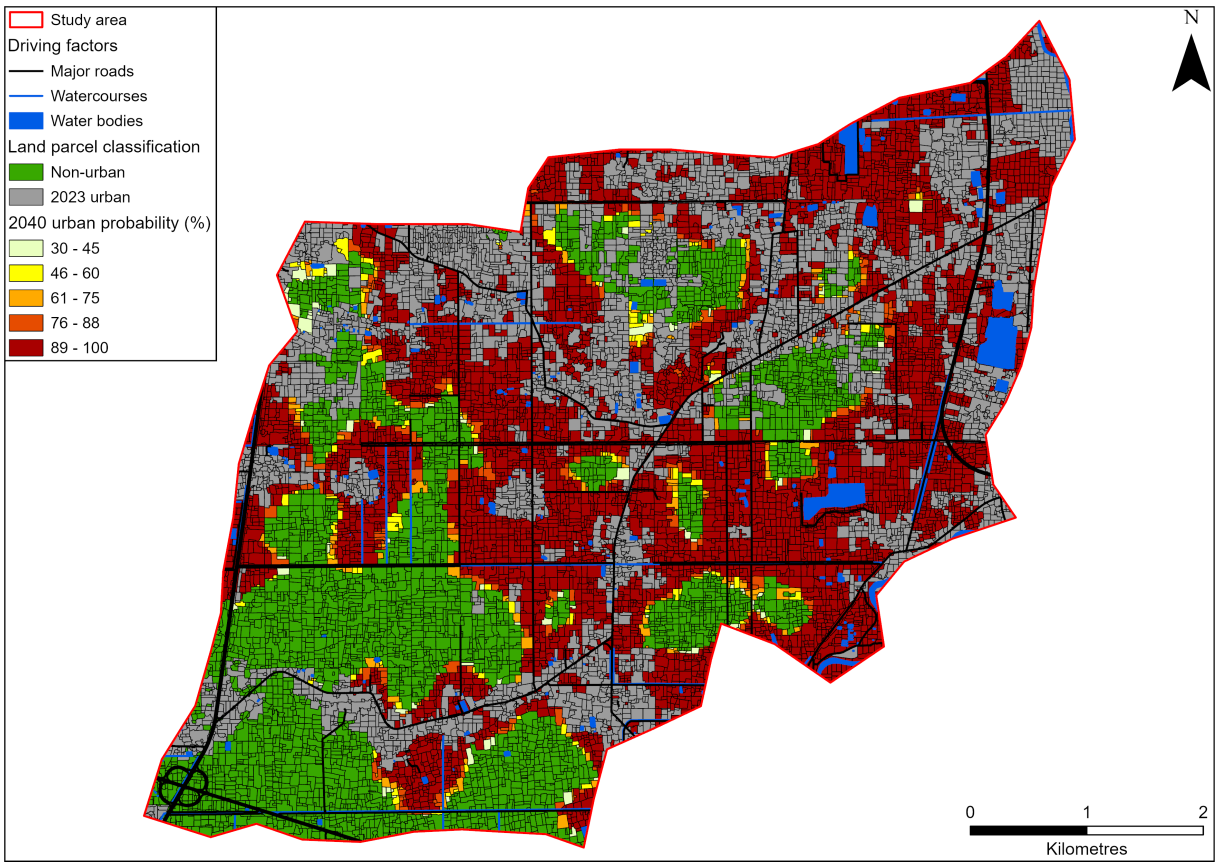

Figure 5.17: EB 2023-2040 urban growth scenario, displaying ABM-simulated probability of non-urban parcels being converted to urban.

5.6.6 New urban extent and probability statistics

Table 5.7 shows the statistics associated with the scenarios which provide further insights into the urban growth projections. In 2023, urban area in the study area compromised 10.41 km^2 across 3,993 parcels, while the non-urban area covered 22.63 km^2 and 10,126 parcels.

In the BAU scenario, new urban area is projected to be 18.04 km^2 and parcel count 7918. Very high probability urban parcels (89-100%) dominate in this scenario with an area of 6.09 km^2 and 2197 parcels. However, the other probability classes are at most 3.83 km^2 smaller than this class, but these four classes share similar parcel counts.

Under the ER scenario, urban expansion is minimal compared to BAU, with new urban area at 7.45 km^2 and a parcel count of 3198. Only 0.13 km^2 and 39 parcels fall into the highest urban probability (89-100%), with most parcels concentrated in lower probability categories (30-45% and 46-60%), indicating limited new urbanisation due to economic downturns.

Table 5.7: 2040 scenario statistics across the new urban probability classes. All area values are in km^2.

Scenario metrics	Total	30-45%	46-60%	61-75%	76-88%	89-100%
BAU scenario						
Parcel area	18.04	2.26	1.99	2.79	4.91	6.09
Parcel count	7918	1036	975	1461	2249	2197
ER scenario						
Parcel area	7.45	2.41	2.07	1.96	0.88	0.13
Parcel count	3198	1123	950	777	309	39
SP scenario						
Parcel area	13.85	1.7	2.21	3.18	3.97	2.79
Parcel count	6091	897	1059	1506	1694	935
LB scenario						
Parcel area	6.28	0.18	0.24	0.3	0.54	5.02
Parcel count	3032	84	87	154	258	2449
EB scenario						
Parcel area	13.96	0.27	0.34	0.41	0.71	12.24
Parcel count	6448	109	143	181	315	5700

The SP scenario has a slightly lower increase in urban extent compared to BAU. The new urban area is projected to be 13.85 km^2 and parcel count is 6091 by 2040. The highest probability areas (89-100%) cover 2.79 km² with 935 parcels, suggesting that the SP scenario encourages concentrated urban zones of high probability. Other parcels are more evenly distributed among lower probabilities, reflecting a balanced approach to controlled urban expansion.

The LB scenario features the lowest growth with new urban area at 6.28 km^2 and parcel count at 3032. High-probability zones (89-100%) occupy 5.02 km^2, mainly as isolated pockets (2449 parcels), suggesting a dispersed and unplanned urban growth pattern driven by market speculation.

In the EB scenario, new urban area is 13.96 km^2 and parcel count is 6448. Most new urbanised areas fall into the highest probability category (89-100%), suggesting intensive development in almost all available land in response to economic demand, with much lower area values and little variation in the other probability classes.

Overall, all scenarios produce distinct outputs, which colloquially shows the ABM is reactive to changes in model variables. For example, the LB scenario outputs low

and unplanned growth compared to the SP scenario which is planned and high growth shows polar opposite forecasts, in terms of total urban extent, probability distribution, patterns and locations. These different probability distributions and patterns have been demonstrated numerically and visually in this section. Interpretations of the scenarios and reasons for their difference is addressed and explained in Chapter 7.

5.7 Summary

The ABM has been developed based on techniques from existing urban growth models, theories, and grey literature on the Cambodian spatial planning process. The ABM uses irregular land parcel GIS data as its spatial unit and involves four submodels that replicate the spatial planning process. The stochasticity of the model has stabilised after 240 runs, the ABM has been calibrated and validated according to the FKS coefficient value, which has been weighted by the area accuracy between observed and simulated parcels in calibration and validation. Sensitivity analysis using Sobol and Morris of the model variables has identified their importance, direction of influence, and quantification their interactions, using the FKS as the output of interest.

Five scenarios have been simulated up to the year 2040. The BAU scenario is based on the calibrated and validated ABM variables. The other four scenarios were developed using a scenario matrix with axes of growth rates and regulations based on socioeconomic forecasts that Phnom Penh may experience documented in the grey literature. Overall, these scenarios provide distinct patterns of urban growth probability, location, and quantity up to the year 2040. The scenario outputs are of strong relevance to help decision makers understand future growth and change hotspots within the study area. The scenario outputs from the ABM are used in the next chapter to perform the urban form generation.

Chapter 6

URBAN FORM GENERATION

6.1 Introduction

Having established the methodology and results for simulating urban growth locations in Chapter 5, this chapter describes the methodology to generate future urban form for each ABM scenario. The ABM land parcel outputs are converted into a building demand raster which is applied to the pix2pixHD GAN. Each scenario has a unique dasymetric mapping technique applied to reflect varying densities and styles of urban development.

As described in Section 4.2.4, the first step in urban form generation is the creation of a future synthetic road network. The 2,400 cells with the highest density type 2 road network density extracted in Section 4.4.1 are converted into images and applied to 12 GAN experiments that translate an input road density image into a target road network image. The FID score is used to assess the quality of the generated images, then following post-processing of the images to vector format, the total road length, average block area and node degree count are used to evaluate the overall performance of the GANs on the test dataset. The best performing experiments are then applied to the study area to determine the most suitable for future road network generation based on the amount of roads generated.

Following road network generation, synthetic future building footprint generation requires the synthetic road network layout (Section 4.2.4). Using the 2,400 cells at 500 m^2 and 1000 m^2 with the highest building density in Cambodia extracted in Section 4.4.2, building footprint generation is performed in 4 GAN experiments that translate an road network input image into a target road network with building footprints image. The FID score is used to assess the quality of the generated images, then following post-processing of the images to vector format, the building count, average area, and average perimeter are used to evaluate the overall performance of the GANs on the test dataset. The best performing GAN experiments are then applied to the 2040 BAU road network to determine the best one based on which generated the highest quantity of building footprints.

6.2 Agent-based model and generative adversarial network integration

Pix2pixHD, described in Section 4.2.4, requires detailed images to perform image synthesis and the land parcel maps of future urban extent alone in Section 5.6 are not sufficient information to train a GAN on image-to-image translation. To explicitly link the ABM outputs with the GAN inputs, the development probability of the land parcels was used as a proxy for building development demand to generate a density raster of 2040 building density. Building quantity and distribution in the study area were calculated differently for each scenario using an approach similar to multi-class dasymetric mapping.

6.2.1 Establishing development density

Intellectually, this part of the research is exploring unchartered territory with no literature having previously developed an approach to generate future urban form based on simulated future urban expansion. A pragmatic and logical approach must be taken to use the ABM output to feasibly generate future urban form. In previous urban growth modelling literature, simulated development probability for each spatial unit has been used as a proxy for demand, for example, to represent a parcel's suitability or propensity to convert from one land use to another (e.g. Long et al., 2021; Brown et al., 2002; Liu and Long, 2016). The more frequently a parcel is selected in simulation, the higher its inferred development demand and land price which will maximise returns (Brown et al., 2002). Therefore, Urban growth probabilities generated by the ABM can be translated to represent the expected likelihood that a parcel will be converted from non-urban to urban during the simulation period (2023-2040). These probabilities become expectations of development intensity: a parcel that is simulated as urban with an 80% probability across the 240 runs is four times more likely to accommodate new buildings than one with a 20% probability. Treating the probability surface as a proxy for development demand acts a fundamental step to bridge the gap between the ABM's stochastic land parcel conversion probabilities and the detailed image-based inputs required by the GAN for urban form generation.

There were strong practical reasons for using the development probability as a proxy for development density. Pix2pixHD requires detailed and granular input images to perform synthesis. Continuous rasters (e.g. elevation, density) have been commonly used to guide road generation in previous DGM studies (e.g. Birsak et al., 2022; Fang et al., 2022; Gu et al., 2024). The ABM simulation maps showing simulated urban probability at the parcel scale were insufficient to provide the nuanced information needed to train the GAN for realistic road network generation. Converting the development probabili-

ties into a development density raster provided the explicit input necessary for the GAN to learn from and generate urban form effectively. In data-scarce environments, where detailed spatiotemporal land parcel data is frequently unavailable, the integration of more widespread data sources for GAN training is necessary. Attempting to classify and map urban parcel change across all of Southeast Asia would have been infeasible. Instead, road density rasters from OSM, were used to guide the GAN training, allowing the extraction of 2,400 training and testing images. The dasymetric mapping method used to convert probabilities into density is changeable across different scenarios, with variations in how buildings are distributed. This practical flexibility allows for further investigation into how density inputs affect outputted urban form patterns. Overall, translating development probability to a density raster was a viable approach, as the GAN could only be trained on inputs that were spatially granular and sufficiently available across a vast geographical region, such as road density derived from OSM data. This necessitated that future urban demand in the study area was represented in a comparable density format so the trained GAN could effectively interpret this input and generate plausible future road networks. It was a fundamental step that enabled the entire integrated modelling workflow to function effectively.

6.2.2 Dasymetric mapping for business as usual, economic recession and land bubble scenarios

The following approach for building distribution mapping was applied to the BAU, ER and LB scenarios as development patterns and trends are expected to stay consistent with historic urban change in the study area. The distribution approach focused on buildings and urban area change within the study area only. In these three scenarios the status quo of urban planning is being maintained for different reasons. In BAU, there is not an expectation for future development densities and patterns to differ from the historic ones since 2014. In ER and LB, developers and urban planners do not want to exacerbate the current scenario conditions and therefore keep development densities consistent with historic development.

The land parcels that changed to urban between 2014 and 2023 were identified and a count of the number of Google buildings which had their centre in these parcels was taken. The average curtilage was calculated from the total parcel area and number of buildings. Table 6.1 shows the specific statistics. Then the number of buildings that could feasibly fit into the new 2040 urban area was calculated and the total parcel area was subdivided into equal area polygons. The subdivided polygons were converted to point centroids and a density raster at 10 m resolution was generated from them. This created a uniform distribution of development throughout the study area that is consistent with historic growth as Boreys have been developed, which are gated housing

communities of equal density featuring gridded housing units (Figure 4.3).

6.2.3 Dasymetric mapping for the sustainable planning and economic boom scenarios

For the SP and EB scenarios, a different approach was taken using the Phnom Penh EO4SD data (Section 3.2.4). The low, medium and high density EO4SD parcels (Figure 3.3) were dissolved in each classification and Google building polygons were counted that had their centre in the dissolved polygon and the average building area with curtilage was calculated (Table 6.1).

To distribute the different densities of buildings in the 2040 urban area for the SP scenario, the probability that a parcel would turn urban from the ABM was used as a proxy for building demand. This probability was then divided into three classes based on natural breaks in the data to identify high, medium, and low demand parcels. Each demand classification was subdivided into equal areas based on the curtilage area from the EO4SD data (Table 6.1). The subdivided polygons were converted to point centroids and a 10 m density raster was generated. This approach is intended to create a varied distribution of building density in the study area. Urban planners in the SP scenario are interested in generating a sensible urban structure with differing levels of density, essentially creating preplanned cities to effectively accommodate the future population growth.

For the EB scenario, the EO4SD high density building curtilage value (Table 6.1) was used to generate subdivided polygons within the study area and then convert to centroids. As developers are trying to maximise profits in this scenario due to good economic conditions, it was considered appropriate to use the highest density option from EO4SD so as many buildings as possible can be developed.

Table 6.1: Building and parcel statistics from study area and EO4SD data. Area values in m^2.

Statistics	Study area	Low density	Medium density	High density
Total parcel area	6,800,341	34,021,227	15,292,141	30,947,666
Building quantity	12,649	143,340	48,982	152,918
curtilage area	537	237	312	202

6.2.4 Scenario dasymetric mapping outputs

Figures 6.1 and 6.2 show the 2040 building density rasters for the five scenarios and Table 6.2 reveals statistics for these density rasters. These figures show visually distinct

patterns that reflect the unique intensities and spatial arrangements of the buildings from each scenario.

The EB scenario has the highest statistics in all scenarios, with 69,116 buildings projected, indicating a high concentration of buildings and significant variability. The SP scenario forecasts 59,134 buildings. The standard deviation in the SP scenario shows more variability, suggesting that the building distribution is less consistent than the EB scenario (Table 6.2). This is expected as the SP scenario should have a more varied density due to using three different building areas when generating the demand raster.

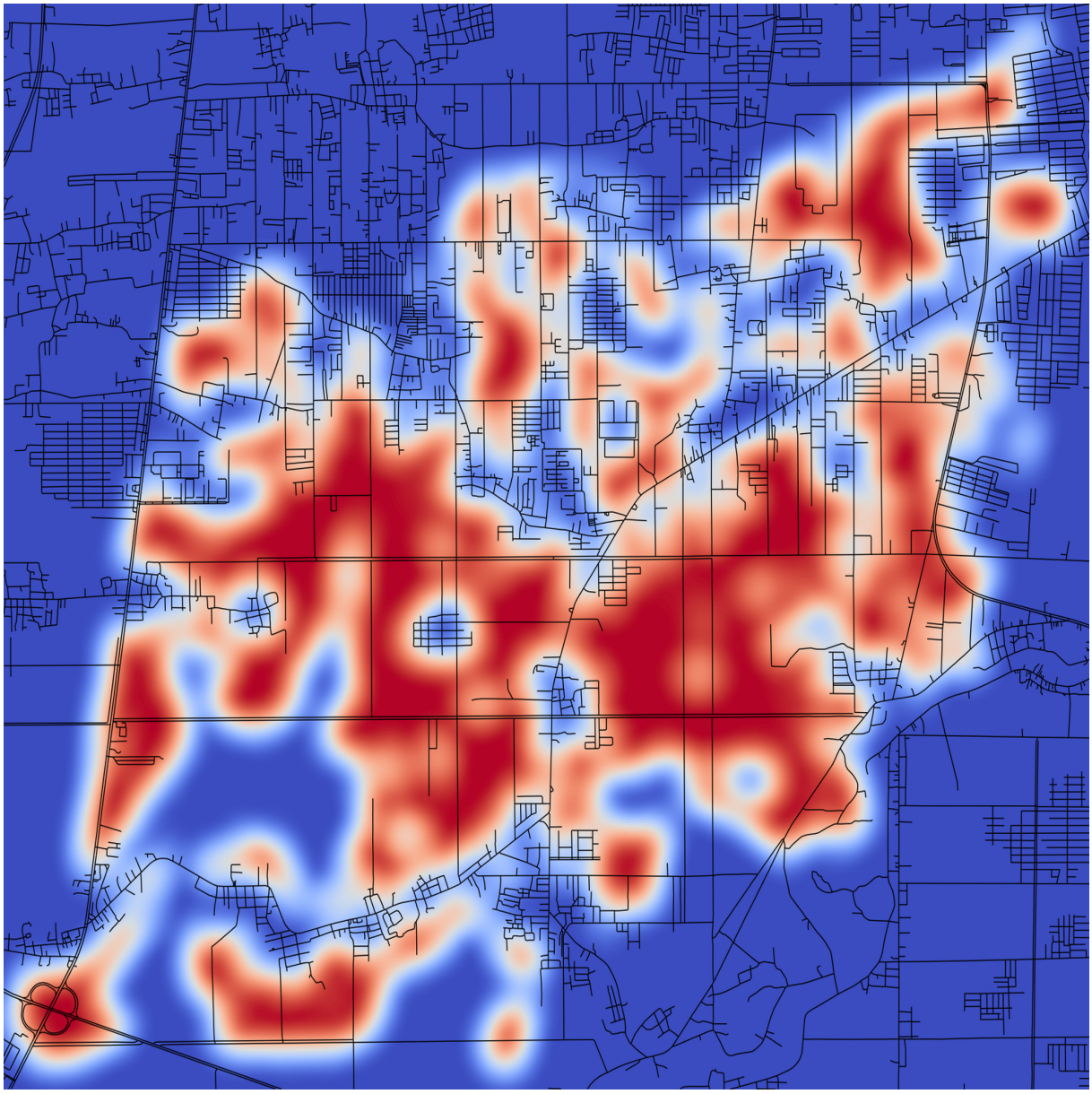

Figure 6.1: BAU 2040 building demand input image for pix2pixHD. Red = higher future building density, blue = low future building density.

In contrast, the BAU scenario forecasts 33,607 buildings and shows a more moderate density pattern, with a maximum and mean density of 1,863.45 and 592.24 buildings per km^2, respectively. The standard deviation of 652.18 in this scenario suggests mod-

erate variability, indicating a relatively uniform spread of building density compared to highly clustered patterns in the EB and SP scenarios.

The ER and LB scenarios exhibit the lowest mean densities at 438.37 and 401.24 buildings per km^2, respectively. Their maximum densities and standard deviations are also the lowest of the five scenarios. The quantity of new buildings is also the lowest at 13,874 for ER and 11,690 for LB. These values suggest that these scenarios have low-density development, as shown by the scattered and less clustered building distributions (Figure 6.2) compared to the other scenarios.

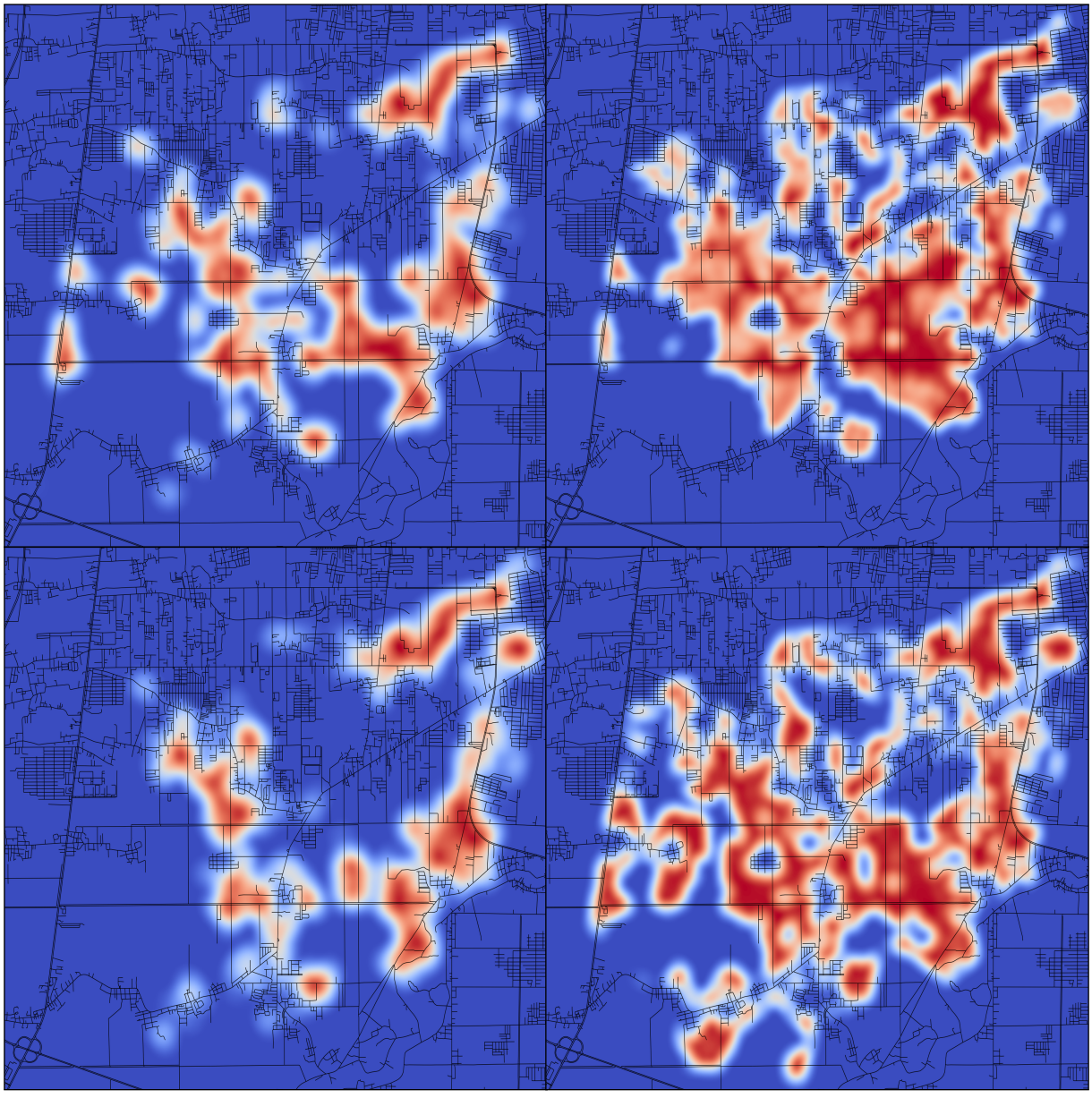

Figure 6.2: 2040 scenarios building demand input images for pix2pixHD. Top row: ER scenario (left) and SP scenario (right). Bottom row: LB scenario (left) and EB scenario (right). Red = higher future building density, blue = low future building density.

In general, EB and SP scenarios indicate clustered high-density development, whereas the BAU, ER, and LB scenarios illustrate lower-density patterns with more dispersed

building arrangements. These variations suggest that different planning priorities, such as prudently accommodating population growth or maximising profitability, directly impact building density and distribution in the study area. A positive sign is that all the scenarios show distinct and unique patterns both visually and statistically. The mean densities vary widely, with the lowest and highest being 206.01 (LB scenario) and 1,163.63 (EB scenario) buildings per km^2. The maximum densities for BAU, ER and SP are all very close, but their mean densities vary much more, whereas the maximum and mean densities of EB and SP show a greater difference compared to the other three scenarios.

Table 6.2: Building density statistics for each scenario in buildings per km^2.

Scenario	Maximum Density	Mean Density	Standard Deviation	Building count
BAU	1,863.45	592.24	652.18	33,607
ER	1,838.68	244.55	438.37	13,874
SP	4,951.90	1,042.39	1,539.85	59,134
LB	1,809.09	206.01	401.24	11,690
EB	4,825.33	1,163.63	1,526.48	69,116

6.3 Pix2pixHD

The next stage in the framework is to use the pix2pixHD GAN to generate synthetic urban form images to depict future road networks and buildings in the new urban area simulated by the ABM and densities generated in the previous section.

6.3.1 Pix2pix baseline

Pix2pixHD by Wang et al. (2018) is an advancement on the original pix2pix model by Isola et al. (2017) as it is designed to operate at higher image resolutions by using a coarse-to-fine generator, multi-scale discriminator architecture, and robust adversarial learning objective function. Pix2pix is a conditional GAN framework for image-to-image translation which uses a U-Net for the generator G and a patch-based fully convolutional network for the discriminator D. It is trained on a pair of corresponding images, s_i and x_i. The objective function of pix2pix is expressed as:

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{s,x}[\log D(s, x)] + \mathbb{E}_s[\log(1 - D(s, G(s)))] \tag{6.1}$$

where G tries to minimise the objective against D which tries to maximise it in this minimax game:

$$\min_G \max_D \mathcal{L}_{\text{GAN}}(G, D) \tag{6.2}$$

6.3.2 Coarse-to-fine generator

The generator architecture displayed in Figure 6.3 is divided into two sub-networks: G_1 (global generator network) and G_2 (local enhancer network). G_2 enhances the resolution of G_1 by 4 times (2 times along each image dimension), for example, a 1024 x 1024 image from G_1 would be enhanced and outputted to 2048 x 2048 by G_2. G_1's architecture is built upon Johnson et al. (2016) which includes a convolutional front-end, a set of residual blocks and a transposed convolutional back-end. The image is passed through these components to produce an image of the same resolution. G_2 also has the same components, but the input to G_2's residual block is the element-wise sum of the feature maps from both G_2's front-end and G_1's back-end, integrating global information from G_1 into G_2. In the training phase, G_1 is trained first followed by G_2 and then all networks are fine-tuned together. This design effectively merges global and local information and enables efficient image synthesis by using a multi-resolution pipeline.

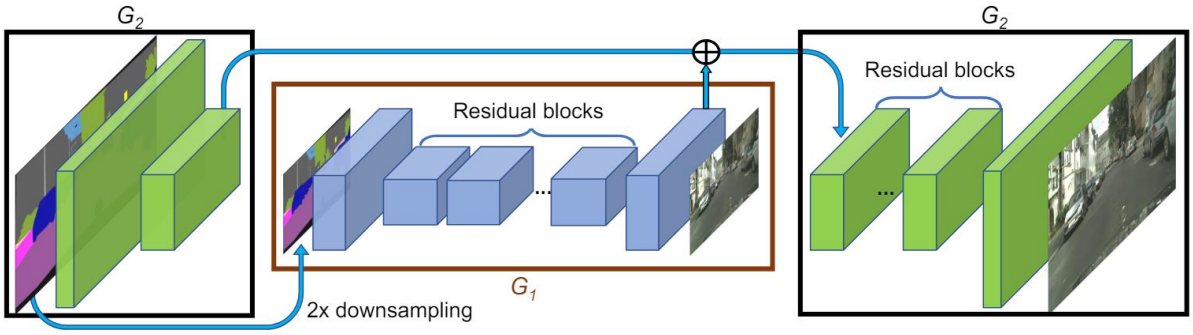

Figure 6.3: Pix2pixHD generator architecture (Wang et al., 2018).

6.3.3 Multi-scale discriminators

A large receptive field is required by the discriminator to differentiate between high-resolution real and synthesised images. Pix2pixHD uses three multi-scale discriminators with the same network structure. The real and synthesised images are downsampled by factors of 2 and 4, creating a three-scale image pyramid. Each discriminator is then trained to differentiate real from synthesised images at their respective scale. The coarsest scale discriminator has the largest receptive field which provides a global view of the image and ensures global consistency. The finest scale discriminator helps the generator produce intricate details. By using this approach, training the coarse-to-fine generator is simplified because extending to higher resolutions only requires adding a fine-scale discriminator, not retraining the entire network. Without these multi-scale discriminators, generated images would exhibit repeated patterns.

6.3.4 Adversarial loss

The original GAN loss from pix2pix (Equation 6.1) is improved by introducing a feature matching loss based on the discriminator (Isola et al., 2017). This addition helps stabilise the training process by requiring the generator to produce realistic statistics across multiple scales. This is done by extracting features from various layers of the discriminator and matching these intermediate representations between real and synthesised images. The feature extractor up to the ith layer of the discriminator D_k is represented as $D_k^{(i)}$, encompassing the computation from input to the ith layer of D_k. The feature matching loss is calculated as:

$$\mathcal{L}_{\text{FM}}\left(G, D_k\right) = \mathbb{E}_{(s,x)} \sum_{i=1}^{T} \frac{1}{N_i} \left[\left\| D_k^{(i)}(s,x) - D_k^{(i)}(s, G(s)) \right\|_1 \right] \tag{6.3}$$

where T is the total number of layers and N_i denotes the number of elements in each layer. This feature matching loss is similar to perceptual loss, which has been useful in previous image super-resolution and style transfer (Johnson et al., 2016; Ledig et al., 2017). The full objective function of pix2pixHD combines the GAN loss and the feature matching loss as:

$$\min_G \left(\left(\max_{D_1, D_2, D_3} \sum_{k=1,2,3} \mathcal{L}_{\text{GAN}}\left(G, D_k\right) \right) + \lambda \sum_{k=1,2,3} \mathcal{L}_{\text{FM}}\left(G, D_k\right) \right) \tag{6.4}$$

where λ controls the importance of $\mathcal{L}_{FM}$ and $\mathcal{L}_{GAN}$.

6.4 Road network generation

6.4.1 Generative adversarial network experiments

Pix2pixHD requires an input image and corresponding target image, so the 2400 training and testing cells of type 1 (major roads with small sets of minor residential roads) and 2 roads (minor residential roads) generated in Section 4.4.1 were converted to images as follows. For the GAN experiments performed, the input image was of the type 1 and 2 roads, with a percentage of the type 2 road sets omitted and replaced by a density raster. The target image contained the same roads as the input along with the omitted roads in red (Figure 6.4). The use of a density raster was intended to provide the GAN with ambiguous or partial information, encouraging it to learn the underlying road patterns. This approach aimed to train the GAN to generalise and mimic realistic road structures based on incomplete inputs, enabling it to later generate meaningful outputs when provided with future building density raster under different ABM scenarios (Figures 6.1 and 6.2).

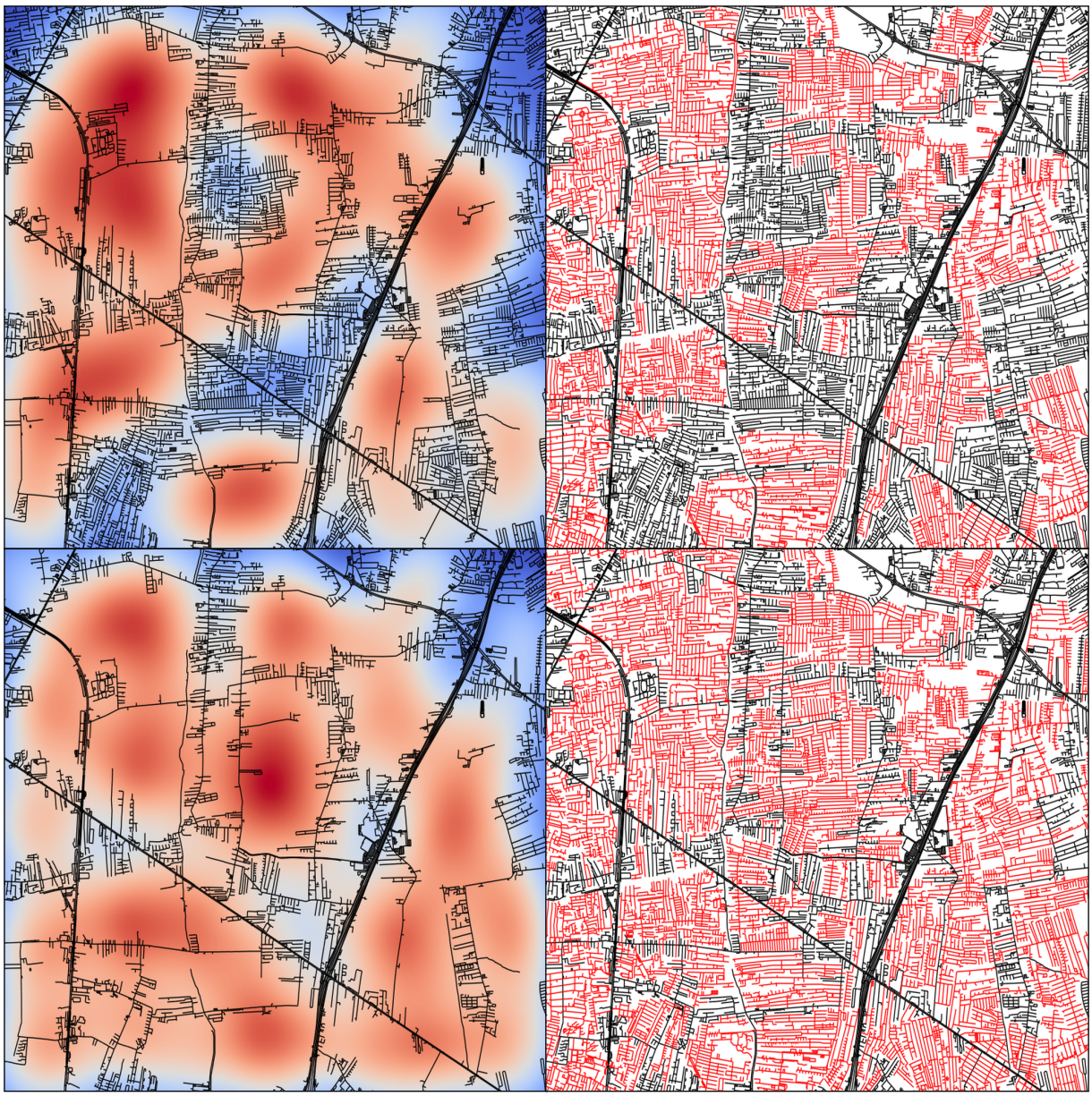

Figure 6.4: Two examples of the percentage-based removal of type 2 road sets, replaced by their density. Input image (left) and the target image (right) at a resolution of 1024×1024 pixels. Top row: 25% of type 2 road sets removed. Bottom row: 75% of type 2 road sets removed.

Twelve experiments were performed for the generation of road networks in three groups. All groups involved removing the same percentages of type 2 roads to assess the sensitivity of the GAN's performance. However, the image resolution and the generator network parameter (global or local) was varied between the three groups:

- Group 1 - generated images at 1024x1024 pixel resolution of the fishnet cells and trained using the local generator network.
- Group 2 - generated images at 1024x1024 pixel resolution of the fishnet cells and trained using the global generator network.
- Group 3 - generated images at 2048x2048 pixel resolution of the fishnet cells and trained using the global generator network.

The global generator network is a fully convolutional network which processes images at their original resolution. The local generator network acts as an additional refinement layer on top of the global generator, adding extra residual blocks to enhance local details in high-resolution outputs. (Wang et al., 2018). Within each group, the four experiments involved removing 25%, 50%, 75%, and 100% of the type 2 road sets (Figure 6.4). The sets of contiguous type 2 networks were randomly selected when removed and weighted by the length of the network. Due to the fishnet polygons being at a relatively large spatial scale for a GAN it is important to measure and assess the performance across variations in pixel size and key model parameters.

Other pix2pixHD parameters were kept at the default values. The learning rate was 2.00×10^{-4} which was kept constant for the first 100 epochs but for the final 100 it was linearly decayed to 0. GAN training progress was saved every 25 epochs. The fineSize parameter was kept at its default value of 512×512. This parameter defines the cropped size of the input image, allowing the model to extract multiple crops from the same image. As a result, it effectively increases the number and diversity of training samples and reduces memory usage, making training more efficient. These experiments were run on an Amazon Web Services (AWS) g5.xlarge instance designed for GPU-intensive tasks. This instance is equipped with a NVIDIA A10G Tensor Core GPU, which provides 24 GB of VRAM (AWS, 2025). This large VRAM make it a suitable AWS instance to use for image-based urban form generation.

6.4.2 Road network post-processing

As the GAN output was in image format, these needed to be converted to geospatial data. The following steps were performed to do this:

1. Red roads in the generated images were identified which represent the generated road networks. A binary mask was created where 0 is no roads and 1 is roads.
2. The image was resized and transformed to align with its respective fishnet cell location.
3. The binary mask was skeletonised, converting it into a thin version retaining its topological properties. A graph is constructed from the skeletonised image.
4. The coordinates of the graph nodes were transformed into geographic coordinates. Each edge of the graph was converted to a line string object representing a segment of the skeletonised feature in geographic space.
5. The line string objects were then converted to a shapefile with the CRS of its fishnet cell.

Once the test images had been converted to shapefile format, tidying was performed based on Hartmann et al. (2017). This tidying resulted in a set of smooth and well-integrated road shapefiles suitable for further analysis:

1. Firstly, some road segment artefacts at junctions were removed based on their length, which all fell within a specific range.
2. This step was repeated multiple times with different length criteria to filter out unwanted or isolated segments and retain only the relevant parts.
3. The sharp angles in the lines were then smoothed with a tolerance of 2 metres. This helped create smoother and more visually appealing road lines by reducing sharp angles and irregularities.
4. After smoothing, the lines were simplified by removing redundant nodes while preserving the overall shape of the roads. The tolerance for simplification was also set to 2 metres.
5. The generated road lines were then snapped to the existing road network if they were within 20 metres of it. This ensured that the generated roads were seamlessly connected to the existing road network, eliminating gaps and misalignments.
6. Finally, all 400 test shapefiles were merged together for the accuracy assessment with the ground truth road network.

6.4.3 Road network accuracy metrics

As is common in GAN research, the FID score introduced by Heusel et al. (2017) was used to assess the quality of the generated images in the 400 test images dataset (Wu and Biljecki, 2022, 2023; Gu et al., 2024). The mIoU score was omitted as this metric is based on pixel overlap between images. The aim of urban form generation here is to capture the overall structure and layout of urban form, not to produce an exact match between generated and ground truth images.

However, as identified in Section 2.5.3, the pixel-based metrics alone are not appropriate to assess the connectivity, topology, and completeness of the network against the ground truth. Vector-based accuracy metrics are therefore needed to more effectively evaluate the completeness, topology, and connectivity of road networks. Converting outputs to vector and using further metrics is needed to fully assess the accuracy of generated road networks.

An additional three vector-based road accuracy metrics were used based on the previous urban GAN literature. These metrics were selected from the literature for their

efficiency in capturing the core characteristics of the road network. The calculations for these metrics are shown in Eq. 6.5, 6.6, and 6.7 where $_{GN}$ and $_{GT}$ refer to generated and ground truth road networks, respectively. Firstly, road length compares the generated and ground truth total road length (RL) (Moradi et al., 2022):

$$\Delta\ RL\ (\%) = \left(\frac{RL_{GN}}{RL_{GT}} - 1\right) \times 100. \tag{6.5}$$

City blocks (spaces bounded by roads on all sides) were extracted and the generated and ground truth average block area (BA) were compared as shown in Eq. 6.6 below (Hartmann et al., 2017):

$$\Delta\ BA\ (\%) = \left(\frac{BA_{GN}}{BA_{GT}} - 1\right) \times 100. \tag{6.6}$$

Node degree is defined by the number of direct links a node has and reflects the importance of nodes within their network (Yang et al., 2023a). The generated and ground truth node degree counts (ND) were compared as shown in Eq. 6.7 to determine the count accuracy per node degree:

$$\Delta ND_d = \frac{ND_{GN,d}}{ND_{GT,d}} - 1. \tag{6.7}$$

Where d is a specific node degree. The resulting absolute node degree accuracy values were then weighted by the ground truth count for their respective node degree and averaged to produce the overall node degree accuracy with Eq. 6.8:

$$ND\ (\%) = \left(\frac{\sum_d |\Delta ND_d| \times w_d}{\sum_d w_d}\right) \times 100. \tag{6.8}$$

Where $w_d = ND_{\mathrm{GT},d}$ is the weight for node degree d, which is the count of nodes with degree d in the ground truth network. This weighting ensures that node degrees with fewer nodes had less influence on the overall result. For example, a node degree of six might only occur in one node for the ground truth data, but if the generated data has no nodes with a degree of six, this could disproportionately skew the results.

Road length assesses the network completeness, whereas block area and node degree count focus on the network structure and connectivity, respectively. These metrics provide a comprehensive approach to quantifying the GAN experiment performance in road generation. The mean absolute error (MAE) was calculated from the three vector metrics to determine the best performing GAN to use as it is a more direct and unambiguous measure of the magnitude of the average error than other calculations, such as the root mean square error (Willmott and Matsuura, 2005). Additionally, using absolute values prevents the cancellation of positive and negative errors. For example, averaging -50% and 50% would result in a perfect accuracy of 0%, which does not reflect the true accuracy. This is also why the absolute values were used in Eq. 6.8.

6.4.4 Study area road network assessment

Although the above accuracy metrics are a good indicator of performance using the test data, an assessment should also be performed on the BAU building demand raster (Figure 6.1). The top performing group of experiments were selected and tested on the study area with the BAU raster because the ambition of the research is to have sufficient future urban form information whilst retaining realism of current urban form. The experiment which produced the highest total road length, number of blocks, and total node degree was defined as the best experiment to use for generating future road networks.

6.5 Building footprint generation

6.5.1 Generative adversarial network experiments

The 2,400 training and testing 500 m^2 and 1000 m^2 cells generated in Section 4.4.2 for Cambodia were converted to images as follows. This part of urban form generation involved an input image of the road network and a target image of the road network and building footprints in red (Figure 6.5). Due to the compactness of buildings, their outline was made white so that the individual footprints could be clearly delineated and identified as a singular building. Density rasters were not used as inputs for the building GAN because existing GAN studies have demonstrated effective approaches to generate building footprints directly from road network layouts (e.g. Wu and Biljecki, 2022, 2023). The development of a conditioning strategy for road generation based on density and existing road layouts required several months of research, a more established and readily available approach was needed for the building footprint generation due to time constraints. Consequently, the approach from Wu and Biljecki (2023) that used road layouts as the conditioning input was adopted. In addition, the density inputs of the road GAN are implicitly represented through the road layouts in the input image.

Four GAN experiments were performed for buildings in two different groups. One group performed GAN generation using the 500 m^2 cells with one using the local generator and the other using the global generator. In each group, the two experiments involved training the model with the 500 m^2 and 1000 m^2 cells. All experiments were performed at an image resolution of 1024x1024 because the spatial scale of the cells was much smaller than that of the road network cells. The building footprints are capturing a maximum area of 1 km^2 in their images, compared to the 57 km^2 area in the road network images. GAN training at 2048x2048 resolution on a smaller area was unnecessary as these had a long runtime and subsequent computational cost to them, meaning an exponential increase in cost for a minimal gain in performance.

As this was a more straightforward task for the GAN than the road network generation, some of the parameters were adjusted. The learning rate was lowered to 2.00×10^{-5} and was kept constant for the first 50 epochs and the linearly decayed to 0 for the subsequent 50 epochs. GAN training progress was saved every 25 epochs. The g5.xlarge AWS instance used for the road network experiments was used for building footprint experiments (AWS, 2025).

Figure 6.5: 500 m^2 (top row) and 1000 m^2 (bottom row) spatial scale of the building footprint images. Input image (left) and target image (right) at 1024x1024 resolution.

6.5.2 Building footprint post-processing

The test images were resized and transformed to align with their respective fishnet cell location. A binary mask was generated where the red buildings were 1 and every other

colour were zero. The binary mask was converted into a shapefile to get polygons of the building footprints for each test cell. The shapefiles of each cell were then merged into a single shapefile. Once in shapefile format, the footprints were normalised to remove undesirable artefacts and correct distortions in their geometry by using a polyline compression algorithm (Gribov, 2019). The building footprints with holes were filled. Only footprints with an area greater than 10 m^2 and less than 5000 m^2 and within 25 metres of existing roads were retained. This avoided spuriously small or large buildings or those too far from the roads, so these did not impact the accuracy metrics.

6.5.3 Building footprint accuracy metrics

The FID score was used again for building footprint accuracy, as well as three vector building footprint accuracy metrics which are taken from Wu and Biljecki (2023). In Eq. 6.9, 6.10, and 6.11, the $_{GN}$ and $_{GT}$ symbols refer to generated and ground truth, respectively. The change in building count (Eq. 6.9) compares the total generated building count and the ground truth building count (BC):

$$\Delta\ BC\ (\%) = \left(\frac{BC_{GN}}{BC_{GT}} - 1\right) \times 100. \tag{6.9}$$

Eq. 6.10 compares the average generated and ground truth footprint area (FA):

$$\Delta\ FA\ (\%) = \left(\frac{FA_{GN}}{FA_{GT}} - 1\right) \times 100. \tag{6.10}$$

Eq. 6.11 compares the average generated and ground truth footprint perimeter (FP):

$$\Delta\ FP\ (\%) = \left(\frac{FP_{GN}}{FP_{GT}} - 1\right) \times 100. \tag{6.11}$$

Similarly to road accuracy metrics, the building count assesses the completeness of the generated building, whereas the footprint area and the perimeter focus on the general size and shape of those generated, respectively. These distinct metrics provide a comprehensive approach to quantifying the performance of the GAN experiments in building generation. The MAE was calculated from the three building footprint vector metrics to determine the best performing GAN to use.

6.5.4 Study area building footprint assessment

The best performing group of experiments were applied to the BAU scenario generated road network and the GAN which produced the highest quantity of building footprint metrics was ranked the best and used in all scenarios. A 500 m^2 and 1000 m^2 fishnet was generated and filtered, so only cells that intersect the BAU road network were fed into the GANs to populate the road networks with buildings footprints.

The GAN 2040 building outputs for each scenario were converted to density rasters to statistically and visually assess how accurately they reflect the density patterns from the dasymetric inputs described in Section 6.2.4. Each scenario's GAN building density raster was compared to its corresponding dasymetric density raster shown in Figures 6.1 and 6.2. This assessment aimed to determine how well the GAN outputs reflect the density characteristics of the input data used for urban form generation.

For each scenario, two statistical tests were applied: a t-test to compare the mean densities of the rasters, and the Kolmogorov–Smirnov (KS) test to compare the pixel value distributions of the rasters (Kim, 2015; Berger and Zhou, 2014). For both tests, 500 pixel values were randomly sampled from the GAN and dasymetric rasters. These tests were only applied to building density rasters, as both the GAN and dasymetric building densities were in the same units of measurement (buildings per km^2). In contrast, the GAN 2040 road outputs was excluded from this comparison because it differed in its unit of measurement (km of road per km^2) and variable type, violating key assumptions of the statistical tests and rendering the results invalid if applied (Kim, 2015; Berger and Zhou, 2014). The statistical hypotheses for both tests were:

- Null hypothesis (H_0): There is no significant difference between the two groups (i.e., the means or distributions are equal).
- Alternative hypothesis (H_1): There is a significant difference between the two groups.

If the p-value is less than 0.05, the null hypothesis is rejected, indicating a statistically significant difference between the two density rasters. If the p-value is greater than 0.05, the null hypothesis is not rejected, suggesting insufficient evidence to conclude a significant difference in either the means or the distributions.

6.6 Road network generation results

6.6.1 Road network accuracy outputs

Table 6.3 shows the accuracy values for the road network generation. Lower FID score values are more favourable. For the road length and block area vector metrics, values closer to zero indicate better performance, with negative values representing underestimation and positive values indicating overestimation. For the node degree and MAE columns, values closer to zero are better because this indicates less deviation from the ground truth values, irrespective of overestimation or underestimation. The initial insight drawn from Table 6.3 highlights a notable issue regarding GAN underestimation

in the comparison of generated and ground truth road networks, as reflected in the road length and block area vector metrics. Specifically, all experiments consistently demonstrate an underestimation in the quantity of roads when compared to the ground truth.

The 25% experiments across the three groups have the most favourable road length, block area, and node degree values. For the 1024x1024 local group, the best MAE is the experiment where 25% of type 2 roads were removed, with a score of 36.12%. The FID score is also the best at 29.52 in all the groups and experiments. As type 2 road removal increases, the overall accuracy score consistently decreases, reaching a very high value of 81.61%. This reduction corresponds to an increase in FID values, indicating that the experiments struggled to maintain performance when generating greater quantities of road networks. This pattern of decline also occurs in the 1024 global and 2048 global experiments.

Similarly, the 1024x1024 global group's 25% experiment have the best overall accuracy at a slightly higher value of 38.25% and FID of 25.99 compared to the local group. In addition, this experiment has the best road length, block area and node degree accuracy values within the group, but the road length and node degree metrics outperform the 25% 1024x1024 local experiment. As percent removal of road networks increase, both FID and average accuracy drop sharply, but the 100% value remains lower here than in the global 1024x1024 experiment. In this group of experiments, the average accuracy values of the 25%, 50% and 75% are higher than the local 1024 group. This shows that using the local option in pix2pixHD does improve vector metrics performance, with the exception of the 100% experiment.

Average accuracy for the 25% experiment in the 2048x2048 global group achieves the best score at 34.46% and second-best FID of 28.17 for the 12 experiments conducted. Within the 2048x2048 group, the 25% experiment has the best road length and node degree values, but the 50% experiment has a better block area value at -1.45%, which is the best value for this metric in all 12 experiments. In this group, the block area accuracy metrics have extremely high values showing these experiments capture the average area of road blocks well. However, comparing the groups, the 25% 1024x1024 global experiment has the best road length and node degree values across all 12 experiments. As with the other two groups of experiments, in the 2048x2048 globabl group the MAE and FID deteriorates as road removal percentages increases.

Table 6.3: Accuracy results for the road network generation GAN experiments. RL = road length, BA = block area, ND = node degree, and MAE = mean absolute error (of vector metrics). All values are in % apart from FID.

Experiment	FID↓	△ RL	△ BA	ND↓	MAE↓
1024x1024 resolution - local					
25%	29.52	-42.22	-20.13	46.02	36.12
50%	37.40	-54.96	-25.21	60.01	46.73
75%	43.32	-57.32	-36.93	60.82	51.69
100%	131.99	-84.46	-81.71	78.67	81.61
1024x1024 resolution - global					
25%	25.99	-39.29	-33.65	41.80	38.25
50%	35.58	-50.75	-42.76	54.56	49.36
75%	82.86	-70.91	-63.96	64.27	66.38
100%	91.41	-74.30	-53.69	72.13	66.71
2048x2048 resolution - global					
25%	28.17	-46.08	-4.42	52.87	34.46
50%	32.67	-47.59	-1.45	59.48	36.17
75%	33.30	-47.74	-16.45	58.68	40.96
100%	98.97	-84.04	-92.96	73.65	83.55

The 100% removal of type 2 roads has spuriously high MAE values in the 1024x1024 local and 2048x2048 global groups. This is an indication that the 100% GAN experiments in all groups have experienced problematic training due to inconsistent values between the groups. Based on only the 25%, 50% and 75% overall accuracy values for the groups, the 2048x2048 global group has the lowest average accuracy at 37.20%, whereas the 1024x1024 local and global has 44.85% and 51.33%, respectively. Although the 25% value in the local group does feature a good accuracy of 36.12% which can compete with some of the 2048 resolution values. The standard deviation of the 1024 local averages is 6.49, 1024 global average is 11.56 and the 2048 global average is 2.75. This indicates a clear pattern of higher resolution and better overall and more consistent performance from 2048 GAN group. As a result, the 25%, 50%, and 75% experiments in the 2048x2048 group were assessed on the study area.

The FID scores do not align well with the overall accuracy score from the vector-based metrics in Table 6.3. For example, the 25% 1024x1024 global experiment results in the highest FID at 25.99, whereas it's vector-based average accuracy is the fourth lowest at 38.25%. When looking at the same GAN experiment between groups the FID scores show a larger difference than the overall vector-based accuracy metrics. In the 75% experiments between the local and global 1024x1024 groups, the FID value has a difference of 39.54 but the average vector-based metrics have a difference of only

14.69% (Table 6.3). Although direct comparisons between FID and vector metric values are challenging, there is a trend of conflict between the values and more variability in the FID scores.

6.6.2 Node degree outputs

Table 6.4 shows further information about the node degree performance of the 25%, 50%, and 75% 2048x2048 GAN experiments. This table summarises the accuracy, generated quantities, and ground truth quantities of node degree counts for GAN-generated data at 25%, 50%, and 75% completion levels. The accuracy values are based on Equation 6.7 and values closer to zero are more favourable.

All GAN experiments in Table 6.4 achieve a higher accuracy in capturing node with degrees of 1, 2, and 3 but severely underperform in capturing node degree 4,5 and 6. The 25% experiment has the node degree 1, 2, and 3 accuracy values closest to zero at -42.07%, -57.65% and -31.04%, respectively. The 75% GAN achieves the best node degree 4 at -94.35% and the 50% GAN achieves the best node degree 5 at -96.26%, but both of these accuracy values are very far from zero. None of the GANs are successful in generating a node with degree 6, although the ground truth road network only has one node degree 6 present.

Table 6.4: Node degree outputs for 2048 25%, 50% and 75% GANs. ND = node degree and accuracy is in %.

ND metrics	ND 1	ND 2	ND 3	ND 4	ND 5	ND 6
25% GAN						
Accuracy	-42.07	-57.65	-31.04	-95.11	-97.65	-100.00
Generated	102,663	502,337	151,639	1,214	2	0
Ground Truth	177,234	1,186,059	219,879	24,823	85	1
50% GAN						
Accuracy	-49.69	-64.28	-37.54	-94.72	-96.26	-100.00
Generated	119,872	561,096	180,644	1,695	4	0
Ground Truth	238,262	1,570,910	289,197	32,126	107	1
75% GAN						
Accuracy	-48.31	-63.99	-34.30	-94.35	-96.75	-100.00
Generated	143,315	649,771	216,147	2,073	4	0
Ground Truth	277,258	1,804,432	328,973	36,709	123	1

There is a significant underestimation issue across all the GAN experiments, with the accuracy closest to zero being -31.04% where the 25% GAN generates 151,639 nodes with a degree of 3 compared to the ground truth of 219,879. With the exception of node

degree 2 in all experiments, there is a trend of accuracy deterioration as the quantity of ground truth node degree counts decreases. For example, the 50% GAN achieves an accuracy of -49.69% for node degree 1 with 238,262 nodes with this degree in the ground truth, but node degree 4 has an accuracy of -94.72% with much fewer nodes at 32,126. The results for node degree 4 across all experiments are unsatisfactory. Considering that all the ground truth node degrees are in the tens of thousands, this score should be higher, especially when node degree 4 is an indication of gridded networks. This evidence suggests that the GAN generally struggles to generate more complex junctions, possibly due to its image-based format and post-processing steps when converting to vector.

6.6.3 Road network visual outputs

Whilst the accuracy metrics reveal the best performing GAN experiments, visual inspection of the test data also reveals useful insights to further verify conclusions made. As noted in the previous section, all 100% experiments have high average accuracy values and high FID scores that indicate erroneous generation of roads networks. Figure 6.6) shows the results of these experiments and supports this argument that shows that the GAN is producing a limited and repetitive set of generated networks. In addition, the 1024x104 global 75% experiment is also a high value compared to its counterpart experiments and similar to the 100% experiment in this group. Visualisation of this experiment's results in Figure 6.7 also show the same issues as all the 100% experiments. From this assessment, it can be argued that any FID score greater than 50 or average accuracy greater than 60 is too deteriorated a performance and that the GAN model has become erroneous.

Figures 6.6 and 6.7 exhibit mode collapse, a well-known and common issue in GANs. Mode collapse occurs when the generator produces a limited variety of outputs, often repeating the same patterns (Goodfellow, 2016). If the generator discovers an especially plausible output, it may learn to generate only that output, as its goal is to fool the discriminator. Ideally, the discriminator should reject such repetitive samples, forcing the generator to diversify. However, if the discriminator gets stuck in a local minimum and fails to adapt, the generator can exploit this weakness, leading to a cycle where it over-optimises for the current discriminator without learning a diverse distribution. This results in the generator producing a narrow set of outputs rather than capturing the full data distribution (Zhang et al., 2018).

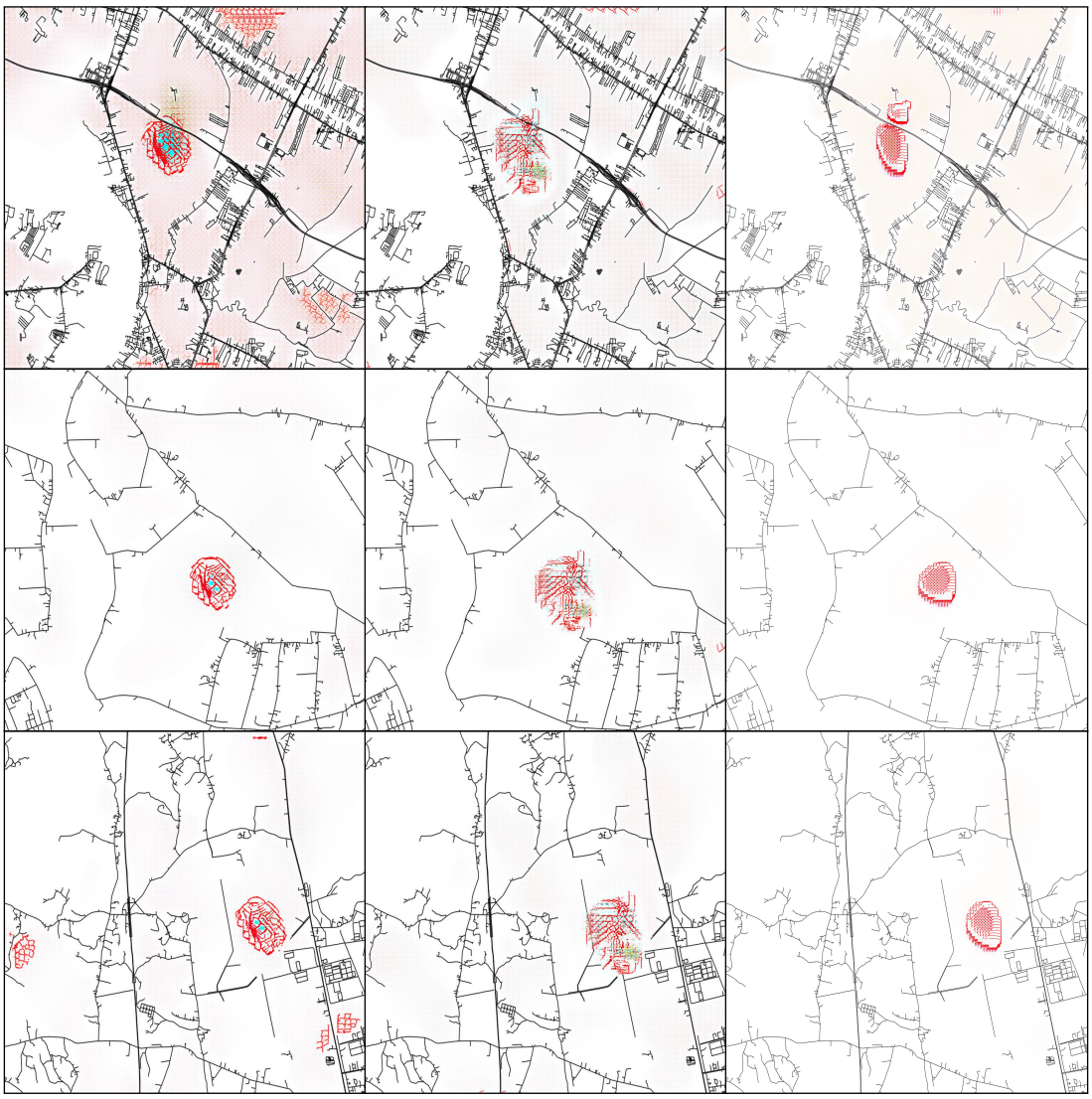

Figure 6.6: The 100% GAN output. Left: 1024 global, middle: 1024 local, right: 2048 global.

Figure 6.7: Sample of 1024x1024 75% global experiment outputs.

Figure 6.8 visualises the GAN outputs from the 2048x2048 experiments. It shows that the GAN performs well in capturing road network patterns against the ground truth, although their accuracy scores do vary slightly. This qualitatively verifies that the three experiments are all appropriate to use in the study area assessment, especially considering that the 1024 groups are much more susceptible to GAN issues when training,

such as the 75% and 100% experiment in the 1024x1024 global group. The 2048 group provides a much more consistent and stable set of experiments based on the visual outputs in Figure 6.8 and accuracy scores in Table 6.3.

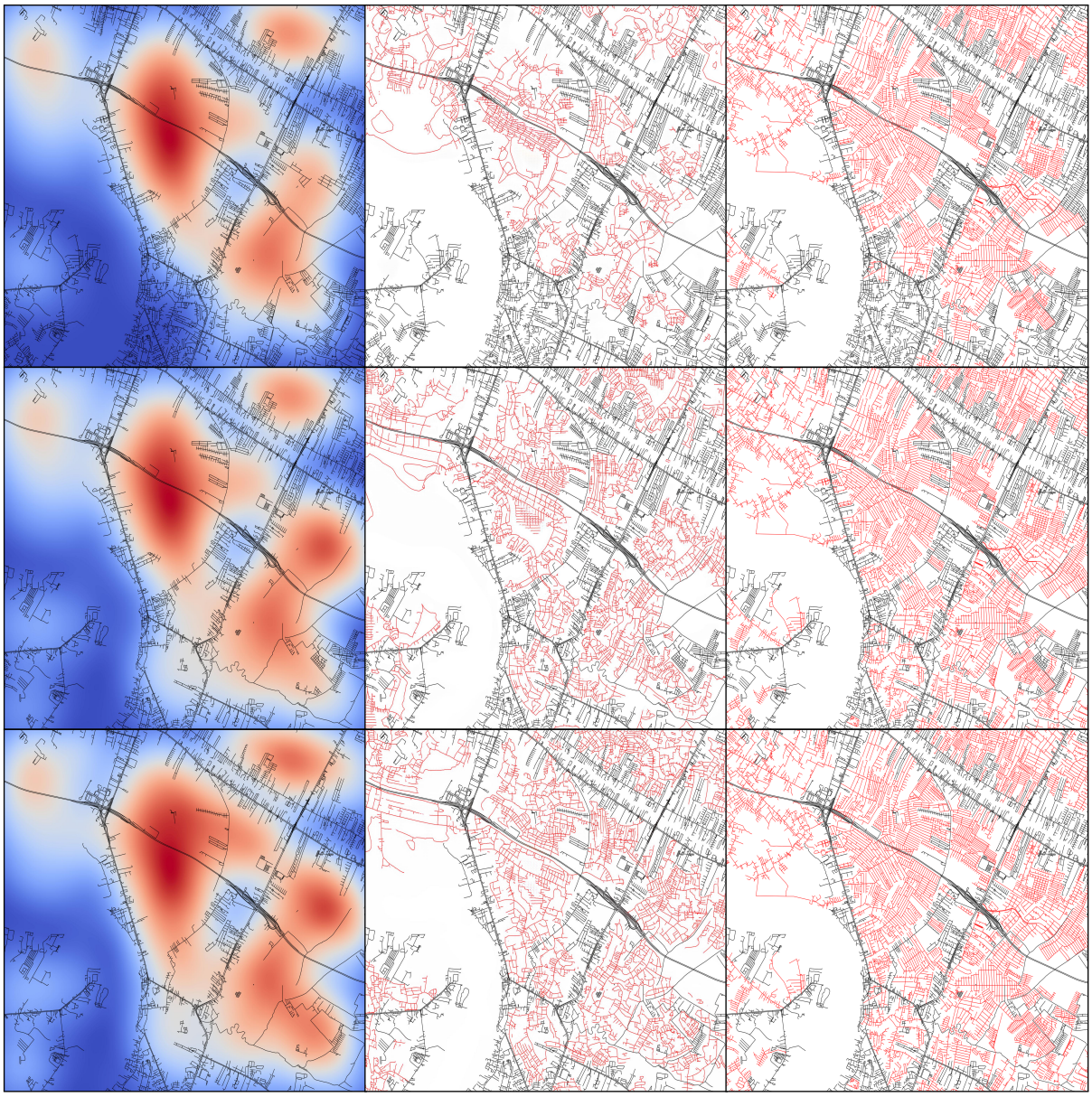

Figure 6.8: The 2048 GAN output. top row: 25%, middle row: 50%, bottom row: 75%. Left column: input, middle column: generated, right column: ground truth.

6.6.4 Road network generation in study area

When the 25%, 50% and 75% GANs were applied to the study area, the 75% achieved the highest values by a large margin across the metrics (Table 6.5). By removing 75% of the type 2 roads it maintains a good accuracy score whilst generating sufficient content of synthetic future road networks by 2040. Figure 6.9 shows that the 25% GAN generates a much sparser network with less gridded residential roads in the areas expected to transition to urban by 2040. The 50% and 75% outputs generate realistic residential neighbourhoods in the high demand areas with some major roads connecting them.

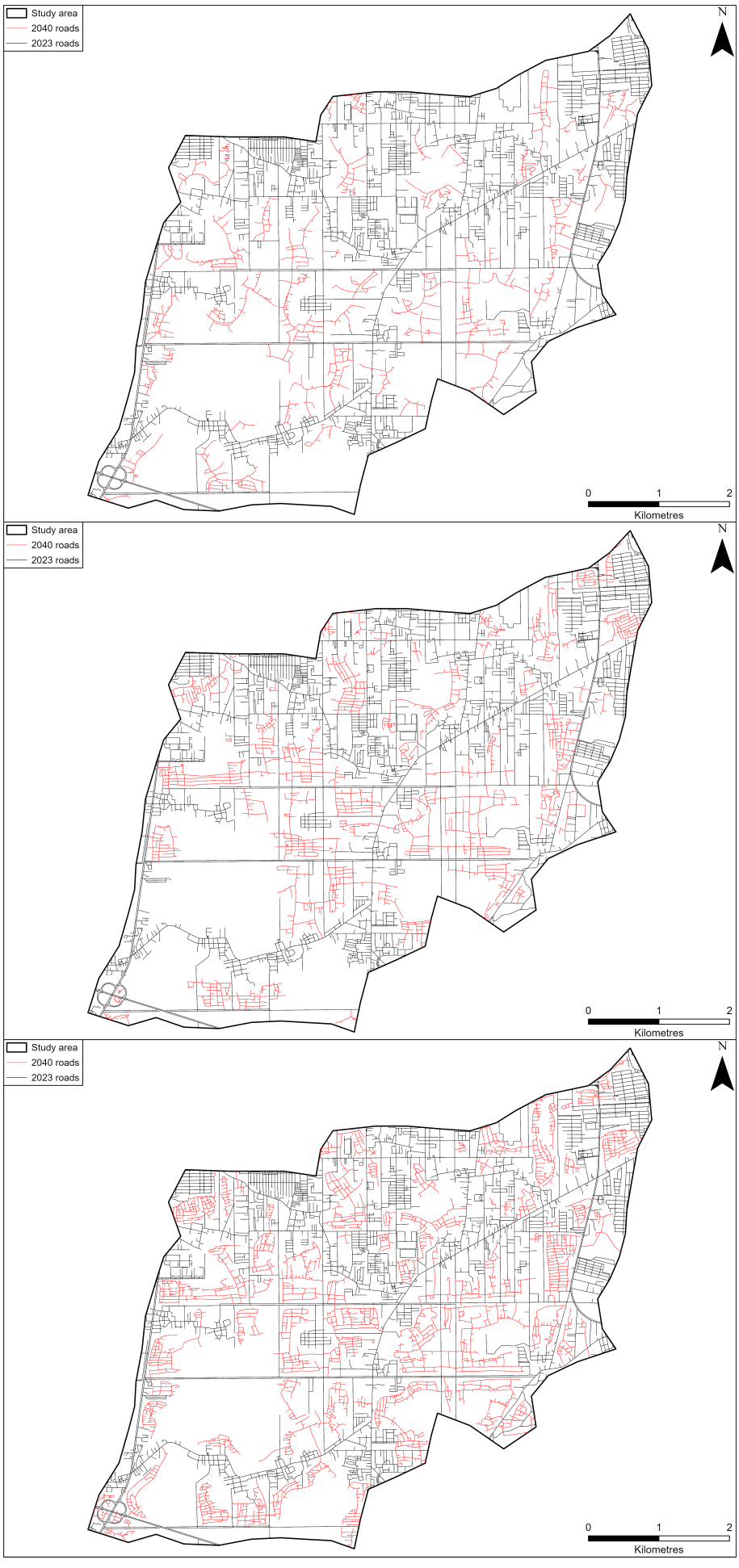

Figure 6.9: The GAN outputs on the study area (25% top, 50% middle, 75% bottom).

Table 6.5: Total length, number of blocks, and node degree on AOI for 2048x2048 outputs.

Percent	Road length↑	Block quantity↑	Total node degree↑
25%	92,265	109	6,464
50%	149,613	284	8,592
75%	227,683	579	16,648

6.7 Building footprint generation results

6.7.1 Building footprint accuracy scores

Table 6.6 shows the accuracy scores for the GAN building experiments. These results provide a clearer answer to the best performing group to apply to the study area compared to the road network scores in Table 6.3. The local group has the best accuracy, with 1000 m^2 experiment achieving the best average score at 29.48%. The generated mean footprint area and perimeter across all the experiments are overestimated against the ground truth, whereas the count of buildings is underestimated. This suggests that the GAN produced fewer but larger building footprints. Furthermore, the building count accuracy values are much further from zero compared to the area and perimeter metrics. In general, the overestimation of the area and perimeter metrics is relatively minor compared to the larger underestimation of the building count of the GAN experiments (Table 6.6).

Table 6.6: Accuracy scores for the building footprint generation. BC = building count, FA = footprint area, FP = footprint perimeter, and MAE = mean absolute error (of vector metrics). All values are in % apart from FID.

Experiment	FID↓	△ BC	△ FA	△ FP	MAE↓
local					
500 m^2	44.74	-63.32	59.76	39.78	54.29
1000 m^2	48.69	-60.85	11.70	15.90	29.48
global					
500 m^2	42.38	-63.76	69.23	46.63	59.87
1000 m^2	49.35	-63.32	28.50	28.38	40.07

The FID score has contradictory results with the average accuracy of the vector-based metrics in Table 6.6. If the FID score was used solely, then the global 500 m^2 GAN experiment would be considered the most accurate. However, the more robust vector metrics reveal that the 1000 m^2 local GAN experiment has the highest overall accuracy. When comparing the FID values across the two groups, the values are more consistent

in the 1000 m^2 experiments in the local and global groups, only varying by a difference of 0.66. These FID scores are generally higher than the road network FID scores in Table 6.3, with the 1000 m^2 local and global experiments very close to 50 which was considered too high for a reasonable output in the road network GAN experiments.

Neither the FID, vector metric or average scores in Table 6.6 are spurious relative to the other values, suggesting that none of the GAN models completely failed while being trained on the images. The 1000 m^2 local experiment gets the best vector-based metric scores out of all the experiments. The building count metric is similar in all the experiments, with variation in the average scores likely due to how well the GAN experiments capture the footprint area and perimeter.

6.7.2 Building footprint visual outputs

Visualisation of the best-performing local GAN experiments in Table 6.6 is necessary for additional verification. Figure 6.10 shows an example output from the local group. There is a clear underestimation issue from the GAN experiments across the three selected images. The GAN does capture the shape of building footprint well, with the majority being rectangular with 4-5 sides, although some do appear to be irregular shapes which are not as well defined as the building footprint shapes in the ground truth images. The ground truth building footprints do not always follow the road layout perfectly, with some buildings being very offset from the roads. The GAN struggles to generate these offset buildings as the generation is based on the road network layout. This is particularly true for the top row in Figure 6.10. The GAN appears to perform better when provided with highly dense road networks (middle row in figure 6.10), but still suffers from underestimation by not being able to delineate densely packed buildings and instead generating larger single building footprints.

The 1000 m^2 local GAN outputs in Figure 6.11 presents some contrasting results to the 500 m^2 experiment in 6.10. Given the larger spatial area, the GAN effectively generates buildings along roads in suitable and probable areas and not along major highways where buildings would not be constructed in reality (top row in Figure 6.11). However, in the bottom row of Figure 6.11 there are building footprints along some of the major roads that the GAN does not generate. The GAN struggles to effectively capture the unpredictable nature of building development in Cambodia, with building development patterns difficult to recognise and identify. Underestimation is an issue because the GAN does not generate buildings guided by the road network as in the 500 m^2 experiments. Larger building generation appears to be problematic, with the footprint being less well defined compared to smaller building footprints (bottom row, Figure 6.11).

Figure 6.10: 500 m^2 local experiment outputs. Left: input image, middle: generated image, right: ground truth image.

Although the FID scores for these local experiments are in their 40s, the visual output shows that they do produce viable outputs which can be used for building footprint generation. This contrasts with the GAN FID scores for the road network in Table 6.3, the best performing being in the 20-30 region and an upper limit of anything 50 or higher being considered too high for good results. This supports the idea that the FID score alone is insufficient to assess the GAN accuracy, as it is sensitive to both spatial scale and pixel resolution of the images, leading to variations in its values. Therefore, vector accuracy metrics are essential to assess GAN performance as they enable a consistent comparison between GAN experiments on a common basis.

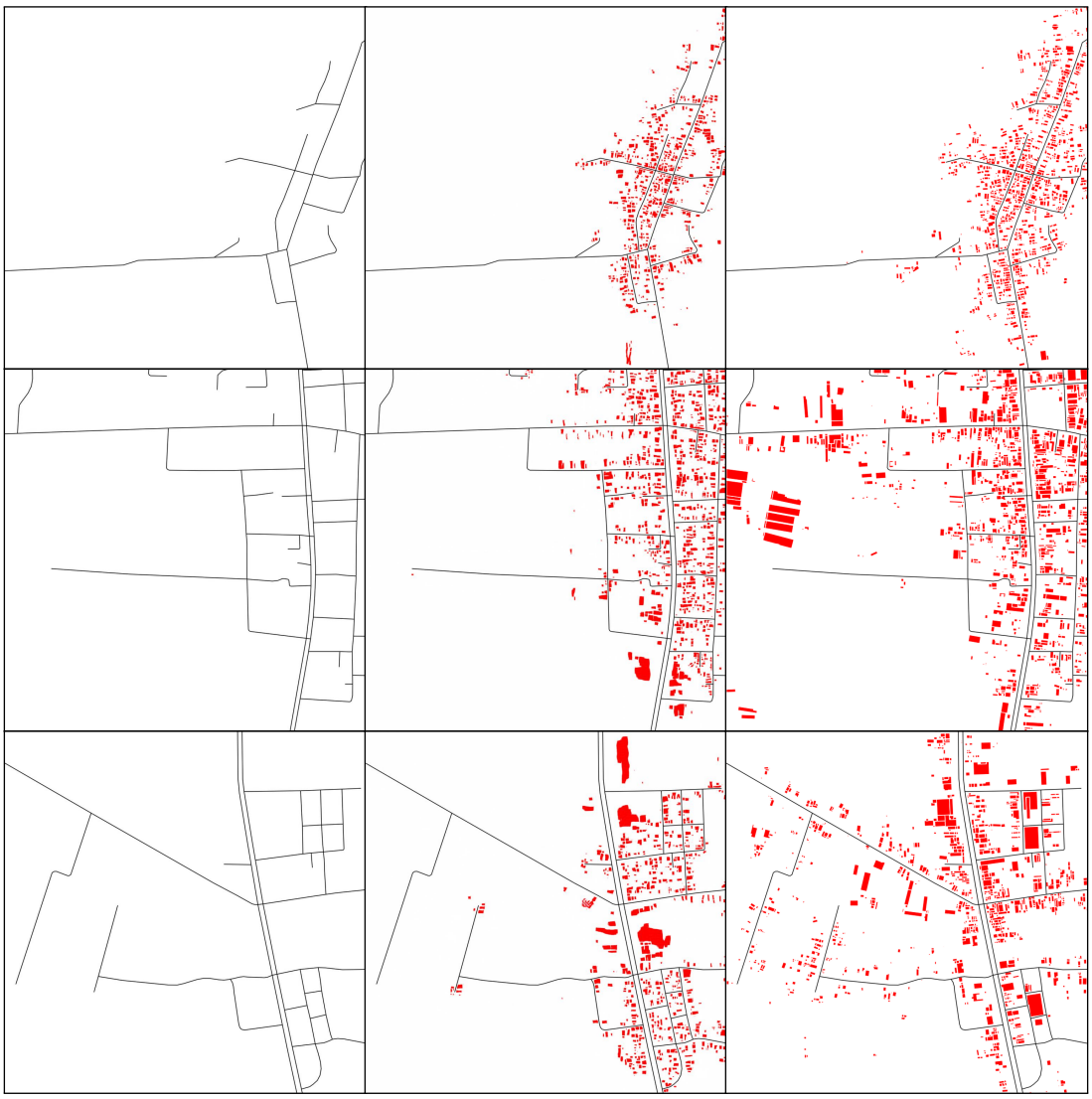

Figure 6.11: 1000 m^2 local experiments outputs. Left: input image, middle: generated image, right: ground truth image.

6.7.3 Study area building footprint generation

The two local GANs were applied to the 2040 road networks and the 500 m^2 experiment generated 17,862 buildings and the 1000 m^2 experiment generated 16,628 buildings. This difference is relatively small, with the difference being 1,234 buildings and Figure 6.12 reveals that the two GANs perform better on some parts of the 2040 BAU road network depending on the location. Therefore, the 500 m^2 building footprints were used to supplement the 1000 m^2 data as both datasets had gaps in them meaning they would compliment each other well. The 1000 m^2 data were prioritised as they had the highest accuracy score (Table 6.6), so in any case where there was an intersection of building footprints, the 500 m^2 footprints were removed.

Another benefit of taking this approach of using both GANs was it helped reduce the underestimation of building demand between the GAN and dasymetric approach. Under

the BAU, the dasymetric distribution approach forecasts 28,952 buildings by 2040 and merging the footprints together results in 21,376 buildings, representing a 26% reduction. The underestimation at 14,324 or 13,453 buildings would be much more severe if only one GAN was used, showing a 51-53% reduction and combining both approaches yield a more accurate forecast.

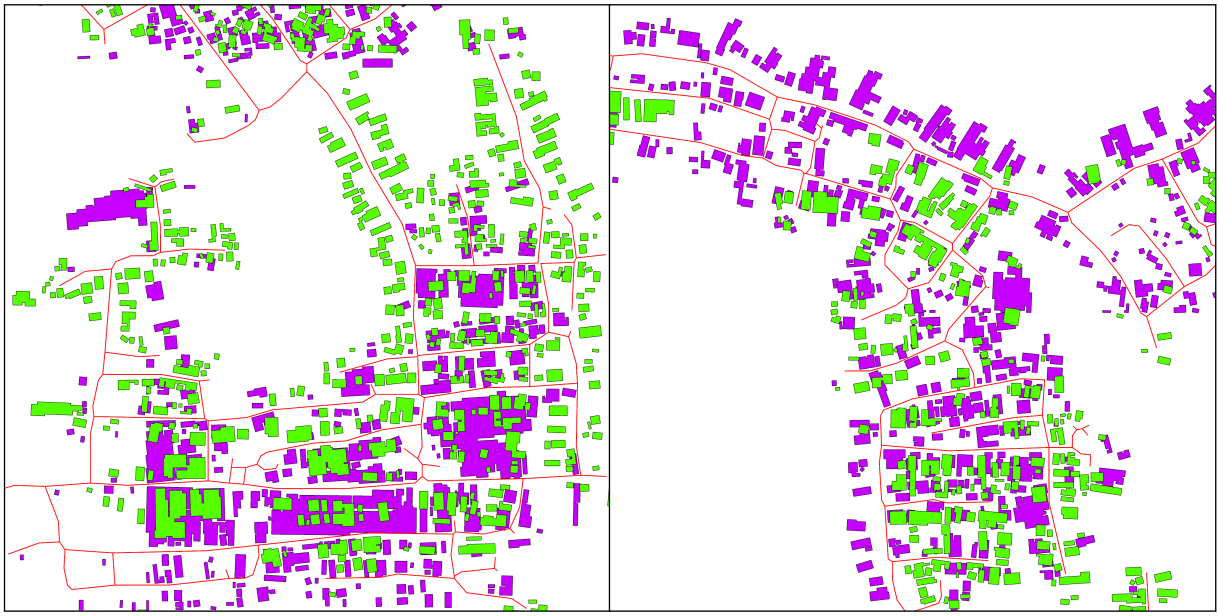

Figure 6.12: Examples of difference in generation between the 500 m^2 and 1000 m^2 local GANs. Green buildings are 1000 m^2 GAN and purple buildings are 500 m^2 GAN.

6.8 Urban form scenario results

6.8.1 Business as usual urban form

In 2023, the total road length in the study area was 318.34 km and the building count was 24,873. Figure 6.13 displays the BAU scenario of urban form by 2040. The GAN simulated 26,521 new buildings displaying an 106.63% increase from 2023 and new roads increased by 71.52% with an additional 227.68 km constructed. However, the future building count generated by the GAN is underestimated by 21.08% in this scenario compared to the dasymetric projection which projected 33,607 buildings. The 2040 roads generated appear realistic and integrate seamlessly with the existing road network which reveals a logical and believable extension of urban form. There are also numerous residential neighbourhoods and major roads connecting these areas, reflecting cohesive urban form with well-connected transportation routes.

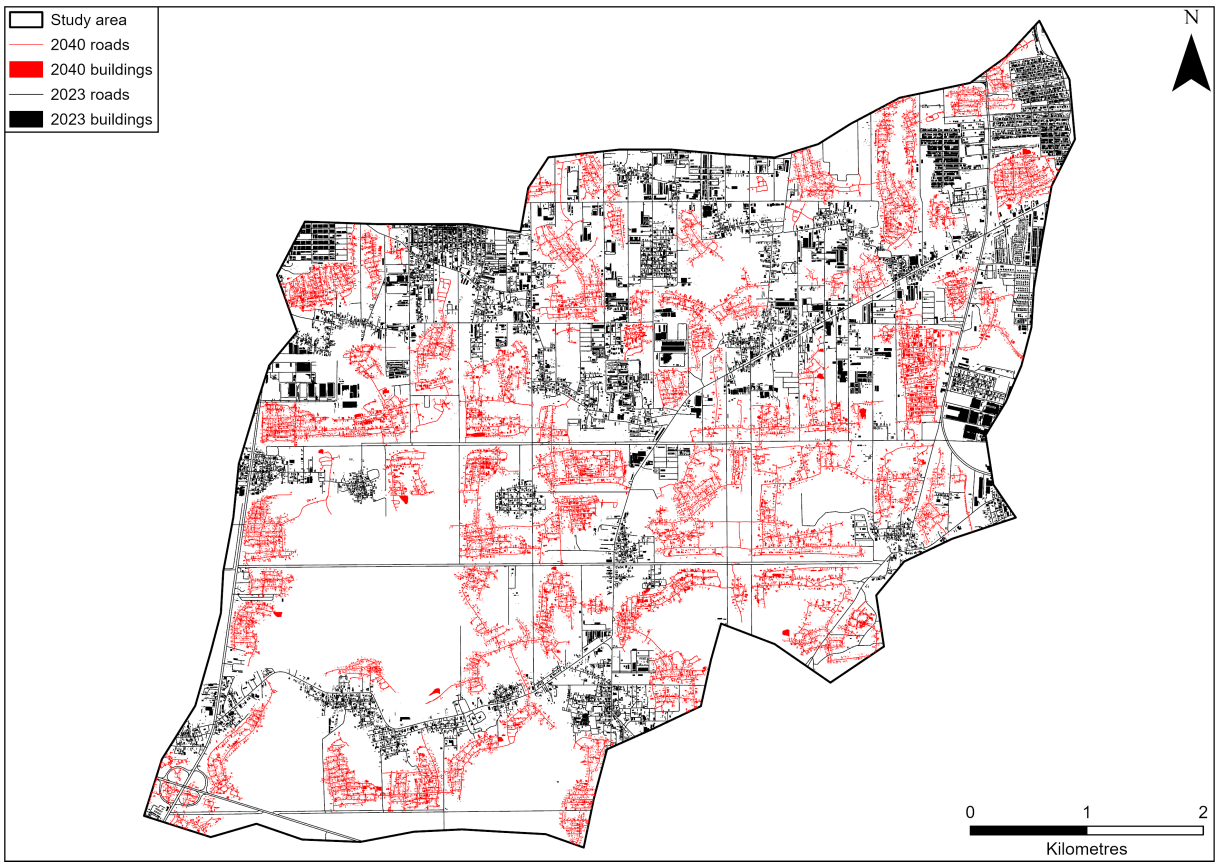

Figure 6.13: 2040 urban form for BAU scenario, generated from ABM urban growth outputs.

6.8.2 Economic recession urban form

Figure 6.13 displays the ER scenario of urban form. The GAN produced 11,634 new buildings (46.77% increase) and 83.63 km of new roads (26.27% increase). In this scenario, the dasymetric projection was 13,874 buildings so the underestimation of the GAN is slightly less at 16.14% compared to BAU. The roads and building increase is much less as is expected when fewer land parcels are being converted to urban under this scenario. However, the road networks appear less logical with long stretches running adjacently to existing major roads. There are a lot less residential neighbourhoods and building footprints are more sporadically distributed, with larger patches of undeveloped land alongside sections of the new roads.

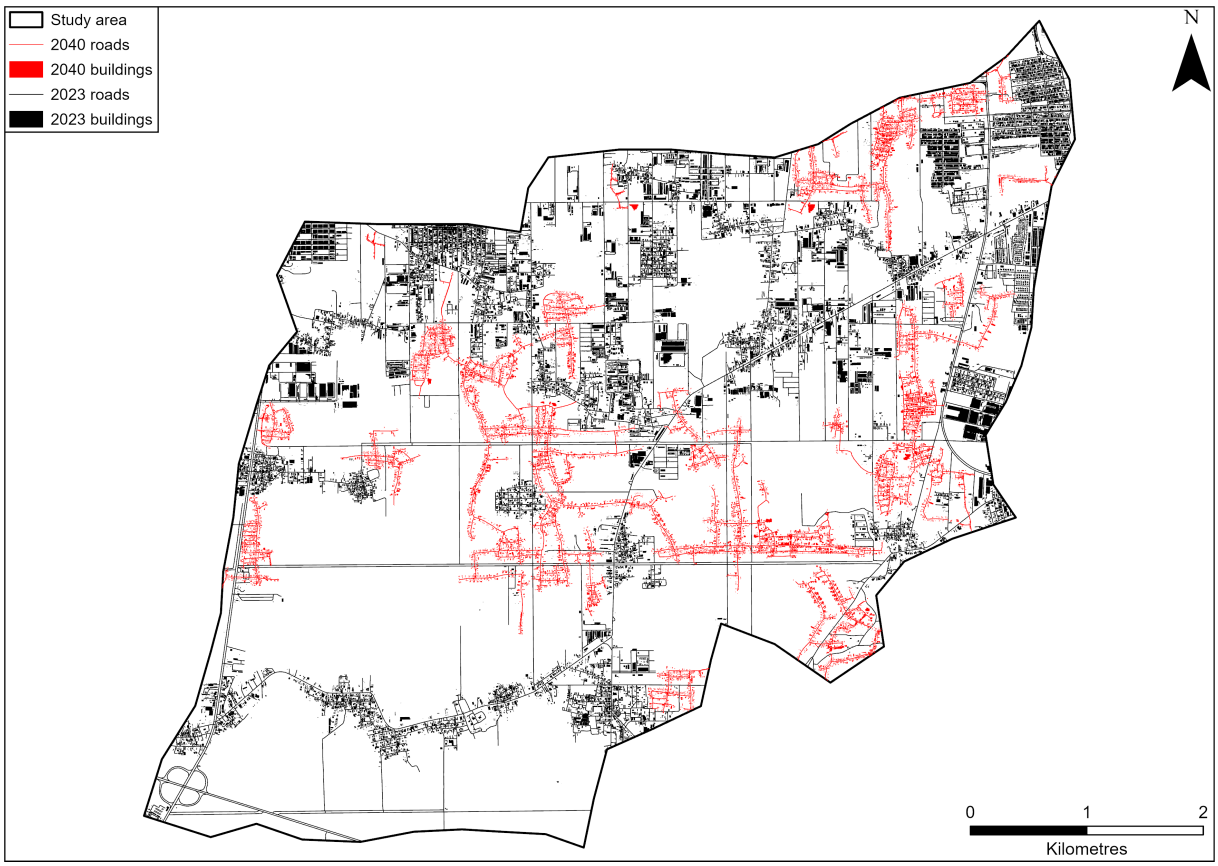

Figure 6.14: 2040 urban form for ER scenario, generated from ABM urban growth outputs.

6.8.3 Sustainable planning urban form

Figure 6.15 displays the SP scenario of urban form. The GAN produced 18,497 buildings (74.37% increase) and 153.41 km of new roads (48.19%) from 2023. The dasymetric approach projected 59,134 buildings in this scenario meaning the underestimation is much greater than the previous two scenarios at 68.72%. However, this scenario produced the third highest building footprint count out of all scenarios. The urban form patterns are visually similar to the BAU scenario in Figure 6.1, with a cohesive and highly dense future urban form. Despite the variable building density raster generated for this scenario having a much higher maximum density than the BAU, the actual impact on the generated buildings is minimal, with the difference being 35.65% in building count between these scenarios.

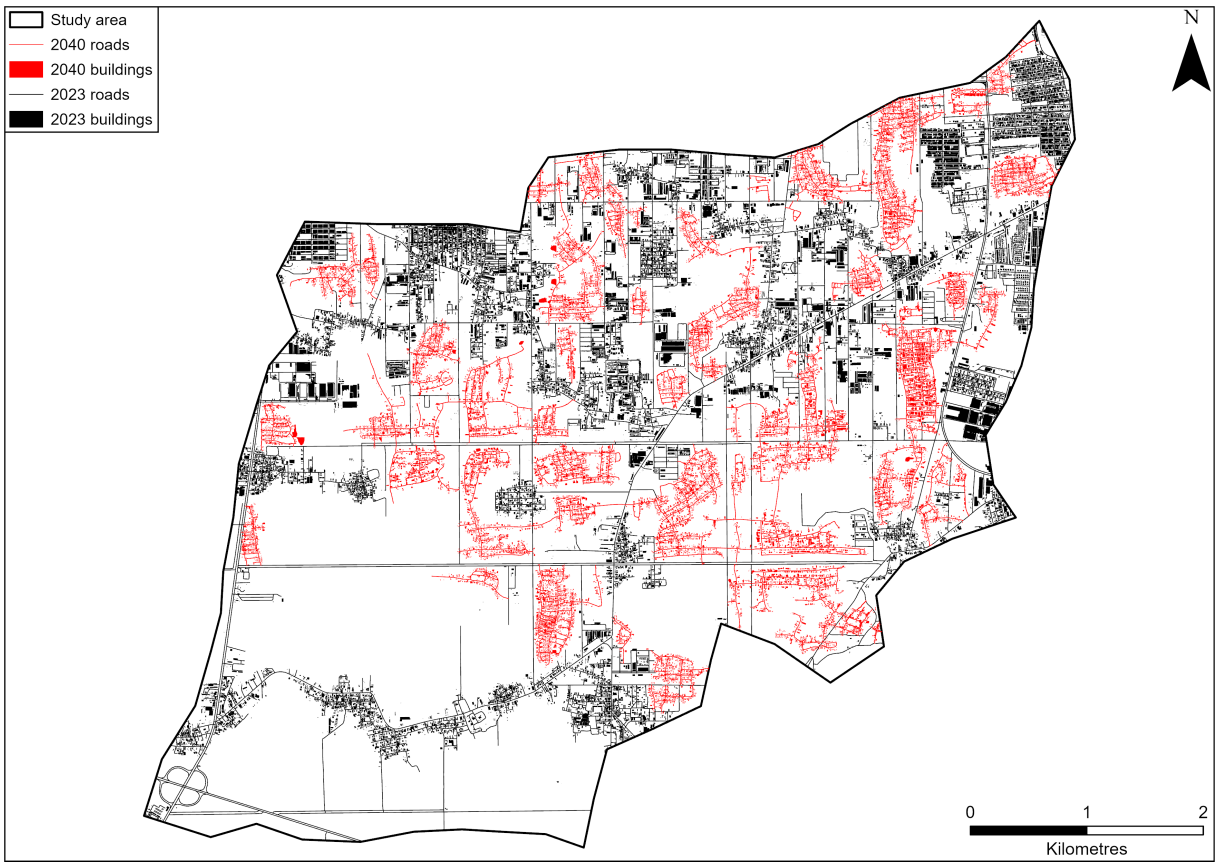

Figure 6.15: 2040 urban form for SP scenario, generated from ABM urban growth outputs.

6.8.4 Land bubble urban form

Figure 6.16 displays the LB scenario of urban form. The GAN produced 9,688 new buildings (38.95% increase) and 69.64 km of new roads (21.88% increase). The underestimation of buildings is 17.13% with the dasymetric projection being 11,690 buildings for this scenario which is the second closest GAN generation to dasymetric projection of all the scenarios. Compared to the ER scenario (Figure 6.14), this contrasts with it and features numerous residential neighbourhoods and sufficient building footprint filling of the road networks. The overall urban form output therefore appears more logical and cohesive as in the BAU and SP scenarios. However, the urban form quantity is the lowest in this scenario due to the compact nature of land development in the LB scenario.

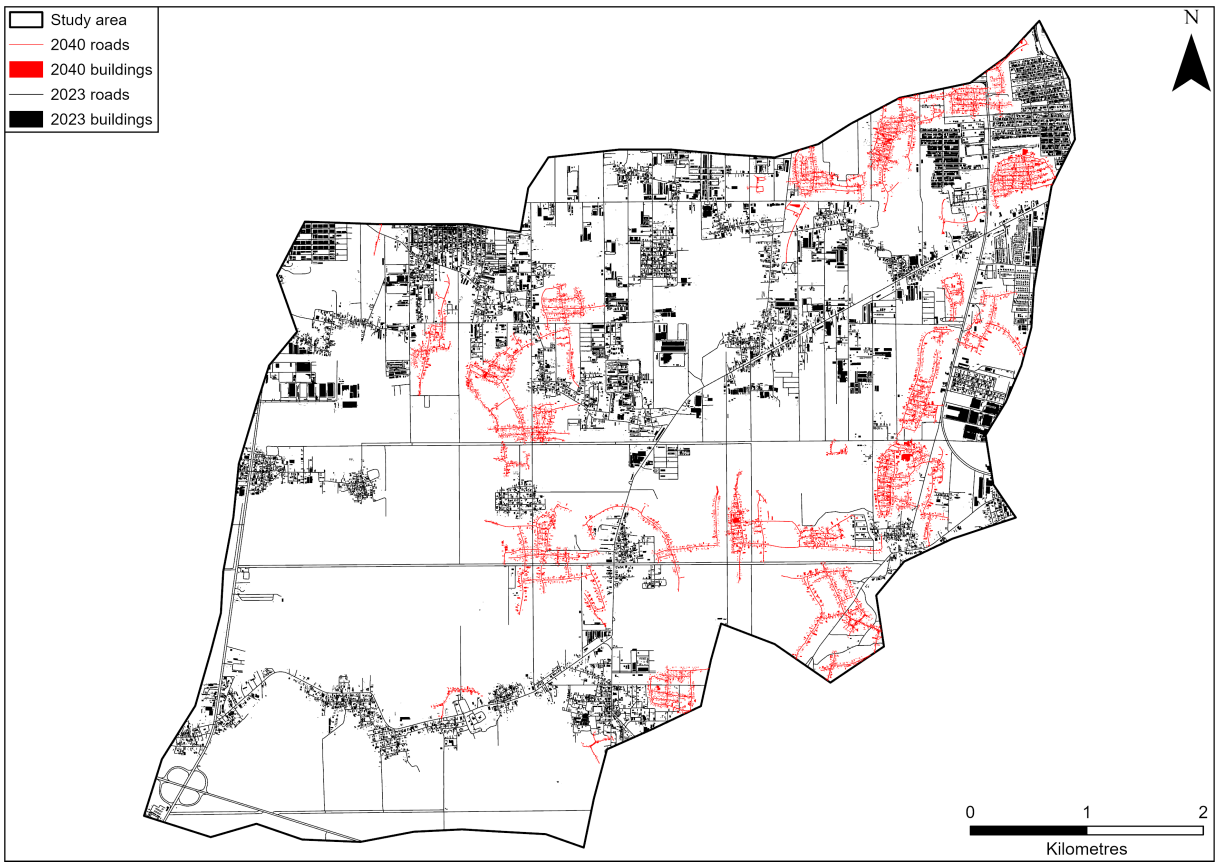

Figure 6.16: 2040 urban form for LB scenario, generated from ABM urban growth outputs.

6.8.5 Economic boom urban form

Figure 6.17 displays the EB scenario of urban form. The GAN produced 19,305 buildings (77.61% increase) and 154.31 km of new roads (48.47% increase). As the dasymetric approach projected 69,116 buildings, this scenario has the highest underestimation at 72.07%. Again, the EB scenario shares very similar urban form patterns to the BAU and SP scenarios in common places of high demand between the scenarios. The urban form output appears more cohesive and logical when compared with the ER scenario (Figure 6.14). The difference between the quantity of urban form in this scenario is not massively different from the urban form quantity of the SP scenario. The amount of land developed and dasymetric building demand projections vary greatly between these scenarios.

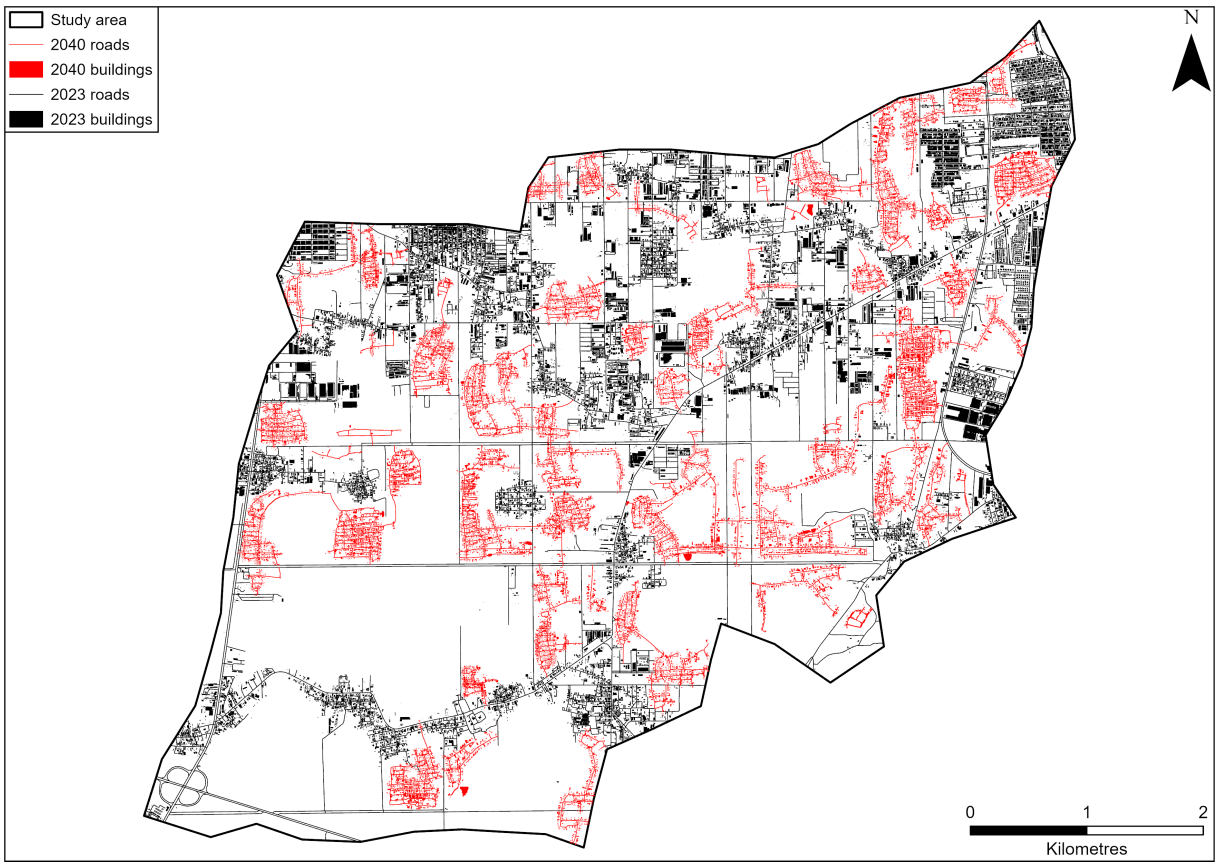

Figure 6.17: 2040 urban form for EB scenario, generated from ABM urban growth outputs.

Another observation for all scenarios is that the distribution of the input rasters used for road network generation appears to have little impact in influencing road network patterns. Figures 6.1 and 6.2 show how visually the scenarios provide different density projections. However, the road patterns are very similar in high urban demand urban areas with different density values. The input rasters effectively define the extent and location of the 2040 road networks and are responsive to the input density raster provided, generating urban form within those hotspots of development.

6.8.6 Dasymetric and generative adversarial network density comparison

Table 6.7 shows the results of the t-tests and KS tests used to determine if the spatial patterns in the density maps are the same. Across all scenarios, there is a statistically significant difference between the raster's mean (t-test) and pixel value distributions (KS test). All the values are below 0.05 and therefore the null hypothesis is rejected, indicating a statistically significant difference between the dasymetric and GAN density rasters across all scenarios. This demonstrates that the GAN building density outputs do not statistically reflect the development density inputs fed into the GAN.

Table 6.7: Results of the statistical tests used to assess whether the GAN density raster and dasymetric density raster are statistically different across the five scenarios.

Scenario	T-test p-value	ks test p-value	Mean/Dist Different?
BAU	$1.78x10^{-19}$	$1.26x10^{-10}$	yes/yes
ER	$2.67x10^{-10}$	$1.88x10^{-3}$	yes/yes
SP	$9.012x10^{-62}$	$5.92x10^{-62}$	yes/yes
LB	$2.41x10^{-10}$	$1.89x10^{-4}$	yes/yes
EB	$1.17x10^{-72}$	$2.86x10^{-42}$	yes/yes

To further exemplify this disparity in density patterns, Figure 6.18 shows both rasters for the BAU scenario. The spatial pattern similarities between these rasters is not visually compelling, the distribution of density pixels appears to vary widely between each raster. The GAN output density raster appears to be much more fragmented with isolated areas of development when compared to the more contiguous dasymetric input. However, the high-density building placements are similar, as the high-density area (red) in the GAN density raster mostly matches with high-density regions in the input density, but have a more confined area of highly dense development (Figure 6.18).

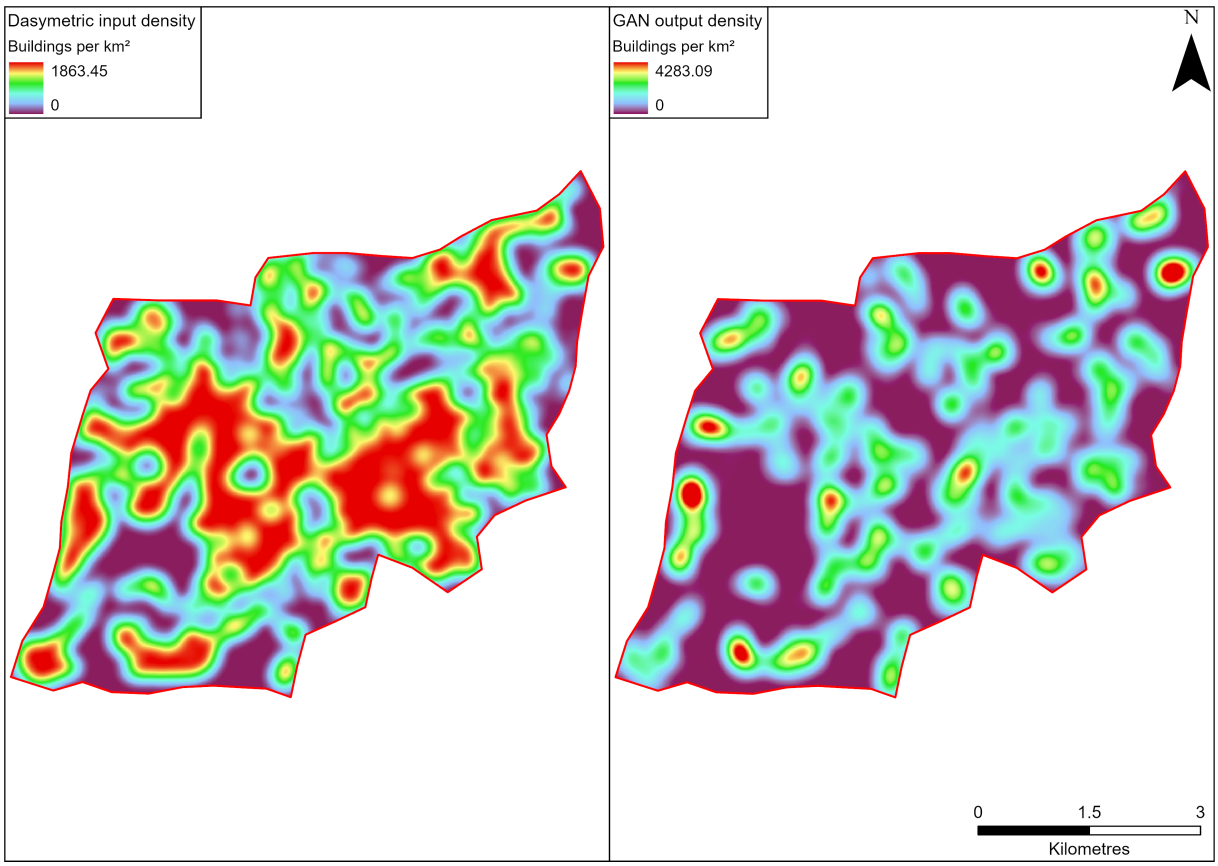

Figure 6.18: BAU density comparison between dasymetric input density raster (left) and the GAN output density raster (right).

Figure 6.19 shows the SP scenario results, which also support the inference of high density areas in the GAN output matching with the dasymetric input as seen in the BAU scenario (Figure 6.18). It also reveals that the GAN building outputs are generated strictly within the dasymetric density extent, as evidenced by the absence of spurious development outside areas defined in the dasymetric input, such as the lack of buildings in the south-west corner of the study area (Figure 6.19). However, GAN building development density has much less presence compared to the dasymetric input. Like the BAU density rasters in Figure 6.18, the SP rasters are not visually similar and it is easy to discern differences between their spatial patterns (Figure 6.19).

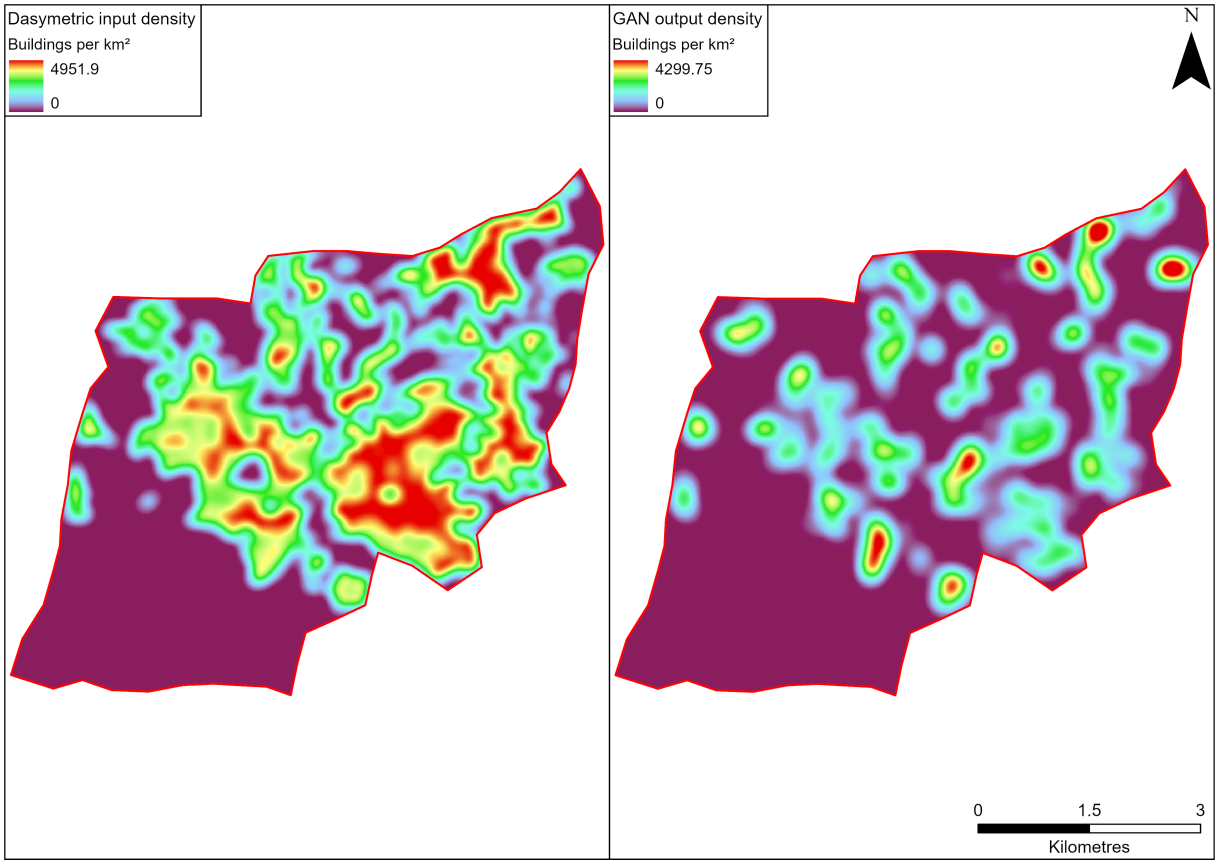

Figure 6.19: SP density comparison between dasymetric input density raster (left) and the GAN output density raster (right).

Overall, it can be argued that the GAN results do not effectively reflect the dasymetric development density that was used as input into the DGM. Two statistical tests proved that, across all scenarios, the mean and distribution of raster pixel values were statistically different. Density map visualisation of two scenarios also show that the spatial pattern similarities are not convincing. While the GAN density maps underestimate development extent, they exhibit implicit and relative location, distribution, and intensity similarities to the dasymetric input, with correctly located hotspots and alignment of high-density areas. This suggests that although the relative spatial patterns are consistent, the absolute values differ.

6.9 Summary

For road network generation, the 2048x2048 global road network experiments, with the exclusion of the 100% type 2 road removal experiment, were identified as the best performing GANs using the MAE from the vector metrics. As a result, the 25%, 50% and 75% type 2 road removal experiments were applied to the BAU density raster. The 75% experiment produced the highest number of roads, blocks, node degree, and was used for the generation of future road networks in all scenarios.

The results of the building footprint GAN experiments provided a clearer indication of performance, with the MAE from the local 1024x1024 group largely outperforming the global group. Therefore, both the 500 m^2 and 1000 m^2 were applied to the BAU road network. However, the GANs produced very similar quantities of building footprints, with the 500 m^2 experiment producing slightly more. Upon visual inspection, the 500 m^2 and 1000 m^2 outputs were better in different areas. Therefore, both were used together, with priority given to the 1000 m^2 to ensure a sufficient quantity of building footprints were generated along the 2040 road networks.

Finally, dasymetric mapping was an appropriate technique to generate the future urban form. The GANs were responsive to the different density areas of the scenarios produced and produced varied locations of urban from depending on the scenario. The actual urban form layout did not seem to be affected when different densities values were used. This chapter has established and proven that using pix2pixHD can feasibly generate realistic future urban form in the study area.

Chapter 7

DISCUSSION

7.1 Introduction

This chapter reviews and critiques the geospatial modelling framework presented in Chapters 5 and 6. The four stages of the framework are discussed: the ABM, evaluation protocol, dasymetric mapping, and GAN. It compares these stages with similar work performed in the academic literature to justify approaches and identify shortcomings. It benchmarks the performance of ABM and GAN against what other modelling studies achieved. Multidisciplinary applications of the modelling framework outputs are detailed. Overall, it places the modelling framework within the wider academic context and evaluates its contributions.

7.2 Review of the agent-based model

7.2.1 Scenario interpretations and insights

The five scenarios simulated reveal useful insights into the ABM. The BAU scenario simulates 18.04 km^2 of urban growth, but according to the inputted five-year land demand value of 4.17 km^2 this should actually be 14.18 km^2 over the 17-year period of future simulation. This overestimation is a product of the longer model runtime in the future simulation when compared to the evaluation period of nine years and the 240 model runs to capture the stochasticity of the model. As the urban planner performs a random weighted selection, this results in parcels being hit more frequently for development over the 17-year period and 240 model runs, meaning that more parcels than anticipated exceed the 30% probability threshold. Increasing the threshold to a higher value, such as 40%, changes the simulated value to 16.57 km^2, a value closer to the anticipated land change. Extending the temporal span of calibration and validation would prevent overestimation, as this study only used a nine-year period for both, while the future simulation covered nearly twice that duration. Calibrating and validating over a 17-year period would better tune the model variables to simulate urban growth to 2040.

The SP scenario (Figure 5.15) shows that the ABM responds to the weights of the urban planner's driving factors, since urban growth is focused in the north-east corner of the study area, with no developments occurring in the opposite south-west corner. Even though the planner ultimately does a random selection which impacts results, this weighting has a significant influence over 240 model runs and captures the stochasticity stabilising and development being pulled towards the centre of Phonm Penh. Figure 7.1 shows how the distance from Phnom Penh increases the further from the north-east corner of the study area.

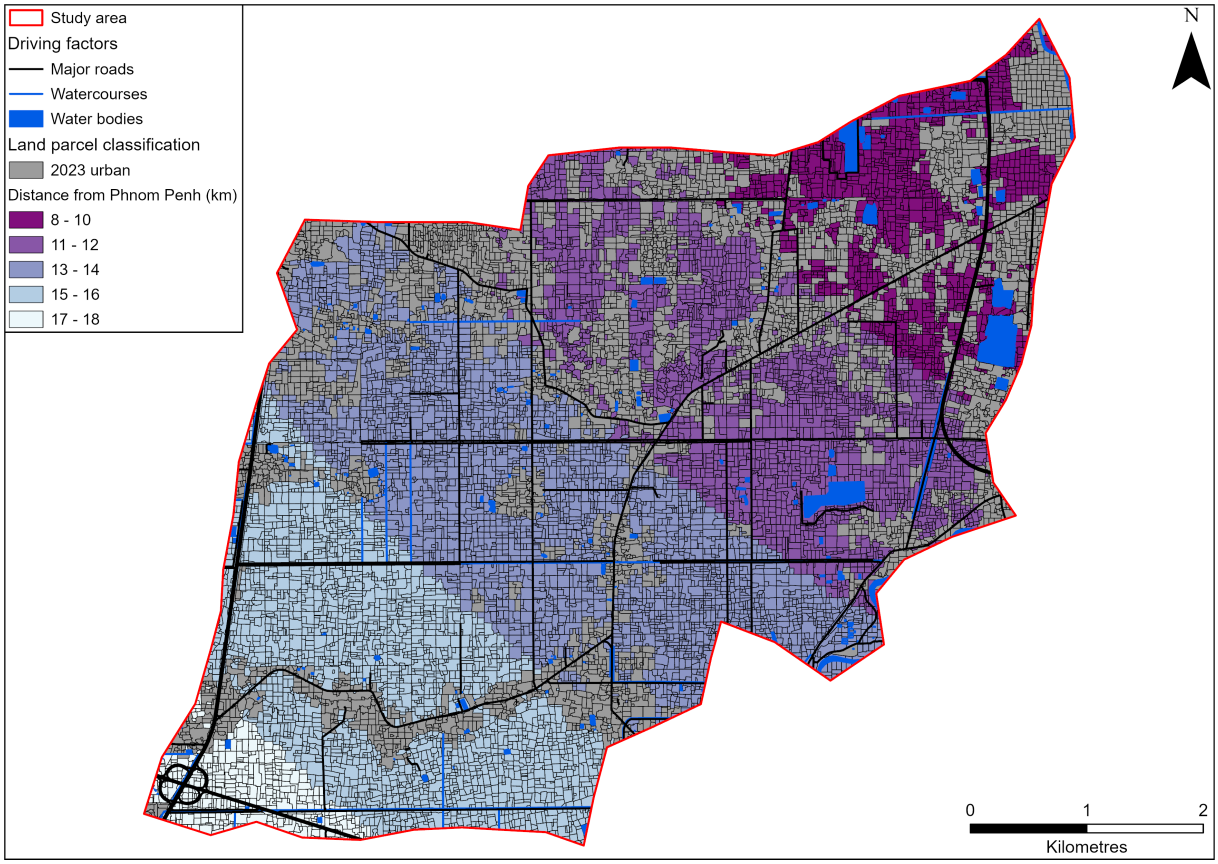

Figure 7.1: Land parcels distance from the centre of Phnom Penh.

The willingness to sell model variable clearly has a strong influence on the progress of urban growth, with higher values impeding the development from occurring. This is reflected in the EB scenario where the very high five-year land demand of 6.74 km^2 should result in the whole area turning to urban. However, this is not the case as willingness to sell is compared equally to the random pool, making an owner much less likely to sell. In the EB and LB scenarios, developers have full autonomy to select the high-utility parcels, so they keep returning to the best parcels in each cycle until they get a favourable sale from the owner. A higher willingness to sell value (set at 1.0 in EB) slows the rate at which a developer can acquire a parcel which aligns with the findings of Koch et al. (2019) who also determined that lower values result in a higher number of new urban cells. The LB scenario also exemplifies the developer's preferences in the

eastern and central regions of the study area (Figure 5.16), which is not found in the BAU, ER, or SP scenarios where the urban planner defines the development locations and the developer has to select parcels in these zoning areas.

The urban planner's master plan selection adds a lot of stochasticity and regulation to the model. This is demonstrated through the contrasts between the BAU scenario (Figure 5.13) and the EB scenario (Figure 5.17), the EB scenario being dominated by 89-100% probability parcels taking up 88.4% parcels simulated, while the BAU scenario having a much more uniform spread of parcel counts and areas in each probability class (Table 5.7). This also reveals that the random weighted selection is a stochastically stronger process in the model than the landowner's random decision when deciding to sell their parcel.

7.2.2 Assumptions

It is important to identify and consider the key assumptions of the ABM to provide complete transparency and guide future research. Firstly, the ABM only simulates non-urban or urban parcel change in the study area without considering simulating different types of land use, such as industrial, commercial, or green spaces. The key assumption is that all growth in the study area was residential development. This is in line with what was historically observed in the study area between 2014-2023 and other CA and ABM urban growth studies also adopt this approach of simulating binary classifications of new urban growth (e.g. Agyemang et al., 2023; Mustafa et al., 2017).

Another assumption is using theories, such as bounded rationality and bid-rent theory, to justify agent behaviour and decision-making. This is due to the absence of empirical surveys and detailed census data which resulted in simplifications when designing urban planner, developer, and landowner agents. Using surveys would have provided stronger support for agent behaviour. In addition, the inclusion of resident agents alongside the other agents used may have been beneficial in adding more model realism in terms of capturing demand-side dynamics of the urban development process. However, due to limited OSM data, information on residential preferences, such as amenities, was not possible to obtain and so they were omitted.

The urban planner agent has complete control on urban development locations. This is the law in Cambodia, but since the implementation of these spatial planning laws was disputed by the World Bank (2024), The LB and EB scenarios removed the planner to see the impact on the ABM. Alternatively, the urban planner could have a development control parameter, such as in Agyemang et al. (2023), which explores weak to strong development control to examine how this impacts urban development in the model.

There is a common issue of data latency in urban growth ABMs, especially when working at local spatial scales and in rapidly developing regions. Data latency refers to the delay between when real-world changes occur and when they are reflected in the dataset (Xu et al., 2020). For example, the ABM performs historical simulations from 2014-2023, but the major OSM roads are data from 2023. Major roads were used to try and avoid latency issues as minor roads were frequently added to the study area during the historical period, which could skew ABM outputs. However, Ring Road 3, a 53 km highway, was constructed towards the end of the historical period (Figure 7.2). This is one of the challenges in working in rapidly developing regions, as Cambodia is also investing heavily in major infrastructure projects to improve transportation corridors within the country. Although older 2014 OSM data for the study area were considered to obtain the original major road network, this was not possible due to discrepancies and incomplete coverage. This is why the urban footprint was captured in 2014, 2019, and 2023 to minimise data latency impacts on the ABM.

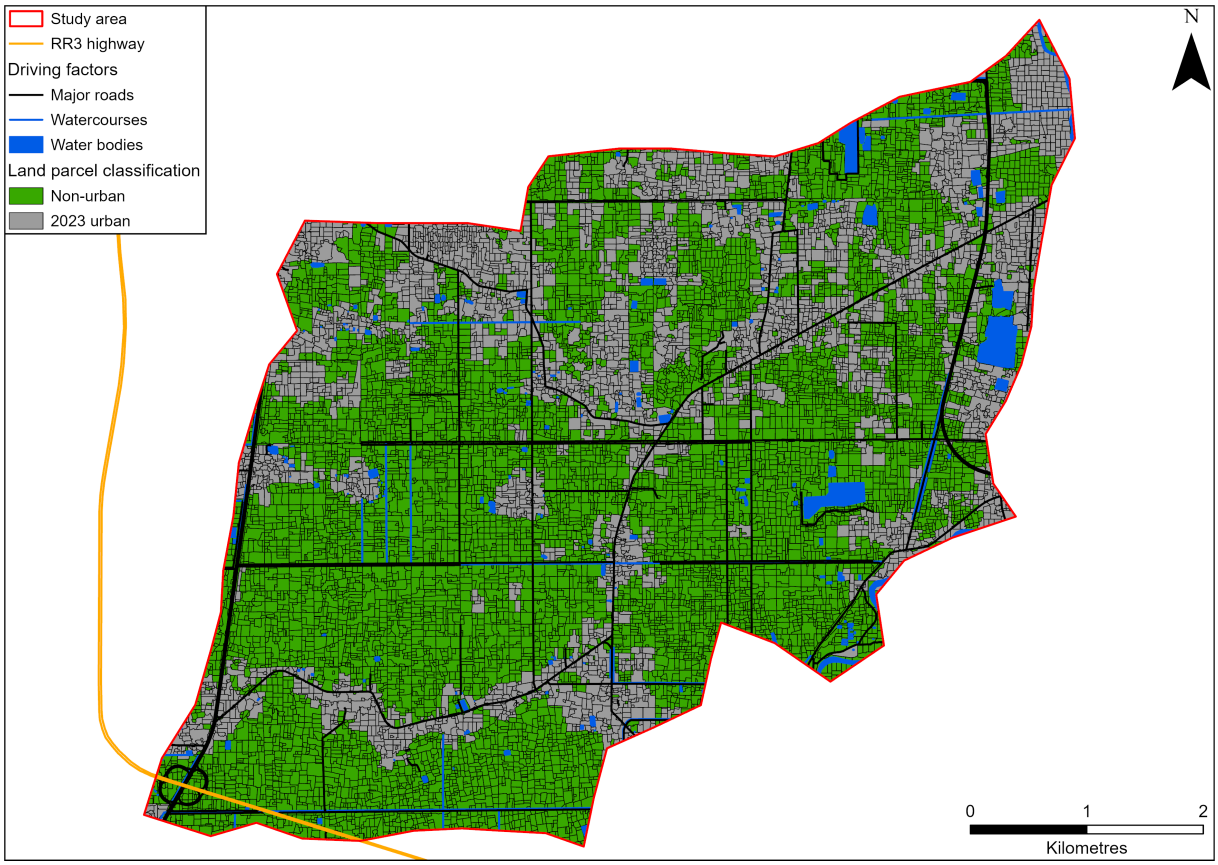

Figure 7.2: Construction of Ring Road 3 highway near the study area took place between 2019-2023.

The developer and landowner driving factor weights are equal in all the scenarios. The urban planner driving factor weights were treated as variables and optimised in calibration. This is because the developer and the landowner do not play a direct role in influencing land use policy. Their decisions are primarily driven by economic considerations, such as land value and demand, rather than regulatory law. Since they operate

within the constraints set by the urban planner, assigning equal influence to their driving factors ensures that the model reflects their market-driven behaviour, rather than implying a disproportionate role in shaping urban growth policy. In addition, it would have been computationally demanding to introduce another 11 driving factor weight variables for exploration in the evaluation process. In addition, the ABM model variables explored in Table 5.2, such as willingness to sell, five-year land demand, and distance from major roads, are kept at a constant value throughout the simulation.

7.2.3 Model computational runtime

It took 19 seconds to run a single model when running 10 in parallel (3.15 minutes every 10 runs) during the evaluation period of nine years. The simulations were run on an Alienware machine with an Intel i9 processor and 64GB of RAM. The 1800 runs for stabilisation analysis took approximately 9 hours and 27 minutes. Identifying 240 runs significantly reduced the modelling time to 1 hour and 15 minutes. This indicates that the stabilisation analysis was very helpful in ensuring efficiency. Kang et al. (2022) set the number of Monte Carlo runs to 250, instead of the 100 used in this thesis. If this value had been used, stabilisation would have been exponentially longer at 23 hours, 37 minutes under 4500 runs.

The GA ran through 194 variable configurations which equalled 46,560 runs which took approximately 5 days and 2 hours, as calibration was half of the evaluation period. This is why the population size parameter of GA was set to the lowest possible value for robust results at 20. Lin et al. (2023) set their population to 100, but under the runtimes provided, doing this with the ABM in this thesis could theoretically have increased the duration of calibration fivefold. A runtime of over 25 days would have been impractical. This is why the parameter values for GA and the number of Monte Carlo runs were set to lower values. To help address this challenge, AWS with high-performance CPU and RAM could be used for more efficient model runtimes and higher parameter values for enhanced robustness. This comes with an added monetary cost and the values selected are within appropriate ranges based on previous optimisation experiments by Lin et al. (2023) and Tang and Jia (2014). When selecting parameter values for GA or Monte Carlo, modellers should take into account the runtime of simulations and what is practical within time constraints.

7.3 Review of the stochastic urban growth model evaluation protocol

The SOLV protocol, described in Section 5.3, uses the FKS to evaluate the urban growth ABM developed in this research. Firstly, it performed a stabilisation analysis of the

ABM using the CV to determine when its stochastic processes have converged. It then carried out calibration and validation to identify optimal variable values that improve the ABM's performance. Finally, it applied Morris and Sobol sensitivity analyses to understand how the ABM output responds to changes in model variables, identifying variable importance, direction of influence, and potential interactions.

7.3.1 Stabilisation analysis

Although Lorscheid et al. (2012) and Lee et al. (2015) suggested that CV differences become negligible when consecutive runs fall below a fixed threshold of 0.01, the knee detection algorithm offers a more flexible and generalisable approach to identify stabilisation points without relying on this assumption. Since CV values can vary depending on the output of interest, a fixed threshold may not always be suitable. The knee algorithm accommodates outputs with different value ranges, such as FKS values between -1 and 1 or simulated area quantities greater than zero, making it more adaptable across different contexts. This flexibility removes the constraint of a predefined threshold and enhances the robustness of the approach for diverse outputs. Moreover, the algorithm allows for customisation based on the curve's shape (concave or convex) and direction (increasing or decreasing), providing more nuanced control to identify stabilisation points. By standardising this detection process, it effectively balances the ABMs stochastic outputs with computational efficiency. In this research, it avoided the need for all 1800 stabilisation runs by identifying that, after 240 runs, the CV difference had dropped significantly, beyond which additional runs would have exponentially increased runtime with diminishing returns in stabilisation.

7.3.2 Calibration and validation

7.3.2.1 Threshold

Setting the probability threshold of 30% (parcels converted to urban at least 72 times) to determine which parcels were simulated as new urban by 2040 across the 240 runs was necessary to reduce stochastic noise in the model. This is supported by Figure 7.3 which shows a large overestimation of simulated parcels at 18.32 km^2 when omitting the threshold compared to the observed change at 4.17 km^2 between 2019 to 2023.

During validation, the model converted 8,032 parcels to urban, of which 1,526 correctly matched the 1,688 observed urban parcels. These 1,526 parcels hit the observed parcels 90,258 times, whereas all 8,032 parcels were hit 315,256 times in total. Only 28.6% of all simulated urban instances actually hit observed urban, indicating a significant level of noise. When applying a probability threshold of 30%, the number of simulated parcels dropped to 1,227, of which 536 parcels correctly matched the 1,688

observed parcels a total of 52,068 times. Among the 1,227 parcels, there were 114,447 total hits. This increases the proportion of observed hits to 45.5%, demonstrating a substantial improvement in model performance by filtering out low-probability simulated parcels. This comparison highlights that the 30% threshold significantly improved the alignment between the simulated and observed urban areas. Furthermore, the number of false positives (simulated new urban but observed non-urban) was significantly reduced from 6506 parcels to 691 (2.04 km^2) using the 30% threshold.

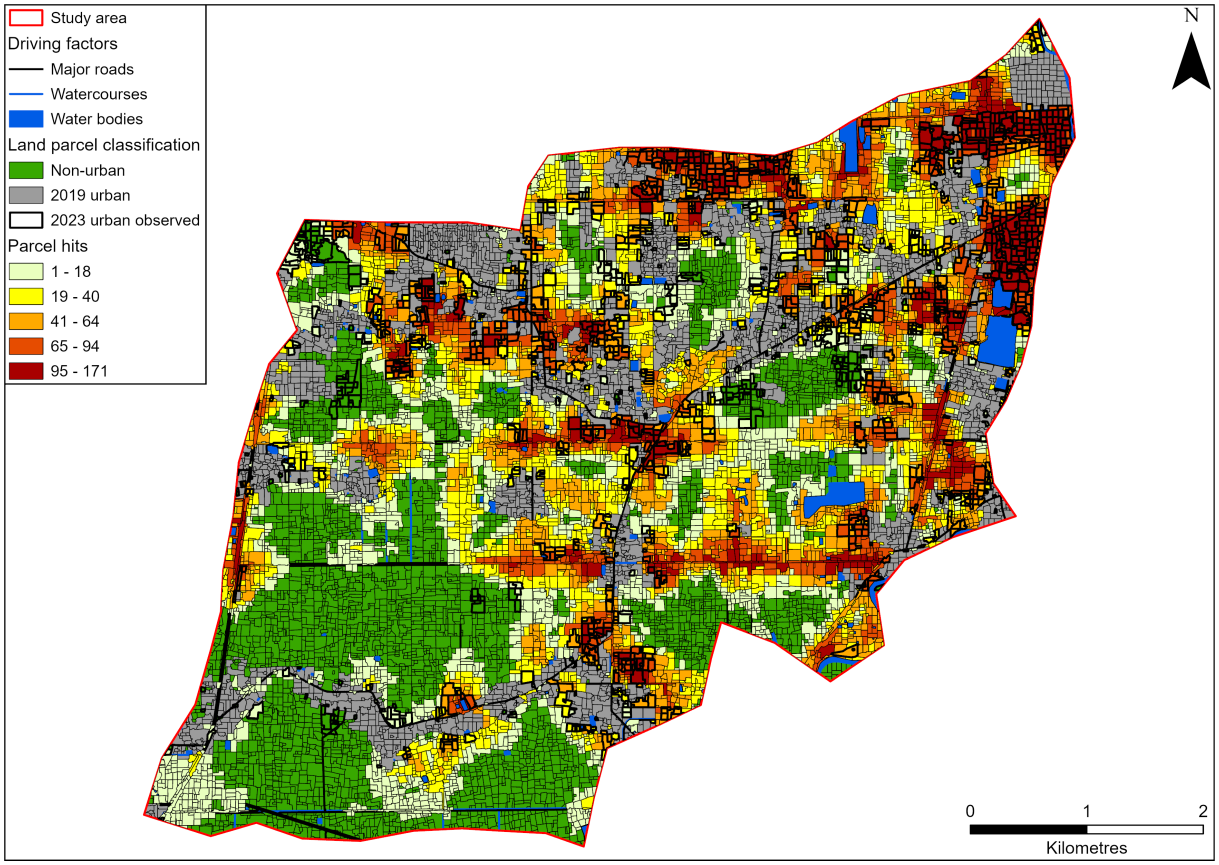

Figure 7.3: Simulated parcels in validation with using threshold, highlighting all the parcels that were simulated as urban at least once in the 240 runs.

However, this raises concerns about using thresholds on cell-by-cell agreement metrics, such as PCM, which can artificially inflate the score. For example, applying a 10% threshold yields a PCM of 70.3%, which increases to 88.3% under a 50% threshold during validation. While this might suggest improved performance, the increase is largely due to the exclusion of many simulated parcels. At the 10% threshold, there are 6,349 true negatives (simulated and observed persistence of non-urban), compared to 10,085 at the 50% threshold. The increase in PCM is primarily driven by the higher count of true negatives, thus biasing the metric in favour of higher thresholds and creating a misleading impression of better performance. Using a 50% threshold also drastically reduces the simulated urban area to just 0.46 km^2, compared to 4.17 km^2 in the validation data, an underestimation of nearly 90%. In contrast, the FKS can handle true negative more accurately as it takes into account the initial land use map. It achieves a value of 0.453

at the 10% threshold, increasing to the 30% threshold, before declining to 0.188 at 50% (Table 5.5) in validation. This highlights the importance of using threshold-based approaches alongside techniques like the area weighted FKS (Eq. 5.10) that incorporate the initial land use map and simulated urban area, as done in this research, to avoid the limitations of cell-by-cell metrics alone.

7.3.2.2 Performance

The ABM achieved an area-weighted FKS score of 0.241 (0.340 non-weighted) in calibration and 0.509 (0.550 non-weighted) in validation (Table 5.5). Both of these values present strong agreement with observed urban change when compared with existing literature. Guzman et al. (2020) achieved an FKS score of 0.274 and van Vliet et al. (2013) obtained an FKS score of 0.005, 0.348, and 0.400 for three CA models. Du et al. (2018) achieved FKS values between 0.338-0.435 when using a fuzzy distance value of 3x3 cells, but when using a fuzzy distance of 11x11 cells, the score increased to between 0.345-0.501 cells as this allowed an even more lenient agreement. Considering the non-weighted values in the Phnom Penh ABM fall within and even exceed some of the highest scores in literature, this shows the model is suitable when compared to other studies. This is especially clear compared to Du et al. (2018) who achieved a maximum FKS of 0.501 when using a much larger neighbourhood area than the ABM's 200 m fuzzy distance (approximately a 4x4 parcel neighbourhood).

Typically, urban growth modelling studies achieve kappa coefficient values greater than 0.8 (e.g. Mozaffaree Pour and Oja, 2021; Du et al., 2018; Liu et al., 2020). This raises the question of why the FKS scores are significantly lower than the traditional kappa scores. Du et al. (2018) attributes this to the fact that the FKS coefficient excludes urban persistence by incorporating the initial urban footprint in the calculation. This results in FKS being much stricter and more robust than the original metric. However, this research has identified a limitation of the FKS metric - its tendency to overestimate the extent of change in order to improve the agreement score, especially when the actual observed change in the study area is minimal.

7.3.2.3 Evaluation period

Splitting calibration and validation into two equal temporal periods should be avoided. The evaluation period division should ensure that an equal amount of urban change has occurred in calibration and validation. This will avoid underperformance at either stage and help fit the model parameters more naturally to ensure robust future simulation. This is important for rapidly developing cities, such as Phnom Penh, which are experiencing sporadic growth. It is challenging for urban growth models to capture these growth rates because there could be large variations in urban change historically.

As noted in Section 7.2.2, one of the assumptions of the ABM is that the model variables listed in Table 5.3 remain constant throughout the simulation. However, when applying the SOLV protocol, it may be worthwhile to explore whether dynamically adjusting these variables at each simulation cycle could improve model performance, particularly given the sporadic nature of urban growth in Phnom Penh. This was not pursued in the current study, primarily due to the substantial data requirements and investigation that dynamic calibration would entail for each year of the evaluation period. Moreover, it raises the question of how such dynamic adjustments would be justified in future simulations. One possible approach could be to assume a cyclical pattern in the historical data, derive annual values for the variables, and then reapply these in projecting future scenarios.

7.3.2.4 Error propagation

As identified, the validation results are stronger than the calibration results. Although the area-weighted FKS value for calibration (0.241) remains within the range reported by other modelling studies, the visual outputs in Figure 5.9 are unconvincing compared to the observed change. The visually poor calibration results can be attributed to the following factors:

- The quantity of observed urban change during the calibration period was relatively low (2.57 km^2), providing limited information for the GA to converge on an optimal parameter set.
- FKS scores are sensitive to overestimation when observed change is sparse.
- The GA method implemented (Section 5.3.3) had a fixed number of generations, without adaptive stopping criteria based on convergence or fitness improvements. This may have constrained parameter optimisation and limited the ABM's ability to fit calibration period dynamics effectively (Lin et al., 2023).
- The five-year land demand parameter (Table 5.2) was a pre-defined value based on observed growth in calibration. Exploring this a range of values for this parameter alongside the eight other variables explored in calibration (Table 5.3) could have helped reduce the overestimation of simulated urban in calibration.

These calibration limitations reduce confidence in the model's ability to learn from historical data. The ABM behaviours may not fully capture the drivers of early development, though this could also reflect real-world inconsistencies, such as the weak enforcement of zoning regulations in Phnom Penh during the calibration period. This is visible in Figures 5.9 and 5.10, where observed development in calibration is more dispersed and fragmented than the observed patterns in validation. The superior area-weighted FKS value in validation (0.509) suggests it generalises better than it learns. This indicates

that the model's behavioural rules align more closely with the spatial dynamics of the validation period, when urban expansion became more structured and policy-driven.

This raises the following implications for the modelling framework. The ABM is fundamental in providing spatial inputs for the subsequent GAN stages. Errors in parcel conversion during calibration can propagate forward, misallocating road network layouts and building distributions. While stochastic stabilisation in SOLV helps reduce random variation, systemic misallocation could introduce bias in the ABM-GAN workflow. For example, the GAN may generate urban form in misclassified growth areas due to inflated future urbanisation probabilities. As a result, there is likely a minor urban form misallocation in the final outputs as even the validation outputs (Figure 5.10) only capture the larger-scale observed development projects, showing a discrepancy with what occurred historically.

This error propagation could be viewed as reducing confidence in the modelling framework. However, it is helpful to consider the reverse case: if calibration had been strong but validation poor, it would suggest overfitting to training data and poor generalisability. This would severely undermine confidence in the ABM's conceptual foundations, the relevance of the chosen driving factors, and the reliability of scenario outputs. This would also exacerbate error propagation in the subsequent stages to a much greater degree. Future urban form from the GAN could appear visually realistic while being spatially incorrect, leading to serious misinterpretation.

While underperformance in calibration is an area of uncertainty, the model's good performance in validation represents the lesser of two methodological risks. Nonetheless, it limits the model's credibility as a predictive tool, as both good calibration and validation performance would have been the best outcome. By acknowledging the implications of these methodological issues, the ABM-GAN workflow could be interpreted as a scenario-based visualisation framework rather than a precise predictive tool. Its strength lies not in generating exact predictions, but in exploring a range of plausible future projections based on clearly defined assumptions. The final urban form results should be treated as potential and plausible change that can inform policy and planning by highlighting development hotspots and patterns.

7.3.3 Sensitivity analysis

Performing Morris and Sobol sensitivity tests on the ABM has proven its credibility and demonstrated that it features non-linear effects and there is significant interaction between model variables. Specifically, it has quantified the importance of variables and associated confidence levels, the variable directions of influence, and the interaction of variables. Morris and Sobol are conventional sensitivity techniques that have led to

these enhanced insights and trust in the ABM.

The SA approach used in this study could be further refined by incorporating spatiotemporal SA into the evaluation process. For instance, sensitivity indices could be computed for each spatial unit within the ABM or at each simulation time step (Ligmann-Zielinska, 2013). Ligmann-Zielinska et al. (2024) developed bivariate sensitivity maps that applied both Morris and Sobol methods to individual spatial units (Figure 7.4).

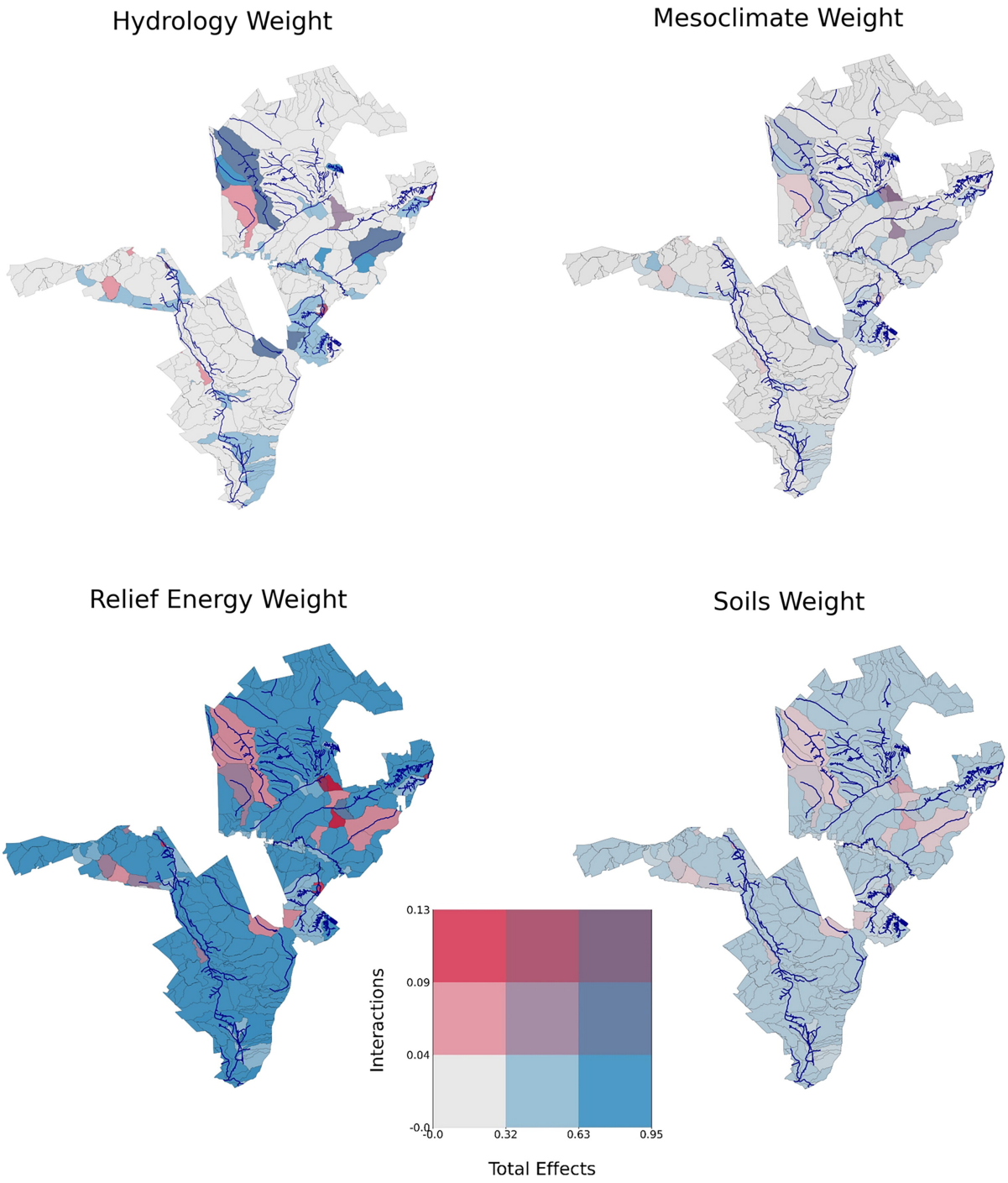

Figure 7.4: Sobol bivariate sensitivity maps. The x axis represents the overall effect of weight factors to the total variance and the y axis represents interactions between weight factors (Ligmann-Zielinska et al., 2024).

By visualising SA results, this approach goes beyond simply identifying key input variables. It also reveals their influence geographically, leading to deeper and more actionable insights. This highlights potential correlations, patterns, or trends between variables that may not be evident when analysed separately or statically. Consequently, it improves the understanding of the dynamic relationships that shape the sensitivity within the study area. Additionally, it helps pinpoint regions with differing levels of sensitivity, an aspect that would be overlooked if only aggregated and non-spatial metrics were considered.

SA are also only typically computed at the end of the model simulation, which assumes that input influence is stable or changing at consistent rates throughout model execution. However, when simulating emergent behaviours in an ABM, this is not usually the case. The final state of the model results does not provide information on the relative and varying importance of variables throughout the execution of the model (Ligmann-Zielinska, 2013). Temporal SA (Figure 7.5) allows for greater understanding of how variables affect the dynamics of the model throughout and at the end of the simulation (Ligmann-Zielinska et al., 2020). It essentially has the same benefits as spatial SA previously described but in the temporal dimension of the model.

Only static Moris and Sobol SA tests are available within the GAMA software used in this research, however, and obtaining observed urban change for every year across the entire study area would have involved retraining the computationally intensive deep learning models required. Spatiotemporal SA were not within the scope of this research and requires further investigation to determine its feasibility within the SOLV protocol.

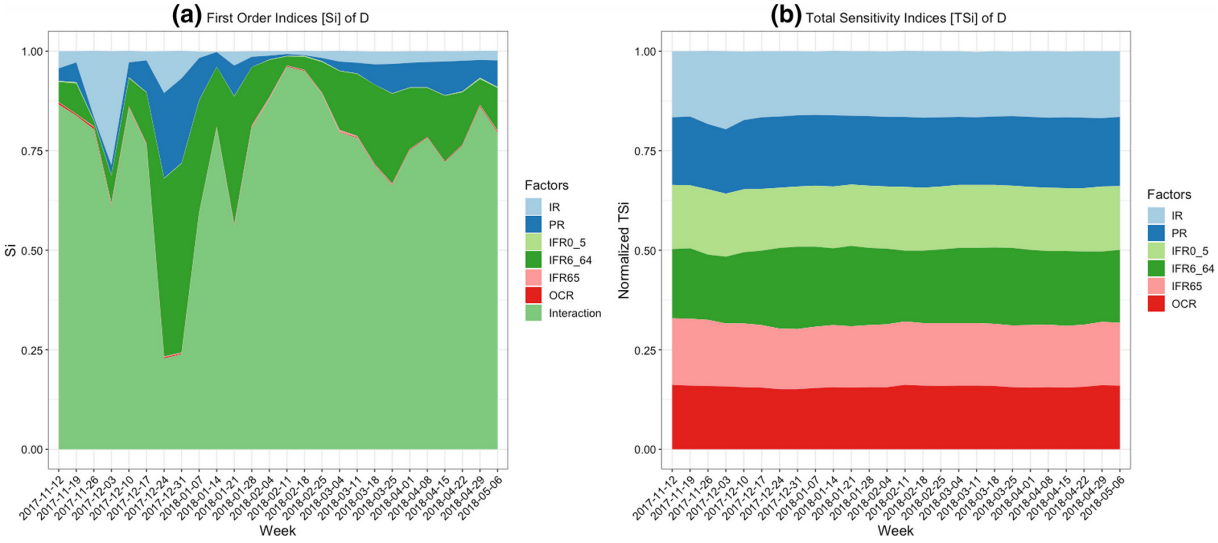

Figure 7.5: GSA applied temporally to ABM of influenza transmission (Kang et al., 2022).

7.3.4 Protocol assessment and comparison

The SOLV protocol encourages selecting a single output of interest, which is standard in existing literature (Ligmann-Zielinska et al., 2020; Lee et al., 2015). Questions arise if multiple outputs of interest should be utilised to gain a more comprehensive and thorough understanding of ABM sensitivities. However, Radchuk et al. (2016) argued that it is more practical to focus on one well-chosen output of interest rather than trying to interpret multiple outputs of interest simultaneously. Multiple outputs of interest can lead to additional challenges to difference in each parameter's impact on the outputs. Not all output behaviours are straightforward, as some may exhibit complex and non-linear behaviours due to interactions among multiple inputs (Radchuk et al., 2016). Essentially, the link between the parameters and the output could be lost or unclear, making it difficult to analyse and interpret them effectively. It is therefore more straightforward and efficient to focus on a single output of interest but report other important metrics, such as the area of new urban land. It would also make it easier for stakeholders to understand the model performance.

The GSA-CAL framework proposed by Kang et al. (2022) places greater emphasis on calibration than on stochastic stabilisation. For each unique configuration of model inputs, 10 simulation runs are conducted to account for stochastic variability, offering broad coverage of the model's randomness. However, the SOLV protocol presents a more robust method for evaluating the effects of stochasticity on model outputs. GSA-CAL may provide a more efficient calibration strategy by performing sensitivity analysis first and then excluding irrelevant parameters from calibration. This efficiency could be considered rudimentary because the selection of key parameters involved drawing an arbitrary line to define importance. Stabilisation in SOLV is conducted across all variables which could influence the identified stabilisation point and, consequently, the reliability of the results. Nonetheless, the efficiency of GSA-CAL may be advantageous, particularly given the runtime demands of the ABM used in this thesis.

Ligmann-Zielinska et al. (2020) outlines three distinct roadmaps for conducting SA on the ABM's purpose. Roadmap 1 is suited for conceptual ABMs and focuses on identifying the main effects of individual inputs, detecting potential interactions, and analysing the nature and strength of these interactions. In contrast, the ABM developed in this thesis aligns with Roadmap 2, which is intended for empirical or application-driven ABMs. The goal of this roadmap is to build trust in the model for scenario analysis and real-world application. It assumes that interactions between inputs are likely and therefore recommends a GSA approach that jointly assesses the individual and combined contributions of stochastic inputs to output variability. This justified the use of both the Morris method and Sobol indices in tandem, as they offer complementary insights into model sensitivity. Roadmap 3, which focuses on complexity-oriented SA, goes beyond

the practical application focus of this research. It seeks to understand the internal dynamics of the model by evaluating spatiotemporal outputs, analysing the topology of input–output relationships, and assessing the sensitivity of non-parametric elements. While this approach offers a deeper and more nuanced understanding of the model's internal processes, it was beyond the scope of this thesis, which prioritised credibility, predictive performance and practical relevance in line with Roadmap 2. Applying Roadmap 3 to a real-world ABM remains a potential avenue for further research.

7.4 Review of dasymetric mapping

The dasymetric mapping to convert ABM outputs to future development density demand has proven to be an effective technique to generate future urban form. The varied density rasters all generate feasible road networks which are reactive to the density locations. However, their impact on the generating different urban form types appears limited. Having higher density vs lower density of future development density appears to have minimal impact on the road layout density generated (Figure 7.6). This suggests that to generate more varied urban form layouts, specific layouts of urban form need to be trained in the GAN.

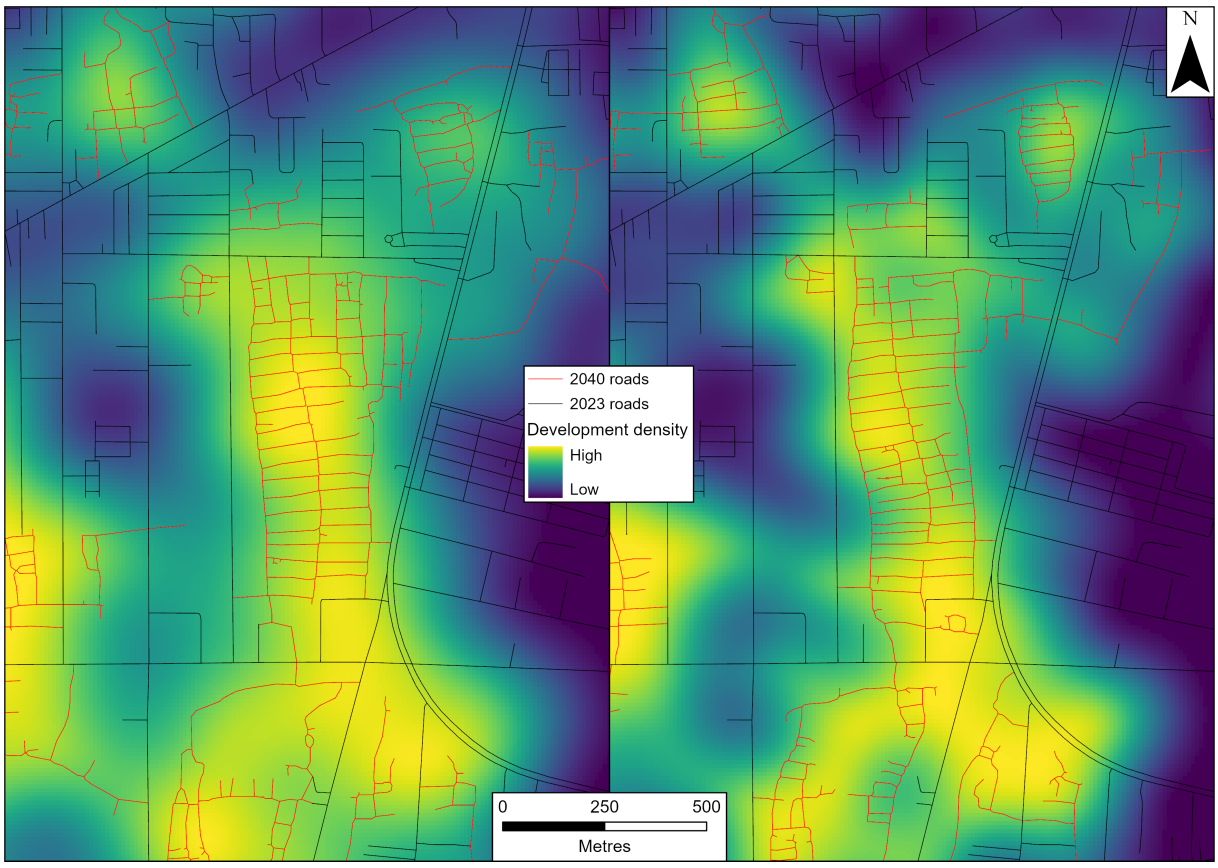

Figure 7.6: BAU density output, maximum of 1,830 buildings per km^2 (left) IP density output (right) with maximum of 4,910 buildings per km^2.

A key assumption is that future urban form follows existing patterns, as there is limited practice or initiative in Southeast Asia to promote new or more sustainable urban form. Moreover, given the large spatial extent of the study area, it was not feasible to capture a specific road patterns (e.g. gridded or organic) in a single fishnet cell as there was no information on OSM road year in Southeast Asia and even if there was, the training images would have been too few to train a GAN because of the large spatial scale employed. As a result, each generated image instead adapts to the surrounding road network pattern present in the input image (Figure 6.8). Another challenge is the mismatch between the quantity of building developments produced by the GAN and those projected using the dasymetric approach. This disparity was unavoidable when linking the ABM outputs to the GAN inputs, due to the data-driven nature of GANs. Unlike rule-based models, GANs cannot be explicitly instructed on how many buildings to generate, instead relying solely on the data patterns they are trained on.

7.5 Review of generative adversarial networks

7.5.1 Generative adversarial network behaviour

The 2048x2048 road network GAN with 75% of type 2 roads removed and the 500 m^2 and 1000 m^2 1024x1024 local building footprint GANs were the most appropriate to use in the study area based on the results in Chapter 6. As an additional verification for these GANs, the generator and discriminator loss has been plotted to understand GAN training behaviour. As Figure 7.7 shows, the discriminator and generator losses overlap at the beginning of training, but gradually diverge from each other. The discriminator loss is generally lower than the generator loss. This typically indicates that the discriminator outperforms the generator by correctly distinguishing between real and fake samples. This shows balanced GAN training where oscillations become more controlled around specific levels, suggesting that neither the generator nor the discriminator is dominant and the GAN is stabilising (Goodfellow et al., 2014).

Unlike conventional deep learning models which are trained using an optimisation algorithm that seeks a low value of a cost function, such as DiceLoss or softmax confidence (e.g. Booth and Burke, 2023; Sirko et al., 2021). The GAN training process is inherently unstable, with fluctuations in both losses. The high variance in generator loss suggests that it is constantly adjusting to counter the discriminator (Goodfellow, 2016). GANs have non-convergence and oscillatory behaviours during training due to their adversarial nature. GANs are unlikely to achieve a stable equilibrium due to the dynamics of the generator and discriminator and the competition game between these two players (Goodfellow et al., 2014).

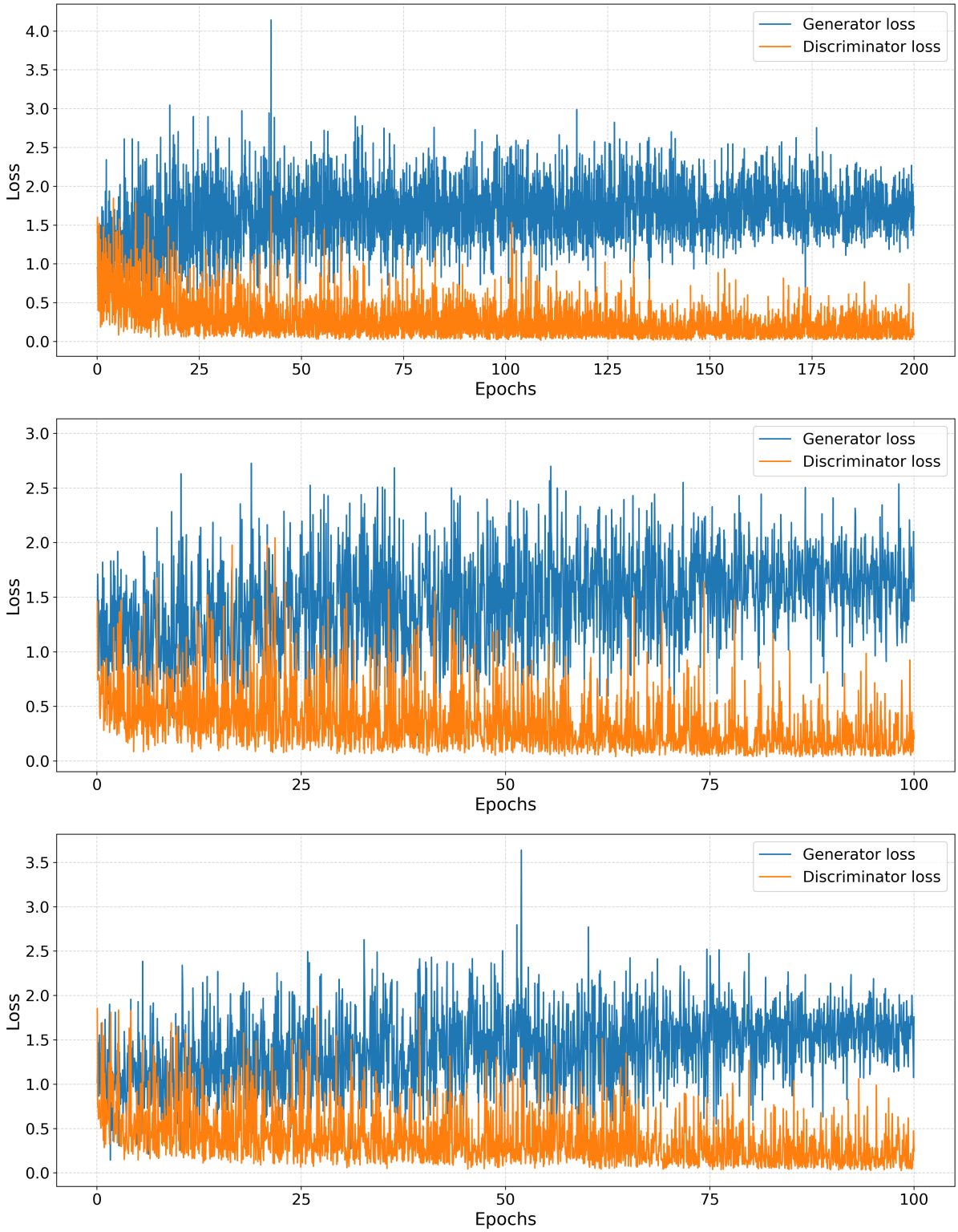

Figure 7.7: GAN loss logs. Top: 75%, middle: 500 m^2, bottom: 1000 m^2.

Figure 7.7 plots exhibit typical GAN behaviour and to support this two other GAN loss logs have been obtained. Figure 7.8 displays the GANmapper loss log experiment for Singapore and Figure 7.9 presents a pix2pixHD loss log from an experiment to predict the heat maps of street vendors from a pedestrian heat map to understand their relationship (Shou et al., 2021). Figures 7.8 and 7.9 display similar behaviour to Figure 7.7 with oscillations narrowing and diverging from each other during training.

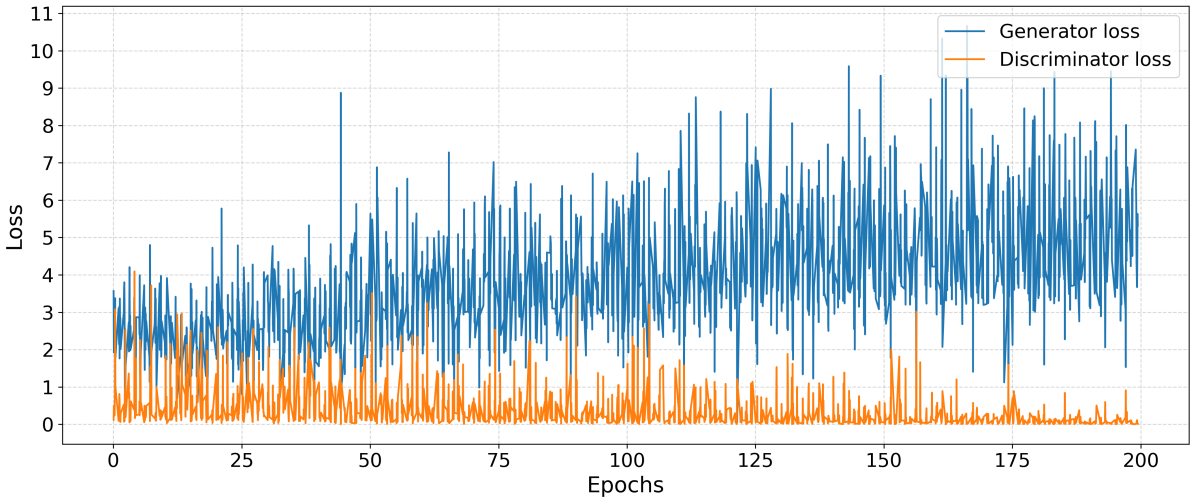

Figure 7.8: Loss log from GANmapper Singapore experiment (Wu and Biljecki, 2022).

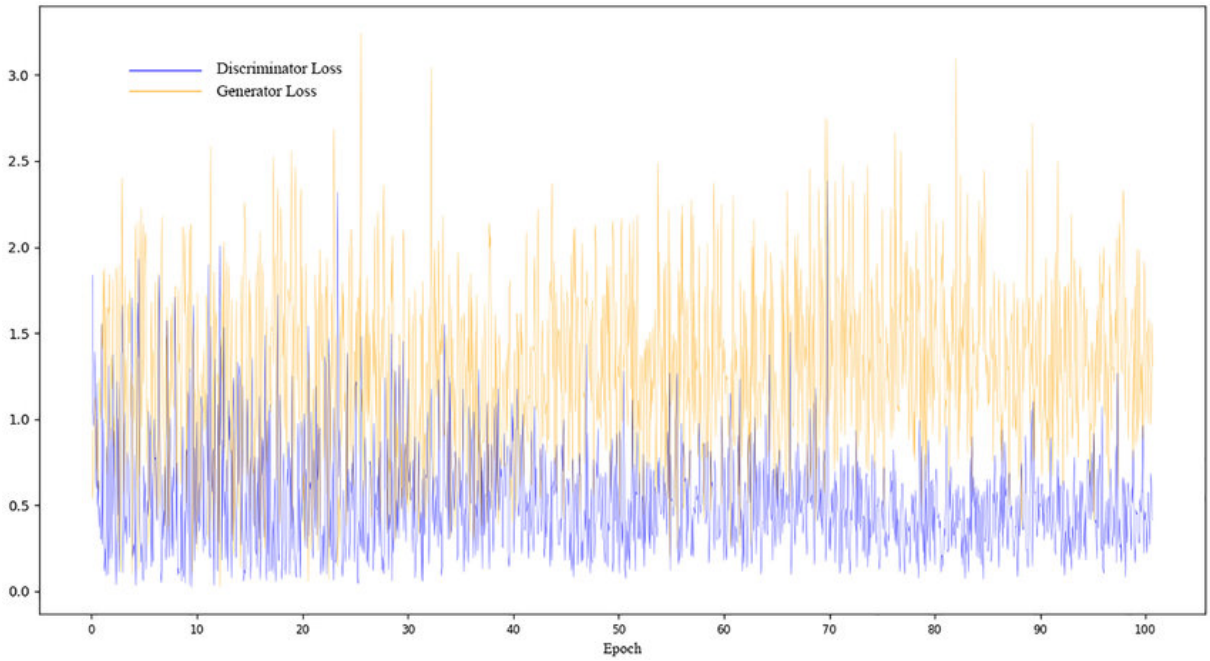

Figure 7.9: Pix2pixHD Loss log from Shou et al. (2021).

To visualise atypical GAN behaviour, Figure 7.10 displays the loss logs from the erroneous 100% experiments in the three groups of the GAN road network experiments that suffer from mode collapse.

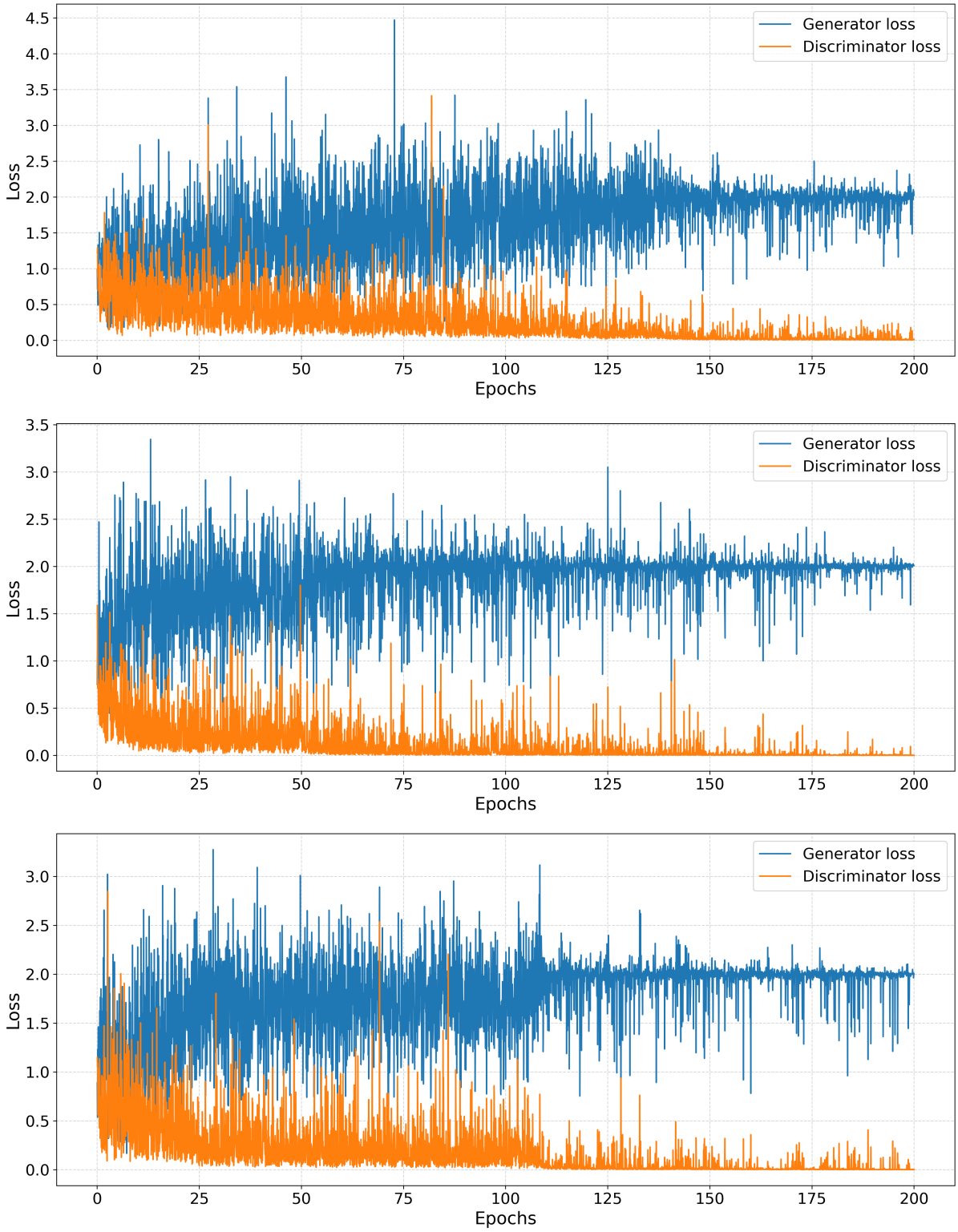

Figure 7.10: The GAN experiments which had 100% of type 2 roads removed. Top: 1024x1024 local, middle: 1024x1024 global, bottom: 2048x2048 global.

The loss patterns of the generator and discriminator are different from those in Figure 7.7. In all three loss plots in Figure 7.10, the discriminator loss appears to drop close to zero and remains very low during most training. This indicates that the discriminator is too powerful and is able to distinguish real from generated samples very easily. The generator loss oscillates within a high range, even as the discriminator loss stabilises

near zero. Typically, as Figure 7.7 shows, both losses should oscillate within a greater range over time, rather than having the generator loss stay consistently high with little variation, indicating irregular GAN behaviour. The limited variation in generator loss suggests that the generator is struggling to produce convincing samples, possibly because the discriminator is not providing meaningful feedback. All three plots in Figure 7.10 show normal GAN behaviour at the beginning of training but seem to encounter problems from the 50th epoch to the 125th epoch. This suggests that 200 epochs for GAN experiments was too excessive in training and led to GAN mode collapse.

7.5.2 Unsuccessful generative adversarial network experiments

Numerous unsuccessful GAN experiments were conducted for the generation of road networks and building footprints, as it involved trial and error to condition the GAN to successfully generate the respective features of the urban form in the study area.

7.5.2.1 Road network experiments

For road network generation, the first experiment attempted was to train a GAN on a 1024x1024 input image of major OSM roads with the density of minor roads and then a target image of the major and minor roads. The major roads and the minor ones were classified according to the fclass column in the OSM shapefile as follows (Birsak et al., 2022):

- Major - motorway, trunk, primary, secondary, motorway link, trunk link, primary link, secondary link.
- Minor - tertiary, tertiary link, residential, living street, road, unclassified.

This experiment performed normally in training and achieved good visual results in the test images (Figure 7.11). When the trained GAN model was applied to the study area, it generated future minor roads in the areas simulated to change to urban by 2040, but completely erased all existing roads (Figure 7.12). The presence of minor roads in the image of the study area was identified as causing the GAN to make this error and the GAN had not been trained to receive minor roads in the input images. This led to the logic outlined in Section 4.4.1 of performing preprocessing to obtain cohesive sets of minor roads and training the GAN to generate these based on keeping some minor roads in the image. This ensured that the GAN was appropriately conditioned to generate future road networks in the study area, which already had a presence of minor roads.

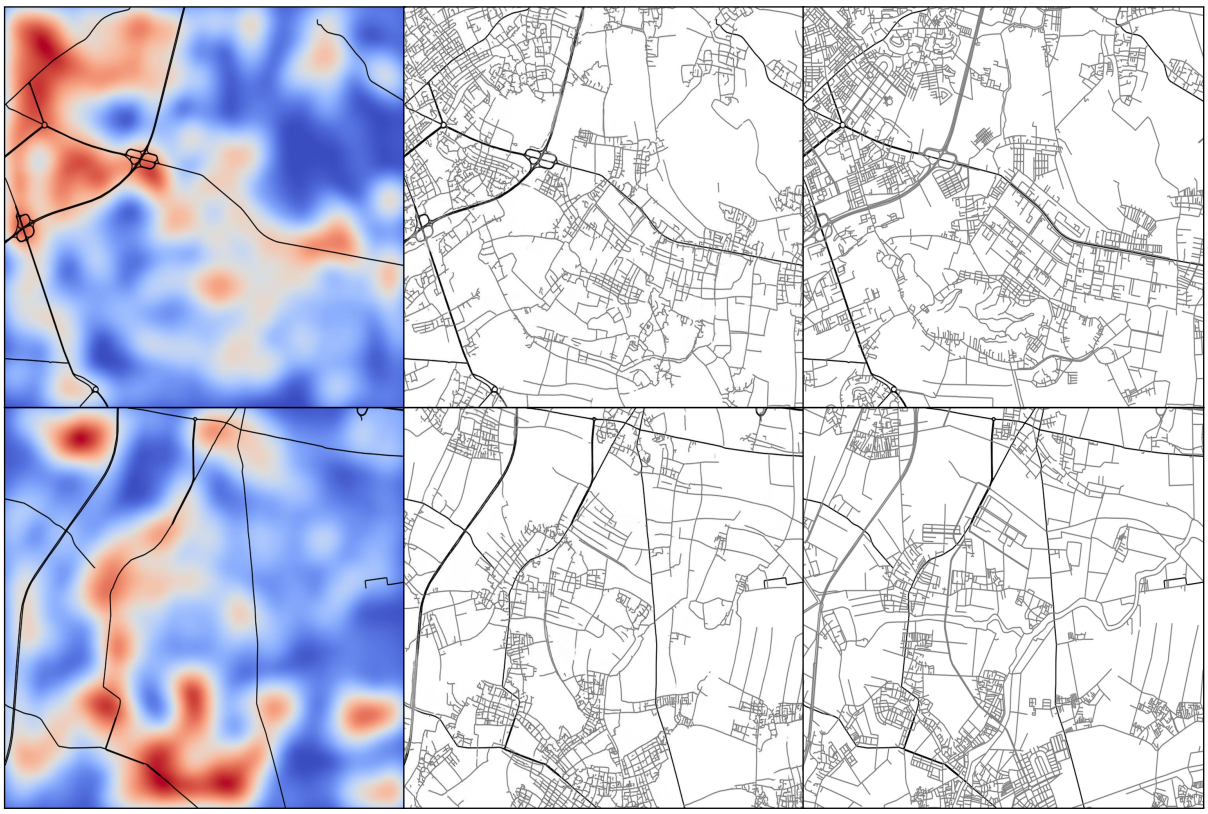

Figure 7.11: Two examples of the test images from the major and minor road experiment. left column: input, middle column: generated, right column: ground truth.

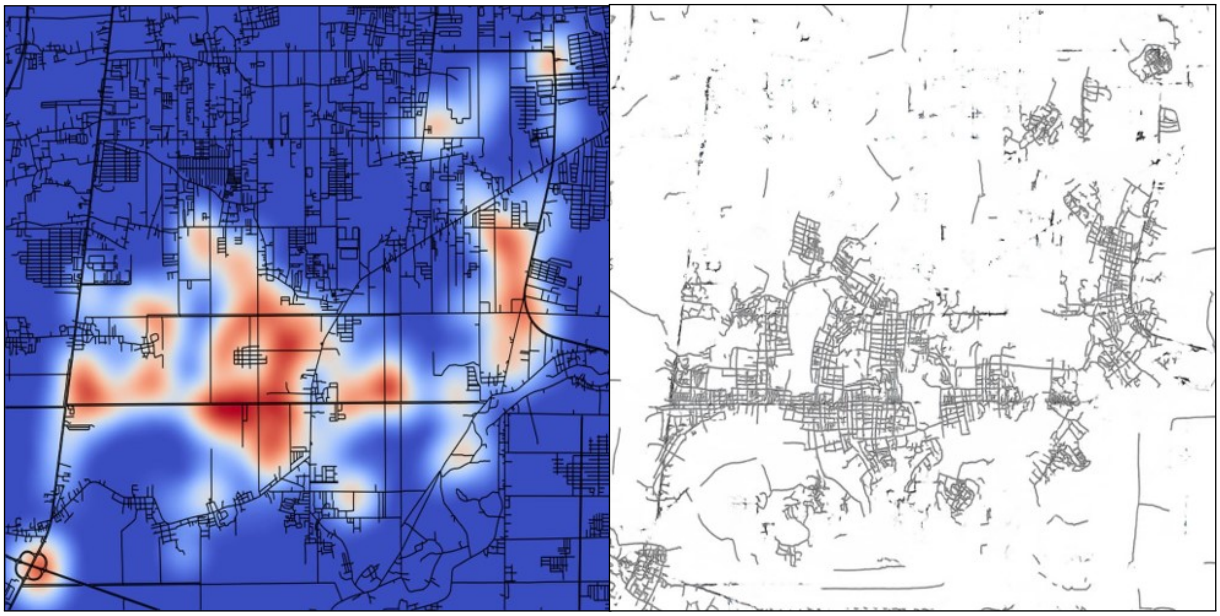

Figure 7.12: Study area future development density input image (left) and output road network generated from the GAN (right).

7.5.2.2 Building footprint experiments

Building footprint generation required a different approach to the generation of road networks. The issue here was that the GAN was failing to train to generate the footprints due to the number of epoch runs and the learning rate. Figure 7.13 shows the GAN

behaviour for building footprint at the 500 m^2 scale using 200 epochs and the learning rate of 2.00×10^{-4}, the same as the road network experiments. The generator loss and discriminator loss are atypical when considering the previous plots of GAN loss. In addition, the generated images are unclear and blurry (Figure 7.13).

Figure 7.13: Building footprint GAN loss log with two visual outputs on training data.

As Figure 7.13 shows, the losses have even fewer oscillations than Figure 7.10. It was identified that pix2pixHD does not need the full 200 epochs to generate building footprints, as they are significantly simpler shapes to capture as opposed to the complex road network layouts. Therefore, the learning rate was lowered to 2.00×10^{-5} and the number of epochs to 100, and as Figure 7.7 indicates, this resulted in much more natural GAN loss plots. Therefore, the failure of GAN was likely due to the excessive number of epochs and the high GAN learning rate, resulting in a convergence failure during the training process.

These unsuccessful building footprint experiments indicate a failure to converge, which means that the GAN training process is not reaching a stable state where the generator and discriminator are effectively learning from each other, resulting in poor quality generated outputs and often showing signs of the generator not improving or even getting worse over time; essentially, the two networks are not reaching a balanced Nash equilibrium in their adversarial game, leading to a lack of meaningful progress in the training process (Goodfellow, 2016).

The GANmapper loss log (Figure 7.8) contrasts with this training approach taken as it used 200 epochs and the default learning rate of 2.00×10^{-4}. The difference in different GAN parameter values for the training is because pix2pix was used for GANmapper whereas the GANs of the building footprint in this research used pix2pixHD. Due to enhancements and differences between pix2pix and pix2pixHD, it requires slightly different approaches. The loss log from Shou et al. (2021) supports this, as the author ran their GAN for 100 epochs, indicating that the pix2pixHD additions of a coarse-to-fine generator, multi-scale discriminators, and an improved adversarial loss result in a different training procedure which requires adjustments to the number of epochs. The loss logs for InstantCITY by Wu and Biljecki (2023) that uses pix2pixHD could not be obtained to compare with the GAN experiments in this thesis.

7.5.3 Generative adversarial network performance

7.5.3.1 Road network performance

Benchmarking the 2048x2048 global GAN experiments against other studies is challenging, as other road network GAN studies only visually show their metrics in histograms, making comparison difficult. In addition, the input to these models vary widely, some using topography, land use, junction points, and road density layers and vary widely in their pixel resolution, which would make the comparison unfair against the GAN developed here designed to create synthetic road networks with minimal input of a density raster (e.g. Yang et al., 2023b; Hartmann et al., 2017).

Gu et al. (2024) did report FID values for their road network diffusion model, which were compared and contrasted with the FID values of the 2048x2048 global experiments. The work performed a set of experiments across five North American cities and achieved FID values ranging between 34.79 and 299.29 in 25 experiments. Considering the 25%, 50% and 75% 2048x2048 GANs achieved FIDs of 28.17 - 33.30, this would suggest that they are fit for purpose and the GANs can still compete against diffusion models for the same FID performance.

Another performance issue that was highlighted in the road network generation results was the consistency of node degree accuracy. It is important to investigate why when the node degrees are greater than three, the accuracy decreases significantly, as this could be due to the tidying of the road network based on the Hartmann et al. (2017) approach. Figure 7.14 shows the vectorised road network before and after tidying. This reveals that while the tidying process is very capable of tidying nodes at t-junctions but at cross-roads it is challenging to maintain a node degree of 4 whilst having a sensible junction inputs. This has resulted in gridded networks being challenging to tidy. Other work has also omitted appropriate ways to tidy cross-roads in the road network and focusses on a maximum of 3 junctions (e.g. Yang et al., 2023b). Despite this shortcoming, the road networks are still contiguous and do not have missing vertices meaning junctions connect and make topological sense meaning they still have a strong level of realism and are appropriate to use for other applications, such as transport modelling.

Additionally, the node degree accuracy decline is due to the GAN's reduced exposure to rarer, higher-degree nodes during training, leading to under-representation in generated data. For instance, node degrees 1 and 2, which have the highest ground truth counts, achieve relatively high accuracy across all GAN training levels, while node degrees 5 and 6, with significantly fewer occurrences in the ground truth, show consistently low accuracy. This indicates that the GAN struggles to learn from and accurately generate less frequent node degrees, which could be due to the limited number of examples in the training data. Thus, while increased GAN training (from 25% to 75%) improves accuracy for more common node degrees, the model still underestimates rare, higher-degree nodes.

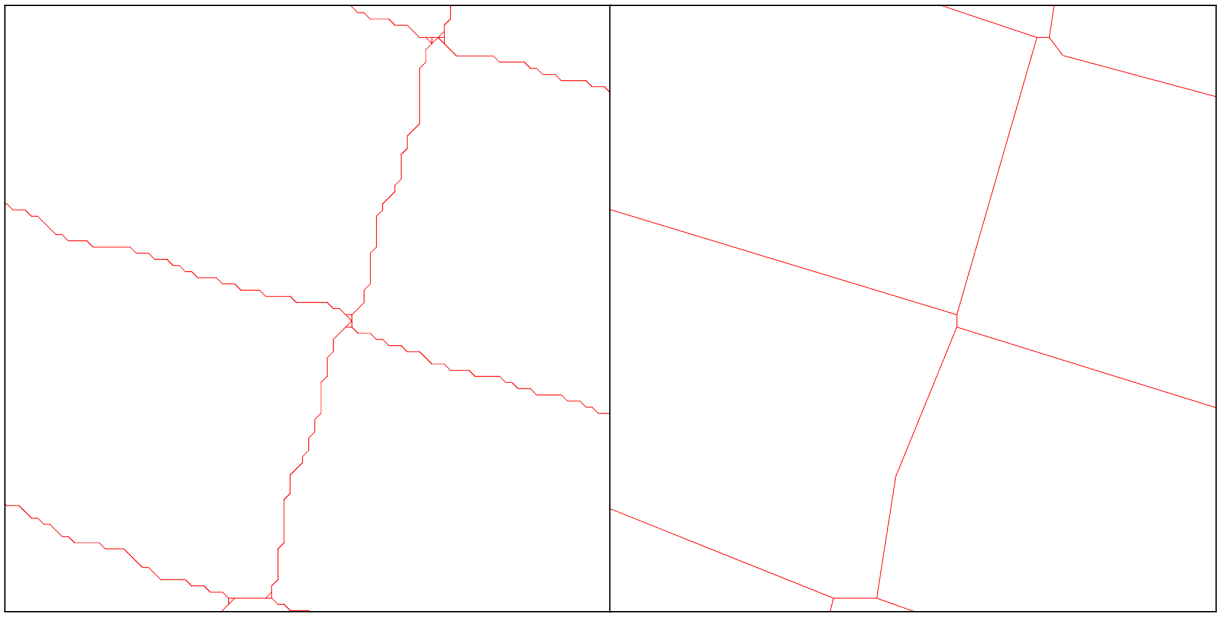

Figure 7.14: Impact on road tidying from pixel outputs. Left generated network before tidying, right generated network after tidying.

Another potential solution to this issue of developing more robust road networks is to use a deep generative model on the road graphs instead. Owaki and Machida (2020) developed RoadNetGAN which uses random walks to generate road networks with roads represented as graphs (Figure 7.15). However, this was not undertaken in this thesis, because of time constraints, lack of an open source repository, and high computational complexity of working with graphs. Instead, the road network was generated with images as this had an established track record in the literature of being feasible. RoadNetGAN also acknowledges the need to investigate the scalability of the proposed method when applied to areas of various sizes. Ensuring the model can handle larger or more complex networks remains an open challenge, as the paper focusses on an area of 1.7 km^2. Transformation of real-world road networks into graphs involves simplifications such as converting curved roads into straight roads and exempting motorways. These simplifications might limit the level of detail and realism that can be captured and generated by RoadNetGAN and pix2pixHD was more appropriate given its ability to generate roads across large areas.

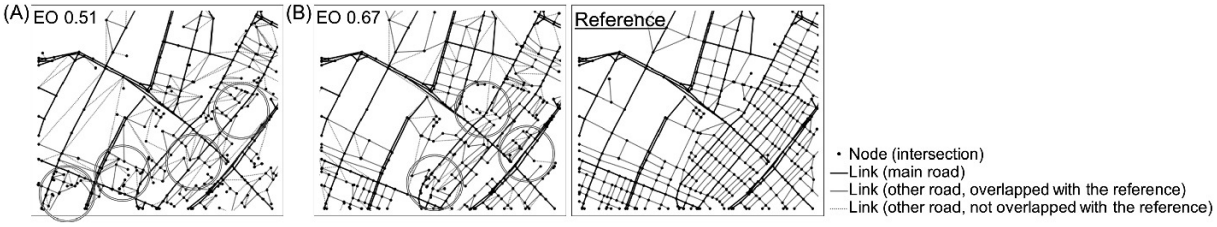

Figure 7.15: Example output of GAN road graphs from Owaki and Machida (2020).

7.5.3.2 Building footprint performance

The performance of the building footprint GANs is much easier to compare, as it used similar vector metrics and FID as InstantCITY, which also used pix2pixHD and used the exact same input of a road layout (Wu and Biljecki, 2023). Beginning with FID, the building footprint local GANs in this thesis outperform the average InstantCITY values, with 44.74 and 48.69 for 500 m^2 and 1000 m^2, respectively. InstantCITY achieved an FID of 81.8 and 75.2 for 500 m^2 and 1000 m^2 indicating an better performance from the GANs in this thesis as values closer to zero are more favourable. However, FID alone is not sufficient enough as mentioned previously to assess GAN performance.

The vector metrics for building footprints measured the difference between generated and ground truth footprints: therefore values closest to zero indicated exact agreement and are more favourable. InstantCITY took the same approach. Beginning with the building count, the local 500 m^2 and 1000 m^2 achieved -63.32% and -60.85%. In InstantCITY, the building count accuracy achieved an underestimation of -25.7% for the 1000 m^2 and -50.3% for the 500 m^2 experiments in the 8 cities. For building area, the local 500 m^2 and 1000 m^2 achieved 59.76% and 11.70%, respectively, whereas Wu

and Biljecki (2023) achieved 19.4% for 1000 m^2 and 30.2% for 500 m^2. The 1000 m^2 GAN outperforms both of these values which is a positive sign. For footprint area, the 500 m^2 achieved 39.76% and 1000 m^2 achieved 15.90%. InstantCITY achieved 8.7% for 1000 m^2 and 7.6% for 500 m^2.

Although the scores in this thesis generally do not meet that accuracy, especially for the building footprint count, this can be explained by the underlying geospatial data. InstantCITY performed experiments in Beirut, Frankfurt, Jakarta, London, New York, Rotterdam, Singapore, and Seattle using OSM building footprints. All of these cities are more developed than Phnom Penh and have a much better coverage of OSM data, which is why they were selected by Wu and Biljecki (2023). As this research relied on Google Buildings data that was extracted using automated approaches, it is unlikely that it meets the same quality assurance standards as OSM data and does not have perfect coverage. In addition, Phnom Penh has a much more random and compact development which does not always logically follow the layout of the roads compared to cities such as London (Figure 7.16). This adds an additional challenge to the GAN learning and explains the issue of underestimation in creating building footprints.

It should be noted that the trends between both studies are the same, for example, both the footprint area and perimeter are overestimated against the ground truth while the building count is underestimated. In addition, the MAE for the 1000 m^2 GAN in this thesis is 29.48% and for the corresponding InstantCITY GAN it is 17.93%, which shows that the average difference between the scores is not extreme.

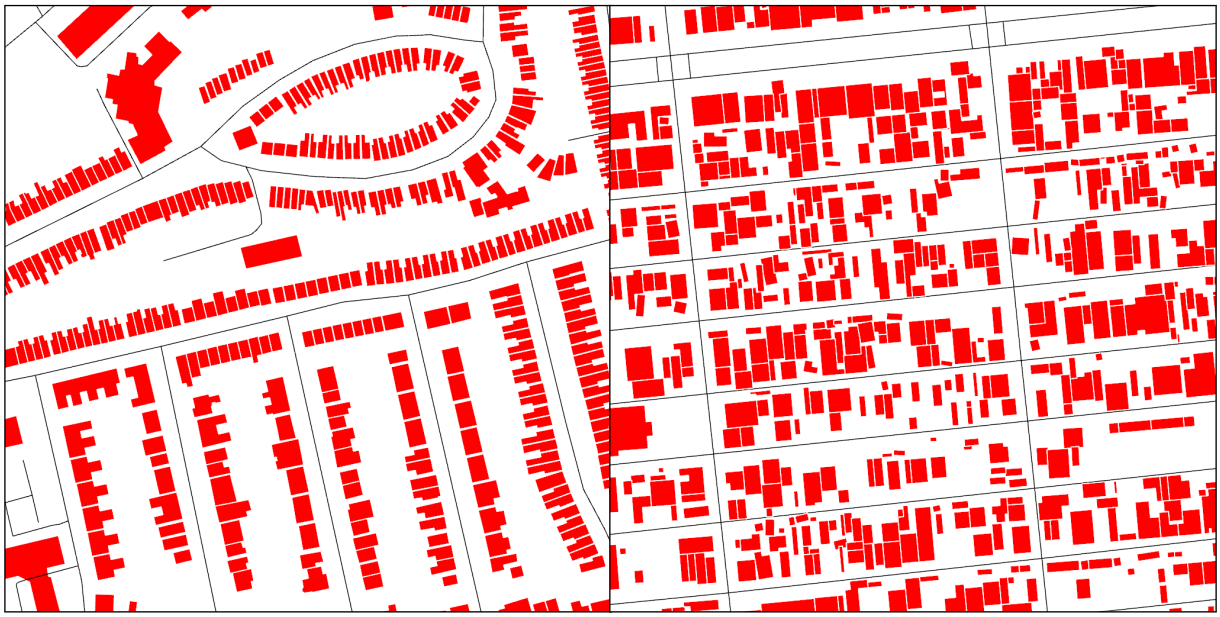

Figure 7.16: OSM buildings in London (left) and Google Buildings in Phnom Pemh (Right).

Another issue in GAN training for building footprint generation is there was no data availability on building age, therefore building footprint patterns similar to that shown in Figure 7.16 were used for training. This does not explicitly follow the patterns of new residential developments identified in the study area previously (Figure 4.3). This is why the simulated building footprints aren't as uniform in their urban form as expected and have more randomness (Figure 7.17). If the building age data was available the GAN could have been trained exclusively on building footprints in borey's exclusively leading to better consistency with new developments.

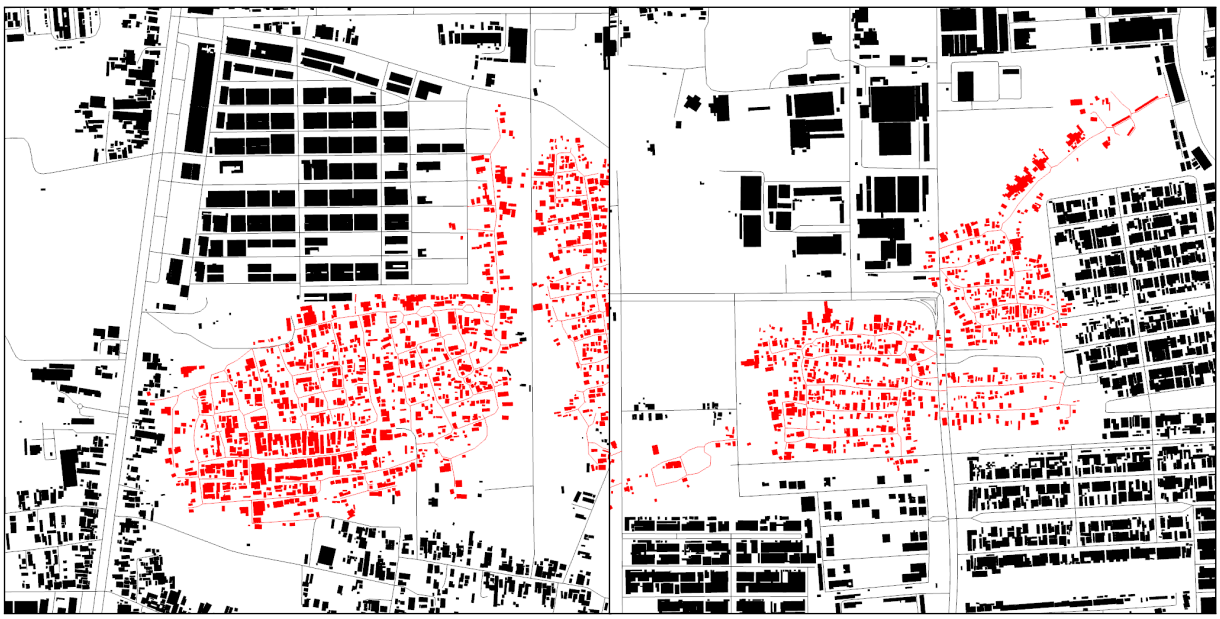

Figure 7.17: Generated urban form (red), existing urban form (black), from BAU scenario.

Another important note is that in the InstantCITY GAN, the 1000 m^2 achieves better average performance (17.93%) across the metrics than 500 m^2 average (29.36%). This also supports the findings in the GAN experiments performed in this thesis, although the gap between performance in the GANs here is bigger in this thesis. Despite this, it is crucial to perform multiple GAN experiments on the area being understood to identify the optimal GAN on a case-by-case basis. Whilst the results in InstantCITY are impressive, the majority of cities are ultimately outperformed in the 1000 m experiments compared to the 500 m^2 experiments. New York gets an average of 22.3% for 500 m^2 and 26.3%, showing that depending on the city type, the smaller scale can sometimes be more appropriate.

ControlCity by Zhou et al. (2024) used a diffusion model to generate building footprints based on road layouts, text prompts, metadata, and landuse information. In the paper, ControlCity was directly compared with pix2pixHD and achieved overall better FID and vector-based metrics than pix2pixHD. It actually overestimated the building footprint count in the 10 cities examined while pix2pixHD was underestimated in 7/10 cities.

This was attributed due to the diffusion model being more capable of generating more buildings due to the additional inputs in the diffusion model. This indicates that diffusion models could be used to provide enhanced performance over GANs and should now be considered for the task of urban form generation.

Another important consideration is the merging of outputs from the 500 m^2 and 1000 m^2 experiments into the future scenarios to create a more complete building footprint layer. The 500 m^2 experiment generated 17,862 buildings, while the 1000 m^2 experiment produced 16,628. However, neither GAN experiment achieved full coverage when generating building footprints. This is partly due to limitations in the Google Buildings dataset, which does not offer complete coverage across all areas. Additionally, the sparser road networks in the 1000 m^2 cells likely made it more difficult for the GAN to infer and generate buildings. In contrast, the 500 m^2 cells tend to contain denser road and urban features, providing the model with better contextual information. This likely explains the higher number of buildings generated in the 500 m^2 experiment. The GAN appears to perform better in areas with denser, more structured urban form, where it can more easily learn the spatial relationships necessary to place buildings realistically.

7.5.4 Planning implications of urban form scenarios

It is important to interpret the five urban form scenario outputs from an applied planning perspective. As previously noted in Section 7.4, the varying dasymetric density rasters for the scenarios had minimal impact on varying the 2040 specific urban form patterns (e.g. road layout type) and instead only controlled the overall location of new urban form. However, the extent and quantity of new urban form in the high growth scenarios (BAU, SP, and EB) and low growth scenarios (ER and LB) provide useful information from a planning perspective.

Among the high-growth scenarios, the BAU scenario showed the highest increase in urban form (Table 7.1). However, as illustrated in Figure 6.13, this expansion results in relatively limited sprawl. This is because the high volume of new development occurs in a cohesive and contiguous pattern, avoiding fragmented or leap-frogging growth (Scheer, 2001). The BAU scenario does pose a high risk of agricultural land loss as its vast growth would result in the urban conversion of productive or ecologically valuable land, posing threats to food security and natural buffers (Wiedenhofer et al., 2018). In contrast, the SP and EB scenarios (Figures 6.15 and 6.17) are more prone to dispersed, fragmented, and isolated development patterns which presents greater inefficiency risks to infrastructure costs and service. Their dispersed growth patterns increase per capita infrastructure costs and complicate service delivery for water, waste, and transit systems (Dempsey et al., 2010). All high-growth scenarios risk infrastructure strain because the addition of new major travel corridors is limited and primarily rein-

forces the existing grid and car-dependent travel. This indicates minimal investment in upgraded mobility systems and a lack of new primary road development (Wiedenhofer et al., 2018; Dempsey et al., 2010). These projections could influence planning strategies by encouraging a greater emphasis on infill development, optimising the use of underutilised central areas. They also highlight the need for targeted infrastructure corridors to accommodate expected growth, diversifying transport options, and stricter development controls to prevent dispersed, inefficient expansion as seen in the EB and SP scenarios.

Table 7.1: Statistics of 2040 simulated urban form outputs (buildings and roads) across all scenarios, compared to 2023 baseline values.

Scenario	Building Count	Building increase (%)	Road Length (km)	Road increase (%)
2023	24,873	–	318.34	–
BAU	51,394	106.63%	546.02	71.52%
ER	36,507	46.77%	401.97	26.27%
SP	43,370	74.37%	471.75	48.19%
LB	34,561	38.95%	387.98	21.88%
EB	44,178	77.61%	472.65	48.47%

For the ER and LB low-growth scenarios (Figures 6.14 and 6.16), further insights emerge from a planning perspective. These scenarios exhibit sparse and fragmented spatial patterns, with new buildings in dispersed clusters and long, linear road segments which are often redundant or parallel to existing infrastructure. There are noticeably fewer new residential neighbourhoods than the high-growth scenarios and significant gaps of undeveloped land remain between built-up areas. Infrastructure inefficiency and high per capita costs is a bigger risk in these scenarios, with many new roads being underutilised, lacking sufficient connections to the existing network and adjacent building development (Kundu et al., 2020). Stalled urban growth is also evidence, as minimal expansion of cohesive neighbourhoods suggests a slowdown in investment due to weak supply, land speculation, and constrained formal development capacity of these scenarios. Dispersed and limited new developments would undermine the viability of service delivery (Dempsey et al., 2010). These conditions imply that planners may need to adopt incremental and flexible planning strategies, such as adaptive zoning and densification of brownfield sites to address reduced fiscal capacity and slower urban development (Yasin et al., 2020; Wiedenhofer et al., 2018). It would become important for planners to avoid the long-term lock-in of inefficient road layouts and building patterns that could hinder further urban expansion after 2040. Planners could prioritise densification of core urban areas and protect strategically located land for future urban development when economic conditions improve, rather than permitting scattered

development. The ER and LB scenarios highlight the vulnerabilities of inefficient development and demonstrate how inefficiencies of urban sprawl can persist even with limited building activity.

Overall, these projections inform future planning approaches by revealing the infrastructural consequences of each urban form scenario, necessitating differentiated strategies for compact growth, infrastructure investment, and regulatory controls. These scenario outputs are useful applied planning tools by illustrating spatial development consequences and enabling evaluation of trade-offs in growth patterns, infrastructure strain, and service delivery efficiency.

7.6 Future urban form applications

The future urban form outputs generated in this thesis have numerous multidisciplinary applications across various models. This section outlines how future road networks and building footprints can be utilised to improve forecasts related to flooding, mobility, evacuation, and energy consumption. The following examples represent only a few potential applications, but the generated urban form is not limited to these use cases.

7.6.1 Flood modelling

CityCAT is a state-of-the-art hydrodynamic model for simulating urban pluvial and fluvial flooding, developed by Glenis et al. (2018). The model distinguishes between permeable and impermeable surfaces, incorporates building footprints as obstacles to flow, and simulates flow pathways, water depths, and velocities during rainfall events (Figure 7.18).

Incorporating the future buildings generated alongside the existing buildings will shed light on future flood sources, pathways and receptors in greater detail. These high-resolution urban growth and flooding outputs will inform improved mitigation and protection measures against future flooding. For example, it opens up the potential for advanced flooding analysis, such as topological analytics and neighbourhood vulnerability and exposure to flooding. Understanding how flooding events affect the topological properties of the simulated and existing road network provides new insight into the damaged road networks and the impact on traffic capacity in the road networks during the inundation event. Building type and density can indicate how the modelled flood events impact communities because the number of buildings that are inundated can be calculated.

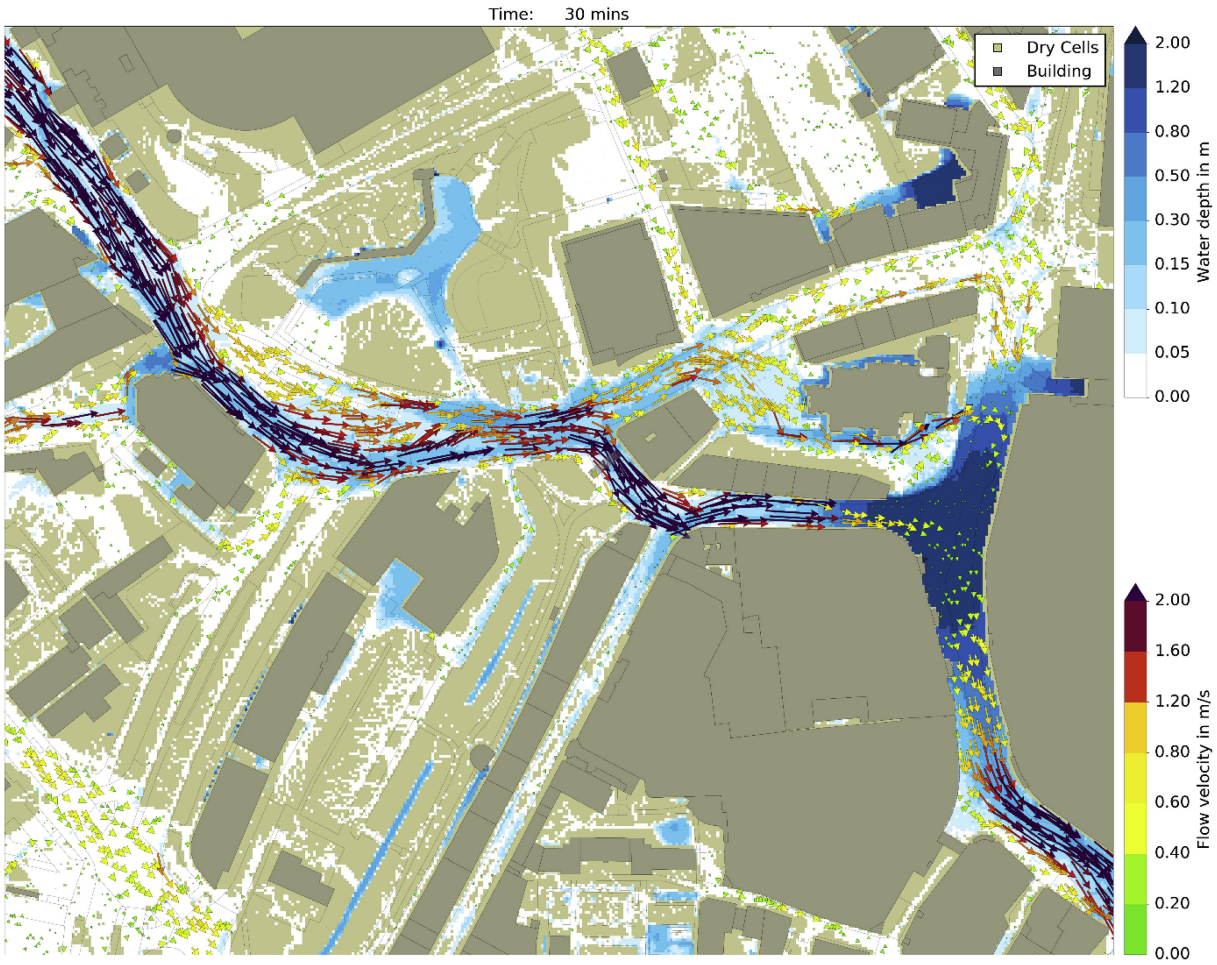

Figure 7.18: CityCAT 30 minute flooding event with 100 year return period in Newcastle Upon Tyne (Glenis et al., 2018).

The addition of new urban form plays a critical role in shaping future climate risk, offering insight into how socioeconomic change intersects with climate change. The Intergovernmental Panel on Climate Change (IPCC) conceptualises risk as a function of hazard, exposure, and vulnerability (Figure 7.19. Future urban development influences all three sides of the triangle. New buildings increase exposure by placing more assets and people in areas potentially affected by climate hazards, such as flooding. These new developments also exacerbate the hazard because densification and increased impermeable surfaces can increase the likelihood pluvial flood events. The form, density, and spatial distribution of the built environment directly shape vulnerability, affecting access to infrastructure, emergency services, and adaptive capacity. Urban form is an underutilised tool for representing hazards, vulnerability, and exposure, making it important in assessing and mitigating future climate risk.

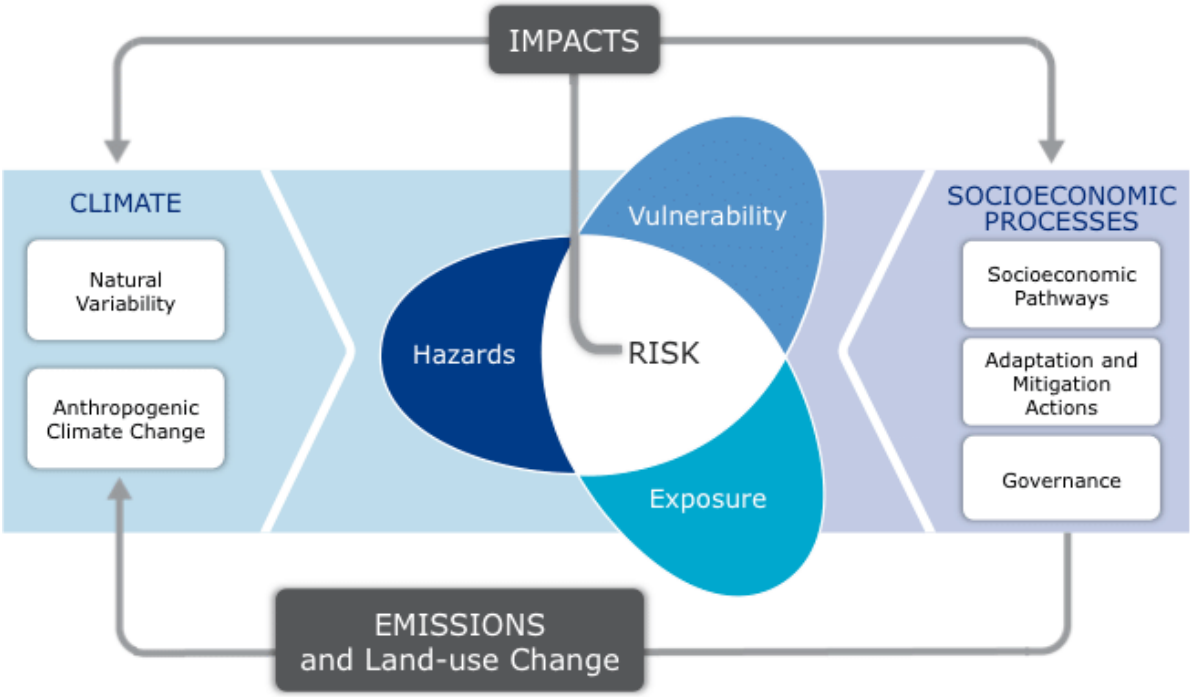

Figure 7.19: IPCC conceptual risk framework.

7.6.2 Evacuation modelling

There are numerous ABMs that simulate evacuations of people based on incidences, such as flooding or bomb threats. Dawson et al. (2011) developed a dynamic ABM of flood incident management that used road networks and buildings and coupled it with a hydrodynamic model to estimate the vulnerability of individuals to flooding under different storm surge conditions, defence breach scenarios, flood warning times, and evacuation strategies. Similarly, Chen and Kennedy (2022) used an agent-based model to simulate what happens when emergency responders respond to a dirty bomb exploding in a public metropolitan place.

The layout of future road networks and the distribution of building footprints can be integrated into such models to assess how evacuation routes may be impacted by new additions of urban form and whether there could be unforeseen evacuation challenges. For example, the different urban form scenarios generated in this thesis may lead to bottlenecks or congestion points in different areas of existing major roads. The scenarios could reveal more appropriate locations which minimise congestion based on this criteria. This would be beneficial to emergency planners and directly support the development and evaluation of emergency plans. Identifying these issues through modelling enables better planning of evacuation pathways and the strategic placement of resources like evacuation shelters. Effective incident management, including well-planned evacuations, is crucial for minimising loss of life and mitigating significant damage during flood events (Dawson et al., 2011; Chen and Kennedy, 2022).

7.6.3 Mobility and transport modelling

Castro et al. (2024) developed an ABM using MATSim and applied a novel open-source methodology to generate key input datasets. The two key inputs of the model were population which represented the demand of individuals living in the study area and the road network that represents the supply that serves as the geospatial representation of the transportation infrastructure used by the simulated agents.

The generated road network and building footprints could be integrated into this MATSim model to identify future travel demand by using the building footprints and average household size to estimate future population distribution. This would allow an assessment of how future population growth and urban development might impact existing and newly generated road networks. This approach also enables an exploration of how changes in urban form influence agents' spatial cognition, daily routines, and choices of transport modes and routes compared to current urban form layouts (Manley and Cheng, 2018). Furthermore, quantifying future transportation network scenarios would be invaluable in understanding their implications and effectiveness for travel patterns and mobility within a city. The modelling framework produces residential urban form, which could be used to examine how future residential developments might affect congestion, whether existing major roads can accommodate future population levels, and how efficient the generated networks would be in supporting urban mobility.

7.6.4 Energy use modelling

Hargreaves et al. (2017) developed a model to forecast residential urban form energy use. The study aimed to overcome the difficulty of reliably estimating the future potential of decentralised supply, which is heavily influenced by the availability of residential space, which is a factor that varies significantly within a city region. Hargreaves et al. (2017) method involved using the dwelling type, floor space, available outdoor space, and plot size of buildings to model energy use.

The building footprint layouts generated could be integrated into the modelling framework developed by Hargreaves et al. (2017) to understand how future residential densities impact future energy systems. Using the dimensions of future dwellings and calculating their associated plot size could be integrated into the energy model to estimate energy consumption and carbon dioxide emissions. The suitability and feasibility of decentralised energy technologies, such as ground source heat pumps, could be explicitly quantified on the basis of the curtilage characteristics using future building footprints. This would offer a more nuanced and accurate assessment of the potential for future decentralised energy systems. The generated urban form offers a strong basis for modelling future energy demand and evaluating future decentralised supply options.

7.7 Summary

Overall, this discussion has critiqued the geospatial modelling framework presented. Through the use of scenarios, the ABM has proven to be responsive to the model variables explored and the assumptions have been listed and justified to ensure the transparency of the ABM developed. The SOLV protocol provides numerous contributions, such as providing an adaptable way to capture the stochasticity of the ABM, ensure the ABM is not overfit when calibrating and validating it historically, and through SA has ensure the model can be applied to a practical application of simulating future urban development. The performance of the ABM is also good compared to other literature.

For the urban form modelling, the dasymetric mapping is an appropriate way to generate future urban form within areas expected to transition to urban, although variations in the density raster have limited impact on the density of road layouts. The GAN performance is close to that of its counterparts such as InstantCITY, but this has been explained by the underlying Google Buildings data which does not have perfect coverage or correlation with the Cambodian road layouts when compared to more developed regions. It is also crucial to perform multiple GAN experiments to test the sensitivity of the model to different inputs and parameters, due to the adversarial training process. Overall, the discussion has asserted that the geospatial modelling framework is robust and can be easily repeated within different regions based on the limited geospatial data in Southeast Asia.

Chapter 8

CONCLUSION AND FUTURE WORK

8.1 Chapter overview

This chapter summarises the research undertaken in this PhD thesis. It links the work back to the aims and objectives outlined in Chapter 1 and justifies how these have been achieved. This is followed by a summary of the key findings and contributions, highlighting the practical relevance of the research, particularly for rapidly developing and data-scarce cities. Research implications are provided which show the broader utility and applicability of the work. Future work is also outlined to guide and enhance further research in this field. Concluding remarks are provided, emphasising the strong research impact and value of this thesis, the challenges of innovation in the framework developed, and the potential of AI.

8.2 Fulfilment of aim and objectives

The aim of this thesis was to develop a geospatial modelling framework capable of simulating, exploring, and characterising future scenarios of urban form, by utilising state-of-the-art techniques. The evidence provided below explains how each objective and ultimately the aim was addressed in this thesis.

The first objective was to simulate future urban growth in an area experiencing rapid development using a spatial behavioural model. Through the literature review, ABMs were identified as an appropriate technique to use to simulate urban growth, as they have been widely used in previous literature and offer a flexible modelling approach to fully understand model behaviour and patterns. In addition, they easily allow for the development and exploration of scenarios compared to other approaches, such as CA. As a result, an ABM was designed for an area of Phnom Penh, based on the urban planning processes from grey literature and existing ABMs of land use change in academic literature.

The second objective was to develop and analyse multiple scenarios of future urban form based on likely socioeconomic forecasts and planning policy changes relevant to the study area. This was achieved through the development of the SOLV protocol, which ensures the credibility and robustness of the model. The protocol performed a thorough and systematic approach to understand the ABM developed by performing a stabilisation analysis to determine when the stochasticity of the model had settled; calibration and validation to ensure optimal model performance; and SA to quantify model variable importance, direction of influence, and interactions between variables. The SOLV protocol was purposefully created to be generic so it could be re-applied to similar stochastic urban growth models to verify them.

The third objective was to generate plausible future urban form using a deep generative model based on the outputs of the urban growth model while considering existing urban form patterns. This was done using the pix2pixHD GAN due to its track record in urban form modelling and its ability to operate at high resolutions, ensuring that it could generate road layouts in the whole study area. The density rasters of smaller residential roads were used to guide the generation of synthetic road networks and condition the GAN to generate road layouts on future development density rasters based on the ABM's simulated development probability outputs. Following this, building footprints were generated based on the road layout patterns. The GAN performance for both urban form elements was assessed using pixel and vector metrics and the best performing ones were applied to the study area.

The fourth objective was to develop and analyse multiple scenarios of future urban form based on likely socioeconomic forecasts relevant to the study area. This was achieved by varying ABM variables explored in the SOLV protocol and dasymetric mapping approaches to reflect the socioeconomic scenarios Phnom Penh may experience in the future scenarios extracted from the grey literature described in Chapter 3. A BAU scenario was developed based on the optimised ABM, and then an additional four scenarios (ER, SP, LB, EB) were created guided by the grey literature, using BAU as a baseline and using a scenario matrix with axes of unregulated to regulated and low to high growth.

8.3 Key findings and contributions

Developing an ABM of urban growth using irregular land parcels as the spatial unit enhanced the realism of the simulation, as these parcels represent the actual units where land transactions occur in reality. However, their irregular shape increased computational demand, leading to longer runtimes. This also highlights the challenge of scalability, such as how localised and smaller scale ABMs using real land parcels can simulate

growth in larger areas, an important consideration for modellers as practitioners and policymakers who may require such simulations. The use of social theories to justify agent behaviour, such as bid-rent and bounded rationality, were appropriate given the lack of available surveys and detailed census data. Scenarios developed have also shed-light on how model variables impact dynamics. In addition, the scenarios were developed with more appropriate means than previous literature, particularly by supporting them with empirical evidence and using a scenario matrix.

The SOLV Protocol has made numerous contributions to both existing urban growth modelling literature. It introduced a structured four-stage evaluation framework: preparation, model stabilisation, calibration and validation, and SA to ensure a comprehensive understanding of model behaviour and accuracy, going beyond the unsystematic evaluation procedures seen in many prior urban growth studies. One of the most important gaps addressed is the thorough model stabilisation process using model runs, CV difference calculations, and student's t-test to ensure the model's stochastic behaviour is adequately captured and understood. This rigorous assessment of stabilisation provides a unique perspective on capturing the stochasticity of urban growth models. Whilst the existing frameworks consider SA and/or calibration, none considered all three components of stabilisation, a separate calibration and validation, and SA. SOLV also uses a dual SA approach using Morris and Sobol which ensures a more comprehensive understanding of variable influences on model outcomes. The SOLV protocol separated calibration and validation periods which ensured the model was not overfit and identified it is more appropriate to split the historic period based on equal quantities of urban growth, not temporal duration. The SOLV protocol adapts and integrates the evaluation processes of broader ABM practices, bridging the gap between general ABM analysis and urban growth modelling.

This research has also demonstrated that GANs are a feasible approach to generate future road network layouts and building footprints. There have been no previous attempts to train a GAN to generate a future urban form, as has been done in this thesis. This required innovative thinking to condition a GAN to do this based on the future density raster in the study area. It has also addressed limitations to generate road layouts in large areas and that GANs involve a challenging training procedure compared to other deep learning techniques. These DGMs are sensitive to input image resolution and the density rasters generated for the scenarios appear to have minimal impact on road layout density. It also identified that GANs can struggle with building footprint generation when given buildings that do not clearly follow road layouts. This will remain an open challenge when performing urban form generation in developing regions.

For Phnom Penh, five potential socioeconomic scenarios have been identified, shedding light on possible future realities for the city. This analysis has enhanced understanding of how urban development may unfold in the coming decades, as well as the key drivers behind these changes. Although the focus is on one area, similar trends are likely to affect other parts of the urban periphery. It is the municipal council's responsibility to manage these developments effectively and devise strategies to mitigate potential negative impacts. Action from the municipal council will be crucial in ensuring that these challenges are addressed proactively. Phnom Penh is at risk to numerous impacts of urban sprawl, such as flooding, higher temperatures and more greenhouse gas emissions. Therefore, this research contributes strongly to helping mitigate and manage these risks through its outputs

8.4 Research implications

This section outlines the impact of the research's methodology and findings, highlighting what it means for urban growth modelling researchers, government policy makers, and urban development practitioners.

8.4.1 Implications for researchers

The geospatial modelling framework developed provides researchers with a robust and adaptable tool to simulate urban growth and form. Its modular design makes it easy to tailor the approach to different cities and research objectives. This framework is not exclusive to data-rich environments. As demonstrated through the Phnom Penh case study, the framework can be effectively applied to similar data-scarce and rapidly developing environments, offering confidence in the transferability of the approach presented. This makes it particularly valuable for researchers working in developing regions where traditional urban models, which require abundant physical and socioeconomic data, often fall short. Beyond urban growth modelling, the urban form outputs generated by the framework are highly transferable. They offer a foundation for cross-disciplinary applications, such as flood risk assessment, evacuation planning, and urban energy consumption modelling. This opens new research opportunities into how the spatial configuration of future cities will influence broader challenges, allowing more granular, integrated, and holistic projections of urban phenomenon.

8.4.2 Implications for policy makers

The results of this research offer valuable insights for government policy makers, particularly at the local and municipal levels. By using detailed land parcel data, the framework enables an accessible understanding of how and where urban growth will occur,

helping policy makers anticipate infrastructure demands in specific neighbourhoods. The generated urban form outputs provide useful visualisations of likely future development, allowing policy makers to identify potential hotspots of sprawl and assess their long-term implications. This could support policy interventions that promote more compact and sustainable development patterns, such as the redevelopment of existing urban areas rather than the continuation of sprawling gated residential housing developments observed in the study area. Furthermore, scenario-based results can help drive regulatory reform, such as the creation or strengthening of laws to guide sustainable urbanisation. For example, policies could be developed to encourage mixed land use, integrate green space, and ensure that future growth is environmentally and socially inclusive. The Ministry of Water Resources and Meteorology in Cambodia, responsible for freshwater supply and flood control, could use this framework to model urban expansion in other cities such as Battambang. This would help identify areas likely to face future water stress and inform proactive investment in water infrastructure, potentially mitigating water-related risks, such as drought.

8.4.3 Implications for practitioners

Urban planners and consultancies can use this geospatial modelling framework as a decision support tool. By simulating and visualising future urban growth and form under different scenarios, the framework supports evidence-based planning and infrastructure development. It would also be beneficial to consultancies, such as the Japan International Cooperation Agency and Mekong Modelling Associates, which work on engineering, infrastructure, and water management projects in Cambodia. The detailed future urban form outputs from the framework could be utilised to examine how they affect processes critical to consultancy projects, such as environmental impact assessments. By integrating the modelling framework into workflows, there is strong potential for practitioners to improve project efficiency and outcomes.

8.5 Thesis future work

The discussion identified shortcomings in this research that can drive future work to fill the gaps in this study. Many areas for future work described below were unable to be investigated due to time and budget constraints of the research project and were not necessarily within the scope of the aims and objectives. In addition, some are not issues with the research but relate to the underlying data used in the work.

8.5.1 Urban growth modelling

The ABM would benefit from empirical surveys to help verify and justify agent behaviour in the model. For example, surveying relevant stakeholders could provide evidence to support the selection of appropriate driving factors for each agent and offer insight into how the agents might realistically respond under different scenario conditions. The data latency issue could also be resolved by better temporal records of when roads and buildings were constructed, which is especially prevalent in rapidly developing regions when performing historical calibration and validation. However, these issues are flexible and can be easily adapted based on the availability of data in the study area. For example, developing an urban growth model in the UK would likely have minimal data latency issues and detailed census information and more up-to-date grey literature to guide agent behaviour.

The SOLV protocol could be improved by utilising spatiotemporal SA which would enable further insight into model dynamics and agent behaviour. In addition, the optimised model variables in calibration stay constant throughout the simulation. This is standard practice in modelling literature, but dynamic changes in the variables when appropriate may prove a useful investigation in the future. There is also an option to perform SA at the beginning of the protocol to help reduce variability dimensionality. An updated FKS metric should be designed to reduce the calculations sensitivity to inflate the score in areas of minimal change, ideally by using area change between simulated and observed, as done in this thesis. Additionally, the use of longer periods for calibration and validation would help ensure more confidence in future simulations, but this is subject to ABM runtimes and high-resolution satellite imagery availability.

8.5.2 Urban form generation

For urban form modelling, more experiments on the impact on development demand density on road network layout would be prudent to better understand the data driven processes in the model. For example, more diverse colour schemes in the density raster may result in more response from the GAN to generate different densities of roads. This suggests a limitation of the GAN model in creating diverse road network layouts, which may influence the effectiveness of the generated data for practical applications. It would be prudent to investigate the underlying causes of this discrepancy, such as the model's ability to represent road density or its sensitivity to specific network features, such as block area, and explore improvements to address this issue. There could also be model coupling between the ABM and GAN. This could involve simulating five years of development in the ABM then generating the urban form and then repeating this process to 2040 could reduce the disparity observed.

As discussed, some of the GAN experiments suffered from mode collapse or convergence failure, due to improper parameter value configurations, such as the number of epochs and learning rate. One potential solution is to implement hyper-parameter tuning and early stopping, as is done in conventional machine and deep learning models, based on FID or the difference between generator and discriminator loss. This approach would help identify appropriate GAN parameter values are set, improving and ensuring GAN training robustness. Alternatively, as the GAN training process is adversarial, a diffusion model created specifically for the urban form generation at a high pixel resolution would be beneficial given that these have better performance than GANs and have a more straightforward training process. There is the potential of developing a graph-based deep generative model for road network generation since this could help improve network connectivity and node degree accuracy as the conversion and tidying from image to vector resulted in road network details being removed.

8.5.3 Other recommendations

Other general recommendations include improving national mapping efforts worldwide by collecting more frequent and accurate temporal records of spatial urban change. This would support the development of more effective and efficient frameworks. The collection of land parcel data and tracking their evolution over time is essential, as is the recording of new building developments and road networks. If this information were readily available and open source, it would have saved time and enabled more sophisticated analyses in this PhD research. More detailed census information in developing countries would also be beneficial because this could allow enhanced agent detail, such as different types of the same agent, thereby increasing model heterogeneity and realism.

In addition, there should be a greater emphasis on the application of generative AI in urban form modelling. This is a rapidly evolving field and the use of GANs has demonstrated their capability to capture the complexities of urban form. There is significant potential for DGMs to capture even more detailed urban form elements, such as road widths, amenities, plot sizes, greenspaces, and semantic land use layers. This would approach the level of detail found in datasets such as Ordnance Survey MasterMap. Also investigating the feasibility of conditioning DGMs to create more varied urban form layouts for the same area and what controls the morphological, functional, and spatial characteristics of the generated outputs in these black-box data-driven models. These advancements would further enhance multidisciplinary applications, generative AI trust and understanding, and scenario development.

8.6 Final remarks

This PhD research has demonstrated the value of future urban form modelling as a critical area of study, especially as the majority of the global population now live in urban areas. Accurate and detailed simulations of how these urban environments will evolve is essential for planning sustainable, resilient, and liveable cities. The key innovations of this work were the creation of an urban growth land parcel scale ABM, the presentation of the SOLV protocol to ensure the transparency and robustness of the urban growth model, and the application of GANs to simulate future urban form. This research demonstrated that future urban form modelling is not only technically achievable, but offers a new direction for urban modelling outputs, with significant implications for urban planning, design, and policy. The framework's implementation in a data-scarce environment, such as Phnom Penh, further demonstrates its adaptability for transferability to other cities.

One of the most challenging aspects was the application of GANs to generate plausible future road networks. This task involved substantial trail and error and should be a strong focus for development in future research. However, this pushed methodological boundaries and affirmed the value of DGMs in capturing the complexity of urban form. Ultimately, this thesis contributes a unique and forward-looking approach to urban form simulation. It highlights the potential of generative AI in geospatial urban modelling and the importance of model quality in producing reliable and actionable insights. As the field continues to evolve, this research offers a springboard for future exploration and interdisciplinary applications, opening up new research opportunities, standards, and developments.

REFERENCES

Achmad, A., Hasyim, S., Dahlan, B. and Aulia, D. N. (2015), 'Modeling of urban growth in tsunami-prone city using logistic regression: Analysis of banda aceh, indonesia', *Applied geography* **62**, 237–246.

Agyemang, F. S. and Sahana, M. (2025), 'Modelling urbanisation in cities in the global south: A review of progress and framework for the future', *Geography Compass* **19**(2), e70019.

Agyemang, F. S., Silva, E. and Fox, S. (2023), 'Modelling and simulating 'informal urbanization': An integrated agent-based and cellular automata model of urban residential growth in ghana', *Environment and Planning B: Urban Analytics and City Science* **50**(4), 863–877.

Ajzen, I. (1991), 'The theory of planned behavior', *Organizational Behavior and Human Decision Processes* **50**(2), 179–211.

Alghais, N. and Pullar, D. (2018), 'Modelling future impacts of urban development in kuwait with the use of abm and gis', *Transactions in GIS* **22**(1), 20–42.

Alonso, W. (1964), *Location and Land Use: Toward a General Theory of Land Rent*, Harvard University Press, Cambridge, MA.

Angel, S., Parent, J., Civco, D. L., Blei, A. and Potere, D. (2011), 'The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050', *Progress in planning* **75**(2), 53–107.

AWS (2025), 'Amazon ec2 g5 instances'. Accessed: 2025-03-04.
URL: *https://aws.amazon.com/ec2/instance-types/g5/*

Ayllón, D., Railsback, S. F., Vincenzi, S., Groeneveld, J., Almodóvar, A. and Grimm, V. (2016), 'Instream-gen: Modelling eco-evolutionary dynamics of trout populations under anthropogenic environmental change', *Ecological Modelling* **326**, 36–53.

Badshah, M. T., Hussain, K., Rehman, A. U., Mehmood, K., Muhammad, B., Wiarta, R., Silamon, R. F., Khan, M. A. and Meng, J. (2024), 'The role of random forest and markov chain models in understanding metropolitan urban growth trajectory', *Frontiers in Forests and Global Change* **7**, 1345047.

Batty, M. (2009), ‘Cities as complex systems: Scaling, interaction, networks, dynamics and urban morphologies.’.

Belgiu, M. and Drăguț, L. (2016), ‘Random forest in remote sensing: A review of applications and future directions’, *ISPRS journal of photogrammetry and remote sensing* **114**, 24–31.

Beneš, J., Wilkie, A. and Křivánek, J. (2014), Procedural modelling of urban road networks, *in* ‘Computer Graphics Forum’, Vol. 33, Wiley Online Library, pp. 132–142.

Berger, V. W. and Zhou, Y. (2014), ‘Kolmogorov–smirnov test: Overview’, *Wiley statsref: Statistics reference online* .

Bhatta, B. (2010), ‘Causes and consequences of urban growth and sprawl’, *Analysis of urban growth and sprawl from remote sensing data* pp. 17–36.

Birsak, M., Kelly, T., Para, W. and Wonka, P. (2022), ‘Large-scale auto-regressive modeling of street networks’, *arXiv preprint arXiv:2209.00281* .

Booth, A. and Burke, R. (2023), A novel application of duck-net for extracting supraglacial lakes in greenland, *in* ‘Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems’, pp. 1–4.

Borgonovo, E., Pangallo, M., Rivkin, J., Rizzo, L. and Siggelkow, N. (2022), ‘Sensitivity analysis of agent-based models: a new protocol’, *Computational and Mathematical Organization Theory* **28**(1), 52–94.

Brasebin, M., Perret, J., Mustière, S. and Weber, C. (2018), ‘3D urban data to assess local urban regulation influence’, *Computers, Environment and Urban Systems* **68**, 37–52.

Brown, D. G., Goovaerts, P., Burnicki, A. and Li, M.-Y. (2002), ‘Stochastic simulation of land-cover change using geostatistics and generalized additive models’, *Photogrammetric engineering and remote sensing* **68**(10), 1051–1062.

Brown, D. G., Page, S., Riolo, R., Zellner, M. and Rand, W. (2005), ‘Path dependence and the validation of agent-based spatial models of land use’, *International journal of geographical information science* **19**(2), 153–174.

Brunsdon, C., Fotheringham, A. S. and Charlton, M. E. (1996), ‘Geographically weighted regression: a method for exploring spatial nonstationarity’, *Geographical analysis* **28**(4), 281–298.

Burke, R., Sengupta, R. and Ford, A. (2024), ‘A 3d agent-based model for simulating urban densification in toronto, canada’, *Environment and Planning B: Urban Analytics and City Science* p. 23998083241261762.

Campolongo, F., Saltelli, A. and Cariboni, J. (n.d.), 'From screening to quantitative sensitivity analysis. a unified approach', *Computer physics communications* **182**(4), 978–988.

Cao, H., Liu, J., Chen, J., Gao, J., Wang, G. and Zhang, W. (2019), 'Spatiotemporal patterns of urban land use change in typical cities in the greater mekong subregion (gms)', *Remote sensing* **11**(7), 1–28.

Castro, D. A., Ford, A., James, P., Palacín, R. and Ziemke, D. (2024), 'A matsim model methodology to generate cycling-focused transport scenarios in england', *Journal of Urban Mobility* **5**, 100078.

Cattaneo, G., Flocchini, P., Mauri, G., Vogliotti, C. Q. and Santoro, N. (1997), 'Cellular automata in fuzzy backgrounds', *Physica D: Nonlinear Phenomena* **105**(1-3), 105–120.

Chakraborty, A., Sikder, S., Omrani, H. and Teller, J. (2022), 'Cellular automata in modeling and predicting urban densification: Revisiting the literature since 1971', *Land* **11**(7), 1113.

Chen, E. Q. and Kennedy, W. G. (2022), Investigating emergency responders' roles in a dirty bomb event with an agent-based model, *in* 'Conference of the Computational Social Science Society of the Americas', Springer, pp. 77–91.

Chicco, D., Warrens, M. J. and Jurman, G. (2021), 'The matthews correlation coefficient (mcc) is more informative than cohen's kappa and brier score in binary classification assessment', *Ieee Access* **9**, 78368–78381.

Clarke, K. C., Hoppen, S. and Gaydos, L. (1997), 'A self-modifying cellular automaton model of historical urbanization in the san francisco bay area', *Environment and planning B: Planning and design* **24**(2), 247–261.

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

Crooks, A., Heppenstall, A., Manley, E. and Malleson, N. (2018), 'Agent-based modelling and geographical information systems: a practical primer'.

Dahal, K. R. and Chow, T. E. (2014), 'An agent-integrated irregular automata model of urban land-use dynamics', *International Journal of Geographical Information Science* **28**(11), 2281–2303.

Dawson, R. J., Peppe, R. and Wang, M. (2011), 'An agent-based model for risk-based flood incident management', *Natural hazards* **59**, 167–189.

Dempsey, N., Brown, C., Raman, S., Porta, S., Jenks, M., Jones, C. and Bramley, G. (2010), 'Elements of urban form', *Dimensions of the sustainable city* pp. 21–51.

Diepart, J.-C., Koditek, W., Hänert, T. and Rock, F. (2016*a*), '1. introduction to the cambodian spatial planning system'.

Diepart, J.-C., Koditek, W., Hänert, T. and Rock, F. (2016*b*), '2. provincial spatial planning handbook'.

Diepart, J.-C., Koditek, W., Hänert, T. and Rock, F. (2016*c*), '3. district & municipal land use master plan and land use plan handbook'.

Dinda, S., Das, K., Chatterjee, N. D. and Ghosh, S. (2019), 'Integration of gis and statistical approach in mapping of urban sprawl and predicting future growth in midnapore town, india', *Modeling Earth Systems and Environment* **5**, 331–352.

Du, G., Yuan, L., Shin, K. J. and Managi, S. (2018), 'Enhancement of land-use change modeling using convolutional neural networks and convolutional denoising autoencoders', *arXiv preprint arXiv:1803.01159* .

Du, J. and Wang, Q. (2011), 'Exploring reciprocal influence between individual shopping travel and urban form: Agent-based modeling approach', *Journal of Urban Planning and Development* **137**(4), 390–401.

ESA (2019), Eo4sd-urban: Phnom penh operations report, Technical Report v1.2, European Space Agency (ESA). Includes Quality Control.
URL: *https://www.esa.int/Applications/Observing*$_{t}he_{E}arth/EO4SD$

Esri (2021), 'Parcel extraction - usa'. Accessed: 2023-06-10.
URL: *https://www.arcgis.com/home/item.html?id=ee7a5cf6ea4242f7a33014c6d16096ce*

Esri (2024), 'Esri wayback imagery'. Accessed: 2024-02-07.
URL: *https://livingatlas.arcgis.com/wayback/*

Fang, Z., Qi, J., Fan, L., Huang, J., Jin, Y. and Yang, T. (2022), 'A topography-aware approach to the automatic generation of urban road networks', *International Journal of Geographical Information Science* **36**(10), 2035–2059.

Fang, Z., Yang, T. and Jin, Y. (2020), 'Deepstreet: A deep learning powered urban street network generation module', *arXiv preprint arXiv:2010.04365* .

Feng, W., Song, P. and Li, B. (2020), Application of multi-agent system in land use decision-making of industrial park: A case study of tianfo health industrial park, shandong, *in* 'International Conference on Spatial Data and Intelligence', Springer, pp. 139–152.

Freitas, D., Lopes, L. G. and Morgado-Dias, F. (2020), 'Particle swarm optimisation: a historical review up to the current developments', *Entropy* **22**(3), 362.

Frumkin, H. (2002), 'Urban sprawl and public health', *Public health reports* .

Fujita, M. (1989), 'Urban economic theory', *Cambridge Books* .

Galin, E., Peytavie, A., Maréchal, N. and Guérin, E. (2010), Procedural generation of roads, *in* 'Computer Graphics Forum', Vol. 29, Wiley Online Library, pp. 429–438.

Gao, C., Lan, X., Li, N., Yuan, Y., Ding, J., Zhou, Z., Xu, F. and Li, Y. (2024), 'Large language models empowered agent-based modeling and simulation: A survey and perspectives', *Humanities and Social Sciences Communications* **11**(1), 1–24.

García, A. M., Santé, I., Crecente, R. and Miranda, D. (2011), 'An analysis of the effect of the stochastic component of urban cellular automata models', *Computers, Environment and Urban Systems* **35**(4), 289–296.

García, A., Santé, I., Boullón, M. and Crecente, R. (2013), 'Calibration of an urban cellular automaton model by using statistical techniques and a genetic algorithm. application to a small urban settlement of nw spain', *International Journal of Geographical Information Science* **27**(8), 1593–1611.

Gaur, S. and Singh, R. (2023), 'A comprehensive review on land use/land cover (lulc) change modeling for urban development: current status and future prospects', *Sustainability* **15**(2), 903.

GGGI (2018), *Phnom Penh Sustainable City Plan 2018-2030*, Phnom Penh, Cambodia. Developed in collaboration with the Phnom Penh Capital Administration (PPCA) and the National Council for Sustainable Development (NCSD).

Glenis, V., Kutija, V. and Kilsby, C. G. (2018), 'A fully hydrodynamic urban flood modelling system representing buildings, green space and interventions', *Environmental Modelling & Software* **109**, 272–292.

Goodfellow, I. (2016), 'Nips 2016 tutorial: Generative adversarial networks', *arXiv preprint arXiv:1701.00160* .

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014), 'Generative adversarial nets', *Advances in neural information processing systems* **27**.

Gribov, A. (2019), Optimal compression of a polyline while aligning to preferred directions, *in* '2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)', Vol. 1, IEEE, pp. 98–102.

Grimm, V., Augusiak, J., Focks, A., Frank, B. M., Gabsi, F., Johnston, A. S., Liu, C., Martin, B. T., Meli, M., Radchuk, V. et al. (2014), 'Towards better modelling and decision support: Documenting model development, testing, and analysis using trace', *Ecological modelling* **280**, 129–139.

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J. et al. (2020), 'The odd protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism', *Journal of Artificial Societies and Social Simulation* **23**(2).

Groeneveld, J., Müller, B., Buchmann, C. M., Dressler, G., Guo, C., Hase, N., Hoffmann, F., John, F., Klassert, C., Lauf, T. et al. (2017), 'Theoretical foundations of human decision-making in agent-based land use models–a review', *Environmental modelling & software* **87**, 39–48.

Gu, X., Zhang, M., Lyu, J. and Ge, Q. (2024), 'Generating urban road networks with conditional diffusion models', *ISPRS International Journal of Geo-Information* **13**(6), 203.

Guo, Z. and Liu, X. (2024), 'How artificial intelligence cooperating with agent-based modeling for urban studies: A systematic review', *Transactions in GIS* **28**(3), 654–674.

Guzman, L. A., Escobar, F., Peña, J. and Cardona, R. (2020), 'A cellular automata-based land-use model as an integrated spatial decision support system for urban planning in developing cities: The case of the bogotá region', *Land use policy* **92**, 104445.

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

Hargreaves, A., Cheng, V., Deshmukh, S., Leach, M. and Steemers, K. (2017), 'Forecasting how residential urban form affects the regional carbon savings and costs of retrofitting and decentralized energy supply', *Applied Energy* **186**, 549–561.

Hartmann, S., Weinmann, M., Wessel, R. and Klein, R. (2017), 'Streetgan: Towards road network synthesis with generative adversarial networks'.

Henry, F., Herwindiati, D. E., Mulyono, S. and Hendryli, J. (2017), Sugarcane land classification with satellite imagery using logistic regression model, *in* 'IOP conference series: materials science and engineering', Vol. 185, IOP Publishing, p. 012024.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B. and Hochreiter, S. (2017), 'Gans trained by a two time-scale update rule converge to a local nash equilibrium', *Advances in neural information processing systems* **30**.

Houet, T., Aguejdad, R., Doukari, O., Battaia, G. and Clarke, K. (2016), 'Description and validation of a "non path-dependent" model for projecting contrasting urban growth futures', *Cybergeo: European Journal of Geography* .

Isola, P., Zhu, J.-Y., Zhou, T. and Efros, A. A. (2017), Image-to-image translation with conditional adversarial networks, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 1125–1134.

Johnson, J., Alahi, A. and Fei-Fei, L. (2016), Perceptual losses for real-time style transfer and super-resolution, *in* 'Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14', Springer, pp. 694–711.

Kafy, A.-A., Rahman, M. S., Hasan, M. M., Islam, M. et al. (2020), 'Modelling future land use land cover changes and their impacts on land surface temperatures in rajshahi, bangladesh', *Remote Sensing Applications: Society and Environment* **18**, 100314.

Kamusoko, C. and Gamba, J. (2015), 'Simulating urban growth using a random forest-cellular automata (rf-ca) model', *ISPRS International Journal of Geo-Information* **4**(2), 447–470.

Kang, J.-Y., Michels, A., Crooks, A., Aldstadt, J. and Wang, S. (2022), 'An integrated framework of global sensitivity analysis and calibration for spatially explicit agent-based models', *Transactions in GIS* **26**(1), 100–128.

Kaviari, F., Mesgari, M. S., Seidi, E. and Motieyan, H. (2019), 'Simulation of urban growth using agent-based modeling and game theory with different temporal resolutions', *Cities* **95**, 102387.

Kempinska, K. and Murcio, R. (2019), 'Modelling urban networks using variational autoencoders', *Applied Network Science* **4**(1), 1–11.

Kim, T. K. (2015), 'T test as a parametric statistic', *Korean journal of anesthesiology* **68**(6), 540–546.

Kingma, D. P. (2013), 'Auto-encoding variational bayes', *arXiv preprint arXiv:1312.6114* .

Kisamba, F. C. and Li, F. (2023), 'Analysis and modelling urban growth of dodoma urban district in tanzania using an integrated ca–markov model', *GeoJournal* **88**(1), 511–532.

Koch, J., Dorning, M. A., Van Berkel, D. B., Beck, S. M., Sanchez, G. M., Shashidharan, A., Smart, L. S., Zhang, Q., Smith, J. W. and Meentemeyer, R. K. (2019), 'Modeling landowner interactions and development patterns at the urban fringe', *Landscape and Urban Planning* **182**, 101–113.

Korah, A., Koch, J. A. and Wimberly, M. C. (2024), 'Understanding urban growth modeling in africa: Dynamics, drivers, and challenges', *Cities* **146**, 104734.

Kropf, K. (2017), *The handbook of urban morphology*, John Wiley & Sons.

Ku, C.-A. (2024), 'Evaluating the effects of land-use strategies on future flood risk reduction in urban areas', *Cities* **150**, 104989.

Kundu, D., Sietchiping, R. and Kinyanjui, M. (2020), *Developing national urban policies*, Springer.

Law, A. M. and Kelton, W. D. (2007), *Simulation Modeling and Analysis*, 4th edn, McGraw-Hill, Boston.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. et al. (2017), Photo-realistic single image super-resolution using a generative adversarial network, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 4681–4690.

Lee, J.-S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, F., Lorscheid, I., Voinov, A., Polhill, J. G., Sun, Z. and Parker, D. C. (2015), 'The complexities of agent-based modeling output analysis', *Journal of Artificial Societies and Social Simulation* **18**(4).

Li, F., Li, Z., Chen, H., Chen, Z. and Li, M. (2020a), 'An agent-based learning-embedded model (abm-learning) for urban land use planning: A case study of residential land growth simulation in shenzhen, china', *Land Use Policy* **95**, 104620.

Li, F., Wang, R., Lu, S., Shao, M., Ding, J. and Sun, Q. (2021), 'Spatiotemporal simulation of green space by considering socioeconomic impacts based on a sd-ca model', *Forests* **12**(2), 202.

Li, J. and O'Donoghue, C. (2013), 'A survey of dynamic microsimulation models: uses, model structure amd methodology', *International Journal of microsimulation* **6**(2), 3–55.

Li, L., Uyttenhove, P. and Van Eetvelde, V. (2020b), 'Planning green infrastructure to mitigate urban surface water flooding risk–a methodology to identify priority areas applied in the city of ghent', *Landscape and Urban Planning* **194**, 103703.

Li, X., Chen, Y., Liu, X., Xu, X. and Chen, G. (2017), 'Experiences and issues of using cellular automata for assisting urban and regional planning in china', *International Journal of Geographical Information Science* **31**(8), pp.1606–1629.

Li, X. and Gong, P. (2016), 'Urban growth models: progress and perspective', *Science bulletin* **61**(21), 1637–1650.

Li, X. and Yeh, A. G.-O. (2005), 'Integration of genetic algorithms and gis for optimal location search', *International Journal of Geographical Information Science* **19**(5), 581–601.

Ligmann-Zielinska, A. (2013), 'Spatially-explicit sensitivity analysis of an agent-based model of land use change', *International Journal of Geographical Information Science* **27**(9), 1764–1781.

Ligmann-Zielinska, A., Jankowski, P., Najwer, A. and Zwoliński, Z. (2024), 'A streamlined approach to uncertainty and sensitivity analysis for models with spatial outputs with an example from geodiversity assessment', *International Journal of Geographical Information Science* **38**(8), 1531–1554.

Ligmann-Zielinska, A., Siebers, P.-O., Magliocca, N., Parker, D. C., Grimm, V., Du, J., Cenek, M., Radchuk, V., Arbab, N. N., Li, S. et al. (2020), ''one size does not fit all': A roadmap of purpose-driven mixed-method pathways for sensitivity analysis of agent-based models', *Journal of Artificial Societies and Social Simulation* **23**(1).

Lin, B., Jabi, W., Corcoran, P. and Lannon, S. (2024), 'The application of deep generative models in urban form generation based on topology: a review', *Architectural Science Review* **67**(3), 189–204.

Lin, J., Li, X., Wen, Y. and He, P. (2023), 'Modeling urban land-use changes using a landscape-driven patch-based cellular automaton (lp-ca)', *Cities* **132**, 103906.

Liu, D., Zheng, X. and Wang, H. (2020), 'Land-use simulation and decision-support system (landsds): Seamlessly integrating system dynamics, agent-based model, and cellular automata', **417**, 108924.

Liu, Y., Batty, M., Wang, S. and Corcoran, J. (2021), 'Modelling urban change with cellular automata: Contemporary issues and future research directions', *Progress in Human Geography* **45**(1), 3–24.

Liu, Y., Li, L., Chen, L., Cheng, L., Zhou, X., Cui, Y., Li, H. and Liu, W. (2019), 'Urban growth simulation in different scenarios using the sleuth model: A case study of hefei, east china', *PLoS One* **14**(11), e0224998.

Liu, Y. and Long, H. (2016), 'Land use transitions and their dynamic mechanism: The case of the huang-huai-hai plain', *Journal of Geographical Sciences* **26**, 515–530.

Long, H., Zhang, Y., Ma, L. and Tu, S. (2021), 'Land use transitions: Progress, challenges and prospects', *Land* **10**(9), 903.

Long, Y. and Zhang, Y. (2015), 'Land-use pattern scenario analysis using planner agents', *Environment and planning B: Planning and Design* **42**(4), 615–637.

Lorscheid, I., Heine, B.-O. and Meyer, M. (2012), 'Opening the 'black box'of simulations: increased transparency and effective communication through the systematic design of experiments', *Computational and Mathematical Organization Theory* **18**, 22–62.

Lu, Y. and Laffan, S. (2018), The use of particle swarm optimization for a vector cellular automata model of land use change (short paper), *in* '10th International Conference on Geographic Information Science (GIScience 2018)', Schloss Dagstuhl–Leibniz-Zentrum für Informatik, pp. 42–1.

Lu, Y., Laffan, S., Pettit, C. and Cao, M. (2020), 'Land use change simulation and analysis using a vector cellular automata (ca) model: A case study of ipswich city, queensland, australia', *Environment and Planning B: Urban Analytics and City Science* **47**(9), 1605–1621.

Manley, E. and Cheng, T. (2018), 'Exploring the role of spatial cognition in predicting urban traffic flow through agent-based modelling', *Transportation Research Part A: Policy and Practice* **109**, 14–23.

Meentemeyer, R. K., Tang, W., Dorning, M. A., Vogler, J. B., Cunniffe, N. J. and Shoemaker, D. A. (2013), 'Futures: multilevel simulations of emerging urban–rural landscape structure using a stochastic patch-growing algorithm', *Annals of the Association of American Geographers* **103**(4), 785–807.

Mialhe, F., Gunnell, Y., Navratil, O., Choi, D., Sovann, C., Lejot, J., Gaudou, B., Se, B. and Landon, N. (2019), 'Spatial growth of phnom penh, cambodia (1973–2015): Patterns, rates, and socio-ecological consequences', *Land Use Policy* **87**, 104061.

Miao, Y., Koenig, R., Buš, P., Chang, M.-C., Chirkin, A. and Treyer, L. (2017), 'Empowering urban design prototyping: a case study in cape town with interactive computational synthesis methods'.

Mithun, S., Parveen, S., Sahana, M. and Chattopadhyay, S. (2023), 'A systematic review on the application of geospatial technology and artificial intelligence in urban growth modeling', *Advancements in Urban Environmental Studies: Application of Geospatial Technology and Artificial Intelligence in Urban Studies* pp. 15–42.

Monroe, K. R. (2001), 'Paradigm shift: from rational choice to perspective', *International Political Science Review* **22**(2), 151–172.

Moradi, M., Roche, S. and Mostafavi, M. A. (2022), 'Exploring five indicators for the quality of openstreetmap road networks: A case study of québec, canada', *Geomatica* **75**(4), 178–208.

Morris, M. D. (1991), 'Factorial sampling plans for preliminary computational experiments', *technometrics* pp. 161–174.

Mozaffaree Pour, N. and Oja, T. (2021), 'Urban expansion simulated by integrated cellular automata and agent-based models; an example of tallinn, estonia', *Urban Science* **5**(4), 85.

Musa, S. I., Hashim, M. and Reba, M. N. M. (2017), 'A review of geospatial-based urban growth models and modelling initiatives', *Geocarto International* **32**(8), 813–833.

Mustafa, A., Cools, M., Saadi, I. and Teller, J. (2017), 'Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (huem)', *Land Use Policy* **69**, 529–540.

Naghibi, F. and Delavar, M. R. (2016), 'Discovery of transition rules for cellular automata using artificial bee colony and particle swarm optimization algorithms in urban growth modeling', *ISPRS international journal of geo-information* **5**(12), 241.

Nong, Y. and Du, Q. (2011), 'Urban growth pattern modeling using logistic regression', *Geo-spatial Information Science* **14**(1), 62–67.

OpenStreetMap contributors (2024), 'Openstreetmap data extract (roads, buildings, etc.)', https://www.geofabrik.de. Data extracted by Geofabrik, https://www.geofabrik.de.

Owaki, T. and Machida, T. (2020), Roadnetgan: Generating road networks in planar graph representation, *in* 'Neural Information Processing: 27th International Conference, ICONIP 2020, Bangkok, Thailand, November 18–22, 2020, Proceedings, Part IV 27', Springer, pp. 535–543.

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J. and Deadman, P. (2003), 'Multi-agent systems for the simulation of land-use and land-cover change: a review', *Annals of the association of American Geographers* **93**(2), 314–337.

Patel, M. H., Abbasi, M. A., Saeed, M. and Alam, S. J. (2018), 'A scheme to analyze agent-based social simulations using exploratory data mining techniques', *Complex Adaptive Systems Modeling* **6**, 1–17.

Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J. and Rombach, R. (2023), 'Sdxl: Improving latent diffusion models for high-resolution image synthesis', *arXiv preprint arXiv:2307.01952* .

Pontius Jr, R. G., Huffaker, D. and Denman, K. (2004), 'Useful techniques of validation for spatially explicit land-change models', *Ecological modelling* **179**(4), 445–461.

Pooyandeh, M. and Marceau, D. J. (2014), 'Incorporating bayesian learning in agent-based simulation of stakeholders' negotiation', *Computers, Environment and Urban Systems* **48**, 73–85.

Qiu, Y., Zhao, X., Fan, D., Li, S. and Zhao, Y. (2022), 'Disaggregating population data for assessing progress of sdgs: methods and applications', *International Journal of Digital Earth* **15**(1), 2–29.

Radchuk, V., Ims, R. A. and Andreassen, H. P. (2016), 'From individuals to population cycles: the role of extrinsic and intrinsic factors in rodent populations', *Ecology* **97**(3), 720–732.

Raimbault, J., Banos, A. and Doursat, R. (2016), 'A hybrid network/grid model of urban morphogenesis and optimization', *arXiv preprint arXiv:1612.08552* .

Rienow, A., Mustafa, A., Krelaus, L. and Lindner, C. (2021), 'Modeling urban regions: Comparing random forest and support vector machines for cellular automata', *Transactions in GIS* **25**(3), 1625–1645.

Runfola, D., Anderson, A., Baier, H., Crittenden, M., Dowker, E., Fuhrig, S., Goodman, S., Grimsley, G., Layko, R., Melville, G. et al. (2020), 'geoboundaries: A global database of political administrative boundaries', *PloS one* **15**(4), e0231866.

Saeedi, S. (2018), 'Integrating macro and micro scale approaches in the agent-based modeling of residential dynamics', *International journal of applied earth observation and geoinformation* **68**, 214–229.

Salem, M., Bose, A., Bashir, B., Basak, D., Roy, S., Chowdhury, I. R., Alsalman, A. and Tsurusaki, N. (2021), 'Urban expansion simulation based on various driving factors using a logistic regression model: Delhi as a case study', *Sustainability* **13**(19), 10805.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A. and Chen, X. (2016), 'Improved techniques for training gans', *Advances in neural information processing systems* **29**.

Saltelli, A. (2002), 'Making best use of model evaluations to compute sensitivity indices', *Computer physics communications* **145**(2), 280–297.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M. et al. (2004), *Sensitivity analysis in practice: a guide to assessing scientific models*, Vol. 1, Wiley Online Library.

Santé, I., García, A. M., Miranda, D. and Crecente, R. (2010), 'Cellular automata models for the simulation of real-world urban processes: A review and analysis', *Landscape and urban planning* **96**(2), 108–122.

Saxena, A. and Jat, M. K. (2020), 'Analysing performance of sleuth model calibration using brute force and genetic algorithm–based methods', *Geocarto International* **35**(3), 256–279.

Scheer, B. C. (2001), 'The anatomy of sprawl', *Places* **14**(2).

Sent, E.-M. (2018), 'Rationality and bounded rationality: You can't have one without the other', *The European Journal of the History of Economic Thought* **25**(6), 1370–1386.

Shan, Y.-h. and Zhu, X.-y. (2011), 'Scenario analysis of urban residential land use utility based on multi-agents' spatial decision', *JOURNAL OF NATURAL RESOURCES* **26**(11), 1832–1841.

Sharma, S. N. (2019), 'Review of most used urban growth models', *International Journal of Advanced Research in Engineering and Technology (IJARET)* **10**(3), 397–405.

Shou, X., Chen, P. and Zheng, H. (2021), Predicting the heat map of street vendors from pedestrian flow through machine learning, *in* '26th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2021)', The Association for Computer-Aided Architectural Design Research in Asia …, pp. 569–578.

Silva, E. A. and Clarke, K. C. (2002), 'Calibration of the sleuth urban growth model for lisbon and porto, portugal', *Computers, environment and urban systems* **26**(6), 525–552.

Simon, H. A. (1956), 'Rational choice and the structure of the environment.', *Psychological review* **63**(2), 129.

Sirko, W., Kashubin, S., Ritter, M., Annkah, A., Bouchareb, Y. S. E., Dauphin, Y., Keysers, D., Neumann, M., Cisse, M. and Quinn, J. (2021), 'Continental-scale building detection from high resolution satellite imagery', *arXiv preprint arXiv:2107.12283* .

Sobol, I. M. (2001), 'Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates', *Mathematics and computers in simulation* **55**(1-3), 271–280.

Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. and Ganguli, S. (2015), Deep unsupervised learning using nonequilibrium thermodynamics, *in* 'International conference on machine learning', PMLR, pp. 2256–2265.

Song, A. and Whitehead, J. (2019), Townsim: Agent-based city evolution for naturalistic road network generation, *in* 'Proceedings of the 14th International Conference on the Foundations of Digital Games', pp. 1–9.

Srinurak, N., Mishima, N., Fuchikami, T. and Duangthima, W. (2016), 'Analysis of urban morphology and accessibility character to provide evacuation route in historic area', *Procedia-Social and Behavioral Sciences* **216**, 460–469.

Sulistiyo, M. D., Kawanishi, Y., Deguchi, D., Hirayama, T., Ide, I., Zheng, J. and Murase, H. (2018), Attribute-aware semantic segmentation of road scenes for understanding pedestrian orientations, *in* '2018 21st International Conference on Intelligent Transportation Systems (ITSC)', IEEE, pp. 2698–2703.

Taillandier, P., Banos, A., Drogoul, A., Gaudou, B., Marilleau, N. and Truong, Q. (2016), Simulating urban growth with raster and vector models: A case study for the city of can tho, vietnam, *in* 'International Conference on Autonomous Agents and Multiagent Systems', pp. pp.154–171.

Taillandier, P., Gaudou, B., Grignard, A., Huynh, Q.-N., Marilleau, N., Caillou, P., Philippon, D. and Drogoul, A. (2019), 'Building, composing and experimenting complex spatial models with the gama platform', *GeoInformatica* **23**, 299–322.

Tang, W. and Jia, M. (2014), 'Global sensitivity analysis of a large agent-based model of spatial opinion exchange: A heterogeneous multi-gpu acceleration approach', *Annals of the Association of American Geographers* **104**(3), 485–509.

Tang, W. and Yang, J. (2020), 'Agent-based land change modeling of a large watershed: Space-time locations of critical threshold', *Journal of Artificial Societies and Social Simulation* **23**(1).

Thanh Son, N., Thi Thu Trang, N., Bui, X. T. and Thi Da, C. (2022), 'Remote sensing and gis for urbanization and flood risk assessment in phnom penh, cambodia', *Geocarto International* **37**(22), 6625–6642.

Tian, W. (2013), 'A review of sensitivity analysis methods in building energy analysis', *Renewable and sustainable energy reviews* **20**, 411–419.

Tong, X. and Feng, Y. (2020), 'A review of assessment methods for cellular automata models of land-use change and urban growth', *International Journal of Geographical Information Science* **34**(5), 866–898.

Tsagkis, P., Bakogiannis, E. and Nikitas, A. (2023), 'Analysing urban growth using machine learning and open data: An artificial neural network modelled case study of five greek cities', *Sustainable Cities and Society* **89**, 104337.

Udom, D. S., Murg, B. J., Virak, O. and Renfrew, M., eds (2020), *Cambodia 2040: Economic Development*, Future Forum, Phnom Penh, Cambodia.

UN (2024), *The Revision of World Population Prospects*, United Nations, New York, NY.

van Vliet, J., Bregt, A. K., Brown, D. G., van Delden, H., Heckbert, S. and Verburg, P. H. (2016), 'A review of current calibration and validation practices in land-change modeling', *Environmental Modelling & Software* **82**, 174–182.

van Vliet, J., Bregt, A. K. and Hagen-Zanker, A. (2011), 'Revisiting kappa to account for change in the accuracy assessment of land-use change models', *Ecological modelling* **222**(8), 1367–1375.

van Vliet, J., Hagen-Zanker, A., Hurkens, J. and van Delden, H. (2013), 'A fuzzy set approach to assess the predictive accuracy of land use simulations', *Ecological Modelling* **261**, 32–42.

Van Vuuren, D. P., Kriegler, E., O'Neill, B. C., Ebi, K. L., Riahi, K., Carter, T. R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R. et al. (2014), 'A new scenario framework for climate change research: scenario matrix architecture', *Climatic change* **122**, 373–386.

Visser, H. and De Nijs, T. (2006), 'The map comparison kit', *Environmental Modelling & Software* **21**(3), 346–358.

von Neumann, J. and Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

Wang, J., Li, G., Lu, H. and Wu, Z. (2024), 'Urban models: Progress and perspective', *Sustainable Futures* p. 100181.

Wang, L., Omrani, H., Zhao, Z., Francomano, D., Li, K. and Pijanowski, B. (2019), 'Analysis on urban densification dynamics and future modes in southeastern wisconsin, usa', *PloS one* **14**(3), e0211964.

Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J. and Catanzaro, B. (2018), High-resolution image synthesis and semantic manipulation with conditional gans, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 8798–8807.

Wiedenhofer, D., Smetschka, B., Akenji, L., Jalas, M. and Haberl, H. (2018), 'Household time use, carbon footprints, and urban form: a review of the potential contributions of everyday living to the 1.5 c climate target', *Current opinion in environmental sustainability* **30**, 7–17.

Willmott, C. J. and Matsuura, K. (2005), 'Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance', *Climate research* **30**(1), 79–82.

World Bank (2024), *Urban Development in Phnom Penh*, World Bank Group, Washington, D.C.

Wu, A. N. and Biljecki, F. (2022), 'Ganmapper: geographical data translation', *International Journal of Geographical Information Science* **36**(7), 1394–1422.

Wu, A. N. and Biljecki, F. (2023), 'Instantcity: Synthesising morphologically accurate geospatial data for urban form analysis, transfer, and quality control', *ISPRS Journal of Photogrammetry and Remote Sensing* **195**, 90–104.

Wu, A. N., Stouffs, R. and Biljecki, F. (2022), 'Generative adversarial networks in the built environment: A comprehensive review of the application of gans across data types and scales', *Building and Environment* **223**, 109477.

Wu, F. (2002), ‘Calibration of stochastic cellular automata: The application to rural-urban land conversions.’, *International journal of geographical information science* **16**(8).

Xing, W., Qian, Y., Guan, X., Yang, T. and Wu, H. (2020), ‘A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation’, *Computers & Geosciences* **137**, 104430.

Xu, T., Gao, J., Coco, G. and Wang, S. (2020), ‘Urban expansion in auckland, new zealand: A gis simulation via an intelligent self-adapting multiscale agent-based model’, *International Journal of Geographical Information Science* **34**(11), 2136–2159.

Yaagoubi, R., Lakber, C.-E. and Miky, Y. (2024), ‘A comparative analysis on the use of a cellular automata markov chain versus a convolutional lstm model in forecasting urban growth using sentinel 2a images’, *Journal of Land Use Science* **19**(1), 258–277.

Yang, J., Tang, W., Gong, J., Shi, R., Zheng, M. and Dai, Y. (2023a), ‘Simulating urban expansion using cellular automata model with spatiotemporally explicit representation of urban demand’, *Landscape and Urban Planning* **231**, 104640.

Yang, L., Li, L., Chen, Q., Zhang, J., Feng, T. and Zhang, W. (2023b), ‘Street layout design via conditional adversarial learning’, *arXiv preprint arXiv:2305.08186* .

Yao, Y., Jiang, Y., Sun, Z., Li, L., Chen, D., Xiong, K., Dong, A., Cheng, T., Zhang, H., Liang, X. et al. (2024), ‘Applicability and sensitivity analysis of vector cellular automata model for land cover change’, *Computers, Environment and Urban Systems* **109**, 102090.

Yasin, M. Y., Yusoff, M. M., Abdullah, J. and Noor, N. M. (2020), ‘Is urban sprawl a threat to sustainable development? a review of characteristics and consequences’, *Geografia* **16**(4).

Zhai, Y., Yao, Y., Guan, Q., Liang, X., Li, X., Pan, Y., Yue, H., Yuan, Z. and Zhou, J. (2020), ‘Simulating urban land use change by integrating a convolutional neural network with vector-based cellular automata’, *International Journal of Geographical Information Science* **34**(7), 1475–1499.

Zhang, X. Q. (2016), ‘The trends, promises and challenges of urbanisation in the world’, *Habitat international* **54**, 241–252.

Zhang, Z., Li, M. and Yu, J. (2018), On the convergence and mode collapse of gan, *in* ‘SIGGRAPH Asia 2018 Technical Briefs’, pp. 1–4.

Zhou, F., Li, H., Hu, R., Wu, S., Feng, H., Du, Z. and Xu, L. (2024), ‘Controlcity: A multimodal diffusion model based approach for accurate geospatial data generation and urban morphology analysis’, *arXiv preprint arXiv:2409.17049* .