



From heat resilience to sustainable co-benefits: Adaptive urban morphology generation based on multimodal data fusion and a novel generative framework

Shiqi Zhou^{a,g}, Xiaodong Xu^{b,*}, Haowen Xu^c, Zichen Zhao^b, Haojun Yuan^{d,*},
Yuankai Wang^{e,*}, Renlu Qiao^b, Tao Wu^c, Weiyi Jia^c, Mo Wang^f, Waishan Qiu^e,
Zhiqiang Wu^c

^a College of Design and Innovation, Tongji University, Shanghai 200093, China

^b Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, 200093, China

^c College of Architecture and Urban Planning, Tongji University, Shanghai 200093, China

^d College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan, 430070, China

^e Department of Urban Planning and Design, The University of Hong Kong, Pok Fu Lam, Hong Kong, China

^f College of Architecture and Urban Planning, Guangzhou University, Guangzhou 510006, China

^g HAI Design Lab, College of Design and Innovation, Tongji University, Shanghai, 20093, China

ARTICLE INFO

Keywords:

Thermal environment
Urban 3D morphology
End-to-end generative design
Generative adversarial network
Local climate zones
Computational urban design

ABSTRACT

Rapid urbanization and global climate change have intensified the Urban Heat Island (UHI) effect. However, practical implementation is often constrained by limitations in data availability and computational capacity, overlooking the influence of socioeconomic factors and spatial heterogeneity. This study proposed an end-to-end urban 3D morphology generation framework that leveraged multimodal datasets, including Local Climate Zones (LCZ), Land Surface Temperature (LST), and Population Density (POPH) through a novel CycleGAN-Pix2pix (CP-GAN) model chain. Using six representative LCZ areas in Guangzhou as case studies, the research evaluated the Urban Morphology Indicators (UMI), Land Use and Land Cover Change (LUCC), and Points of Interest (POI) across various responsive generation scenarios to identify urban morphologies that balanced cooling effects with socioeconomic and ecological benefits. The results showed that: (1) The CP-GAN achieved robust performance in urban morphology generation, demonstrating stable convergence and high precision, with an average structural similarity index exceeding 0.811, along with high signal-to-noise ratios and low error metrics. (2) Rising temperatures reshaped urban morphology, with every 3°C increase reducing green space by 5.47% while raising commercial activity and impervious surfaces by 2.38% and 2.84%, respectively; (3) Population density drove POI clustering but exhibited weaker morphological control than temperature gradients. (4) LCZ4, LCZ5, and LCZ6 exhibited spatial heterogeneity in UMI, LUCC, and POI responses to temperature and population density variations, necessitating LCZ-specific adaptive strategies. This generative system offers fine-grained 3D morphological solutions to mitigate UHI effects while establishing a transformative framework for sustainable urban development.

Abbreviations

CGANs Conditional Generative Adversarial Networks
CP-GAN CycleGAN-Pix2pix Generative Adversarial Network
FID Fréchet Inception Distance

GANs Generative Adversarial Networks
KID Kernel Inception Distance
LCZ Local Climate Zone
LST Land Surface Temperature
LUCC Land Use and Land Cover Change

* Corresponding authors.

E-mail addresses: zhoushiqi@tongji.edu.cn (S. Zhou), xuxiaodong0901@tongji.edu.cn (X. Xu), tjupxhw@tongji.edu.cn (H. Xu), 2411975@tongji.edu.cn (Z. Zhao), haojuny1202@outlook.com (H. Yuan), u3011546@connect.hku.hk (Y. Wang), qiaorl@tju.edu.cn (R. Qiao), diaohaifeng@tongji.edu.cn (T. Wu), jwy@tongji.edu.cn (W. Jia), landwangmo@outlook.com (M. Wang), waishanq@hku.hk (W. Qiu), wus@tongji.edu.cn (Z. Wu).

<https://doi.org/10.1016/j.scs.2025.106452>

Received 18 February 2025; Received in revised form 21 April 2025; Accepted 12 May 2025

Available online 13 May 2025

2210-6707/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

| | |
|---------|--|
| MSE | Mean Square Error |
| NSGA-II | Non-dominated Sorting Genetic Algorithm II |
| POI | Points of Interest |
| POPH | Population Heat (Population Density) |
| PSNR | Peak Signal-to-Noise Ratio |
| SSIM | Structural Similarity Index |
| UHI | Urban Heat Island |
| UMI | Urban Morphology Indicators |
| VAE | Variational Autoencoder |

1. Introduction

Global climate change has increased the frequency and intensity of extreme weather events, particularly urban heat island (UHI), significantly amplifying thermal stress in urban areas. This phenomenon is further exacerbated by accelerated global urbanization, characterized by uncontrolled spatial expansion and rapid population growth. The combined effects of these factors intensify UHI effects, posing critical challenges to urban sustainability and resilience (Gu et al., 2024). Existing research highlighted the strong correlation between urban morphology and environmental conditions (Su et al., 2022; Xu et al., 2024). Besides, given the localized nature of urban thermal dynamics, the block scale provides a more granular perspective for understanding the spatial heterogeneity of heat distribution and the interactions between urban morphology and microclimate. These findings underscore the urgent need for intelligent, dynamically responsive urban block spatial configurations to address the complex impacts of environmental and population changes effectively.

To assess how variations in urban morphology influence thermal conditions quantitatively, researchers have employed a variety of analytical frameworks and classification methods. Among these, Urban Structure Types (UST), Urban Climate Zones (UCZ) and Global Urban Footprint (GUF) stand out as key tools for examining the spatial and structural diversity of urban areas and their effects on local thermal patterns (Lee et al., 2019; Heiden et al., 2012). However, these methods often fell short in practical applications due to limitations in climate relevance, regional specificity, and scale adaptation (Qiu et al., 2018; Lee et al., 2019). In contrast, the Local Climate Zone (LCZ) system, proposed by Stewart and Oke in 2012, offers a more robust and effective framework for such evaluations (Stewart & Oke, 2012). The LCZ framework effectively captures urban local climates. These climates are shaped by the interaction of local weather patterns, natural geographic features, and the three-dimensional (3D) structure of cities. Given its moderate spatial scale and detailed representation of building structures, the LCZ is widely regarded as an ideal unit for implementing urban morphology optimization strategies to mitigate the UHI effect (Zheng et al., 2018; Rodler & Leduc, 2019). This capability lays a solid foundation for quantitative analysis of the relationship between urban morphological heterogeneity and environmental impacts.

In recent years, the integration of multimodal data has gained increasing attention in urban environmental monitoring and planning, emerging as a critical tool for addressing urban heat environments and dynamic population changes (Dalla Mura et al., 2015; De et al., 2021; Roche et al., 2022). Satellite remote sensing data, with its extensive coverage and high spatiotemporal resolution, effectively captures LCZ and Land Surface Temperature (LST), revealing spatial patterns of urban heat environments. However, these static environmental datasets have limitations in capturing the dynamic interactions between urban microclimates and human activities. To address this gap, mobile signaling data has been introduced as an effective solution. By providing real-time, high-frequency information on population density and activity patterns, mobile signaling data enhances the ability to accurately understand the influence of population distribution on urban heat environments, offering a more comprehensive perspective on the interplay between social behavior and microclimatic conditions.

Current research primarily focused on the impact of urban

morphology on environmental temperatures, often analyzing existing urban environments to diagnose current conditions and summarize observed patterns in urban thermal characteristics. Studies consistently highlighted the significant influence of 3D urban morphology on urban thermal environments. For instance, Cao et al. (2021) demonstrated that in Beijing, 3D morphological metrics explained over 60% of temperature variations across different scales and seasons. Similarly, Qiao et al. (2020) found that 3D morphology, particularly building volume and height, significantly enhanced the explanatory power for the urban heat island effect at larger spatial scales. Hu et al. (2020) revealed that while 2D land-use variables were most strongly correlated with LST during summer and spring, 3D building-related variables had a greater impact in colder weather. Despite these findings, challenges remain in analyzing spatial heterogeneity because of the difficulty of obtaining high-resolution data and the substantial computational resources required for dynamic simulations and modeling. Moreover, current studies were often limited to retrospective analyses rather than forward-looking solutions. Retrospective approaches focus on proposing corrective measures after planning implementation, offering limited efficacy. In contrast, prospective solutions involve the evaluation and refinement of design plans prior to implementation, facilitating the identification of optimal strategies and yielding significantly greater impacts. These limitations hinder comprehensive studies of the dynamic evolution of urban morphology but simultaneously underscore critical directions for future urban planning research. Addressing these gaps requires integrating multimodal data and innovative generative technologies to enhance design performance. Additionally, directly simulating dynamic changes in urban morphology by modifying environmental factors offers a promising avenue for advancing the field.

In recent years, research on automated building morphology generation has grown rapidly. From a methodological perspective, most existing studies relied on rule-based approaches and heuristic algorithms to optimize environmental performance. Rule-based generation methods and evaluation tools, such as shape grammar, automated site planning (ASP) models, Non-dominated Sorting Genetic Algorithm II (NSGA-II) models, and cellular automata (CA) models, iteratively generate building configurations based on predefined rules (Zhou et al., 2025). For example, Sun and Rao (2020) optimized building layouts in a Shanghai neighborhood using the Grasshopper plugin, focusing on building height restrictions, plot utilization, and solar radiation assessments. Their results showed a 35% improvement in performance for the optimal solution compared to the least favorable one. Similarly, Boyukliyski et al. (2022) applied genetic algorithms to optimize urban building areas and green space configurations while accounting for population demands, significantly improving urban planning efficiency. However, rule-based generation methods have notable limitations. First, defining effective rules requires extensive expert knowledge and design experience. Second, the generated design solutions often struggle to balance multiple objectives, such as urban aesthetic and environmental benefits, leading to reduced diversity and practical applicability. Besides, the performance of these algorithms heavily depends on parameter optimization and trade-offs between objectives, resulting in high computational costs and limited generalizability (Zhou et al., 2022).

With the growing application of deep generative methods in design, GAN-based approaches have emerged as one of the most prominent technologies. Unlike rule-based modeling, learning-based methods do not rely on extensive expert knowledge or design experience. They use neural networks to automatically learn connections between image features and urban morphology. This process is based on real-world design cases. Huang et al. (2022) proposed a performance-driven urban design process that combines a GAN-based model with multi-objective algorithms. Similarly, Wu and Biljecki (2023) created a GAN-based model to generate urban morphologies. The model was applied to 16 cities. It achieved high accuracy in simulating ground-truth urban forms. However, most GAN-based studies primarily focused on learning typical urban morphological patterns from sample

data, often failing to incorporate environmental requirements and socioeconomic contexts into the design process. This limitation reduces their practical applicability. These gaps underscore the need for a unified framework that integrates morphology generation with socio-economic and ecological development goals, ensuring adaptability to diverse planning and design requirements.

Despite significant advances in understanding the interactions between urban morphology and environmental conditions, existing studies face critical limitations in their practical application to urban planning and design. These limitations include: (1) insufficient integration of multimodal data; (2) the lack of performance-driven, rather than rule-based, three-dimensional urban morphology generation models; and (3) the lack of end-to-end frameworks that replace retrospective approaches with proactive design strategies. To address these research gaps, this study introduces an improved GAN-based model, designed to automatically generate 3D urban morphology solutions that mitigate the UHI effect while ensuring socioeconomic benefits. The model functioned as the core of an end-to-end generative system, producing environmentally compatible and semantically coherent urban layouts based on integrated preset conditions. The specific objectives of this study include: (1) Integrating static datasets (LCZ, LST) and dynamic datasets (POPH) to construct a multimodal dataset; (2) Developing an automated CP-GAN-based 3D morphology generation model achieving controlled cooling effects and ensuring socioeconomic benefits; (3) Establishing a forward-looking framework to evaluate how land-use patterns and urban functionalities in response to thermal environment and demographic dynamic changes; This study seeks to deliver practical solutions for sustainable development in high-density urban environments by balancing tradeoffs among economic, social and ecological development, ultimately enhancing urban adaptability to environmental changes.

2. Methodology and materials

This study proposed an automated generative framework based on the CP-GAN model for targeted creation of urban morphologies with cooling and socioeconomic benefits. This study was organized into four primary stages, as illustrated in Fig. 1: (1) Collecting satellite remote sensing data and mobile signaling data to construct a multimodal dataset encompassing LCZ, LST, and POPH data. (2) Constructing a novel CP-GAN model by integrating CycleGAN and Pix2pix allowed for dynamic 3D urban morphology generation; (3) Developing a proactive analysis framework that controlled either LST or POPH to evaluate their respective impacts on Urban Morphology Indicators (UMIs), Points of Interest (POI) and Land Use and Cover Change (LUCC); (4) Establishing a multivariate analysis framework, where both temperature and population density were simultaneously changed to evaluate the resulting influence on UMI, LUCC and POI. This study employed the Grasshopper plugin within Rhino software to automate the generation of 3D urban morphology, leveraging 2D images produced by generative adversarial networks (GANs) as the foundational input. The CP-GAN framework as an end-to-end generative system, integrates site conditions and planning constraints to produce design solutions tailored to user needs. Additionally, it evaluates the performance of urban sustainability strategies, enabling informed decision-making during the early design stages.

2.1. Study area

Guangzhou as the study area for this research, is located between 22°26'–23°56' N latitude and 112°57'–114°3' E longitude (Fig. 2). Guangzhou is the capital and megacity of Guangdong Province, China, with an area of about 7434.4 km². With more than 18.73 million residents, Guangzhou is one of the most densely populated cities in the world, with about 2,520 individuals per km². Guangzhou's climate falls under the Cwa category in the Köppen-Geiger classification. This means it has a subtropical monsoon climate. In 2022, the average annual

temperature was 23.1°C, with an average summer temperature of 29.2°C, which is higher than the average level of the whole province, making it one of the famous “Stove Cities” in China. The unique combination of high population density, complex urban morphology, and significant climate challenges makes Guangzhou a prime example for studying the impacts of urban thermal environments and exploring strategies to improve thermal performance in subtropical urban areas. Based on the categories of LCZs, six types of LCZ sites were selected as study areas, representing the typical LCZs in Guangzhou's high-density urban built environment. The locations of these six locations are shown in Fig. 2. These six locations represent LCZ1 (GZ01), LCZ2 (GZ02), LCZ4 (GZ03), LCZ5 (GZ04), LCZ6 (GZ05), and LCZ7 (GZ06), which have different architectural forms and functional characteristics.

2.2. Preparation of multimodal data

Multimodal data refers to the integration of diverse data sources collected through various sensing mechanisms, enabling a more comprehensive analysis of complex urban phenomena. This study leveraged a multimodal data fusion approach to deepen the understanding of microclimate characteristics and population dynamics within high-density urban environments. The dataset utilized in this research comprised three primary data types: LCZ, LST, and POPH.

LCZ and LST data, derived from satellite remote sensing, serve to characterize urban areas in terms of building density, land cover, and temperature distribution. These datasets provide a macro-level, static view of the urban environment, essential for analyzing spatial patterns related to urban heat islands and morphological structures. In contrast, POPH are generated from mobile signaling data, which capture high-resolution temporal changes in population density and movement patterns. This dynamic data offers insights into how human mobility influences the thermal landscape of urban areas. By integrating these diverse data modalities, this study can more precisely identify heat-sensitive urban zones and analyze their interactions with human activities. The complementary nature of static environmental metrics from remote sensing and dynamic social behavior data from mobile signaling enhances the accuracy of predictive models.

In addition, it is important to highlight that the datasets used in this study—LCZ, LST, and POPH—are relatively accessible. Both LCZ and LST data are available globally through public remote sensing platforms, enabling the application of this methodology to a wide range of geographic contexts. In cases where population heatmap (POPH) data might be incomplete or unavailable, alternative proxies, such as nighttime light data, can be utilized as substitutes to reflect urban vitality. While data imputation methods could be applied to address incomplete datasets, these methods might introduce biases or inconsistencies that could affect the final results. The simultaneous availability of all three data channels is optimal, as each offers distinct and complementary information essential for the analysis. The absence of any one channel would lead to gaps in the corresponding results components of the analysis, thereby limiting the comprehensiveness and accuracy of the final results.

2.2.1. Acquisition of LCZ data

This study used LCZ maps generated by the advanced LCZ map generator. The tool is provided by the World Urban Database and Access Portal Tools (WUDAPT, <https://www.wudapt.org/>). WUDAPT, structured around the LCZ framework, provided a consistent method for mapping urban morphology to support climate and environmental research at a global scale (Bechtel et al., 2019). This research employed WUDAPT's Level 0 (L0) tool to produce detailed LCZ maps of Guangzhou. More than 20 training samples were selected for each LCZ category utilizing Google Earth imagery. Building heights and distances between buildings were measured using the ruler tool in Google Earth. This was done during the sample selection process. Local experts identified 17 types of Local Climate Zones (LCZs) in Guangzhou. These data were

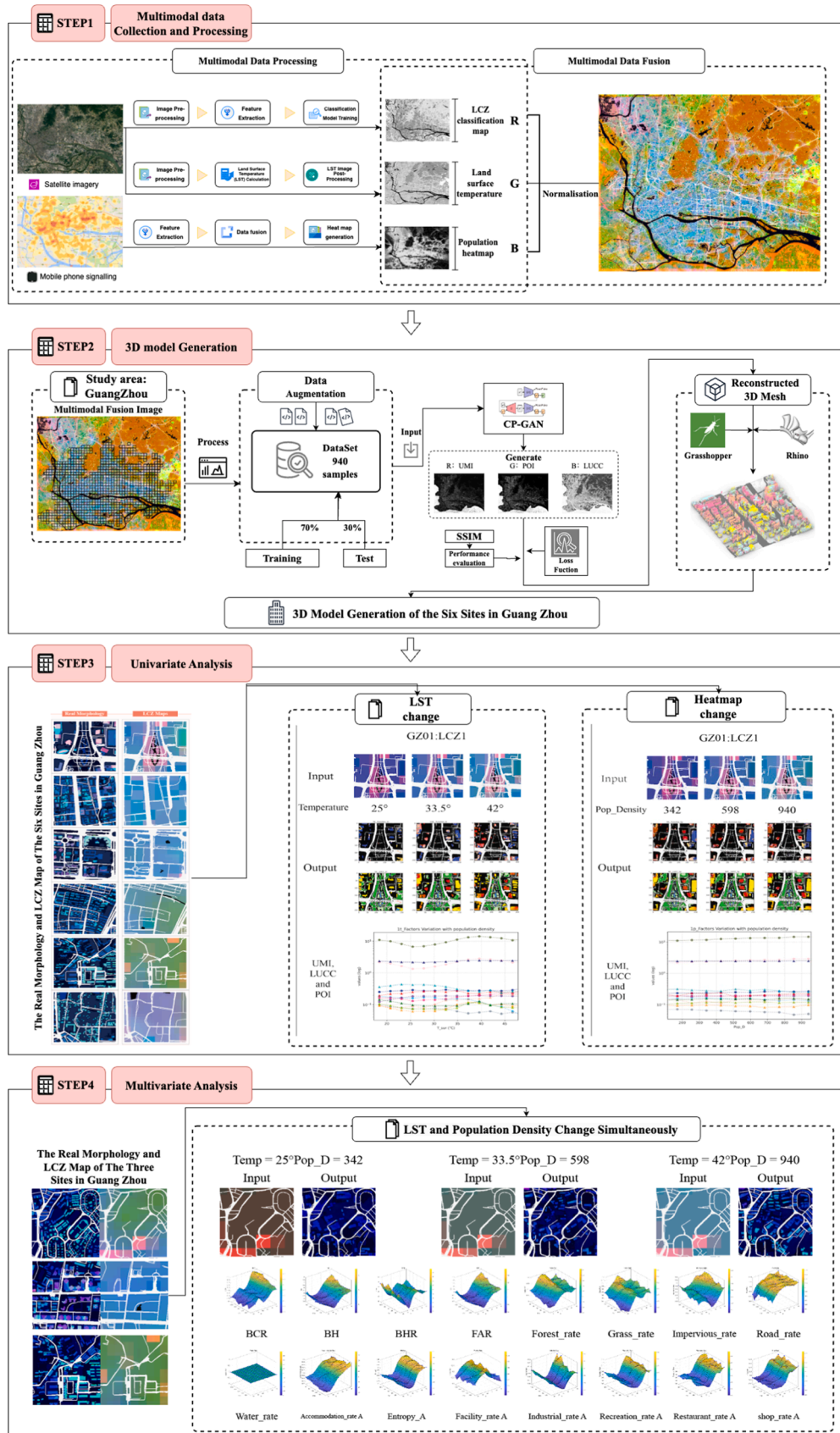


Fig. 1. The workflow of this research. (The relevant data images for the data processing steps are provided in Appendix B for further reference).

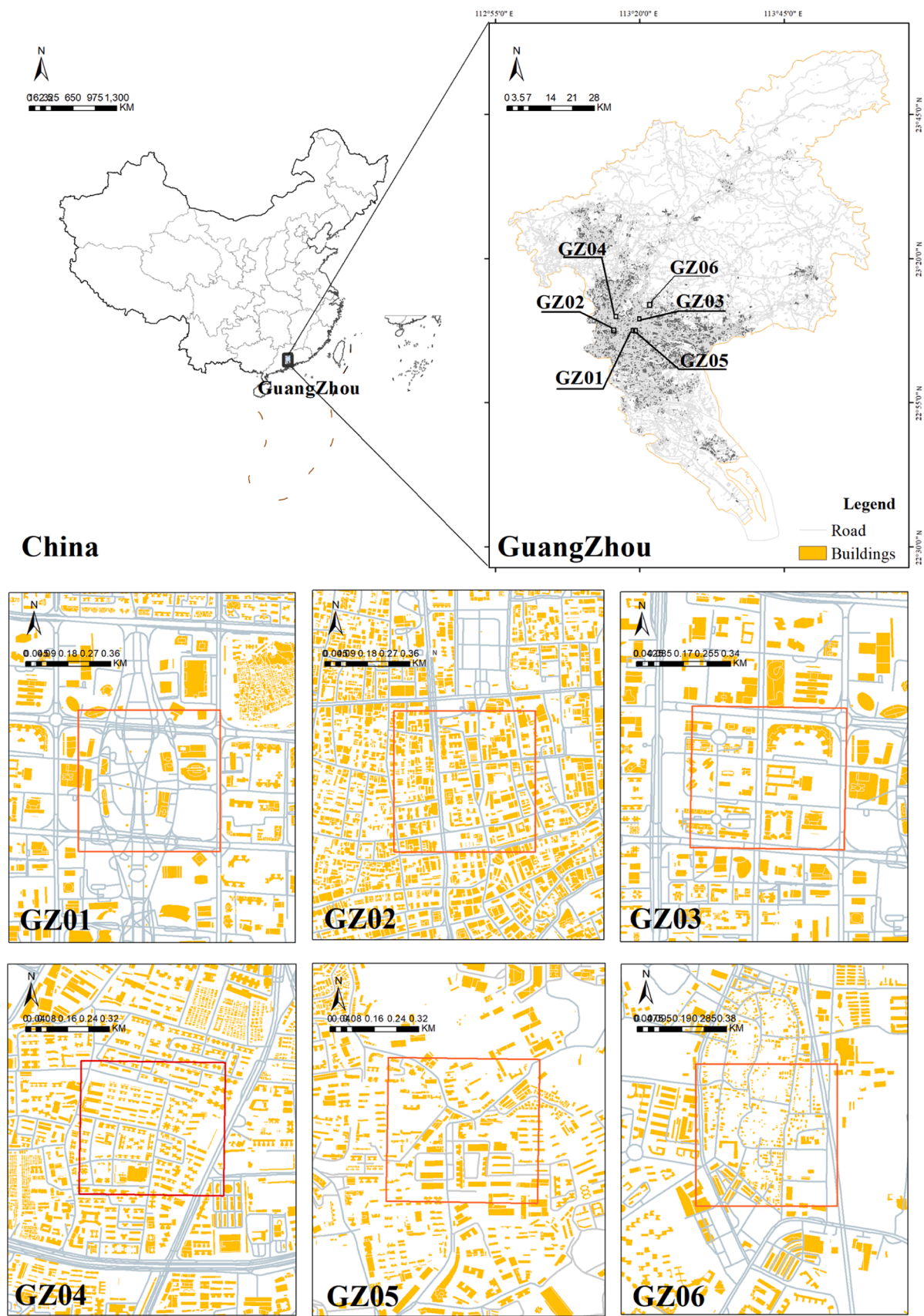


Fig. 2. Locations of Guangzhou and the selected six typical sites.

processed using the WUDAPT Level 0 LCZ generator and classified with an online random forest algorithm at a spatial resolution of 100 meters. To improve sample quality, the training process underwent several iterations until it achieved satisfactory overall accuracy. This effort successfully produced a detailed LCZ map of Guangzhou, illustrated in Fig. 3(a). The analysis showed that vegetated LCZs (LCZ A–D) dominated the land cover, particularly in the Hua Du and Cong Hua districts, while urbanized areas primarily corresponded to LCZ types 1–7. Given the study's emphasis on high-density urban regions, non-urban areas were excluded, and the analysis focused on a key urban center for model training (Fig. 3b).

For a detailed investigation, this study selected six representative sites within the urban core, each defined by a grid of 600m x 600m units. These sites predominantly consisted of LCZs 1–7, with LCZ 3 excluded due to its negligible presence in Guangzhou's urban fabric. The site GZ01 encompassed a mix of commercial and residential areas, primarily categorized as LCZ 1. GZ02 was mostly residential, featuring a blend of LCZ 2 with some commercial elements. GZ03 represented a typical residential community with various services and was dominated by LCZ 4. GZ04 served as a multifunctional zone, combining educational, residential, and commercial uses, predominantly falling under LCZ 5, with some areas classified as LCZ 3. GZ05 was largely a residential area, primarily consisting of LCZ 6. Lastly, GZ06 comprised a mix of older residential buildings and industrial structures, falling mainly within LCZ 7. LCZ maps for these six sites were extracted from the Guangzhou LCZ dataset to facilitate model inference (Fig. 4). The final column represented the POI attributes of each 3D building. In the planar maps, POI attributes were depicted as channel values; however, directly converting these values into 3D models often resulted in suboptimal visual quality. To address this issue and enhance visualization, this study reassigned colors to POI attributes in the 3D morphology. Detailed color information is provided in Table 1.

2.2.2. Acquisition of LST data

This study used the Landsat Collection 2 Science Products Surface Temperature dataset. The dataset was obtained from the United States Geological Survey (USGS, 2013). It is based on U.S. Landsat Analysis Ready Data (USGS, 2023). The data was processed using high scientific standards and derived from remote sensing techniques. It is commonly used for landscape monitoring and change assessment. The dataset provided surface temperature measurements in Kelvin at a 30-meter spatial resolution using the World Geodetic System 1984 (WGS84) datum. However, due to limitations in data availability, nighttime LST

measurements were not included, meaning the analysis in this study focuses solely on daytime LST values.

The process of generating the LST map for Guangzhou is illustrated in Fig. 5. This study first accessed raw LST data from 2023 using the USGS EarthExplorer platform, matching it with built environment datasets. To ensure both the relevance for environmental design and the accuracy, this study only selected LST maps from the summer months (June–September) with minimal cloud cover (less than 1 percent), cloud shadow, and snow or ice coverage. Four maps were obtained, representing each summer month, and the seasonal average LST was determined using the Raster Calculator tool in ArcGIS. The map was clipped to the Guangzhou borough boundaries, and the final output illustrated the average daytime summer LST for Guangzhou. Areas of higher gray value, appearing whiter on the map, indicated higher LST levels. Water bodies outside the main Guangzhou region were excluded from the map due to the study's focus.

2.2.3. Acquisition of population heat map (POPH)

Baidu Heatmap uses real-time geolocation data from smartphone users of Baidu products. It calculates pedestrian activity in various urban areas. This tool captures the dynamic distribution of populations in urban spaces (Fan et al., 2021; Li et al., 2019; Lyu & Zhang, 2019). Therefore, this study used Baidu Heatmap to represent the real vitality values of neighborhoods. Baidu Heatmap data was collected from Guangzhou on 19 to 25 May 2023 (weekdays), avoiding the effects of holidays and extreme hot and cold weather. This study collected heatmap data every 2 hours between 8:00 and 22:00. A total of 40 heatmaps with a spatial resolution of 3.5 m were analyzed. The real vitality value was calculated as the average across different periods on weekdays.

2.2.4. Data encoding

To facilitate the integration of image inputs into surrogate models, we performed encoding on LCZ classification and urban morphology images. This process is essential for preparing the data and establishing the required representations in the surrogate models. The 16 different LCZ categories are represented using distinct colors based on RGB encoding. LUCC, which includes building outlines and vegetation boundaries, is encoded into grayscale images. Additionally, data such as POPH, POI, building height, and LST are normalized and then linearly mapped into the RGB channel value range. Consequently, the input RGB channels represent LCZ, LST, and POPH, while the output channels correspond to building height, POI, and LUCC. The dataset consists of 940 samples, with 70% allocated for training and the remaining 30% for

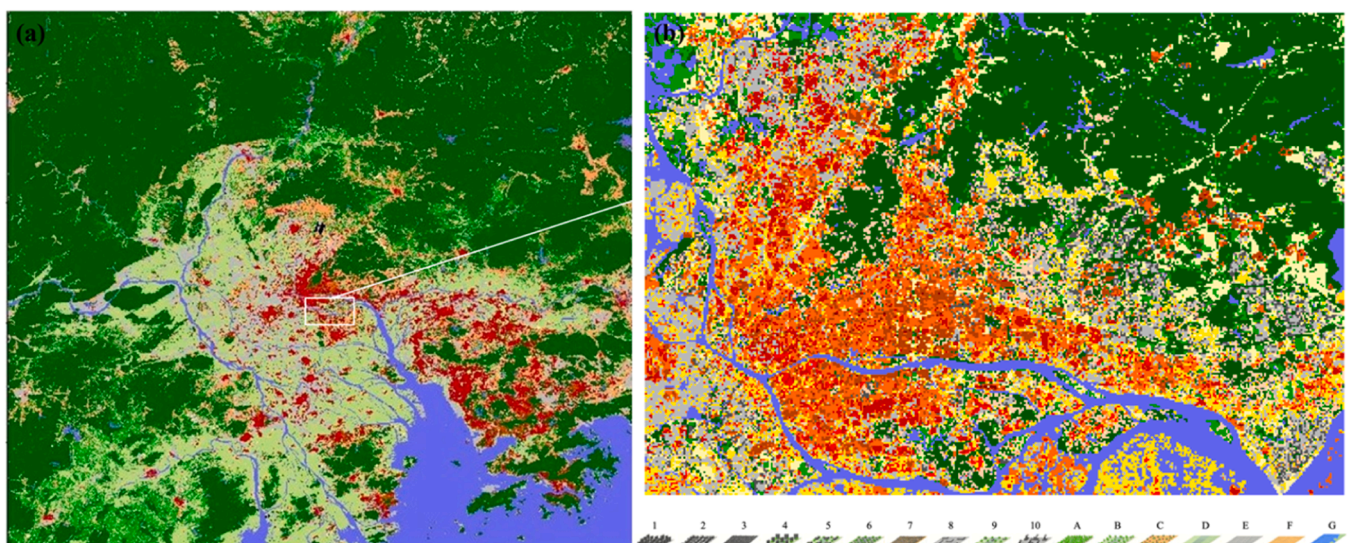


Fig. 3. LCZs map of Guang Zhou.

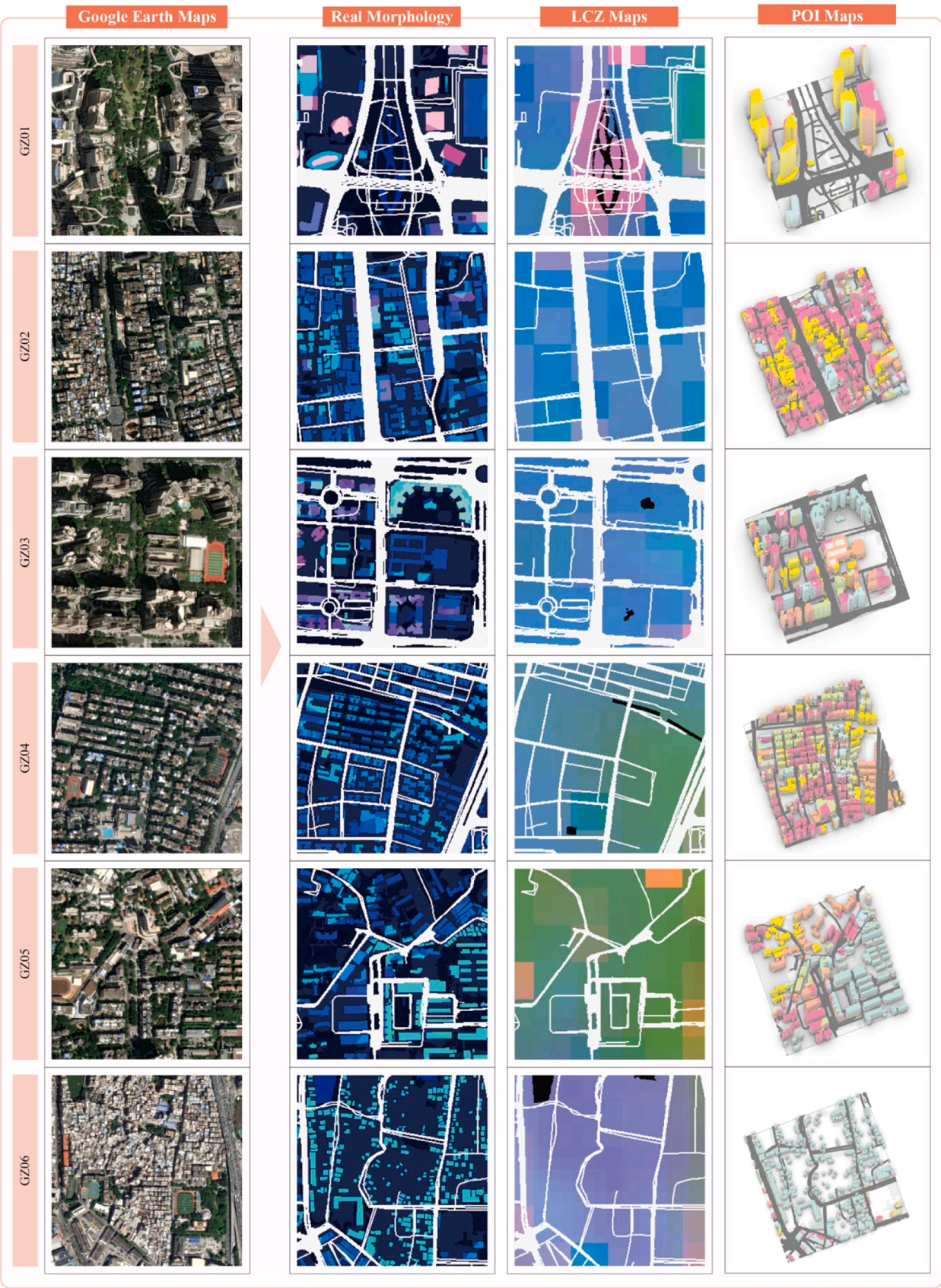



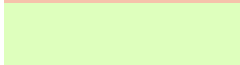
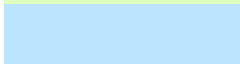

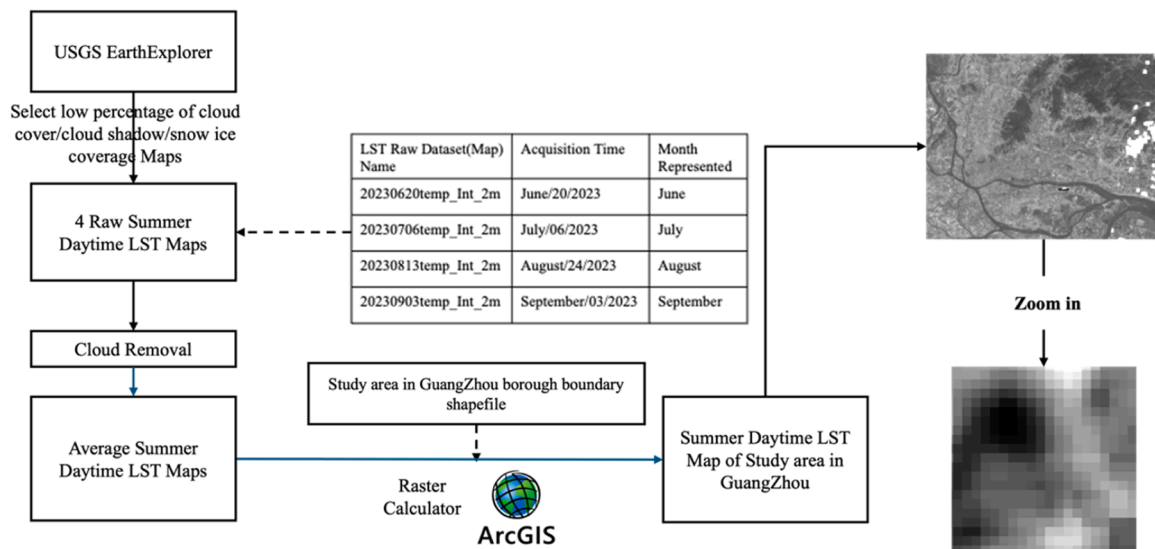


Fig. 4. LCZ and POI maps of the six typical sites.

Table 1

Color comparison of POI attributes for 3D buildings.

| Number | Building types | Color | RGB value |
|--------|----------------|--|-----------------|
| 1 | Facilities |  | (209, 252, 252) |
| 2 | Shop |  | (248, 135, 167) |
| 3 | Restaurant |  | (246, 194, 171) |
| 4 | Recreation |  | (222, 255, 189) |
| 5 | Industrial |  | (186, 228, 255) |
| 6 | Accommodation |  | (251, 251, 135) |
| 7 | No data | | (0, 0, 0) |

**Fig. 5.** Processing of LST datasets.

validation. All samples are processed into paired data for training the Pix2pix and CycleGAN models.

2.3. The framework of CP-GAN model

Generative Adversarial Networks (GANs) divides into two primary elements, including the generator and the discriminator. The generator utilizes convolutional layers to create new samples from input vectors, while the discriminator assesses these samples' authenticity, producing a probability score (Creswell et al., 2018). The principal aim of GANs is to generate data distributions that are indistinguishable from real-world samples. This process is driven by a minimax game where the generator strives to fool the discriminator, which in turn enhances its ability to identify fake data (Zhou et al., 2025). This iterative training ultimately yields highly realistic data. Among Conditional GANs (CGANs), Pix2pix has been found to surpass CycleGAN in terms of predictive accuracy and

model robustness (Zhou et al., 2023). Nevertheless, both algorithms have unique advantages and limitations. The Pix2pix model excels in generating high-quality translations due to its use of paired training data, enabling it to capture essential image information effectively and deliver clear outputs (Cira et al., 2021). However, it struggles with preserving morphological details, particularly in accurately rendering pixel information at the edges (Zhu et al., 2025). In contrast, CycleGAN addresses this limitation by iteratively learning edge features from individual images, allowing for more precise delineation of morphological contours. Its reliance on unpaired data further enhances its versatility (Zhu et al., 2017).

To leverage the strengths of both models, this study introduces a novel hybrid model, CP-GAN, which integrates the methodologies of CycleGAN and Pix2pix. By combining these approaches, CP-GAN overcomes the individual limitations of each model, providing a more accurate simulation and generation of urban morphology. The

architecture of the CP-GAN model is depicted in Fig. 6.

The CP-GAN model integrates the strengths of both Pix2Pix and CycleGAN, enabling efficient utilization of both paired and unpaired training data for improved image translation (Zhou et al., 2025). Its architecture operates in two main stages: the Pix2Pix and CycleGAN stages.

In the Pix2Pix stage, the model employs a U-Net generator G and a PatchGAN discriminator D to manage paired data. the Pix2Pix is a specific implementation of conditional generative adversarial networks (cGANs). The model's performance is enhanced using a combined loss function that combines adversarial loss with $L1$ loss and $L2$ loss. The core of \mathcal{L}_{CGAN} lies in the interaction between the generator and the discriminator. These two components are optimized through a game-like process. The generator aims to maximize the probability of the discriminator making incorrect judgments, while the discriminator strives to improve its ability to correctly distinguish real from generated data. Ultimately, this competitive relationship achieves equilibrium, progressively enhancing the quality of the generated images. The adversarial loss function is expressed as:

$$\mathcal{L}_{CGAN}(G, D) = \mathbb{E}_{x,y}[\log(D(x, y))] + \mathbb{E}_{x,z}[\log(1 - D(G(x, z)))] \quad (1)$$

Where, G is the generator, D is the discriminator, x is the input image, y is the real image and z is the random noise.

Additionally, the $L1$ loss is used to optimize the generator, ensuring that the generated samples are sharper and more visually clear. The $L1$ loss is mathematically expressed as:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1] \quad (2)$$

The $L_{L1}(G)$, while effective in minimizing pixel-wise differences, is prone to overfitting the training images (Aithal et al., 2020). To address this limitation, this study added $L_{L2}(G)$ as an extra term in the total loss function. This study uses $L_{L2}(G)$ loss minimizes the squared differences of each feature, which may result in smoother generated samples. The loss function is mathematically expressed as:

$$\mathcal{L}_{L2}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_2] \quad (3)$$

The combined loss function is expressed as:

$$G^* = \underset{G}{\operatorname{argminmax}}_D \mathcal{L}_{CGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) + \beta \mathcal{L}_{L2}(G) \quad (4)$$

Where λ and β are tuning parameters that control the relative contributions of the two losses. The term $\underset{G}{\operatorname{argminmax}}_D \mathcal{L}_{CGAN}(G, D)$ represents the optimization process. Specifically, \min_G refers to the process of adjusting the generator G parameters. The goal is to reduce the likelihood that the discriminator D can correctly identify the generated images as fake. Conversely, \max_D represents the discriminator's objective to enhance its ability to distinguish real image pairs from generated

ones. This adversarial interaction continues until a Game-theoretic equilibrium (G^*) is achieved, where the G is refined to generate images that accurately reflect the structural features of the ground truth, and the D is no longer able to reliably distinguish real from generated images (Albert et al., 2018).

In the CycleGAN stage, generator F plays a critical role in mapping images from domain Y to domain X , working collaboratively with generator G to establish a bidirectional mapping between domains X and Y . Specifically, G (G_{XY}) and F (work alongside discriminators D_X and D_Y , facilitating image translation between two domains, X and Y , while maintaining cycle consistency. The total loss function in this stage incorporates adversarial loss and cycle consistency loss, ensuring accurate domain translation. It is represented by the following equation:

$$G^*, F^* = \underset{G, F}{\operatorname{argminmax}}_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y) \quad (5)$$

$\underset{G, F}{\operatorname{argminmax}}_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y)$ is finding parameters for the Generator (G, F) to minimize ($\min_{G, F}$) the loss of the generator to produce samples that are as realistic as possible, spoofing the discriminator and the process for the Discriminator (D_X, D_Y) to maximize (\max_{D_X, D_Y}) the loss in order to try to distinguish between real and generated samples, forcing the generator to continuously improve. Similar to the pix2pix model, the G and D continuously compete with each other until reaching a Nash equilibrium (G^*, F^*) is reached.

By combining the conditional generation and $L1$ loss of Pix2Pix with the cyclic consistency loss of CycleGAN through these two stages of task processing, the total loss function for CP-GAN integrates these components. The total loss function for CP-GAN integrates these components:

$$\begin{aligned} L_{CP-GAN} = & L_{GAN}(G, D) + \lambda_{L1} L_{L1}(G) + L_{GAN}(G_{YX}, D_X) + L_{GAN}(G_{XY}, D_Y) \\ & + \lambda_{cyc} L_{cyc}(G_{XY}, G_{YX}) \end{aligned} \quad (6)$$

$L_{GAN}(G, D)$ represents the adversarial loss in the Pix2Pix stage, where G is the generator that transforms input image X into output image Y , and D is the discriminator that distinguishes between real image pairs (X, Y) and fake pairs ($X, G(X)$). The term L_{L1} is a hyperparameter that controls the relative importance of the $L_{L1}(G)$, helping to ensure structural similarity. $L_{GAN}(G_{XY}, D_Y)$ and $L_{GAN}(G_{YX}, D_X)$ represent the adversarial losses for the CycleGAN stage. G_{XY} and G_{YX} are the generators that translate images between domains X and Y , while D_X and D_Y are the discriminators that distinguish between real and generated images in domains X and Y respectively. The cycle consistency loss $L_{GAN}(G_{XY}, G_{YX})$ ensures that an image translated from domain X to domain Y and back to domain X (and vice versa) remains similar to the original image, maintaining semantic consistency. The hyperparameter λ_{cyc} controls the weight of this cycle consistency loss. By optimizing these loss functions, CP-GAN aims to produce high-quality and semantically consistent

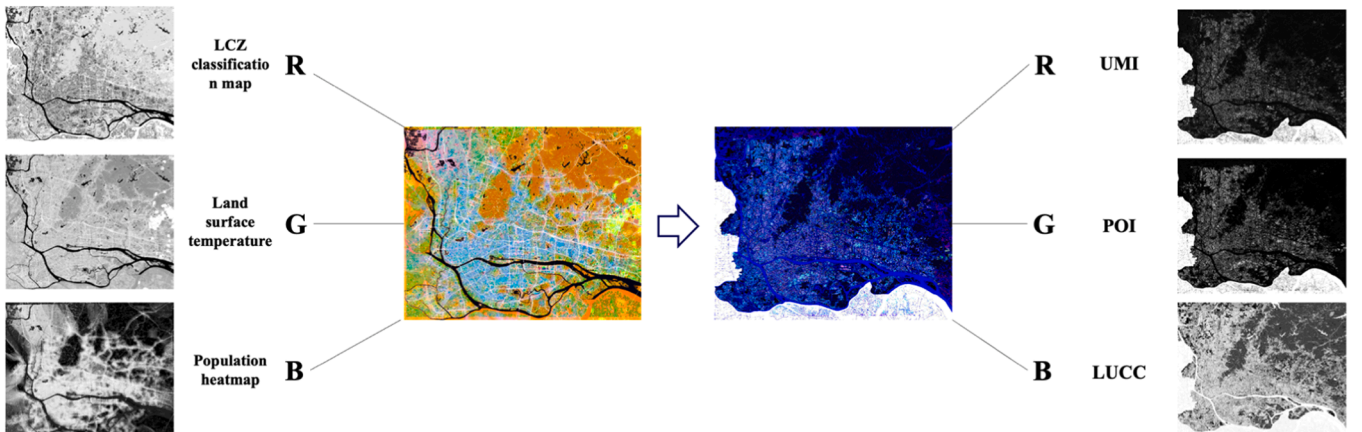


Fig. 6. RGB image channel data fusion.

images across different domains.

2.4. Performance evaluation

To measure the precision of the GAN training process quantitatively, this study employed a set of five commonly used evaluation metrics based on a thorough review of the literature. These metrics include the Structural Similarity Index (SSIM) (Wang et al., 2004), Peak Signal-to-Noise Ratio (PSNR) (Desalegn & Jifara, 2024; Bagavathi et al., 2024; Samreen & Venu, 2024), Coefficient of Determination (R^2) (Lucas et al., 2018; Song et al., 2022), Mean Square Error (MSE) (Liang et al., 2022; Desalegn & Jifara, 2024), Fréchet Inception Distance (FID) (Heusel et al., 2017), and Kernel Inception Distance (KID) (Bińkowski et al., 2021). The specific calculation formula is as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (7)$$

Where μ_x and μ_y are the mean brightness values of images x and y , respectively. σ_x and σ_y are the standard deviations of images x and y , respectively. σ_{xy} is the covariance between images x and y . C_1 and C_2 are constants used to avoid division by zero. Generally, $C_1 = (K_1L)^2$ and $C_2 = (K_2L)^2$, where $K_1 = 0.01$ and $K_2 = 0.03$, and L is the dynamic range (e.g., for 8-bit images, $L = 255$).

$$PSNR = 10 \times \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - \text{Mean}(K)]^2}{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [K(i, j) - \text{Mean}(K)]^2} \quad (9)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (10)$$

Where $I(i, j)$ is the true value of image I , $K(i, j)$ is the predicted value of image K , and $\text{Mean}(K)$ is the average value of pixels of the genuine image K .

$$FID = |\mu_x - \mu_y|^2 - \text{Tr}(\Sigma_x + \Sigma_y - 2\sqrt{\Sigma_x \Sigma_y}) \quad (11)$$

Where $|\mu_x - \mu_y|^2$ is the sum squared difference between the feature-wise mean of both the real and the generated images. $\text{Tr}(\Sigma_x + \Sigma_y - 2\sqrt{\Sigma_x \Sigma_y})$ is the difference calculated using the trace of the matrix, x and y represent the covariance matrices of the genuine and generated images

$$KID = MMD^2(P, Q) = \|\mu_P - \mu_Q\|_F^2 \quad (12)$$

Where P and Q denote the features extracted from the genuine and generated images, respectively. The MMD measures the difference between the feature distributions of P and Q . μ_P and μ_Q denote the mean values of the features P and Q .

3. Results and discussion

All code is written in Python with the help of the TensorFlow library. The experiments were run on a platform built with an Intel Core i5-12400 CPU and an NVIDIA GeForce RTX 3070 GPU. The model ran for over 1200 training sets, and the total training time was approximately 43 hours. It can be integrated as an API for easier access and use. The trained model was deployed on our Transo platform in our lab.

3.1. CP-GAN generation results

3.1.1. Training process and validation of model performance

The loss curves for CP-GAN were depicted in Fig. 7, illustrating the training dynamics of both the generator and discriminator. The generator's loss curve reflects its ability to produce images that deceive the discriminator, while the discriminator's loss curve measures its skill in differentiating between real and generated images. Additionally, the overall GAN loss curve primarily represents the discriminator's performance, whereas the L1 loss curve indicates how accurately the generator reconstructs the target images (Huang et al., 2022). As shown in Fig. 7, the loss curves for both the generator and discriminator exhibited smooth trajectories, suggesting that the model maintained stable performance throughout the learning process (Kastner & Dogan, 2023). The total GAN loss curve demonstrated a gradual decrease, with stabilization occurring after approximately 400 generations. This steady decline highlighted the model's progression toward stability. After 750 epochs, the loss reached its minimum value of 15, indicating convergence towards the optimal solution. Similarly, the L1 loss curve showed a gradual decline during the early training phase, stabilizing around generation 200 and achieving its lowest point at generation 750. The plots below illustrated the progressive learning effect of the CP-GAN model. During the initial 50 iterations, significant changes were observed in both the urban morphology and the distribution of POI, indicating that the model is still refining its understanding of the spatial features. By the 250th iteration, the model began to establish a rudimentary yet recognizable urban form, with the POI types largely settled. At the 1,000th training iteration, the learning process has effectively converged, resulting in a city morphology and POI distribution that closely align with the target patterns. This indicated that the model has successfully captured the essential contours and spatial characteristics of the urban structure.

3.1.2. Validation of model performance

To rigorously evaluate the predictive accuracy of the CP-GAN model, this study conducted a comprehensive assessment of its output quality using both qualitative and quantitative approaches. This dual assessment framework enables a more precise and intuitive analysis of the model's performance, capturing both the visual fidelity and numerical accuracy of the generated urban morphology. Fig. 8 presented the generated images alongside relative error assessment maps for six typical LCZ samples, displaying ground truth maps, urban texture maps generated by CP-GAN, 3D regional models, and relative error maps. The urban texture maps primarily rely on CP-GAN's generation capabilities, while the 3D models are derived from the texture maps using RGB-based automation functions via the Grasshopper plugin in Rhino (S. Zhou et al., 2023).

The relative error maps (Last column in Fig. 8) demonstrated minimal discrepancy between generated textures and models, highlighting the high-quality output of CP-GAN. In the relative error maps, the color gradient from blue to red indicated increasing error levels. The analysis reveals that, overall, the prediction errors were generally low across the selected plots. However, higher error regions (highlighted in red) were observed near road-adjacent areas in plots GZ01, GZ05, and GZ06. This discrepancy is likely due to pixel-level confusion in the input data, particularly in regions where roads intersect with buildings, leading to reduced model accuracy in these areas (Fig. 9).

To ensure the optimal performance of the CP-GAN model, this study first referred to prior research, particularly studies that utilized similar hybrid GAN approaches, to provide an initial range for the λ values. Based on these insights, this study employed a grid search method to explore different λ values, then conducted an ablation study to evaluate the impact of different hyperparameter settings. Specifically, we tested eight different combinations of the discriminator learning rate (learning_rate_D), generator learning rate (learning_rate_G), and batch size. These combinations were selected based on prior studies and empirical

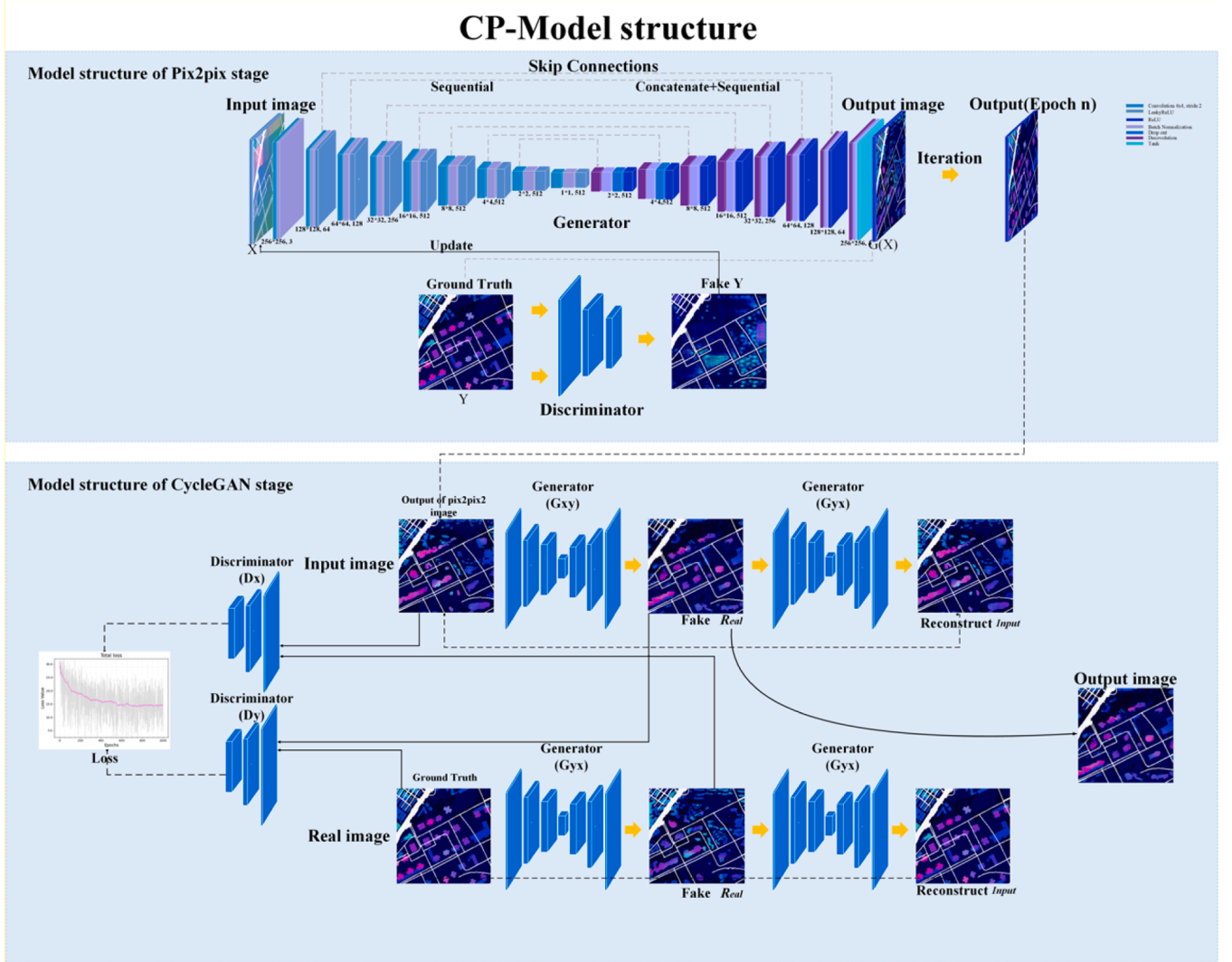


Fig. 7. Model architecture of CP-GAN model.

testing, and the performance of each model was evaluated using SSIM as the primary performance metric. The results of the ablation experiment (Fig. 10) show that Model 4, which had a learning_rate_D of 8.0×10^{-5} , a learning_rate_G of 1.0×10^{-4} , and a batch size of 4, achieved the highest SSIM value of 0.811, outperforming all other configurations. This combination of hyperparameters was selected as the optimal setting for the CP-GAN model in this study. The total loss curve also showed that Model 4 achieved smoother convergence with lower overall loss, as compared to other models. This ablation study not only demonstrated the importance of hyperparameter tuning but also highlighted the robustness of the CP-GAN model for urban morphology generation.

This study evaluated the performance of the CP-GAN model in terms of image quality and predictive accuracy using metrics such as SSIM, PSNR, R^2 , MSE, FID and KID calculated through functions provided by the TorchMetrics Python library (TorchMetrics K.L., 2022). The results revealed that CP-GAN achieves an average SSIM of 0.811 with all individual values exceeding 0.75, reflecting its strong ability to preserve structural fidelity in generated images. The average PSNR of 15.386 demonstrates the model's capability to produce low-noise, high-quality outputs. With a mean R^2 of 0.687, the model reliably reconstructs urban morphological details. Furthermore, the average MSE of 3317.627 indicates acceptable levels of reconstruction error, while the FID and KID averages of 19.315 and 1.922, respectively, highlight minimal divergence between real and generated images. Collectively, these metrics

affirm the CP-GAN model's efficacy and reliability in generating accurate, high-quality urban morphology representations (Table 2).

To comprehensively evaluate the performance of CP-GAN, this study compared it with another advanced generative model, the Variational Autoencoder (VAE) (Fig. 11). VAE is a deep learning model typically used to generate high-quality images by learning the distribution of input data. The VAE model was trained on the same multimodal dataset as CP-GAN. The results showed that VAE had a loss rate of 19.14%, which was significantly higher than that of CP-GAN. Moreover, CP-GAN exhibited superior performance compared to VAE across multiple kinds of evaluated dimensions. Structurally, CP-GAN improved the SSIM score by 18.4%, indicating a stronger ability to preserve spatial patterns. The PSNR also increased by 13.5%, reflecting enhanced suppression of reconstruction noise. From a perceptual standpoint, CP-GAN reduced the FID by 17.2% and KID by 12.9%, suggesting that the generated images more closely resembled real data distributions. In terms of model fit, CP-GAN outperformed VAE with a 15.3% improvement in R^2 , highlighting its enhanced explanatory power. Simultaneously, the mean squared error dropped by 9.2%, pointing to more accurate predictions (Table 2). Visual comparisons of images generated by both models revealed that VAE produced incomplete building forms, lacked accuracy in building edges, and exhibited missing building units compared to CP-GAN (Fig. 11). Based on these findings, CP-GAN demonstrated outstanding capability in generating realistic and structurally consistent

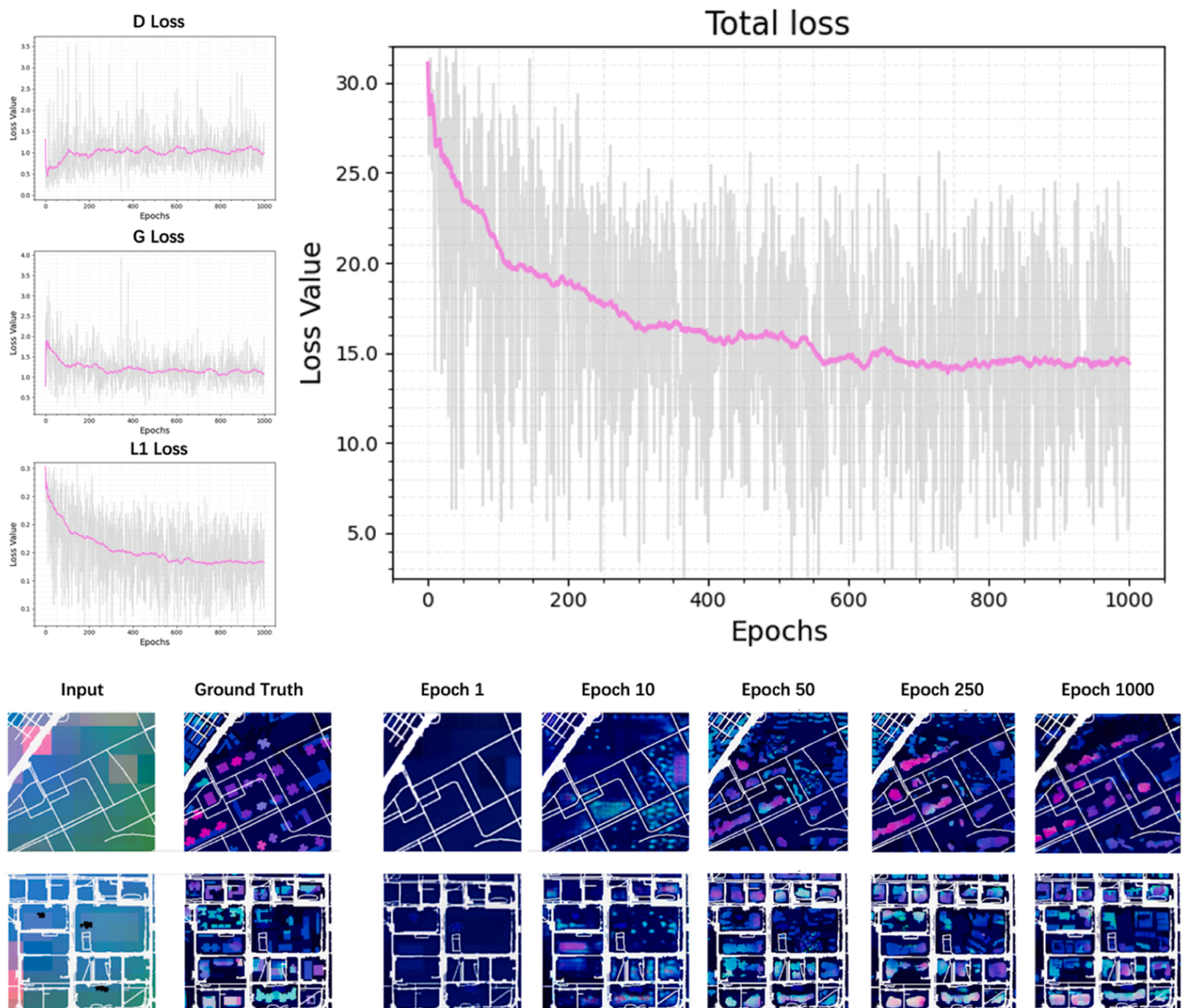


Fig. 8. Training loss curves of CP-GAN.

urban morphology, making it the model of choice for this study, (Figs. 12 and 13).

To evaluate the robustness and responsiveness of the CP-GAN model, this study conducted a structured sensitivity analysis focusing on two key input variables: Land Surface Temperature (LST) and Population Heat (POPH). The analysis was divided into two parts. First, a univariate sensitivity analysis was performed, where LST and POPH were independently adjusted across a range of values to observe their isolated effects on the generated urban morphology. Second, a bivariate sensitivity analysis was conducted, in which both LST and POPH were varied simultaneously to examine their combined influence. These experiments aimed to investigate how fluctuations in environmental and demographic conditions impact the CP-GAN-generated outputs. The detailed results are presented in Sections 3.2 and 3.3.

3.2. Process and results of univariate analysis

Building on the robust generative capabilities of the CP-GAN model, as a step toward prospective research, this study investigated how urban environmental factors influence urban morphology by adjusting the model's conditional inputs. The model used multimodal inputs,

integrating LCZ, LST, and POPH, to generate outputs that depict variations in urban features. Specifically, the outputs included visual representations of changes in POI, LUCC and Build Footprint. Building footprints are further utilized to calculate UMIs, which provide quantitative metrics for analyzing urban form, like building height, building density and floor area ratio. POIs represent significant locations such as amenities, residential areas, restaurants, and shops, while LUCC refers to changes in land use and surface characteristics, like vegetation and impervious surfaces. The subsequent analysis explored the effects of independent variations in LST and POPH on POI, LUCC, and UMI to further understand their influence on urban spatial patterns. For brevity, this study presents an analysis of representative experimental data, with the specific indicators and their calculation methods, as well as the complete dataset for each parcel, detailed in Appendix 1.

3.2.1. Effect of temperature change on UMI, LUCC and POI

Scenario 1 explored the impact of varying land surface temperature (ranging from 18°C to 49°C) on POI, LUCC, and UMI across six key parcels in Guangzhou (GZ01 to GZ06), while keeping regional population density constant. Fig. 12 illustrated the predicted changes in POI, LUCC, and UMI using line charts, and further provided visual outputs at

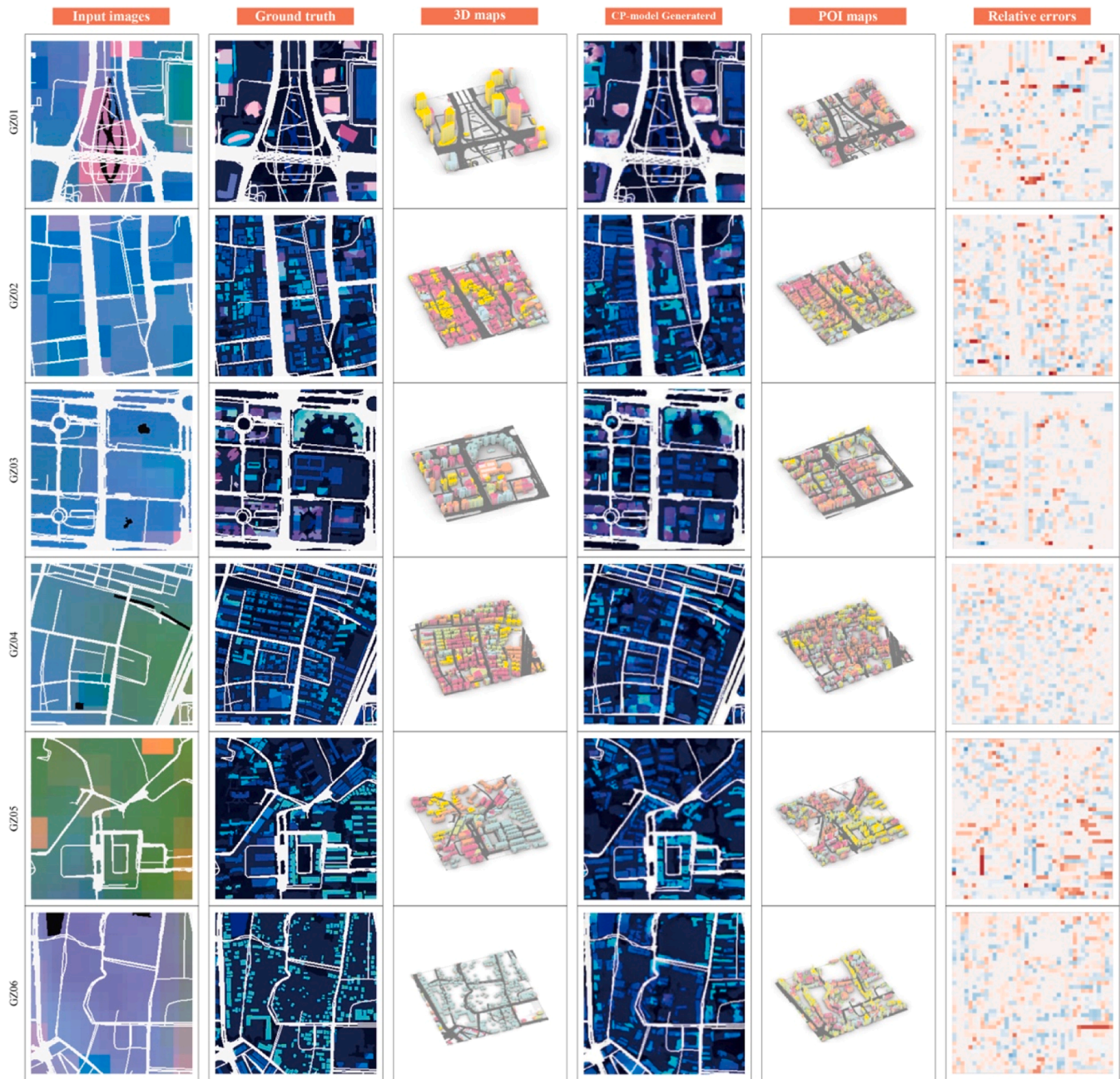


Fig. 9. Generated images for six locations with their corresponding relative error evaluation maps. Note: Dark red and blue pixels highlight major differences between the actual and predicted values, while light-colored pixels indicate accurate predictions.

three critical temperature benchmarks—25°C (25% increase), 33.5°C (50% increase), and 42°C (75% increase). These visualizations offered a clearer understanding of how temperature fluctuations influence urban spatial patterns and land use dynamics.

The analysis of POI changes across the six parcels revealed distinct patterns influenced by rising temperatures. In GZ01 (LCZ1) and GZ02 (LCZ2), leisure spaces, commercial facilities, and dining establishments underwent significant shifts as temperatures increased. Notably, when temperatures exceeded 35°C, there was a sharp rise in the number of commercial and entertainment facilities, with the commercial facility rate in GZ01 increasing from 0.0947 to 0.2121, and in GZ02 from 0.0597 to 0.3019. Conversely, residential and public service facilities declined, suggesting a shift in population mobility toward commercial and entertainment venues under high temperatures, likely due to the cooling benefits provided by air conditioning in these spaces (Zullo et al., 2019;

Yin et al., 2021). However, this shift also implied increased energy consumption in these buildings. A similar trend was observed in GZ03 (LCZ4) and GZ05 (LCZ6), where commercial POIs expanded significantly. Unlike GZ01 and GZ02, these parcels also experienced marked growth in industrial functions. In particular, industrial density surged at 33.5°C and 42°C, indicating a concentration of industrial activities in these areas under elevated temperatures. In GZ04 (LCZ5), the patterns of entertainment and residential functions aligned with those observed in GZ01 and GZ02. Meanwhile, GZ06 (LCZ7), characterized by high green cover, showed a rapid increase in shop rate from 0.1 to 0.174 as temperature rises from 25°C to 35°C.

From the perspective of LUCC, the green rate of all plots decreased with the increase of temperature, but the amplitude was different. A clear trend emerged where forest and grassland cover in GZ01 (LCZ1) and GZ02 (LCZ2) declined sharply after temperature exceeds 30°C, with

| Surrogate model | Learning_rate_D | Learning_rate_G | Batch size | SSIM |
|---------------------------------|-----------------|-----------------|------------|-------|
| Model 1 | 2.0e-5 | 4.0e-4 | 4 | 0.808 |
| Model 2 | 8.0e-5 | 4.0e-4 | 4 | 0.807 |
| Model 3 | 2.0e-5 | 1.0e-4 | 4 | 0.781 |
| Model 4 (The model selected) | 8.0e-5 | 1.0e-4 | 4 | 0.811 |
| Model 5 | 2.0e-5 | 4.0e-4 | 1 | 0.781 |
| Model 6 | 8.0e-5 | 4.0e-4 | 1 | 0.778 |
| Model 7 | 2.0e-5 | 1.0e-4 | 1 | 0.779 |
| Model 8 | 8.0e-5 | 1.0e-4 | 1 | 0.779 |

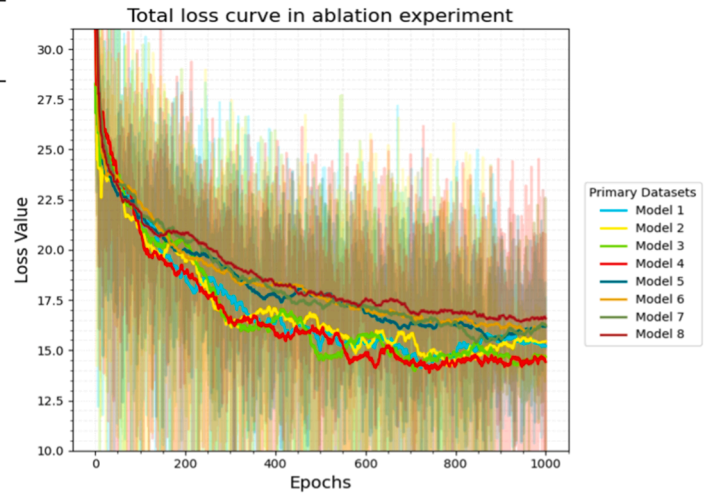


Fig. 10. The results of the ablation experiment.

Table 2

Performance comparison of CP-GAN and VAE using testing datasets.

| Model Name | Sites | SSIM | PSNR | R2 | MSE | FID | KID |
|------------|---------|-------|--------|-------|----------|--------|-------|
| CP-GAN | GZ01 | 0.810 | 15.182 | 0.750 | 3398.508 | 22.360 | 2.240 |
| | GZ02 | 0.761 | 14.201 | 0.542 | 4043.479 | 24.850 | 2.480 |
| | GZ03 | 0.869 | 16.696 | 0.865 | 2599.210 | 14.890 | 1.490 |
| | GZ04 | 0.840 | 16.361 | 0.772 | 2757.957 | 16.790 | 1.680 |
| | GZ05 | 0.836 | 15.227 | 0.656 | 3371.592 | 17.670 | 1.710 |
| | GZ06 | 0.752 | 14.649 | 0.539 | 3735.017 | 19.330 | 1.930 |
| | Average | 0.811 | 15.386 | 0.687 | 3317.627 | 19.315 | 1.922 |
| VAE | GZ01 | 0.694 | 13.345 | 0.592 | 3321.641 | 24.194 | 2.787 |
| | GZ02 | 0.689 | 12.905 | 0.502 | 3441.374 | 28.527 | 1.166 |
| | GZ03 | 0.684 | 13.613 | 0.613 | 4105.117 | 18.739 | 2.702 |
| | GZ04 | 0.684 | 13.518 | 0.640 | 3670.138 | 23.210 | 2.928 |
| | GZ05 | 0.683 | 14.080 | 0.589 | 3786.852 | 23.784 | 1.003 |
| | GZ06 | 0.675 | 13.878 | 0.640 | 3607.490 | 21.467 | 2.647 |
| | Average | 0.685 | 13.556 | 0.596 | 3655.435 | 23.320 | 2.206 |

an almost complete loss of forest observed above 40°C. For instance, in GZ02, the forest rate decreased from 0.2981 to 0.0120, while the grassland rate dropped from 0.1208 to 0.0247. These results aligned with Biao et al. (2012), who found that hot cities exhibit significantly lower green space coverage compared to cooler cities, which in turn increases the proportion of impervious surfaces and exacerbates urban heat island effect. Similarly, GZ03 (LCZ4) and GZ05 (LCZ6) also showed substantial reductions in forest and grass cover under rising temperature. The temperature thresholds for changes in impervious rates varied significantly across different LCZs. For example, GZ01 (LCZ1) and GZ02 (LCZ2) exhibited rapid increases in impervious surface rates when temperatures exceeded 30°C, while GZ04 (LCZ5) and GZ05 (LCZ6) showed a marked acceleration only at temperatures above 42°C. In contrast, GZ06 (LCZ7) experienced a sharp rise in impervious surface rates when temperatures surpassed 25°C. These variations likely contribute to more pronounced spatial heterogeneity in Urban Heat Island (UHI) effects at the local scale (Unger et al., 2018). The underlying causes of these differences are rooted in the distinct characteristics of each LCZ. High-density, high-rise zones face increasing functional demands as temperatures rise, driving the expansion of hardened surfaces to accommodate these needs. In contrast, medium- and low-density areas have lower building densities, leading to slower functional changes and a delayed response to temperature increases. Meanwhile, low-density, green-covered zones are highly sensitive to temperature fluctuations due to their ecological characteristics.

From the standpoint of UMI, the BCR in GZ01 (LCZ1) and GZ03 (LCZ4) initially decreased slightly with rising temperature, followed by

an increase, eventually stabilizing around 37°C. This indicated that the impact of temperature on building coverage is more pronounced in non-high-rise building zones compared to high-rise zones. This difference is primarily because high-rise zones, such as LCZ1 and LCZ4, typically already exhibit building coverage rates near their upper limits, leaving limited room for further increases in response to temperature variations (Xu et al., 2022). Similarly, the FAR in GZ01 (LCZ1), GZ02 (LCZ2), and GZ03 (LCZ4) followed a comparable trend, with a slight initial decline as temperature rose, followed by an increase, and stabilization beyond 37°C. In contrast, the FAR in other parcels showed no significant response to temperature changes, suggesting that mid- to high-rise buildings are more sensitive to thermal variations in their volumetric utilization than low-rise structures.

In conclusion, rising temperatures exert a substantial influence on specific land uses, notably leading to the degradation of green spaces and residential zones, while concurrently amplifying the share of commercial activities. The data indicate that for every 3°C increase in temperature, the average green space coverage across various LCZs decreases by 5.47%, commercial activity increases by 2.38%, and impervious surface rates rise by 2.84%. In contrast, industrial and entertainment functions demonstrate greater resilience to thermal stress, exhibiting reduced sensitivity to temperature fluctuations. This study highlighted that open spaces are more vulnerable to ecological degradation and urban hardening under high-temperature conditions compared to compact areas. These findings underscored the necessity for urban planners to prioritize green space design in open areas to enhance resilience against climate-induced stresses.

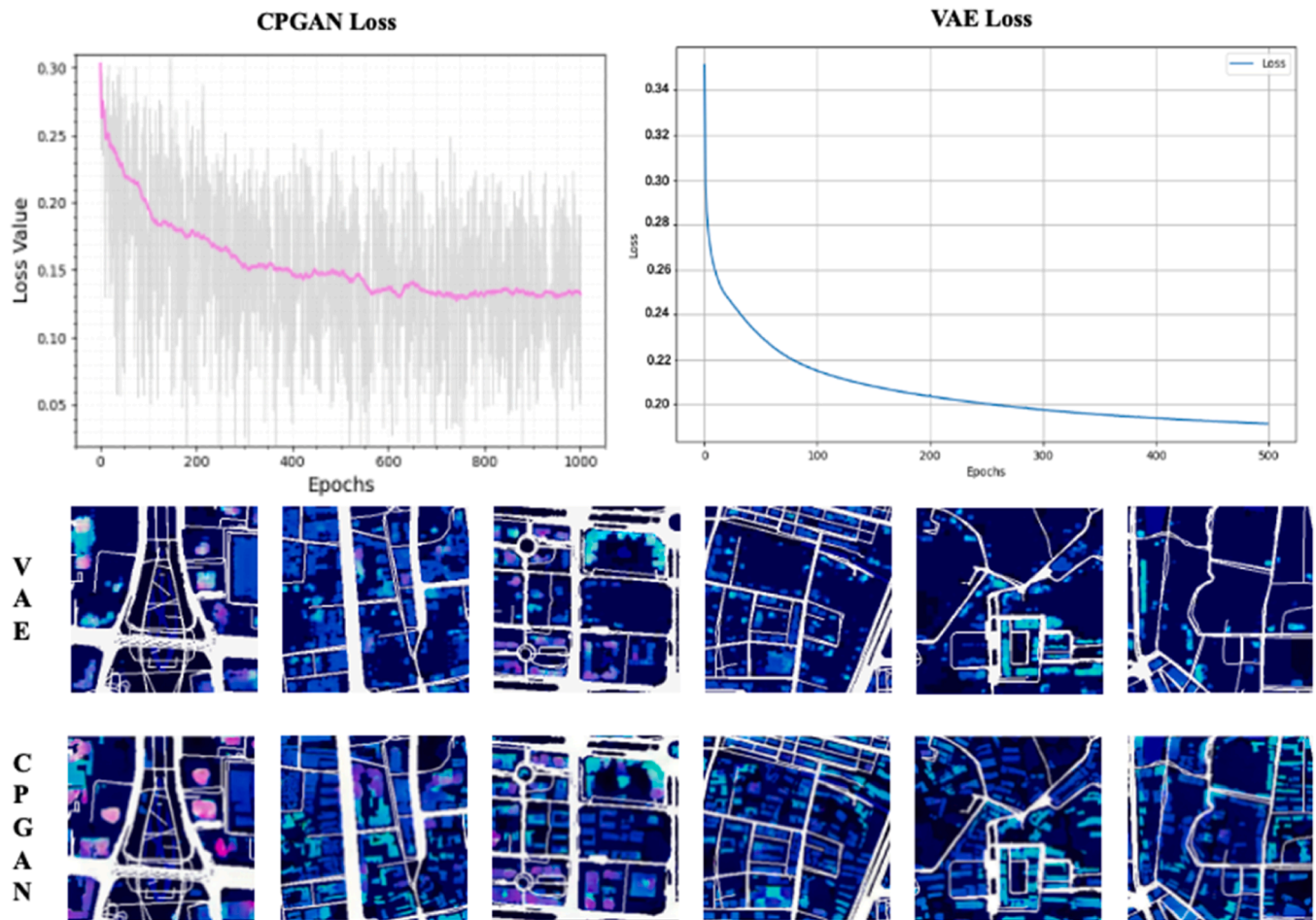


Fig. 11. Comparison of CP-GAN and VAE model performances.

3.2.2. Effect of population density change on POI, LUCC and UMI

Scenario 2 explored the impact of varying population density (ranging from 85 to 1196 people/km²) on POI, LUCC, and UMI across six key parcels in Guangzhou (GZ01 to GZ06), while keeping regional land surface temperature constant. Fig. 13 illustrated the predicted changes in POI, LUCC, and UMI using line charts, and further provides visual outputs at three critical population density benchmarks—342 people/km² (25% increase), 598 people/km² (50% increase), and 940 people/km² (75% increase). These visualizations offered a clearer understanding of how population density fluctuations influence urban spatial patterns and land use dynamics under certain climate conditions.

The analysis of POI distribution revealed that as population density increases, commercial, entertainment, and dining facilities become more concentrated, indicating a clustering effect of economic activities and services in high-density areas (Zhao et al., 2020). For instance, in GZ01 (LCZ1), the concentration of commercial and public facilities increased significantly as population density rises from 342 to 940 people per square kilometer. Although higher population density led to a greater concentration of POI, the land use proportion of POI remained relatively stable, showing much less variation compared to the shifts observed under changing temperature conditions. This suggests that changes in the distribution of POI have a greater impact on the urban thermal environment than on the urban population.

Within the framework of LUCC, increasing population density was associated with a decline in forest and grassland coverage. For example, in GZ01 (LCZ1), forest rate decreased from 0.166 to 0.115, while grassland rate dropped from 0.0695 to 0.045. These findings were consistent with those of Li et al. (2019), who reported that urbanization and population pressure significantly reduce ecological spaces. Notably,

the decline in green coverage driven by population density appeared to follow a more linear and gradual trend compared to the fluctuations observed under changing temperature conditions. Furthermore, impervious rates increased with rising population density in GZ05 (LCZ6) and GZ06 (LCZ7), whereas in GZ02 (LCZ2), GZ03 (LCZ4), and GZ05 (LCZ5), these rates remained relatively stable. This disparity could be attributed to the lower initial development intensity in low-rise building areas, which allows for an increase in impervious surfaces through new infrastructure to accommodate population growth. In contrast, other zones were more fully developed, and population increases were primarily reflected in the optimization of spatial functions rather than additional infrastructure development. Importantly, temperature changes exerted a broader and more significant impact on impervious rates compared to population density. This is because temperature variations drive adjustments in urban infrastructure and functionality at a city-wide scale, whereas the effects of population density are more localized to specific areas (Pan et al., 2023).

Considering UMI, variations in BCR across parcels exhibited distinct patterns in response to changes in population density. In GZ01 (LCZ1), GZ02 (LCZ2), GZ03 (LCZ4), and GZ04 (LCZ5), the BCR increased with rising population density, whereas in GZ05 (LCZ6), it remained relatively stable, and in GZ06 (LCZ7), it even decreased. This variation arose from the spatial heterogeneity of urban structures, reflecting differences in land use policies and planning objectives across regions. In urban core areas, higher land-use intensity is required to accommodate population growth (Zhao et al., 2023). In contrast, open zones such as LCZ6 and LCZ7 experience a decline in BCR due to the constraints imposed by green spaces and lower building densities (Yu et al., 2010). Similarly, the FAR increased with population density in high-rise zones such as

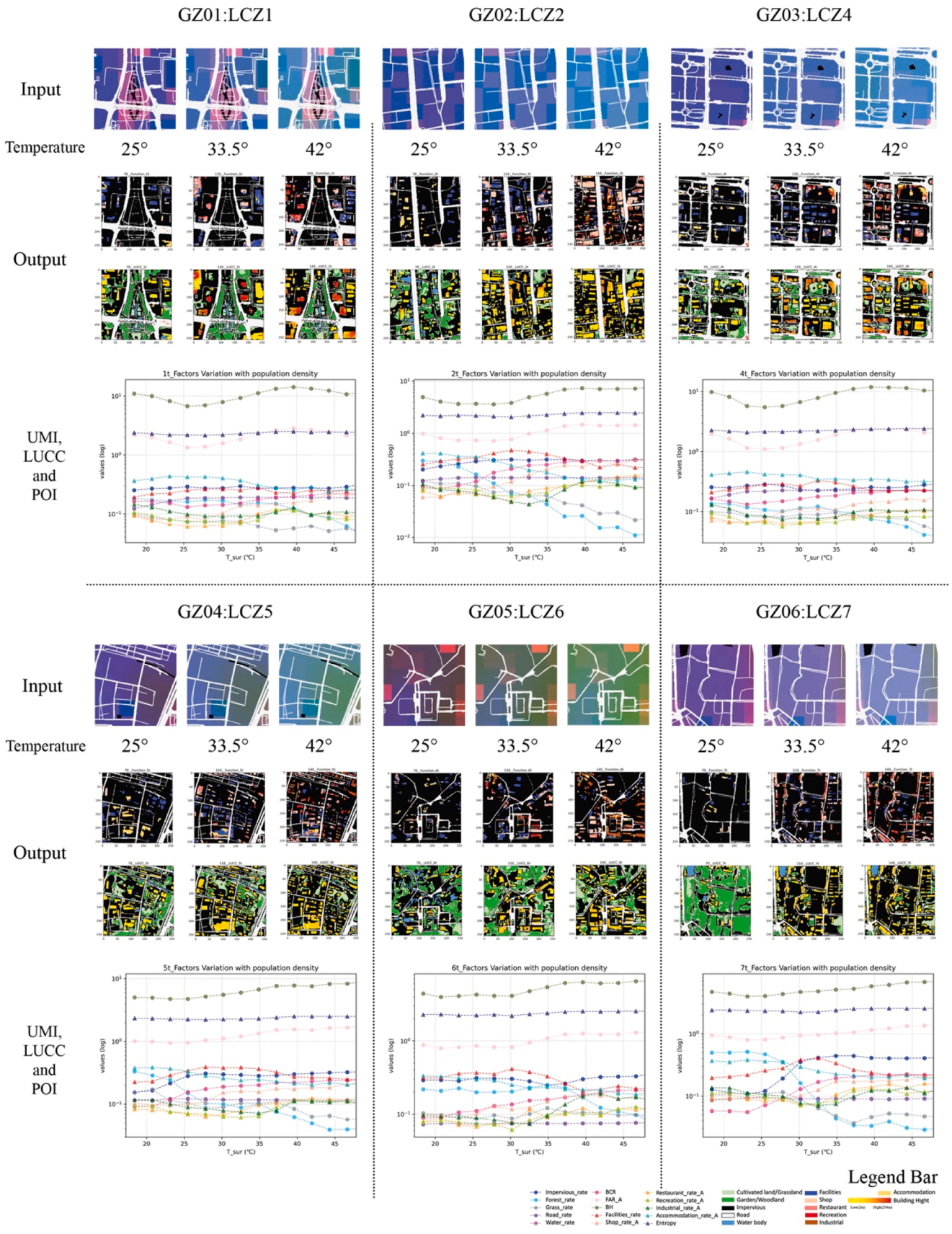


Fig. 12. Simulation analysis of changes in land use and urban functions of the six LCZ parcels under Scenario 1.

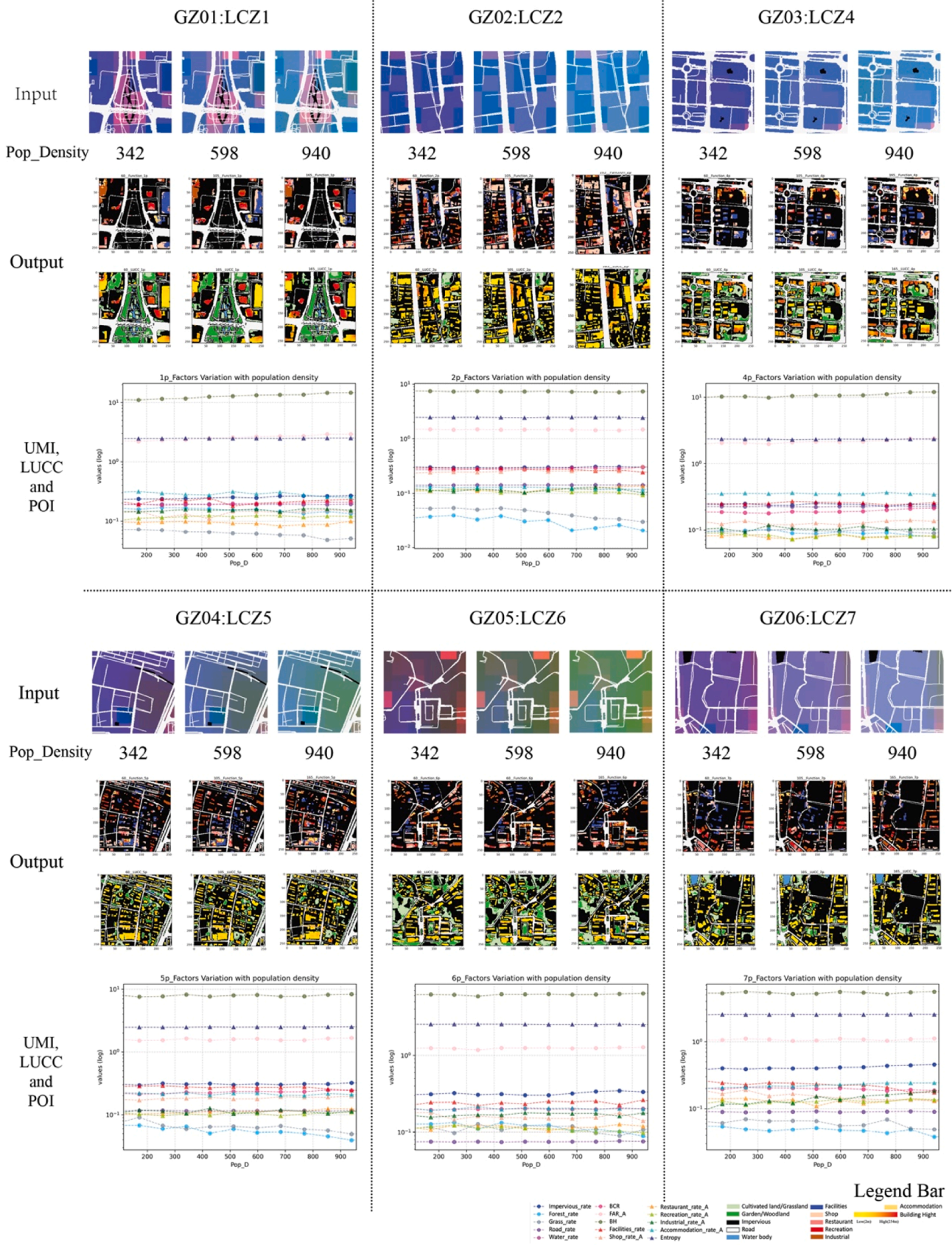


Fig. 13. Simulation analysis of changes in land use and urban functions of the six LCZ parcels under Scenario 2.

GZ01 (LCZ1) and GZ03 (LCZ4), reflecting the trend of vertical expansion in urban centers to meet urbanization demands (Li et al., 2021).

Overall, commercial, entertainment, and dining facilities exhibited a clustering effect in high-density areas, though changes in land use proportions remained minimal. Green space coverage decreased linearly with rising population density, while population density significantly influenced the BCR and FAR in high-rise building parcels. However, its impact on the diversity of POI and UMI across most parcels was relatively limited. Compared to the effects of temperature fluctuations, the

influence of population density changes on land use and urban morphology was considerably smaller.

3.3. Process and results of bivariate analysis

This section employed a dual-factor analysis to investigate the combined effects of temperature and population density on POI, LUCC, and UMI, providing insights into their collective impact on urban spatial morphology. The study focused on parcels within LCZ4, LCZ5, and LCZ6

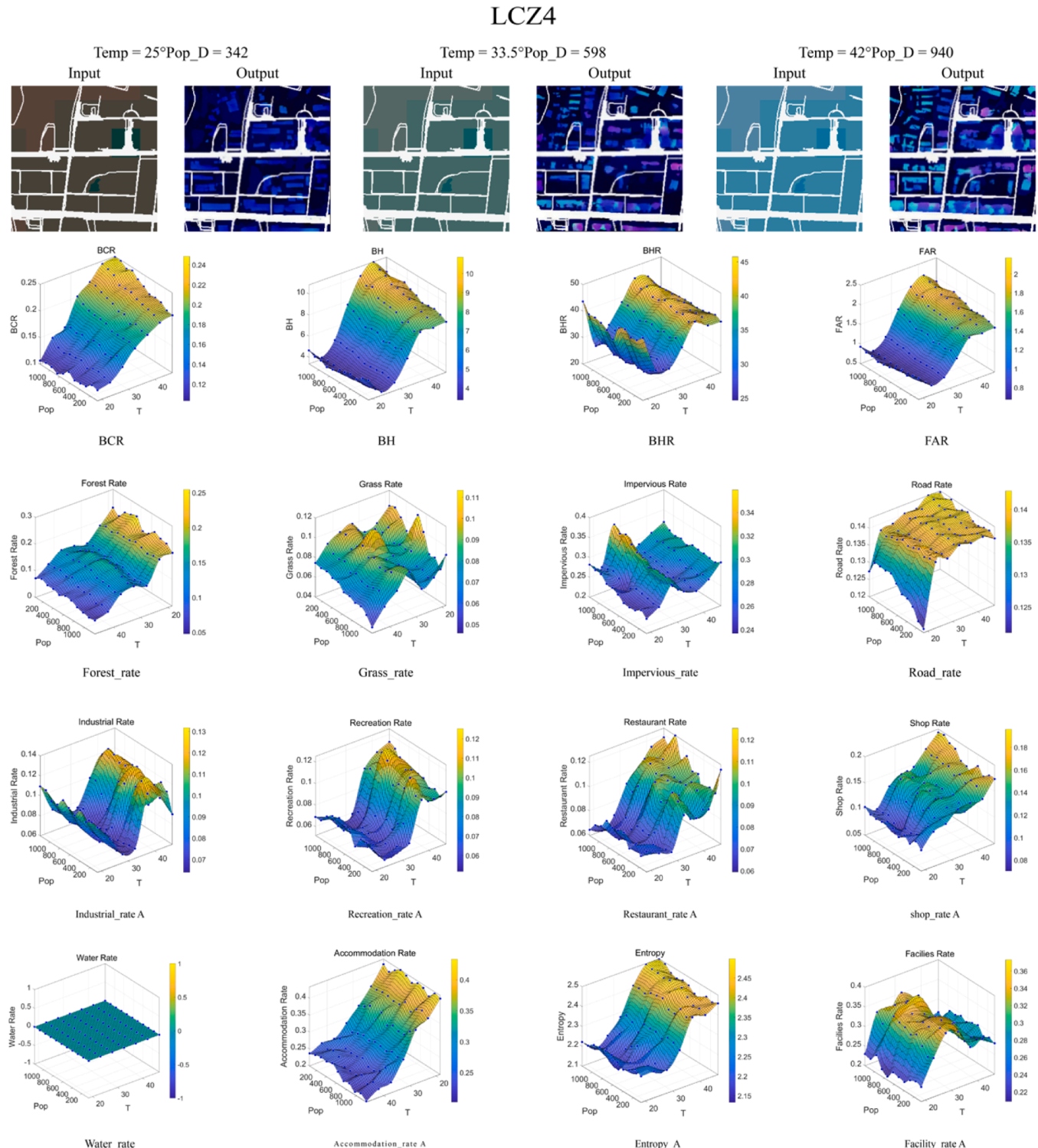


Fig. 14. Simulation analysis of changes in land use and urban functions of the LCZ4 under Scenario 3.

in Guangzhou due to their distinctive urban morphological characteristics, which represent a broad spectrum of density, building height, and land-use diversity. Other LCZ types were excluded due to their limited diversity in POI and LUCC characteristics or lower sensitivity to the combined effects of temperature and population density, making them less relevant to the objectives of this study.

Scenario 3 explored the impact of varying land surface temperature (ranging from 18°C to 49°C) and population density (ranging from 85 to 1196 people/km²) on POI, LUCC, and UMI across these three parcels.

Figs. 14–16 illustrated the predicted changes in POI, LUCC, and UMI of LCZ4, LCZ5, LCZ6 using line charts respectively, and further provided visual outputs at three critical benchmarks—25°C with a population density of 342 people/km² (25% increase), 33.5°C with 598 people/km² (50% increase), and 42°C with 940 people/km² (75% increase). These visualizations offered a clearer understanding of how temperature and population density fluctuations synergistically drive urban spatial patterns and land use dynamics.

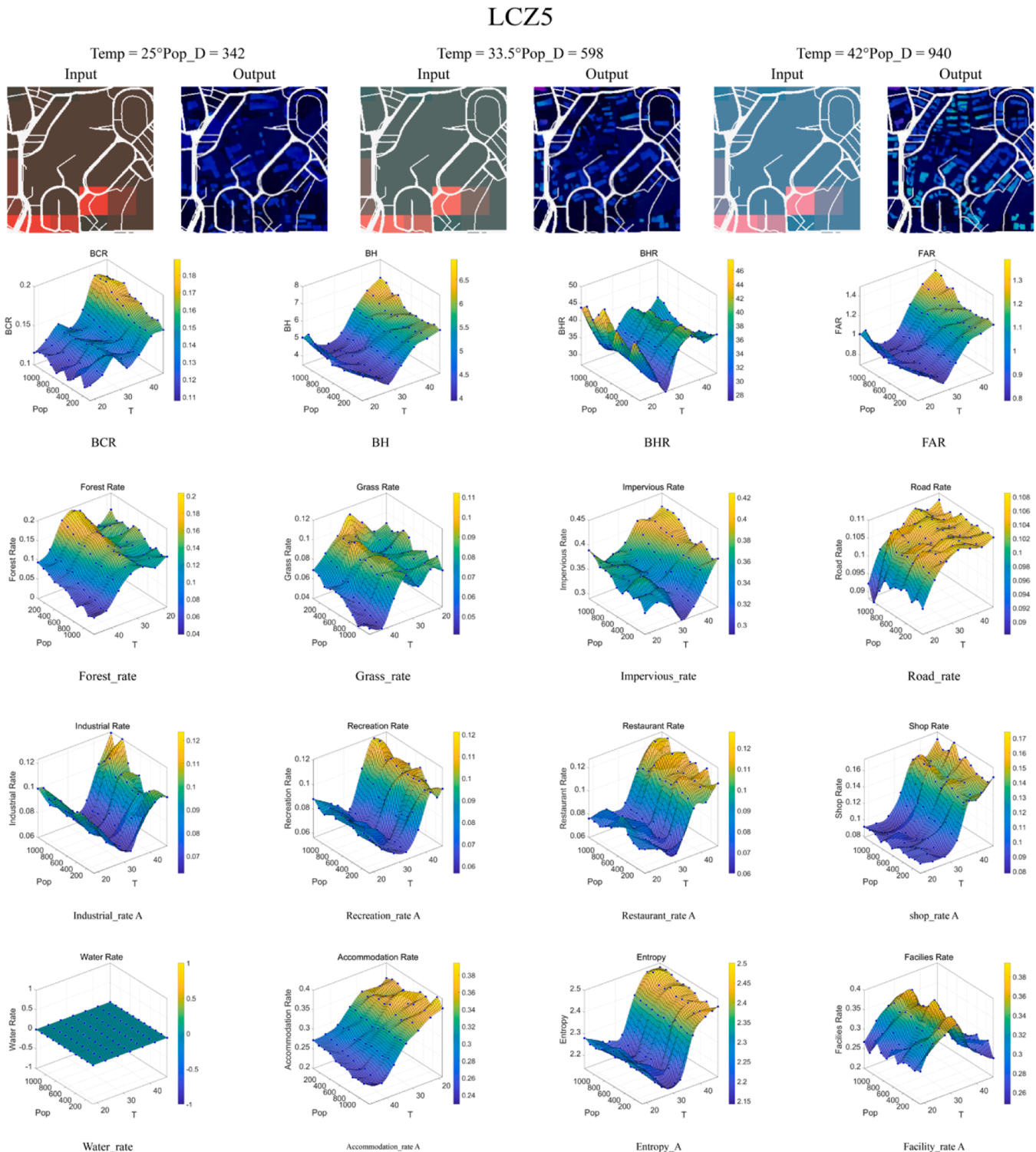


Fig. 15. Simulation analysis of changes in land use and urban functions of the LCZ5 under Scenario 3.

LCZ6

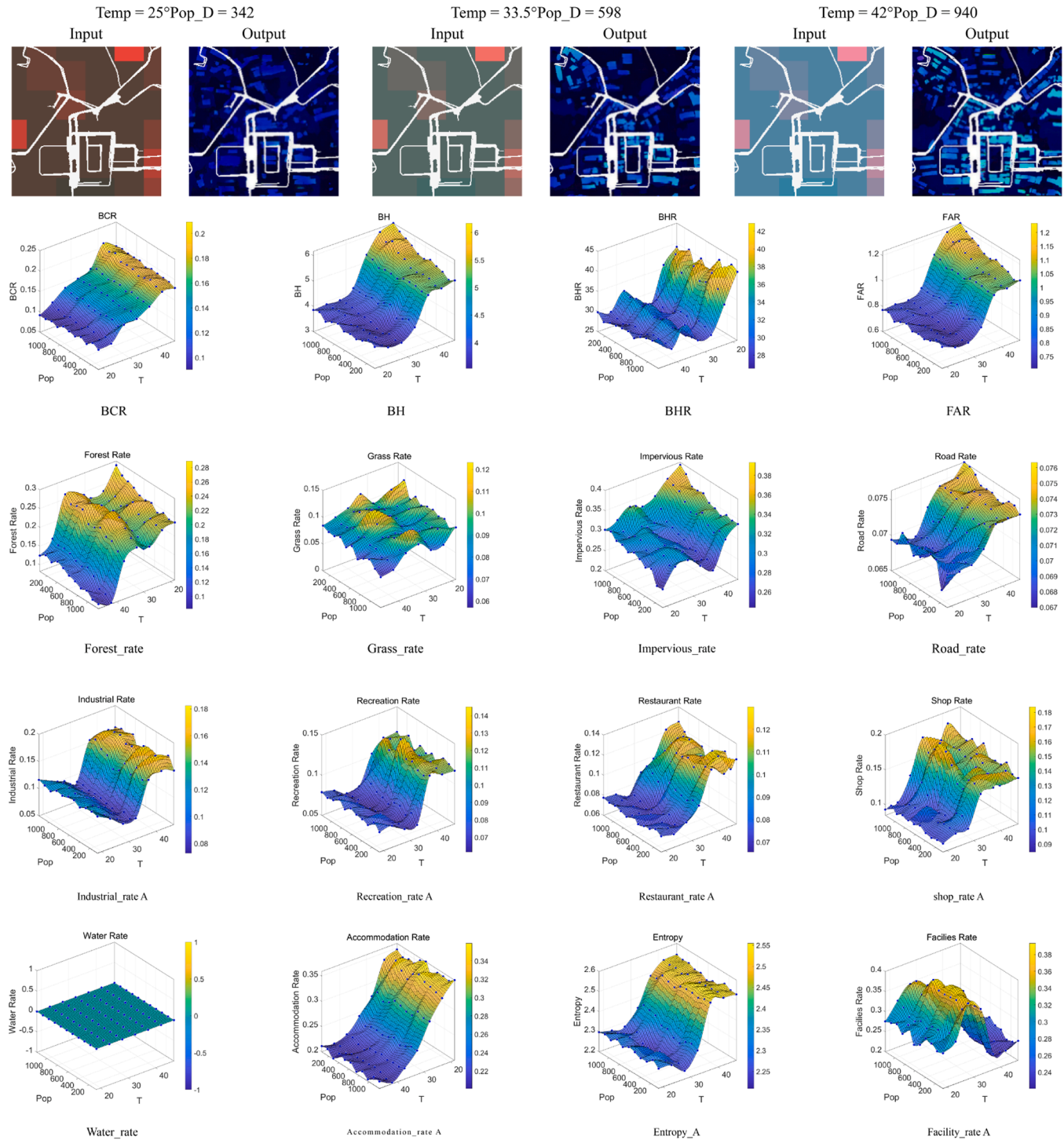


Fig. 16. Simulation analysis of changes in land use and urban functions of the LCZ6 under Scenario 3.

3.3.1. The analysis of the LCZ4

LCZ4 captures high-density urban centers with significant economic activities, and this study can gain insight into the dynamics of space use and resource allocation under different temperature and population density conditions in urban centers by analyzing this site. In terms of POI, Industrial land use under higher population density exhibited a dynamic response pattern characterized by two significant fluctuations. Specifically, the proportion of industrial land decreased from 0.105 at 18.4°C to 0.068 at 27.8°C, rises again to 0.125 at 40°C, and subsequently

declined to 0.097 at 46°C. These fluctuations could be attributed to the complex interplay between environmental constraints and socio-economic demands. Additionally, this pattern highlights the critical role of specific temperature thresholds, such as 27.8°C and 40°C, in influencing industrial land use decisions within LCZ4. Meanwhile, restaurant land use fluctuated even more prominently. The restaurant rate peaked at 0.121 at 27.8°C and a population density of 740 people/km² but decreased slightly to 0.107 at 46°C. This also reflected the phenomenon that with increasing temperature, restaurants first expand

their operations to outdoor spaces, followed by a subsequent contraction back to indoor areas.

Focusing on LUCC, the forest rate declined markedly as urban development intensified. Initially at 0.232 in cooler, low-density conditions (18.4°C, 56.9 people/km²), the forest rate decreased sharply to 0.051 at higher temperatures and densities. Although the grassland rate fluctuated with rising temperature and density, overall, it decreased from 0.113 at low temperatures to 0.062 under high temperature and density conditions. This indicated that urban green spaces in the area play a crucial role in regulating urban temperature and population density, with forests exhibiting a significantly greater impact compared to grasslands (Lei et al., 2014; Vieira et al., 2018; Zhou et al., 2019). The impervious rate also served as a key indicator, demonstrating notable changes in urban development intensity. As both temperature and population density rise, this rate increased to 0.31 at 46.6°C and 1,195 people/km², suggesting High temperature and population growth necessitate a high proportion of impervious surfaces to provide adequate support, consistent with literature on the urban heat island effect and impervious surface dynamics (Sun et al., 2022).

3.3.2. The analysis of the LCZ5

LCZ5 reflects transitional zones with mixed land-use characteristics, and this study can gain insight into the dynamics of space use and resource allocation under different temperature and population density conditions in transitional zones by analyzing this site. In terms of POI, the differentiation in industrial land use under high-temperature conditions highlighted the spatial heterogeneity of urban functions and adaptive responses. In high-density areas, the industrial land use rate accelerated to 0.124, driven by the demand for economic expansion and intensified land-use efficiency to support population growth. In contrast, in medium- and low-density areas, the industrial land use rate stabilized around 0.1, reflecting a focus on maintaining land-use stability and prioritizing environmental sustainability. This divergence underscored the role of population density as a key driver of industrial growth, shaping the varying responses of industrial land allocation to environmental pressures across different urban contexts.

Pertaining to LUCC, impervious rate generally rose with increasing temperature and population density, reflecting intensified construction activity under these conditions (Hua et al., 2020). However, there was a notable dip in the 30–35°C range, suggesting that in moderate temperatures, urban planning may prioritize comfort and ecological considerations over further construction. Exploring UMI, The BCR similarly peaked at 0.185 between 40–45°C and 900 to 1,100 people/km², underscoring heightened land development intensity under extreme conditions (Wang et al., 2024). This ratio slightly declined in the 30–35°C range, due to more green space being allocated at moderate temperatures.

3.3.3. The analysis of the LCZ6

LCZ6 represents low-density urban areas that balance built-up spaces and green cover, and this study can gain insight into the dynamics of space use and resource allocation under different temperature and population density conditions in low-density urban areas by analyzing this site. About POIs of LCZ6 parcel, the industrial land use rate in LCZ6 stabilized at 0.13 after reaching its peak around 40°C, demonstrating the resilience of low-density areas like LCZ6 in maintaining functional balance under high-temperature conditions. This contrasted with the more dynamic industrial land use changes observed in medium- and high-density zones under similar thermal stresses. The stabilization at 0.13 also suggested that, in low-density regions, this value may represent a threshold for industrial land use under such environmental conditions, highlighting the unique adaptive characteristics of LCZ6 in response to thermal and spatial constraints.

Analyzing LUCC, forest cover was better preserved in moderate-density and moderate-temperature areas, which peaked at 0.28 under densities of 400–600 people/km² and around 35°C. Meanwhile, in LCZ6,

the grassland rate remained stable at approximately 0.1, regardless of changes in temperature or population density. This stability could be attributed to the relatively simple ecological function and high adaptability of grasslands, which allow them to be distributed across diverse environmental conditions. Additionally, grasslands were often characterized by dispersed land use and low economic development potential, leading to a lower priority for conversion or development. Besides, the impervious rate increased with both temperature and density, though a slight decline occurs in the 35–40°C range.

3.3.4. Differences of these LCZs

A comparative analysis of Guangzhou's LCZ4, LCZ5, and LCZ6 highlighted both similarities and differences in the performance of UMI, POI, and LUCC under varying temperature and population density conditions. For UMI, the BCR in LCZ4 and LCZ6 showed a linear increase with rising temperatures, while LCZ5 demonstrated a minor peak at approximately 30°C, followed by a fluctuating upward trend. This variability in LCZ5 might result from increased commercial activities or short-term residential demands with this comfortable temperature. In contrast, the BH and FAR exhibited similar patterns across all three LCZ types, indicating consistent responses to temperature and population density variations.

From a LUCC perspective, LCZ4 retained relatively high grassland and forest rates under favorable temperature and low-density conditions (20°C, 200 people/km²), with values of 0.12 and 0.24, respectively. Similarly, LCZ6 maintained high ecological coverage under comparable conditions. In contrast, LCZ5 achieved notable grassland and forest coverage only near 35°C. These variations indicated that green spaces in LCZ4 areas effectively mitigate urban heat island effects. In contrast, urban green spaces in LCZ5 and LCZ6 exhibited limited effectiveness in mitigating urban heat island effects. This outcome might primarily result from the vegetation composition or potential errors inherent in the model's generation process.

On the basis of POI, LCZ6 showed a sharp increase in shop density around 30°C, whereas LCZ4 and LCZ5 exhibited a gradual and steady rise with increasing temperature. This difference was likely due to the higher flexibility of commercial layouts in low-rise areas, which allowed for rapid responses to consumer demand under favorable conditions. Conversely, commercial facilities in LCZ4 and LCZ5 were more concentrated, with longer development cycles and less adaptability to temperature fluctuations. Other POI categorized across the three LCZ types display largely consistent changes.

In conclusion, under the combined influences of temperature and population density, LCZ4, LCZ5, and LCZ6 showed similar trends in urban functions, as reflected in UMI and POI, emphasizing shared characteristics of functional integration and facility distribution. However, significant differences in LUCC revealed the diversity in urban morphology and land development potential. These findings offered critical insights for refining planning strategies across different LCZ types, highlighting the balance between functional consistency and adaptive land development.

3.4. Implementation suggestions

This study employed univariate and bivariate analyses to reveal the nuanced land-use optimization and functional distribution needs of different LCZ parcels in Guangzhou under varying temperature and population density conditions. In compact mid- and high-rise areas, such as LCZ01 and LCZ02, urban POIs tended to shift towards commercial and recreational activities in a high-temperature environment, thereby exacerbating the urban heat island effect. This dynamic underscored the necessity of heat island management strategies, including policies favoring vertical development and heat-resistant materials. Vertical expansion not only preserves ground-level green spaces but also enhances ventilation and moderates thermal conditions through the use of climate-adaptive materials (Coutts et al., 2013).

LCZ04, an open high-rise zone dominated by dense commercial and industrial activities, relies critically on green spaces for effective regulation of urban temperature. Planners can propose targeted cooling strategies, such as shaded green spaces, green roofs, and cooling corridors, to mitigate the heat island effect while enhancing user comfort and commercial viability (Bowler et al., 2010). Additionally, multifunctional green spaces that accommodate both commercial and industrial needs could be promoted within the area to counteract the expansion of impervious surfaces, improving thermal comfort and operational stability.





LCZ05, sensitive to both density and temperature, demonstrates strong resilience in transportation and commercial functions under high temperature, accompanied by reduced demand for building land and public facilities. This behavior presents a strategic opportunity to enhance green space ratios, such as the development of urban wetlands and parks, to bolster urban ecological resilience during extreme heat (Bowler et al., 2010). Furthermore, increasing the floor area ratio in extremely high-temperature and high-density scenarios, supported by

green infrastructure, can accommodate additional residential and commercial functions. This integrated approach balances development through green layouts, mitigating the impacts of high temperature and density.

Due to its open and low-density characteristics, LCZ6 demonstrates the most significant role of urban green spaces in regulating urban temperature among all types. Planners could promote vertical industrial development to minimize ecological loss while maintaining POI diversity. Prioritizing permeable surfaces and vegetative buffers would enhance thermal regulation and improve adaptability to rising population density (Fini et al., 2017; Liu et al., 2020). Additionally, government policies could implement low-density zoning to curb the expansion of impervious surfaces, preserving ecological resilience under escalating environmental pressures (Chu et al., 2022; Song et al., 2022).

In conclusion, the findings suggest a multi-layered planning approach that integrates green infrastructure, adaptive zoning, and responsive land-use policies across all LCZ types. High-density areas should prioritize green space expansion, adopt heat-resistant materials,

Table 3
Optimization implementation suggestions for different LCZ areas.

| LCZ | Design Suggestions | Schematic Diagram |
|----------------|--|---|
| LCZ01 LCZ02 | Adopt heat-resistant materials, prioritize vertical development to preserve ground-level green spaces, and improve thermal conditions. |  |
| LCZ04 | Targeted cooling strategies like shaded green spaces, green roofs, cooling corridors, and multifunctional green spaces. |  |
| LCZ05 | Develop urban wetlands and parks, increase green space ratios, and integrate green infrastructure with high-density development. |  |
| LCZ06 | Promote vertical industrial development, prioritize permeable surfaces, vegetative buffers, and implement low-density zoning. |  |

and implement vertical development to safeguard ecological functions and counteract thermal effects. Conversely, ecologically rich parcels with extensive vegetation coverage should focus on conserving native greenery to enhance resilience to rising temperatures and density. These strategies provide a balanced framework for sustainable urban environments, enabling cities to maintain economic vitality and ecological resilience amid changing environmental conditions (Table 3).

4. limitations and future research

The integrated CP-GAN-based optimization framework developed in this study demonstrates promising adaptability in adjusting urban form to optimize environmental conditions. However, certain limitations remain that warrant attention in future research. This study's development of CP-GAN by coupling Pix2pix and CycleGAN improved the model's precision. However, emerging alternatives like diffusion models, autoregressive models, normalized flow, transformer-based models, and energy-based models (EBM) show considerable promise for 3D building generation and urban planning applications, meriting further exploration as these models advance.

Besides, uncertainties in the input datasets may significantly affect the performance of the CP-GAN model. First, the LCZ classification is subject to potential misclassification errors due to limitations in remote sensing resolution and the inherent subjectivity of training sample selection. Such errors can lead to inaccuracies in representing local urban textures, which may ultimately skew the generated urban morphology. Second, LST measurements, which are derived from satellite data, can be affected by factors such as sensor noise, atmospheric interference, and the exclusion of nighttime data. These inaccuracies might compromise the fidelity of the simulated thermal environment, potentially affecting how the model interprets and responds to temperature variations. Third, the population heatmap data, derived from mobile signaling, exhibit variability due to temporal fluctuations and sampling biases. Although we mitigate this issue by averaging data collected over multiple time intervals, residual variability could still lead to inconsistencies in depicting urban vitality patterns. Future work should incorporate a more systematic uncertainty propagation analysis, possibly through sensitivity analysis or probabilistic modeling, to quantify how these input variations affect the generated urban morphologies and the subsequent urban design decisions.

In terms of model evaluating, the effectiveness of the generated urban morphology was primarily evaluated using model-based metrics such as SSIM and PSNR, which reflect structural accuracy and image quality. However, the evaluation overlooked key indicators used by urban designers to assess the practical rationality of a plan—such as functional coherence (e.g., how well the spatial layout supports mixed-use development or pedestrian connectivity) and spatial usability (e.g., walkability, accessibility, and human-scale considerations). This may lead to scenarios where the generated outcomes pass algorithmic validation but lack practical value for real-world application. Future research could integrate both model-driven evaluation metrics and design-oriented assessment criteria to provide a more comprehensive evaluation of the generated plans, ensuring that they meet not only technical standards but also the holistic demands of climate-adaptive and human-centered urban design.

Meanwhile, the generalizability of the CP-GAN model has not been fully validated across different urban environments. However, we have previously published a study that explored the applicability of the Pix2Pix model in six major global cities—London, New York, Paris, Shanghai, Sydney, and Tokyo (Zhou et al., 2023). This study demonstrated that Pix2Pix performed well across various urban environments, including different climate zones and population densities, showing promising generalizability. Given that CP-GAN integrates all the strengths of the Pix2Pix model, it is expected to exhibit similar, if not improved, generalizability. In future research, we plan to conduct additional experiments to further evaluate and validate the

generalization capabilities of CP-GAN in diverse urban contexts.

Moreover, this study prioritized optimizing urban form to enhance regional environmental resilience, focusing on surface temperature and population density. However, urban design is inherently complex, and a more holistic and accurate assessment framework for urban form is needed. Environmental factors such as urban wind flow and water currents should be incorporated into future evaluations. In a densely populated city like Guangzhou, buildings often serve multiple functions, such as residential, commercial, and hospitality. Generating accurate function-based building classifications is crucial for energy consumption estimations and represents a key area for further research.

5. Conclusion

This study developed a novel CP-GAN-based generative model chain leveraging multimodal data, including LCZ, LST, and POPH as a combination of static and dynamic datasets. By integrating these diverse data sources, the model enabled comprehensive analysis and simulation of urban dynamics, producing detailed insights into key urban metrics such as LUCC, UMI, and POI. Using Guangzhou—a subtropical, high-density city—as a case study, the model applied these parameters to six representative 600×600m parcels, capturing diverse urban characteristics aligned with specific LCZ types. The high average SSIM and FID values, exceeding 0.811 and 19.315, respectively, confirmed the model's efficacy in generating high-precision urban morphologies. This research also presented an automated multi-objective assessment and optimization framework that dynamically simulates urban morphology changes under varying temperature and population density conditions. The results indicate that (1) Rising surface temperatures profoundly reshape urban spatial morphology, with the urban heat island effect significantly influencing commercial, recreational, and industrial facilities. Establishing green spaces emerges as a critical strategy for mitigating urban heat island effects and enhancing thermal resilience in urban areas. (2) LCZs exhibit varied temperature response thresholds, underscoring the necessity of integrating thermal resilience into urban planning frameworks to effectively address temperature-sensitive urban transformations and ensure adaptive capacity. (3) The impact of population density on urban morphology is highly localized, necessitating tailored urban planning strategies that balance localized density dynamics with broader climate adaptation measures. (4) LCZ4, LCZ5, and LCZ6 share functional integration in urban morphology and POI responses to temperature and population density, yet exhibit significant differences in LUCC and ecological performance. These findings underscore the necessity of adaptive planning strategies to enhance ecological effectiveness and mitigate urban heat island effects, tailored to the unique characteristics of each LCZ type. This approach is essential for fostering sustainable and climate-resilient urban development. The study identified adaptive strategies, including green infrastructure integration, vertical expansion, and responsive zoning, to balance ecological preservation with economic activity, thereby enhancing urban resilience in densely populated areas. This research contributed an innovative, adaptable framework with valuable insights for sustainable urban development, offering practical solutions to mitigate environmental pressures in high-density cities.

CRedit authorship contribution statement

Shiqi Zhou: Writing – original draft, Software, Methodology, Formal analysis. **Xiaodong Xu:** Visualization, Validation, Formal analysis. **Haowen Xu:** Software, Data curation. **Zichen Zhao:** Software, Formal analysis. **Haojun Yuan:** Visualization, Methodology. **Yuankai Wang:** Writing – review & editing, Supervision, Formal analysis. **Renlu Qiao:** Formal analysis. **Tao Wu:** Software. **Weiye Jia:** Visualization, Software. **Mo Wang:** Supervision, Software. **Waishan Qiu:** Supervision, Formal analysis. **Zhiqiang Wu:** Supervision.

Declaration of competing interest

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

Acknowledgement

The work was supported by Basic Theory of Sustainable Urban Planning, Construction, and Governance' under the 14th Five-Year Plan of the State Key Research and Development Program of the People's Republic of China (2022YFC3800205).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.scs.2025.106452](https://doi.org/10.1016/j.scs.2025.106452).

Data availability

The data that has been used is confidential.

References

- Aithal, B. H., Das, S. K., & Subrahmanya, P. P. (2020). Urban 3D structure reconstruction through a generative adversarial network model. *Arabian Journal for Science and Engineering*, 45(12), 10731–10741. <https://doi.org/10.1007/s13369-020-04850-7>
- Albert, A., Strano, E., Kaur, J., & González, M. (2018). Modeling urbanization patterns with generative adversarial networks. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 2095–2098). <https://doi.org/10.1109/IGARSS.2018.8518032>
- Bagavathi, C., Gadde, U. K., Immadisetty, R. G., & Vasireddy, A. K. (2024). Augmented super resolution GAN (ASRGAN) for image enhancement through reinforced discriminator. In D. Pastor-Escuredo, I. Brigui, N. Kesswani, S. Bordoloi, & A. K. Ray (Eds.), *The Future of Artificial Intelligence and Robotics* (pp. 109–118). Switzerland: Springer Nature. https://doi.org/10.1007/978-3-031-60935-0_11
- Bechtel, B., Alexander, P. J., Beck, C., Böhner, J., Brousse, O., Ching, J., Demuzere, M., Fonte, C., Gál, T., Hidalgo, J., Hoffmann, P., Middel, A., Mills, G., Ren, C., See, L., Sismanidis, P., Verdonck, M.-L., Xu, G., & Xu, Y. (2019). Generating WUDAPT level 0 data – Current status of production and evaluation. *Urban Climate*, 27, 24–45. <https://doi.org/10.1016/j.uclim.2018.10.001>
- Biao, Z., Gaodi, X., Bin, X., & Canqiang, Z. (2012). The effects of public green spaces on residential property value in Beijing. *Journal of Resources and Ecology*, 3(3), 243–252. <https://doi.org/10.5814/j.issn.1674-764x.2012.03.007>
- Bińkowski, M., Sutherland, D. J., Arbel, M., & Gretton, A. (2021). *Demystifying MMD GANs* (arXiv:1801.01401). arXiv. <https://doi.org/10.48550/arXiv.1801.01401>
- Bowler, D. E., Buyung-Ali, L., Knight, T. M., & Pullin, A. S. (2010). Urban greening to cool towns and cities: A systematic review of the empirical evidence. *Landscape and Urban Planning*, 97(3), 147–155. <https://doi.org/10.1016/j.landurbplan.2010.05.006>
- Boyukliyski, S., Petrova-Antonova, D., Hristov, E., & Hristov, K. (2022). Multi-objective optimisation of urban design using a genetic algorithm. In *2022 International Conference Automatics and Informatics (ICAI)* (pp. 345–350). <https://doi.org/10.1109/ICAIS5857.2022.9959987>
- Cao, Q., Luan, Q., Liu, Y., & Wang, R. (2021). The effects of 2D and 3D building morphology on urban environments: A multi-scale analysis in the Beijing metropolitan region. *Building and Environment*, 192, Article 107635. <https://doi.org/10.1016/j.buildenv.2021.107635>
- Chu, M., Lu, J., & Sun, D. (2022). Influence of urban agglomeration expansion on fragmentation of green space: A case study of Beijing-Tianjin-Hebei Urban agglomeration. *Land*, 11(2). <https://doi.org/10.3390/land11020275>. Article 2.
- Cira, C.-I., Manso-Callejo, M.-Á., Alcarria, R., Fernández Pareja, T., Bordel Sánchez, B., & Serradilla, F. (2021). Generative learning for postprocessing semantic segmentation predictions: A lightweight conditional generative adversarial network based on Pix2pix to improve the extraction of road surface areas. *Land*, 10(1). <https://doi.org/10.3390/land10010079>. Article 1.
- Coutts, A. M., Daly, E., Beringer, J., & Tapper, N. J. (2013). Assessing practical measures to reduce urban heat: Green and cool roofs. *Building and Environment*, 70, 266–276. <https://doi.org/10.1016/j.buildenv.2013.08.021>
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1), 53–65. <https://doi.org/10.1109/MSP.2017.2765202>. IEEE Signal Processing Magazine.
- Dalla Mura, M., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., & Benediktsson, J. A. (2015). Challenges and opportunities of multimodality and data fusion in remote sensing. *Proceedings of the IEEE*, 103(9), 1585–1601. <https://doi.org/10.1109/JPROC.2015.2462751>. Proceedings of the IEEE.
- De, S., Wang, W., Zhou, Y., Perera, C., Moessner, K., & Alraja, M. N. (2021). Analysing environmental impact of large-scale events in public spaces with cross-domain multimodal data fusion. *Computing*, 103(9), 1959–1981. <https://doi.org/10.1007/s00607-021-00944-8>
- Desalegn, L., & Jifara, W. (2024). HARA-GAN: Hybrid attention and relative average discriminator based generative adversarial network for MR image reconstruction. *IEEE Access*, 12, 23240–23251. <https://doi.org/10.1109/ACCESS.2024.3364699>. IEEE Access.
- Fan, Z., Duan, J., Lu, Y., Zou, W., & Lan, W. (2021). A geographical detector study on factors influencing urban park use in Nanjing, China. *Urban Forestry & Urban Greening*, 59, Article 126996. <https://doi.org/10.1016/j.ufug.2021.126996>
- Fini, A., Frangi, P., Mori, J., Donzelli, D., & Ferrini, F. (2017). Nature based solutions to mitigate soil sealing in urban areas: Results from a 4-year study comparing permeable, porous, and impermeable pavements. *Environmental Research*, 156, 443–454. <https://doi.org/10.1016/j.envres.2017.03.032>
- Gu, X., Wu, Z., Liu, X., Qiao, R., & Jiang, Q. (2024). Exploring the nonlinear interplay between urban morphology and nighttime thermal environment. *Sustainable Cities and Society*, 101, Article 105176. <https://doi.org/10.1016/j.scs.2024.105176>
- Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T., & Mueller, A. (2012). Urban structure type characterization using hyperspectral remote sensing and height information. *Landscape and Urban Planning*, 105(4), 361–375. <https://doi.org/10.1016/j.landurbplan.2012.01.001>
- Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & 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. <https://proceedings.neurips.cc/paper/2017/hash/8a1d694707eb0fefe65871369074926d-Abstract.html>
- Hu, Y., Dai, Z., & Guldmann, J.-M. (2020). Modeling the impact of 2D/3D urban indicators on the urban heat island over different seasons: A boosted regression tree approach. *Journal of Environmental Management*, 266, Article 110424. <https://doi.org/10.1016/j.jenvman.2020.110424>
- Hua, L., Zhang, X., Nie, Q., Sun, F., & Tang, L. (2020). The impacts of the expansion of urban impervious surfaces on urban heat islands in a coastal city in China. *Sustainability*, 12(2). <https://doi.org/10.3390/su12020475>. Article 2.
- Huang, C., Zhang, G., Yao, J., Wang, X., Calautit, J. K., Zhao, C., An, N., & Peng, X. (2022). Accelerated environmental performance-driven urban design with generative adversarial network. *Building and Environment*, 224, Article 109575. <https://doi.org/10.1016/j.buildenv.2022.109575>
- Kastner, P., & Dogan, T. (2023). A GAN-based surrogate model for instantaneous urban wind flow prediction. *Building and Environment*, 242, Article 110384. <https://doi.org/10.1016/j.buildenv.2023.110384>
- Lee, D., Oh, K., & Jung, S. (2019). Classifying urban climate zones (UCZs) based on spatial statistical analyses. *Sustainability*, 11(7). <https://doi.org/10.3390/su11071915>. Article 7.
- Lei, W., Shuwen, Z., & Yunlong, Y. a. O. (2014). The impacts of green landscape on urban thermal environment: A case study in Changchun city. *Geographical Research*, 33(11). <https://doi.org/10.11821/dlyj201411009>. Article 11.
- Li, J., Li, J., Yuan, Y., & Li, G. (2019). Spatiotemporal distribution characteristics and mechanism analysis of urban population density: A case of Xi'an, Shaanxi, China. *Cities*, 86, 62–70. <https://doi.org/10.1016/j.cities.2018.12.008>
- Li, Z., Jiao, L., Zhang, B., Xu, G., & Liu, J. (2021). Understanding the pattern and mechanism of spatial concentration of urban land use, population and economic activities: A case study in Wuhan, China. *Geo-Spatial Information Science*, 24(4), 678–694. <https://doi.org/10.1080/10095020.2021.1978276>
- Liang, Z., Huang, J. X., & Antani, S. (2022). Image translation by Ad CycleGAN for COVID-19 X-ray images: A new approach for controllable GAN. *Sensors*, 22(24). <https://doi.org/10.3390/s22249628>. Article 24.
- Liu, Y., Li, T., & Yu, L. (2020). Urban heat island mitigation and hydrology performance of innovative permeable pavement: A pilot-scale study. *Journal of Cleaner Production*, 244, Article 118938. <https://doi.org/10.1016/j.jclepro.2019.118938>
- Lucas, T., Tallec, C., Ollivier, Y., & Verbeek, J. (2018). Mixed batches and symmetric discriminators for GAN training. In *Proceedings of the 35th International Conference on Machine Learning* (pp. 2844–2853). <https://proceedings.mlr.press/v80/lucas18a.html>
- Lyu, F., & Zhang, L. (2019). Using multi-source big data to understand the factors affecting urban park use in Wuhan. *Urban Forestry & Urban Greening*, 43, Article 126367. <https://doi.org/10.1016/j.ufug.2019.126367>
- Pan, T., Li, B., & Ning, L. (2023). Impervious surface area patterns and their response to land surface temperature mechanism in urban-rural regions of Qingdao, China. *Remote Sensing*, 15(17). <https://doi.org/10.3390/rs15174265>. Article 17.
- Qiao, Z., Han, X., Wu, C., Liu, L., Xu, X., Sun, Z., Sun, W., Cao, Q., & Li, L. (2020). Scale effects of the relationships between 3D building morphology and urban heat island: A case study of provincial capital cities of mainland China. *Complexity*, 2020(1), Article 9326793. <https://doi.org/10.1155/2020/9326793>
- Qiu, C., Schmitt, M., Mou, L., Ghamisi, P., & Zhu, X. X. (2018). Feature importance analysis for local climate zone classification using a residual convolutional neural network with multi-source datasets. *Remote Sensing*, 10(10). <https://doi.org/10.3390/rs10101572>. Article 10.
- Roche, J., De-Silva, V., Hook, J., Moencks, M., & Kondoz, A. (2022). A multimodal data processing system for LiDAR-based human activity recognition. *IEEE Transactions on Cybernetics*, 52(10), 10027–10040. <https://doi.org/10.1109/TCYB.2021.3085489>. IEEE Transactions on Cybernetics.
- Rodler, A., & Leduc, T. (2019). Local climate zone approach on local and micro scales: Dividing the urban open space. *Urban Climate*, 28, Article 100457. <https://doi.org/10.1016/j.uclim.2019.100457>
- Samreen, S., & Venu, V. S. (2024). Enhanced image super resolution using ResNet generative adversarial networks. *EBSCOhost*. <https://doi.org/10.18280/ts.410432>
- Song, B., Liu, P., Li, J., Wang, L., Zhang, L., He, G., Chen, L., & Liu, J. (2022). MLFF-GAN: A multilevel feature fusion with GAN for spatiotemporal remote sensing images.

- IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–16. <https://doi.org/10.1109/TGRS.2022.3169916>. IEEE Transactions on Geoscience and Remote Sensing.
- Song, Y., Xia, S., Xue, D., Luo, S., Zhang, L., & Wang, D. (2022). Land space change process and its eco-environmental effects in the Guanzhong Plain urban agglomeration of China. *Land*, 11(9). <https://doi.org/10.3390/land11091547>. Article 9.
- Stewart I. D., & Oke T. R. (2012). *Local climate zones for urban temperature studies*. <https://doi.org/10.1175/BAMS-D-11-00019.1>.
- Su, Y., Wang, Y., Wang, C., Zhao, Q., He, Y., Zhou, D., & Zhou, N. (2022). The influence of spatiotemporal radial road on the outdoor thermal environment. *Urban Climate*, 45, Article 101280. <https://doi.org/10.1016/j.uclim.2022.101280>
- Sun, C., & Rao, J. (2020). Study on performance-oriented generation of urban block models. In P. F. Yuan, Y. M. Mike Xie, J. Yao, & C. Yan (Eds.), *Proceedings of the 2019 DigitalFUTURES* (pp. 179–188). Springer. https://doi.org/10.1007/978-981-13-8153-9_16.
- Sun, Z., Li, Z., & Zhong, J. (2022). Analysis of the impact of landscape patterns on urban heat islands: A case study of Chengdu, China. *International Journal of Environmental Research and Public Health*, 19(20). <https://doi.org/10.3390/ijerph192013297>. Article 20.
- TorchMetrics K.L. (2022). *Welcome to TorchMetrics—PyTorch-Metrics 1.6.1 documentation*. <https://torchmetrics.readthedocs.io/en/latest>.
- Unger, J., Skarbit, N., & Gál, T. (2018). Evaluation of outdoor human thermal sensation of local climate zones based on long-term database. *International Journal of Biometeorology*, 62(2), 183–193. <https://doi.org/10.1007/s00484-017-1440-z>
- USGS. (2013). *Landsat 8-9 Operational Land Imager /Thermal Infrared Sensor Level-2, Collection 2*. U.S. Geological Survey. <https://doi.org/10.5066/P9OGBGM6>. Dataset.
- USGS. (2023). *Landsat U.S. analysis ready data | U.S. geological survey*. <https://www.usgs.gov/landsat-missions/landsat-us-analysis-ready-data> /Dataset.
- Vieira, J., Matos, P., Mexia, T., Silva, P., Lopes, N., Freitas, C., Correia, O., Santos-Reis, M., Branquinho, C., & Pinho, P. (2018). Green spaces are not all the same for the provision of air purification and climate regulation services: The case of urban parks. *Environmental Research*, 160, 306–313. <https://doi.org/10.1016/j.envres.2017.10.006>
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612. <https://doi.org/10.1109/TIP.2003.819861>. IEEE Transactions on Image Processing.
- Wang, Z., Peng, Y., Li, Y., Zhou, X., & Xie, Y. (2024). Exploration of influencing factors of land surface temperature in cities within the Beijing–Tianjin–Hebei region based on local climate zone scheme. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 9728–9744. <https://doi.org/10.1109/JSTARS.2024.3396167>. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
- Wu, A. N., & 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. <https://doi.org/10.1016/j.isprsjprs.2022.11.005>
- Xu, Y., Yang, J., Zheng, Y., & Li, W. (2024). Impacts of two-dimensional and three-dimensional urban morphology on urban thermal environments in high-density cities: A case study of Hong Kong. *Building and Environment*, 252, Article 111249. <https://doi.org/10.1016/j.buildenv.2024.111249>
- Xu, Y., Zhang, C., & Hou, W. (2022). Modeling of daytime and nighttime surface urban heat island distribution combined with LCZ in Beijing, China. *Land*, 11(11). <https://doi.org/10.3390/land11112050>. Article 11.
- Yin, Y., Grundstein, A., Mishra, D. R., Ramaswamy, L., Hashemi Tonekaboni, N., & Dowd, J. (2021). DTEx: A dynamic urban thermal exposure index based on human mobility patterns. *Environment International*, 155, Article 106573. <https://doi.org/10.1016/j.envint.2021.106573>
- Yu, B., Liu, H., Wu, J., Hu, Y., & Zhang, L. (2010). Automated derivation of urban building density information using airborne LiDAR data and object-based method. *Landscape and Urban Planning*, 98(3), 210–219. <https://doi.org/10.1016/j.landurbplan.2010.08.004>
- Zhao, F., Tang, L., Qiu, Q., & Wu, G. (2020). The compactness of spatial structure in Chinese cities: Measurement, clustering patterns and influencing factors. *Ecosystem Health and Sustainability*, 6(1), Article 1743763. <https://doi.org/10.1080/20964129.2020.1743763>
- Zheng, Y., Ren, C., Xu, Y., Wang, R., Ho, J., Lau, K., & Ng, E. (2018). GIS-based mapping of Local Climate Zone in the high-density city of Hong Kong. *Urban Climate*, 24, 419–448. <https://doi.org/10.1016/j.uclim.2017.05.008>
- Zhou, S., Diao, H., Wang, J., Jia, W., Xu, H., Xu, X., ... Wu, Z. (2025). Multi-stage optimization framework for synergetic grey-green infrastructure in response to long-term climate variability based on shared socio-economic pathways. *Water Research*, 274, 123091.
- Zhou, S., Jia, W., Diao, H., Geng, X., Wu, Y., Wang, M., ... Wu, Z. (2025). A CycleGAN-Pix2pix framework for multi-objective 3D urban morphology optimization: enhancing thermal performance in high-density areas. *Sustainable Cities and Society*, 106400.
- Zhou, S., Liu, Z., Wang, M., Gan, W., Zhao, Z., & Wu, Z. (2022). Impacts of building configurations on urban stormwater management at a block scale using XGBoost. *Sustainable Cities and Society*, 87, 104235.
- Zhou, S., Wang, Y., Jia, W., Wang, M., Wu, Y., Qiao, R., & Wu, Z. (2023). Automatic responsive-generation of 3D urban morphology coupled with local climate zones using generative adversarial network. *Building and Environment*, 245, Article 110855. <https://doi.org/10.1016/j.buildenv.2023.110855>
- Zhou, W., Cao, F., & Wang, G. (2019). Effects of spatial pattern of forest vegetation on urban cooling in a compact megacity. *Forests*, 10(3). <https://doi.org/10.3390/f10030282>. Article 3.
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 2242–2251). <https://doi.org/10.1109/ICCV.2017.244>
- Zhu, Y., Wang, Y., Li, J., Song, Q., Chen, D., & Qiu, W. (2025). BikeshareGAN: Predicting dockless bike-sharing demand based on satellite image. *Journal of Transport Geography*, 126, 104245.
- Zullo, F., Fazio, G., Romano, B., Marucci, A., & Fiorini, L. (2019). Effects of urban growth spatial pattern (UGSP) on the land surface temperature (LST): A study in the Po Valley (Italy). *Science of The Total Environment*, 650, 1740–1751. <https://doi.org/10.1016/j.scitotenv.2018.09.331>