



Original Article

Development of an artificial intelligence model for CFD data augmentation and improvement of thermal environment in urban areas using nature-based solutions

Junghyeon Ahn ^{a,b,1}, Jaekyoung Kim ^{c,1}, Junsuk Kang ^{a,d,e,*}

^a Department of Civil and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea

^b JH Solution Co., Ltd., No.1701, 135 Gasan digital 2-40, Geumcheon-gu, Seoul 08504, Republic of Korea

^c Department of Environmental Landscape Architecture, Gangneung-Wonju National University, Gangneung 25457, Republic of Korea

^d Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul 08826, Republic of Korea

^e Department of Landscape Architecture and Rural Systems Engineering, Seoul National University, Seoul 08826, Republic of Korea

ARTICLE INFO

Keywords:

Urban Climate Resilience
Machine Learning
Urban Greening Plans
Urban Heat Islands
Computational Fluid Dynamics
Green Space Optimization

ABSTRACT

Heatwaves have a significant impact on urban areas, driving efforts to mitigate the urban heat island (UHI) effect through green infrastructure and sustainable planning. By integrating computational fluid dynamics (CFD) with digital twin technology, this study evaluates the effectiveness of climate adaptation infrastructures in urban areas. However, applying digital twin technology for UHI analysis and integrating data into actionable insights faces challenges due to long simulation times and focus of analysis. This study aimed to mitigate the societal impacts of urban heat islands and address the gaps in existing research and technology. A new machine learning model was developed to improve the urban thermal environment by optimizing green spaces and combating urban heat islands in densely populated cities, by integrating artificial intelligence (AI) and digital twin technology. Combining the strengths of Random Forest and XGBoost, the model was trained and tested on a dataset derived from CFD simulations to identify effective strategies for urban green spaces allocation. The primary results of the study are divided into three parts. First, a high-precision model for data augmentation and green space optimization was developed using machine learning. Second, the developed model reduced the time required for CFD simulation analysis from over 400,000 h to less than 1 h. Finally, the study found that the strategic placement of green spaces could result in approximately 1 % of the total urban area temperature. The results highlight the importance of strategic planning in the distribution of urban green space for effective mitigation of heat islands. The proposed model can be used as an efficient tool for sustainable urban development and is consistent with the overall goal of creating more livable and climate-resilient cities.

1. Introduction

Climate change has triggered a variety of global impacts, manifested in changes to cities, ecosystems, health, and socioeconomic structures, with Earth's average temperature increasing by approximately 0.89 °C (Bosello et al., 2007; Clarke et al., 2022; Gasper et al., 2011; Grimm et al., 2013; Kim and Kang, 2022; Mimura, 2013; Pearce-Higgins et al., 2022). The IPCC highlights the need for climate protection and adaptation strategies in its AR5 and AR6 reports (Arias et al., 2021; Masson-Delmotte et al., 2021; Pachauri et al., 2014; Shukla et al., 2019). Innovations such as digital twins, AI, big data, and sensor technologies

are pivotal in improving urban climate resilience (Beckett, 2022; Fu et al., 2022; Henriksen et al., 2022; Kim and Kang, 2023a; Kim et al., 2023; Rowland et al., 2022; Ye et al., 2023). Heatwaves have a significant impact on urban areas, driving efforts to mitigate the urban heat island (UHI) effect through green infrastructure and sustainable planning, as is the case in cities worldwide, including South Korea (Bae and Lee, 2020; Depietri et al., 2012; Hambrecht et al., 2022; Jang and Kim, 2021; Kim and Kang, 2020, 2023b; Kim et al., 2022; Rentschler et al., 2023).

These strategies, aimed at improving urban sustainability, include adopting green buildings, utilizing sustainable materials, and

* Corresponding author at: Department of Civil and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea.

E-mail address: junkang@snu.ac.kr (J. Kang).

¹ Equal contribution as a first author

emphasizing increasing focus on Nature-based Solutions (NbS) for urban resilience (Fang et al., 2023; Kim et al., 2020). Adaptation efforts for vulnerable populations have introduced a variety of measures, from architectural changes to outdoor cooling solutions, to mitigate the health impacts of heatwaves (Lowe et al., 2011; Parmesan et al., 2022; Thompson et al., 2023; Vu et al., 2019). CFD, integrated with digital twin technology, delivers detailed climate vulnerability assessments and enables the effectiveness of climate adaptation infrastructures, including NbS, without significant costs (Calautit et al., 2013; Hua et al., 2021; Jang et al., 2022; Rodriguez and Velazquez, 2019; Toparlar et al., 2015).

However, applying digital twin technology for UHI analysis and integrating data into actionable insights faces challenges due to long simulation times and focus of analysis, highlighting the need for advances in urban climate analysis and planning methodologies.

To mitigate the societal impacts of UHI and address the gaps in existing research and technology, this study focused on overcoming extended simulation times associated with digital twin technology at the urban scale. By integrating AI and digital twin technology, we aimed to significantly reduce analysis times and thus enable more effective and timely urban planning and climate adaptation strategies. This integration strategy aims to enable the formulation of highly efficient urban climate adaptation plans, leveraging the combined strengths of AI and CFD to provide optimized solutions to improve urban resilience. Through this approach, this research seeks to advance the practical application of digital twin and AI technologies and promote faster and more effective responses to the challenges posed by urban climate phenomena.

The overall goal of this research was to leverage advanced computing and AI methods for urban climate resilience. The objectives of this study were as follows:

- Develop an advanced CFD-based digital twin framework that enables detailed analysis of UHI phenomena in Jeonju by incorporating high-resolution data for realistic urban climate simulations.
- Construct an AI-driven predictive model that accelerates CFD simulation processes, reducing the time required for urban heat analysis and enabling real-time scenario planning.
- Develop an AI algorithm that can analyze CFD data to identify the most effective urban areas for heat island mitigation, with a focus on strategic placement of green infrastructure.
- Conduct a comprehensive assessment of UHI mitigation strategies by simulating the improvement of target green infrastructure areas under extreme climate conditions, providing actionable insights for urban planners and policymakers.

The findings of this study will improve our understanding of how we make cities more livable by addressing heat challenges, which in turn will lead to better practices for building environmentally friendly and climate-smart urban spaces.

The novelty of this study lies in the integration of AI with CFD to create a digital twin framework aimed at the detailed analysis of UHI phenomena. Our approach is innovative by developing a unique AI-driven prediction model that significantly accelerates simulation processes while increasing the precision of urban climate analyses. This model facilitates the strategic planning and implementation of green infrastructure for effective heat mitigation and provides urban planners and policymakers with actionable insights for real-time decision-making.

Furthermore, this study goes beyond the traditional approach of simply identifying UHI hotspots. Instead, it optimizes the placement of urban green infrastructure through a comprehensive process that considers how to minimize UHI effects during the urban design phase. The focus is not merely on locating existing heat concentration areas but rather on determining the locations that will have the greatest impact on mitigating UHI under the same environmental and urban conditions. This approach contributes to improving urban climate resilience and

highlights the development of more livable, ecologically sustainable, and climate-adapted urban environments.

2. Literature Review

To better understand UHI effects and climate vulnerability, this study summarizes key findings from related fields while pinpointing the limitations of current methods. Research trends show that while there is growing knowledge about factors that contribute to UHI and climate vulnerability, a gap exists in applying these findings to develop rapid, actionable urban planning strategies (Li et al., 2020; Hsu et al., 2021; Lee et al., 2022a).

Studies in South Korea have provided insights into the effects of urbanization and population density on heatwave-related mortality (Lee et al., 2022b). However, these studies often lack the integration of advanced simulation techniques that could provide more dynamic urban planning solutions (Kim et al., 2019). The South Korean Ministry of Environment's strategies emphasize the need for tailored, region-specific responses, but there is a clear need for innovative tools that can account for local variations in climate adaptation (Ministry of Environment, 2018).

When examining climate vulnerability, Ford et al. (2018) call for interdisciplinary approaches, which this study aimed to address through a novel integration of CFD simulation and machine learning. While previous CFD studies have shed light on the thermal dynamics of urban designs (Nazarian and Kleissl, 2015; Antoniou et al., 2019), they are often hampered by extensive computational requirements and prolonged simulation times, making them less practical for rapid urban planning needs. Traditional CFD models also tend to focus narrowly on specific environmental factors without integrating socio-economic data, limiting their applicability in comprehensive urban planning.

Similarly, although digital twin technologies hold promises for urban climate modeling, they face challenges in data integration, model accuracy, and scalability, which can reduce their effectiveness in large-scale urban applications (Batty, 2018; Park et al., 2024). The computational intensity and time constraints associated with creating and updating digital twins can hinder their use in scenarios requiring swift decision-making (Shao and Kibira, 2018; Jones et al., 2020). These limitations highlight a critical need for advanced tools that can overcome constraints and offer a more holistic analysis.

The incorporation of machine learning into urban climate studies is still in its infancy. Research shows its potential but also calls for more integrated, sophisticated applications (Kafy et al., 2021, 2022; Lee et al., 2023). This research responds by offering a hybrid analytical approach that blends CFD and machine learning to improve the classification of local climates and simulation times, thus supporting sustainable urban development (Zhang et al., 2021; Oliveira et al., 2022).

In the realm of Nature-based Solutions (NbS), while the benefits to urban resilience are widely recognized (Dhyani et al., 2020; Seddon et al., 2021), there is a lack of quantifiable evidence to guide policy-making (Bayulkun et al., 2021). This review highlights the need for a methodical assessment of multiple benefits of NbS's, which this research addresses by integrating NbS into our digital twin framework to quantitatively assess their impact on urban climate resilience (Augusto et al., 2020; Hayes et al., 2022).

In summary, while CFD simulations and digital twins offer valuable insights into urban thermal dynamics, their current limitations in simulation length and focus of analysis hinder their effectiveness in practical applications. The literature reveals an urgent need for advanced, integrated tools that can rapidly translate complex urban climate data into practical, scenario-based planning. This study aimed to address this need by developing a CFD-based digital twin framework augmented by AI, designed to expedite simulations and broaden the analytical focus, thereby supporting more effective and timely urban environmental planning.

3. Methods

3.1. Workflow

The research process was divided into five phases, as shown in Fig. 1. The first phase included a comprehensive analysis of related studies to ensure the distinctiveness of this research, as discussed in detail in Section 2. The second phase focused on site analysis in the city of Jeonju, South Korea, which is highly vulnerable to heatwaves. The information available on Jeonju City can be divided into spatial and environmental categories, both of which can be accessed through open platforms operated in South Korea. This data serves as the foundation for developing a digital twin, integrating spatial information included land use, building, topography, and location data. Environmental information included annual air temperature, annual relative humidity, wind direction, and wind speed.

The third phase was about creating highly reliable simulations using

CFD, which is a critical component of the digital twin framework. The programs used to create the simulation models were Rhino 7.0 and STAR-CCM+ 2310. Appropriate physical models were selected, and boundary conditions were defined to conduct the simulations, ensuring the digital twin accurately represents the city's environmental dynamics. The first CFD model was subjected to a transient analysis, that included a time-variable analysis. This type of analysis is vital for capturing dynamic changes in environmental conditions over time, allowing for a more comprehensive validation of the simulation against the measured environmental data from August 26, 2023. The second process involves generating AI training data through CFD simulation, primarily to obtain results more quickly. The boundary conditions provided in the simulation include temperature, relative humidity, wind direction, wind speed, and solar energy, ensuring a comprehensive dataset for AI learning within the digital twin.

The fourth phase aimed to identify the most effective areas for land cover improvement using two different AI models: Random Forest and

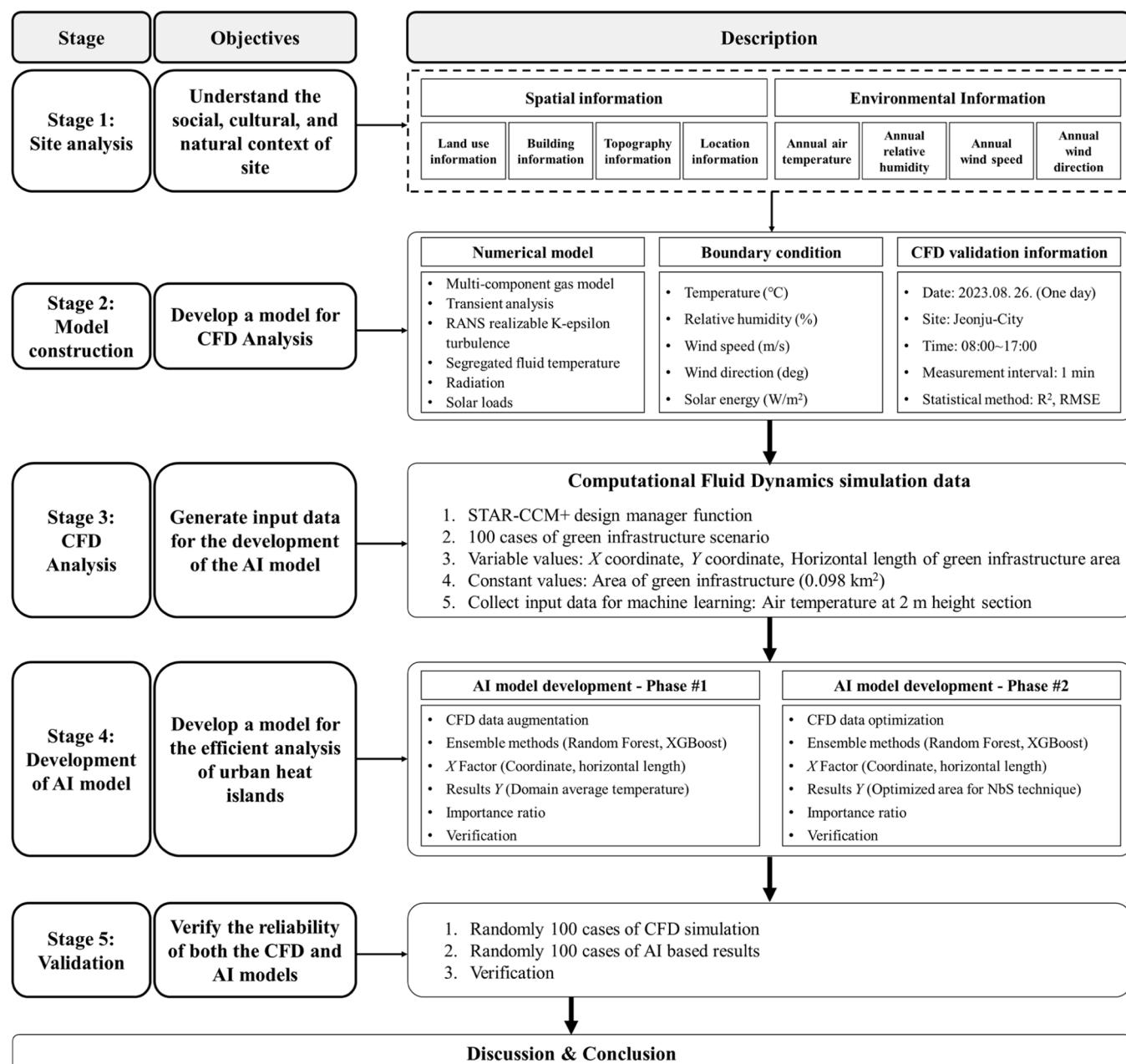


Fig. 1. Schematic sequence of study structure.

XGBoost. The first approach involved developing a data augmentation AI model, utilizing 100 data points for augmentation purposes. The augmentation data generated through this process was validated to ensure its reliability and suitability for further analysis. This model generated hundreds of thousands of additional data points required for AI training from only these 100 data points. Data augmentation was essential to expand the dataset, enhance the diversity of training examples, and improve the models' ability to generalize.

The second approach aimed to identify the exact locations and dimensions where converting impermeable surfaces, other than buildings (i.e., concrete and asphalt), into green spaces would have a significant impact on temperature reduction, while maintaining the shape of these surfaces consistently rectangular. In this study, "green spaces" refers to areas covered with grass, taking into account the assumptions of the Penman-Monteith equation. Specifically, this grass is characterized by a height of 0.12 m and a surface resistance of 70 s/m, which are important parameters for calculating evapotranspiration in the model.

Random Forest (RF), as introduced by Breiman (2001), is a versatile algorithm that can be employed for both classification and regression tasks. The method begins by evaluating all data at the root node, estimating each predictor variable to determine its ability to split the data into different nodes. Typically, tree-based methods involve a pruning process to reduce the tree's size and prevent overfitting, often achieved through cross-validation.

On the other hand, XGBoost is a machine learning system designed to scale tree boosting algorithms, recently becoming one of the most popular and effective methods. It builds a boosting ensemble of weak classification trees by applying gradient descent to optimize the loss function (Cui et al., 2017). This algorithm is known for its efficiency in reducing computation time and its ability to handle both regression and classification problems.

The reason for selecting these two models in an ensemble is that both offer stable predictive performance and effective overfitting prevention through ensemble techniques. Moreover, they are particularly strong in handling non-linear problems. XGBoost's fast training speed makes it well-suited for large datasets, and by combining these models, reliable predictive results were achieved across various datasets.

These models were chosen over others due to their ability to handle large-scale data efficiently, robustness to overfitting—a critical consideration given the potential redundancy in augmented data—and their effectiveness in modeling complex interactions between variables. Alternative models, such as neural networks or support vector machines, were considered. However, they were either computationally less efficient for our dataset size or did not provide the same level of interpretability and robustness needed for practical urban planning applications.

3.2. Target site analysis

The study location was the city of Jeonju in Southwest Korea, with 640,000 inhabitants. During the 2022 summer from June 1 to August 31, the average temperature was 26.2 °C, with the highest recorded temperature reaching 36.1 °C, highlighting significant seasonal temperature extremes. The annual precipitation in 2022 was 1313.1 mm. A wide river flows across the city and over 75 % of the land has terrain with a slope of 15° or less, indicating a generally flat topography. Located at 127°E longitude and 35°N latitude, Jeonju is considered as one of the most affected by UHIs, according to the vulnerability index (Ministry of Environment, 2018). The index is a composite measure weighted in 2:1:1 ratio for climate exposure, sensitivity, and adaptability, respectively. Within the city, Wansan District had the highest vulnerability in the country with a value of 0.61.

Jeonju City has implemented policies such as IoT-based Digital Forest for managing the Garden City of "Ten Million Trees," low impact development, and smart irrigation for rainwater management in response to heatwaves. Among them, the urban street tree planting

project known as the "Ten Million Trees Planting Project" has attracted attention as one of the important policies for heatwave mitigation.

Four criteria were established to select the detailed area for the CFD model in Wansan District,

the most vulnerable area of Jeonju City:

- The target area must be at least 2 km².
- The urban population density should be > 3000 inhabitants/km², indicating a high degree of urbanization.
- The new and old parts of the city within the target site should have a clear distinction.
- The impermeability rate should exceed 70 %, making it susceptible to UHIs and similar problems.

The 2 km² size criterion for the study area is based on previous studies on UHI, which typically range between 30,000 m² and 250,000 m² (Acero and Herranz-Pascual, 2015; Jang and Kim, 2021; Nasrollahi et al., 2017a). The scale chosen here is approximately 8–65 times larger compared to previous studies and ensures that the simulation encompasses sufficient environmental diversity, such as green spaces, impermeable surfaces, and the surrounding context. This scale is optimized for simulation efficiency; sizes larger than 3 km² can significantly increase the simulation time and the potential for errors due to a higher grid density. A population density of over 3000 inhabitants/km² indicates highly urbanized areas, which are critical for UHI analysis. A clear distinction between the new and old parts of the city within the target site is essential to map the different urban planning approaches and their respective impacts on the local microclimates, and to enable a comprehensive analysis of the UHI effects. An impermeability rate above 70 % targets regions most vulnerable to UHI, as lower rates may not demonstrate the full impact.

The region defined according to these criteria is shown in Fig. 2. The size of the target area is 2.7 km², and is located at 35° 48' 07.0"-35° 49' 08.6"N and 127° 05' 37.3"-127° 06' 49.2"E (Fig. 2(b)) with a population density of 3676.2 people/km². The target site includes key public institutions such as the Jeollabuk-do Provincial Office, Council, and Board of Education, and is home to many residents, making the region with a high climate exposure and sensitivity. Of the total area of 2.7 km², 77.3 % is impermeable, and 22.7 % is permeable.

3.3. CFD

In this study, Rhino 7.0 was used to create the CFD model (McNeel et al., 2023). In this study, the Grasshopper code was used to convert 2D CAD information into a 3D model. The CFD program used to achieve the research objectives was STAR-CCM+ 2310. The STAR-CCM+ 2310 version is the latest program, released on October 17, 2023. STAR-CCM+ 2310 is a comprehensive software tool developed by Siemens for simulation and analysis, offering multidisciplinary simulation capabilities, including CFD, computational solid mechanics, heat transfer, particle dynamics, reactant flow, electrochemistry, acoustics, and rheology (Siemens PLM Software, 2023).

To construct the CFD model domain, the components were broadly categorized into Green, Road, Water, Building, and Concrete pavement, each applying solar radiation reflectance values. Green spaces accounted for approximately 6.4 % of the site area (Fig. 3(a)). In addition, the buffer zones were modeled to accurately represent the thermal flow and reduce the effects of boundary condition (Fig. 3(b)). Buffer zones were established around the target site at intervals of approximately 1 km in each cardinal direction: east, west, north, and south. This approach minimizes airflow reflection and interference at the domain edges, thereby increasing simulation accuracy (Franke et al., 2004; Tominaga et al., 2008). The tallest building within the target site was approximately 170 m tall, and all ground surfaces were modeled taking the actual terrain into account. The CFD model reflected the actual digital elevation model (DEM) and was designed as a closed system, extending

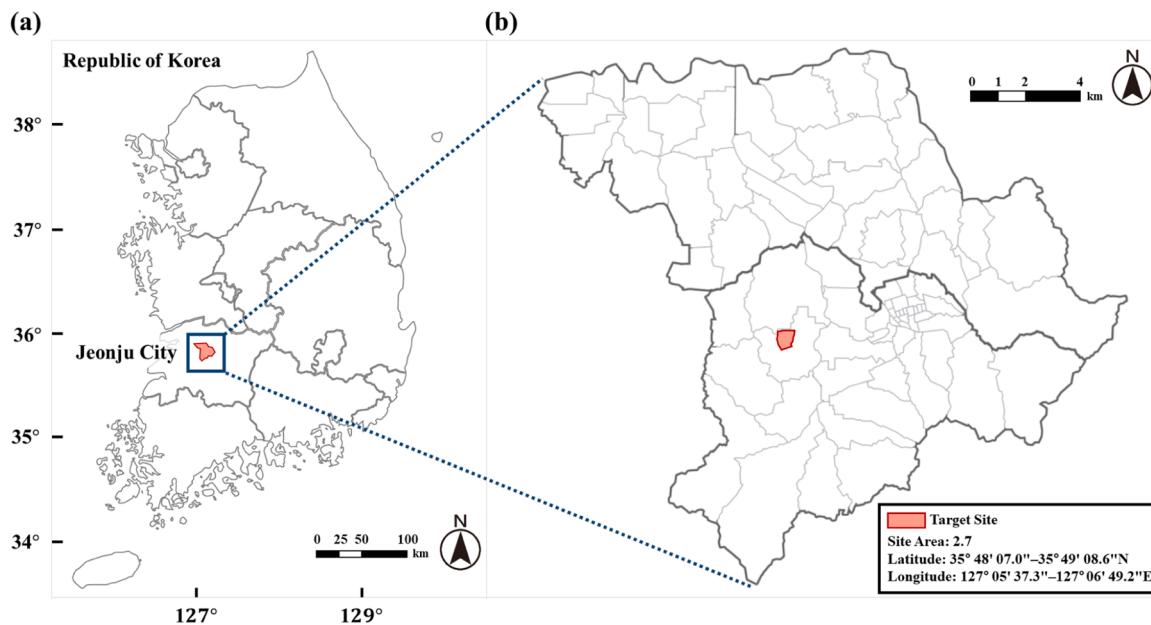


Fig. 2. Detailed scaled site analysis: (a) Location of Jeonju City, (b) Location of target site.

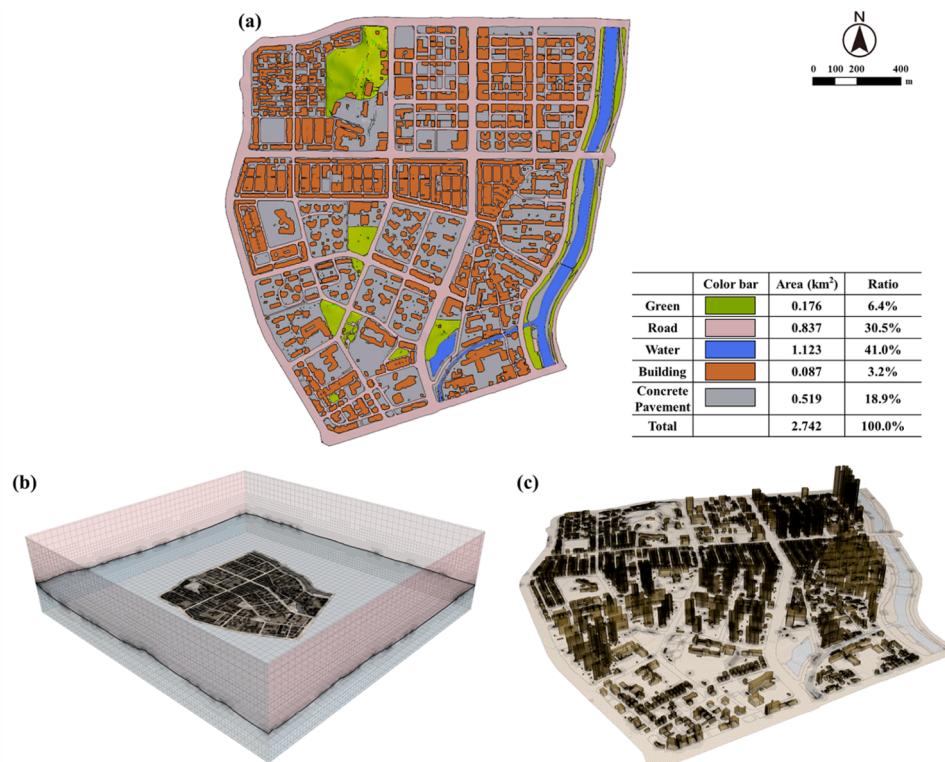


Fig. 3. Computer modeling of the target area: (a) Zoning of the target site, (b) Entire mesh of the domain area, and (c) Aerial mesh of the target site.

up to 650 m above ground and 300 m below ground to consider soil heat storage. The entire domain was modeled as a cuboid, with all sides designated as inlets and outlets to reflect changing environmental conditions. The total size created was $3750 \times 3500 \times 980$ (m), with a simulation mesh comprising 12,807,232 trimmed cells (Fig. 3(c)). In this study, the average temperature was determined at a height of 1.5 m. This height corresponded to the third data point above the cell near the wall cell and consisted of a prism layer, achieving stable simulation results. The detailed mesh configuration values based on previous

studies are listed in Table 1.

To analyze the solar radiation and simulation flow, the main physical models in STAR-CCM+ were selected as shown in Table 2, and the detailed physical model values entered for the topography and buildings, which represent typical values for the case study city, were determined as shown in Table 3 (Kim and Lee, 2024). The interior of a building is assumed to be filled with air. The thermal conductivity of air inside buildings is set at 0.0260305 W/mK to simulate heat retention, while concrete is assigned a thermal conductivity of 1.7 W/mK with a

Table 1

Automated mesh parameters.

Base Size	20.0 m			
Prism Layer Total Thickness	10 (2.0 m)			
Volume Growth Rate	Default Growth Rate	Medium (2 layers per cell size)		
	Surface Growth Rate	Medium (2 layers per cell size)		
Maximum Cell Size	1000.0 (200.0 m)			
Target Surface Size	Building 20.0 (4.0 m)	Outer wall 1000.0 (200.0 m)	Concrete pavement 20.0 (4.0 m)	Topography -
Prism Layer Total Thickness	5.0 (1.0 m)	-	-	75.0 (15.0 m)
Minimum Surface Size	2.0 (0.4 m)	50.0 (10.0 m)	2.0 (0.4 m)	-

Table 2

Main physics model in CFD simulations.

Numerical Model	Description
Cell Quality	Improves solutions for low-quality meshes.
Remediation	
Gradients	Allows customization of gradient computation methods.
Gray Thermal Radiation	Simulates wavelength-independent diffuse radiation.
Implicit Unsteady	Compatible with separate flow and energy models for time-dependent analysis.
k-ε Turbulence	Computes turbulent kinetic energy and dissipation to model turbulence.
Multi-Component Gas	Models mixtures of multiple gases in a single phase.
Reynolds-Averaged	Provides equations for modeling average flow rates in turbulent flow.
Navier-Stokes	
Segregated Flow	Solve mass and momentum conservation equations sequentially.
Solar Loads	Models both direct and diffuse solar radiation loads.
Solution Interpolation	Provides various interpolation methods for solution data during mesh refinement.
Surface-to-Surface	Uses pre-computed view factors for radiant heat exchange calculations between surfaces.
Radiation	
Three Dimensional	Designed for modeling in three-dimensional mesh environments.
View Factors Calculator	Facilitates the calculation of view factors for radiation modeling.

Table 3

Physical and thermal properties of the topography and buildings.

	Topography	Building Wall	Concrete Air
Density	1500 kg/m ³	2400 kg/m ³	1.18415 kg/m ³
Specific heat	1842.0 J/kg - K	882 J/kg - K	1003.62 J/kg - K
Thermal Conductivity	2.6 W/m - K	1.765 W/m - K	0.0260305 W/m - K
Thermal Resistance	-	0.247934 m ² K/W	

thermal transmittance of 0.1765 and an assumed thickness of 0.3 m is used to ensure accurate reflection of the different thermal behavior of these materials. Therefore, the properties of the air inside the building and the values of the thermal resistance of the walls are described. The primary physical model for CFD simulation involves the application of URANS equation with a feasible k-ε turbulence model (Shih et al., 1995). This turbulence model has been successfully used in other validated CFD studies of urban environment simulations (Allegrini and Carmeliet, 2017, 2018; Antoniou et al., 2019; Topalar et al., 2018a, 2018b).

In the CFD simulation, the initial conditions for temperatures in different parts were set as follows: Topography at 16 °C, Buildings at 20 °C, Green areas at 20 °C, Roads at 25 °C, and Concrete pavement at 25

°C. In the land domain, a temperature gradient was observed up to a depth of approximately 15 m, resulting from the temperature difference between the surface and the central area of the land. Below this zone, the temperature stabilizes at approximately 16 °C. To reproduce this condition, a tabular approach was used in the original land configuration: a temperature gradient was established from the surface down to the 15-m mark, while depths beyond that were kept constant at 16 °C. This setup more realistically represented the effects of the land's thermal mass. By incorporating a field function, evapotranspiration was induced in the river and green spaces. The temporal fluctuations in solar radiation and temperature were automatically adjusted using an auto-calculation function. An unsteady model with a spin-up time of approximately 2 h was used for the verification. The spin-up time was set to allow the formation of stable flow patterns within the model and to achieve equilibrium in terms of velocity, pressure, and temperature (Kazdenov et al., 2023; Kim and Kang, 2023a; Pirker and Lichtenegger, 2018; Topalar et al., 2015). The field function equations and boundary conditions used in this study are as follows. In order to initially take moisture aspects into account, the multi-component technology provided by STAR-CCM+ was used. To convert the composition of air and vapor into recognizable forms of moisture, an empirical formula for the necessary saturation vapor pressure was adopted, following the method of Buck (1981).

$$p_{sat} = 611.21e^{\left(\frac{18.678 - \frac{T}{234.5}}{257.14 + T}\right)T}, \quad (1)$$

where e represents the base of the natural logarithm, P_{sat} refers to the saturation vapor pressure (Pa) and T is the temperature (°C).

The most significant factor affecting the temporal temperature fluctuations within the area of interest was solar irradiance, for which the solar load technique provided by the STAR-CCM+ radiation model was used. The necessary parameters for this technique include location-specific information and direct/diffuse solar fluxes. These were determined by applying formulae from the solar calculator provided by the National Renewable Energy Laboratory (Shih et al., 1995). The formulae are as follows.

For the total solar flux q (W/m²);

$$q = C_s \tau_e (\cos(Z_r)) f_s \text{ if } \cos(Z_r) > 0 \quad (2)$$

$$q = 0 \text{ if } \cos(Z_r) \leq 0, \quad (3)$$

where C_s is the solar constant, equal to 1376.0 W/m², f_s is the sunshine factor, τ_e denotes the transmissivity of the Earth's atmosphere, and Z_r is the solar zenith angle at the location (rad).

For the daily angle a_d (deg);

$$a_d = \frac{360^\circ}{365} (d_{year} - 1), \quad (4)$$

where d_{year} is the day of the year.

For Earth radius vector r_E ;

$$r_E = 1.000110 + 0.034221\cos(\alpha_d) + 0.000719\cos(2\alpha_d) + 0.000077\sin(2\alpha_d), \quad (5)$$

where α_d refers to the angle of the day.

The calculation formula only takes into account solar radiation received at the Earth's surface based on the Earth's position over time, using values are the typical for the Jeonju-city. Therefore, adjusting for actual weather conditions using the sunshine factor and modifying the ratio of direct to diffuse solar radiation, which are typical for the case study city, is necessary. In this study, empirical values were adopted with a sunshine factor of 0.5 and a diffuse fraction of 0.2. These values were refined using sensitivity analysis to ensure the highest level of correlation with the actual measured data. The analysis tested values in 0.1 increments, ranging from 0.1 to 1, to determine the settings that most accurately reflected the observed conditions.

The air velocity at the side inlet, $U(z)$ (m/s), was plotted in the form of a logarithmic profile, with the velocity calculated as a function of altitude based on the reference speed of 10 m above sea level (Hong et al., 2012).

$$U(z) = u_{ref} \left(\frac{z}{10} \right)^{0.33}, \quad (6)$$

where u_{ref} refers to the reference velocity at the base height (10 m) and z to the height. Based on the measurement data, the reference speed was entered at 10 m and converted into speed vectors with magnitude and direction components over time. The exponent 0.33 is derived from the logarithmic wind profile standard.

In the atmospheric boundary layer, the generation of turbulence is significantly stronger than usual, requiring an increase in the inflow turbulence intensity and the turbulence dissipation rate beyond the default settings. The equations governing these adjustments are expressed in Eqs. (7–8) (Zhang, 2009).

$$k = (u_{Ref} T I)^2 \quad (7)$$

$$\varepsilon = \frac{1}{\kappa} \frac{u_t^3}{d_w}, \quad (8)$$

where k refers to the turbulent kinetic energy (m^2/s^2), representing the energy contained within the turbulent eddies of the flow, and $T I$ is the turbulent intensity, a dimensionless number indicating that the intensity of the turbulence was set to 10. In Eq. (8), the turbulence dissipation rate ε (m^2/s^3), is calculated using the von Kármán constant κ ($\kappa = 0.42$), the friction velocity u_t (m/s), and the wall distance d_w (m).

The friction velocity u_t (m/s) is calculated as follows (Zhang, 2009).

$$u_t = \frac{\kappa u_{Ref}}{\log \frac{z_{Ref} + z_0}{z_0}}, \quad (9)$$

where Z_{ref} (m) refers to the reference altitude and Z_0 (m) is the roughness height. Eq. (9) was applied to the turbulence terms of the lateral domain. In combination with the 2 km simulation area around the target area, realistic atmospheric currents were included.

To represent the solar load as heat generated at the boundaries, the correct radiation properties must be entered for each boundary. These radiation properties are determined by the specific properties of the material, as listed in Table 4. The boundaries used for the sides of the domain represent the virtual boundaries created to define the domain and the external velocities entered. Since they do not influence space in terms of radiation, both the emissivity and reflectivity are set to 0 to allow all solar loads to pass through. Likewise, the top boundary of the domain is a virtual boundary established to implement a closed system. Therefore, similar to the side boundaries, the emissivity and reflectivity were set to zero to allow the transmission of all solar loads. In addition,

Table 4
Radiation parameters.

Boundary	Material	Emissivity	Reflectivity	Transmissivity
Green	Grass	0.35	0.65	0.0
Concrete Pavement	Concrete	0.5	0.5	0.0
Building	Concrete	0.5	0.5	0.0
Road	Asphalt	0.8	0.2	0.0
Water	Water	0.8	0.1	0.1
Sky, Side		0.0	0.0	1.0
Land	Soil	-	-	-

to eliminate air resistance at the top, the shear condition is set to "slip."

For green spaces, the evapotranspiration of grass differs from that of other materials; therefore, the Penman-Monteith equation for crop evapotranspiration (ET_0) was applied, as shown in Eq. (10) (Allen et al., 1998).

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_{2m} + 273}\mu_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34\mu_2)} \quad (10)$$

where R_n means the net radiation at the crop surface (MJm^{-2}/day), G is the soil heat flux density (MJm^{-2}/day), T_{2m} refers to the air temperature at a 2 m height ($^{\circ}C$), u_2 is the wind speed at a 2 m height (m/s), e_s refers to the saturation vapor pressure (kPa), e_a is the actual vapor pressure (kPa), Δ is the slope of the vapor pressure curve ($kPa/^{\circ}C$), and γ is the psychrometric constant (0.067 kPa/ $^{\circ}C$). The formula for the vapor pressure curve Δ is as follows:

$$\Delta = \frac{4098 \left[0.6108 \text{EXP} \left(\frac{17.27T}{T+237.3} \right) \right]}{(T + 237.3)^2} \quad (11)$$

3.4. Model Validation

In this study, a transient simulation was conducted from 6 a.m. to 6 pm for a total duration of 12 h. This one-day-based environmental data calibration approach has been utilized in numerous studies, demonstrating that data from a single day has sufficient value (Acero and Herranz-Pascual, 2015; Jang et al., 2019; Jang and Kim, 2021; Nasrollahi et al., 2017a). To achieve this, a time step of 60 s and 5 inner iterations were used, resulting in 3600 iterations. To validate the accuracy of the simulation technique, the data measured from three weather stations were compared with the time-dependent simulation results. The three weather stations were installed in different locations-on a pavement, in a green area, and in a residential neighborhood – to observe temperature patterns across various environments. Each station was placed vertically 1 km apart along the same longitude to ensure consistency in data collection. The simulation outputs were extracted at the corresponding points within the domain where the measurements were obtained (Fig. 4).

On August 26, 2023, the measurement equipment was installed on-site, and the temperature was recorded every minute from 6 a.m. to 6 p.m. Each weather station was placed at the same longitude ($127.10^{\circ}E$) but at different latitudes, reflecting different urban environmental conditions. Weather station 1 was located in an area with relatively low building density, weather station 2 was located in a densely built-up area, and weather station 3 was situated in a park area without buildings, effectively detecting temperature fluctuations due to building density and green space configurations. Detailed specifications of the sensors and loggers used at the weather stations are listed in Table 5.

To validate the reliability, statistical methods, namely R^2 for linear regression model, and root mean square error (RMSE), were analyzed. An R^2 value ≥ 0.8 was considered reliable and was based on values from previous studies (Acero and Herranz-Pascual, 2015; Erell and Williamson, 2006; Ghaffarianhoseini et al., 2015; Lee et al., 2016;

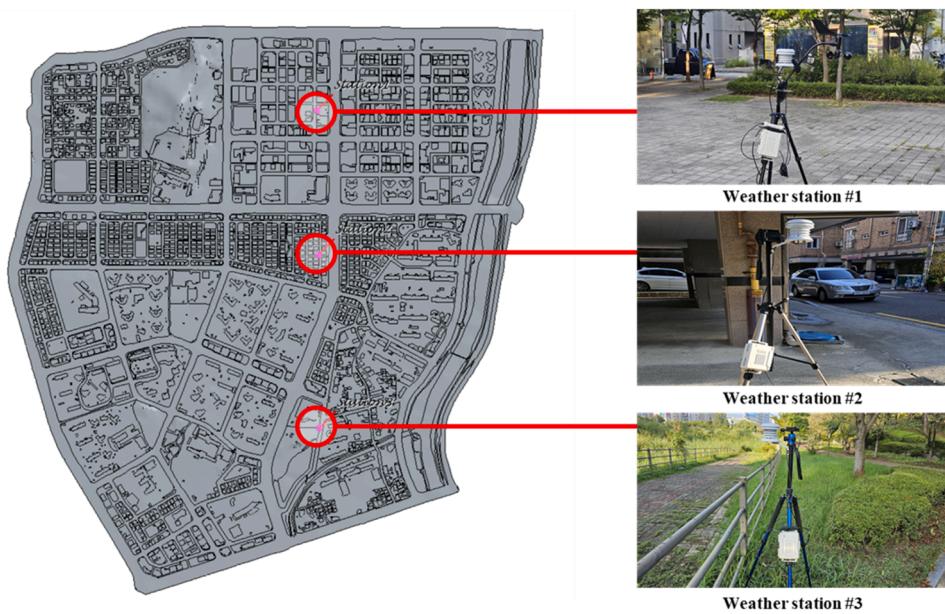


Fig. 4. Location of installed weather stations.

Table 5
Specifications of the sensors.

	Temperature	Humidity	Wind Direction	Wind Speed
Product	S-THC-M002		S-WCF-M003	
Shield	RS3-B		-	
Measuring range	-40–75 °C	0–100 %	0–355°	0–76 m/s
Accuracy	±0.2 °C	±2.5 %	±7°	±1.1 m/s
Resolution	0.02 °C	0.1 % RH	1°	0.5 m/s
Logger	H21-USB			

Nasrollahi et al., 2017a, 2017b; Salata et al., 2016; Taleb and Abu-Hijleh, 2013). The formulae for R^2 and RMSE are as follows:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (12)$$

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}, \quad (13)$$

where y_i is the observed value, \hat{y}_i is the simulated value, \bar{y} is the mean of the observed values, and n is the number of observations.

3.5. Machine Learning Model

As shown in Fig. 1, the objective of the research was to develop two types of machine learning models based on CFD. The first was a data augmentation model and the second was the most effective greening scenario. The CFD simulation results were analyzed to build the dataset necessary for machine learning training. To avoid excessive time consumption in dataset construction, stationary simulation methods, based on the hottest time at 2 pm were performed instead of transient simulation methods. Simulation conditions were used as of August 26, 2023 at the peak temperature, for other aspects, including the grid system, physics, and boundary conditions.

The method used to create the dataset scenarios for CFD simulation involved converting some concrete areas (excluding roads, buildings, green spaces, and water bodies) into green spaces. The aim was to determine the location of the greenery where the overall temperature was minimized at 1.5 m above ground within the domain. The goal of determining the location of greenery was achieved by considering both

environmental and practical constraints. According to a recent study, the current urban green coverage in Seoul is 3.7 %, with a city-wide target to increase it to 15 % (Data Driven Lab, 2015). In the target area for this study, the green coverage is 6.4 %, and while reaching the 15 % goal city-wide poses a challenge, a realistic scenario was considered, aiming to expand the green coverage to about 10 %. The scenario allows for a balanced approach to enhancing greenery while maintaining practical feasibility within the domain. As shown in Fig. 3, the current green area accounted for approximately 6.4 % of the total area, which was increased by 10 % to maximize the effect, assuming an approximate target area of 98,051.36 m² to be converted into green space. The concrete pavement areas are depicted in pink in Fig. 5(a). The process included comparing the average temperature at a 1.5 m cross-section and arranging the green spaces in square shapes.

The input variable X used in both AI models consisted of three factors: the X and Y coordinates of the upper right corner of any rectangular green space and its width. The data augmentation model uses the “itertools.product” function to generate a Cartesian product of these variables, and creating all possible parameter combinations within specified limits. The “itertools.product” function is a tool from Python’s itertools library that generates a Cartesian product of input iterable sets, which is particularly useful in machine learning for parameter optimization. An ensemble model was developed that leverages two basic machine learning models, Random Forest and XGBoost. The data augmentation code captured the predictions for 100 random combinations, which were used for subsequent validation and secondary analysis. For system configuration convenience, Design Manager, an optimization tool built into STAR-CCM+, was used. For our study, Design Manager performed multiple design iterations by adjusting simulation parameters in batch mode and collecting the outputs. Design Manager modifies design parameters (e.g., environmental factors), objectives (e.g., minimizing temperature), and constraints to explore various configurations. Its automated nature ensured that the process was efficient, with each simulation result contributing valuable information to the overall performance landscape. The 100 data points generated represent the most relevant outputs, aligned with the key engineering objectives of our study and selected for use in data augmentation. The range and increment of each parameter are summarized in Table 6. The program optimization process was used to construct 100 datasets using different combinations of variables. The 100 datasets derived from the CFD simulation were extracted into a CSV

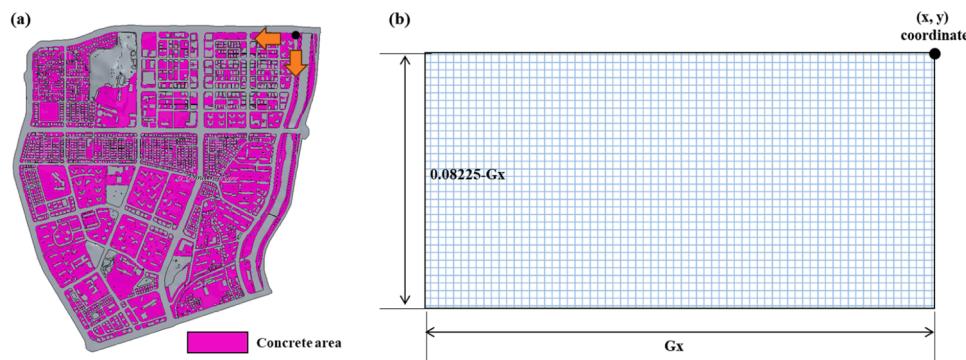


Fig. 5. Methods for urban greening using machine learning: (a) Configuration of the concrete pavement area and (b) schematic diagram of the greening optimization.

Table 6
Dataset parameter for machine learning.

	Minimum (m)	Maximum (m)	Increment (m)
Gx (width)	200	500	10
X	0	1523	10
Y	0	1101	10

table format with three variable parameters and one result, facilitating the machine learning training process.

The X variables used in the second machine learning model were the same as those in the first, with the green space shaped in a rectangular form, as depicted in Fig. 5(b), and the length of the rectangle was automatically calculated based on the target area, with the width as a reference. If the rectangular area extended beyond the domain of the target site, an area equal corresponding to the exceeded portion was added in the y-direction. Coding was added to ensure a uniform green area within the concrete part, regardless of the building area and arrangement. The CFD simulations for the green space arrangement used combinations of these three variable parameters, and the results were reported as an average temperature at a 1.5 m over the entire target site. Among them, the machine learning prediction with the lowest temperature value and the corresponding parameters were encoded as output, along with the predictions and parameters of 100 random cases from all cases, including augmentation. These randomly output 100 cases were then compared with the results obtained using the same parameters in STAR-CCM+ to verify data augmentation and optimization in machine learning.

For both the data augmentation and green space optimization machine learning models, code sets were applied for training and testing, using the extracted results of the CSV table for machine learning training and testing. The dataset was divided into training and test datasets at a ratio of 9, and an ensemble method was used for training which took advantage of the higher outcome of the two models, Random Forest and XGBoost. For testing, a model was applied that selected a prediction with a lower MSE between the two models. The code was written to output the predicted value, MSE, and applied parameters. Subsequently, the training data were analyzed to calculate the important factors and determine which parameter had a greater influence. The important factors for each of the Random Forest and XGBoost techniques were calculated, and a weighted average for the MSE was computed to calculate the final important factor for the ensemble method.

The RandomForestRegressor used default values according to the scikit-learn implementation, with `n_estimators= 100` and `max_depth=None`, while XGBoost was set with `n_estimators= 100`, `max_depth= 6`, and `learning_rate= 0.3`. The machine learning models were written in Google's Colaboratory and included four sets of code. A code set was configured to access Google Drive for convenience in terms of the data input and output. The code for the machine learning model is

described in Appendices A–C.

4. Results

4.1. Thermal Evaluation Using CFD

This section summarizes the comparison results between the measurement data and the CFD simulation for validation. Detailed information about the locations of the measuring stations and simulation methods can be found in Section 3.4. Fig. 6 shows graphs depicting the temporal temperature variations from both the measurements and simulations at each station, as well as the correlation, R^2 , and RMSE values between the measured and simulated results.

Each weather station was placed in a location with different configurations of buildings and green spaces, to capture temperature variations in different environments; in fact, they showed distinct characteristics. Weather station 1 had a maximum temperature of approximately 33°C, with the highest temperatures occurring between 1 p.m. and 3 p.m. The temperatures were highest at weather station 2, which was in a densely built-up area. In contrast, weather station 3, located in a park, showed lower temperatures compared to the other stations.

Looking closer at the simulation results at weather station 1, temperatures remained between 32 °C and 36 °C for 4 h after reaching the peak solar heat at 12 pm. At weather station 2, temperatures began to exceed 34 °C from the peak of solar heat at noon, with the temperature continuing to rise around 36 °C after 2 pm. In contrast to the other two cases, a continuous increase in temperature was observed at weather station 3 until 3 pm. All three weather stations showed a rapid drop in temperature after 3 pm. It is noteworthy that weather station 2 recorded a temperature drop of more than 6 °C.

The measurements and simulation results at all weather stations displayed similar patterns, with the trend line (linear) R^2 values indicating a high level of correlation, approximately 0.95 for all stations, indicating that the simulation accurately reflected the actual phenomena. The correlation coefficients and RMSE values also indicated high reliability: weather station 1 had 0.972 and 0.97, weather station 2 had 0.931 and 0.984, and weather station 3 had 0.975 and 0.975. These results confirm the credibility of the CFD simulation; the simulation settings used in Section 3.3 were adopted as the standard settings for the CFD simulations in this study.

4.2. Data Augmentation

This section summarizes the results obtained from training and testing a dataset from CFD simulations using machine learning techniques. For data augmentation, the average dataset results from CFD simulations at 1.5 m height are shown in Fig. 7. In Fig. 7, "Design number" refers to the simulation case number and represents the

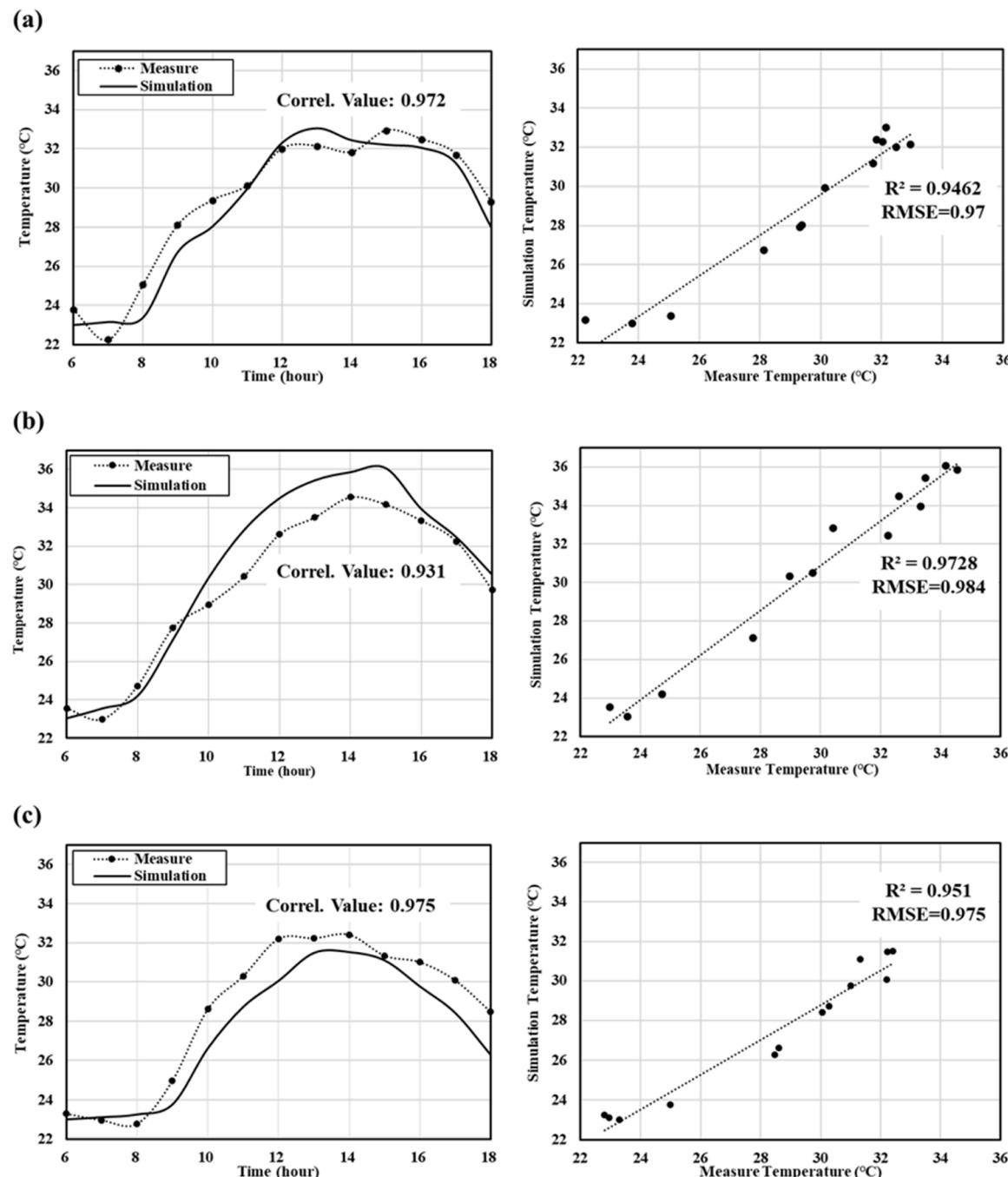


Fig. 6. Comparative analysis of the observation and simulation results: (a) weather station 1, (b) weather station 2, and (c) weather station 3.

average temperature at 1.5 m height in the area generated based on the green space coordinates (X, Y) obtained through the STAR-CCM+ Design Manager program. It took approximately 81 h to compile 100 CFD simulation datasets, each consisting of approximately 12.8 million cells, highlighting the considerable time required to conduct extensive simulations. Considering that the simulation environment had a 3.00 GHz, 48 core configuration, a typical computer build is expected to take significantly more time. Even in the high-performance computing environment used in this study, running simulations for more than 81 h to produce datasets containing more than 100 records was challenging. Therefore, the analyzed 100 datasets were compiled into a CSV file, which was then used in the machine learning training code set for training and testing. As mentioned in Section 3.5, both the Random Forest and XGBoost techniques were used simultaneously, adopting the

ensemble method that selected the highest of the two results. The MSE between the predicted model from training and 10 test results was 3.79e-6, indicating a highly accurate result.

To determine which input elements have a significant impact on machine learning training, the importance of the parameters was evaluated. The importance factors for both the Random Forest and XGBoost techniques were calculated, and a weighted MSE was computed for each method to determine the final importance factor for the ensemble technique. The important factors for each training method are summarized in Fig. 8. Notably, the importance factors predicted by Random Forest and XGBoost were different. Random Forest identified the X-coordinate as the most dominant factor, while XGBoost predicted the length in the X-direction as the most dominant factor. The result of the combined ensemble technique, which integrates both using mass

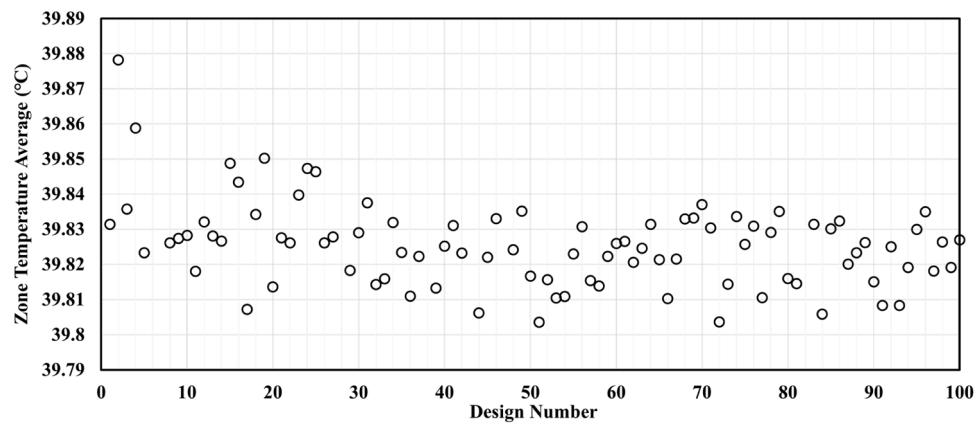


Fig. 7. Training dataset based on CFD simulation.

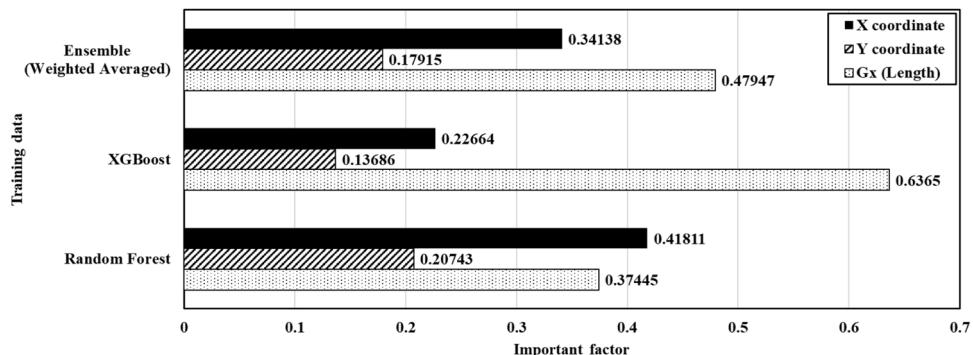


Fig. 8. Important factors identified by the machine learning methods.

weighting, predicts the most accurate result, with an importance order identical to that of XGBoost but with moderate ratios, owing to the inclusion of Random Forest. This suggests that using the ensemble technique as a predictive model is more suitable than using a single model because it can compensate for the weaknesses of each individual model.

The prediction model was utilized to augment the dataset, from which the parameters (x and y coordinates and area) and the predicted temperature were derived for the case that showed the most efficient temperature reduction effect in converting a specific area into a green space. The data augmentation used the same range and increments that

were applied when creating the CFD model dataset. In the CFD simulations, the parameters were selected and the calculations carried out using a developed optimization method (Kim and Kang, 2022). It was determined that the optimization process was unlikely to find an optimized case within 100 simulations. Therefore, data were extracted for all cases, including augmented data. Combining the range and increments of the domain, the total number of possible cases was calculated to be 503,047 and the case with the lowest result was defined as the optimized case. The results and parameters for this case are described in detail in Section 4.3.

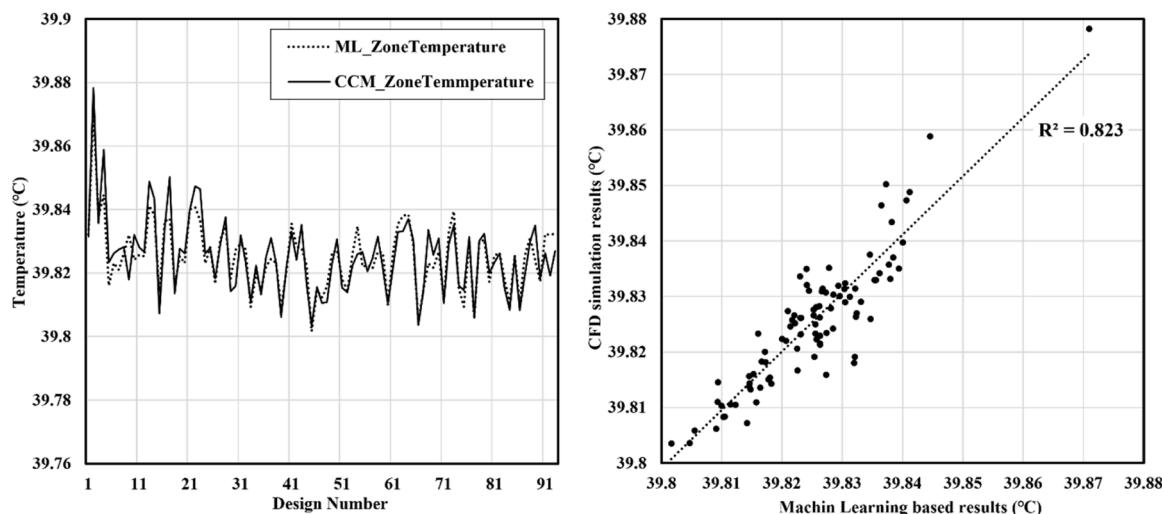


Fig. 9. Comparison of temperatures: Machine learning predictions and CFD simulations.

To validate approximately half a million generated cases, the parameters and results of 100 randomly selected cases were extracted and compared with the CFD simulation results. The validation results are presented in Fig. 9, showing that most of the 100 cases had a similar pattern, and the comparison between the machine learning outcomes and CFD simulation results showed an R^2 value > 0.8 , indicating high consistency. Therefore, using machine learning to augment the data from the CFD simulation results is an innovative method that can compensate for the time-consuming nature of CFD analyses.

4.3. Results of Climate Optimization

The optimized case resulted in a predicted temperature of 39.7999°C . Assuming that the temperature was approximately 40.2°C before the installation of green spaces, the temperature reduction effect was around 1 %. The parameters for this case were $X = -30 \text{ m}$, $Y = 610 \text{ m}$, and $G_x = 380 \text{ m}$. Examining Fig. 10 (a), the location of the optimized case is indicated in green. In the most vulnerable case with the highest temperature, the average temperature reached 39.9°C , with the parameters $X = -200 \text{ m}$, $Y = +1510 \text{ m}$, and $G_x = 870 \text{ m}$. This scenario highlights the importance of green space design in urban areas, as it demonstrates that a strategic layout of greenery can lead to a significant temperature reduction of approximately 0.1 degree. Such a reduction is evidence of the crucial role green spaces play in mitigating UHI effects. Fig. 10 (b–c) compares the temperature distributions from the CFD simulation results before and after the additional placement of green spaces at this location. A CFD simulation was carried out by placing green spaces in the central part of the dense housing area where UHI effects are intensely generated. The analysis revealed that the temperature reduction in the entire domain area was negligible compared to

the temperature value of the optimized case, which validated the research method of identifying optimized cases through machine learning. This finding also indicates that the design of surrounding green spaces may not have a significant impact in areas with very high housing density, where sufficient wind paths are not formed.

5. Discussion

The Discussion section is divided into three parts: (1) a discussion on the AI model developed in this study, focusing on its efficiency, generalizability, and limitations; (2) a discussion of the key findings from the research, particularly the impact of green space planning on UHI mitigation; and (3) an exploration of the current study's limitations and directions for future research.

5.1. Discussion on the AI Model Development

This study presents a machine learning model capable of significantly reducing the time required for CFD analysis. Considering that it took 81 hours to analyze 100 cases in the environment used in this study, analyzing all 503,047 data points would require over 400,000 hours. By using an ensemble technique with the XGBoost and Random Forest algorithms, a balanced model was developed that combines the strengths of both models. This machine learning model demonstrates great potential as an efficient modeling method, applicable not only to architectural environments but also to fields such as structural, industrial, indoor dynamics, and aeronautical engineering.

However, the study's methodology, while effective for the specific urban area analyzed, may have limitations in its general applicability to other regions with different climatic conditions, urban morphologies,

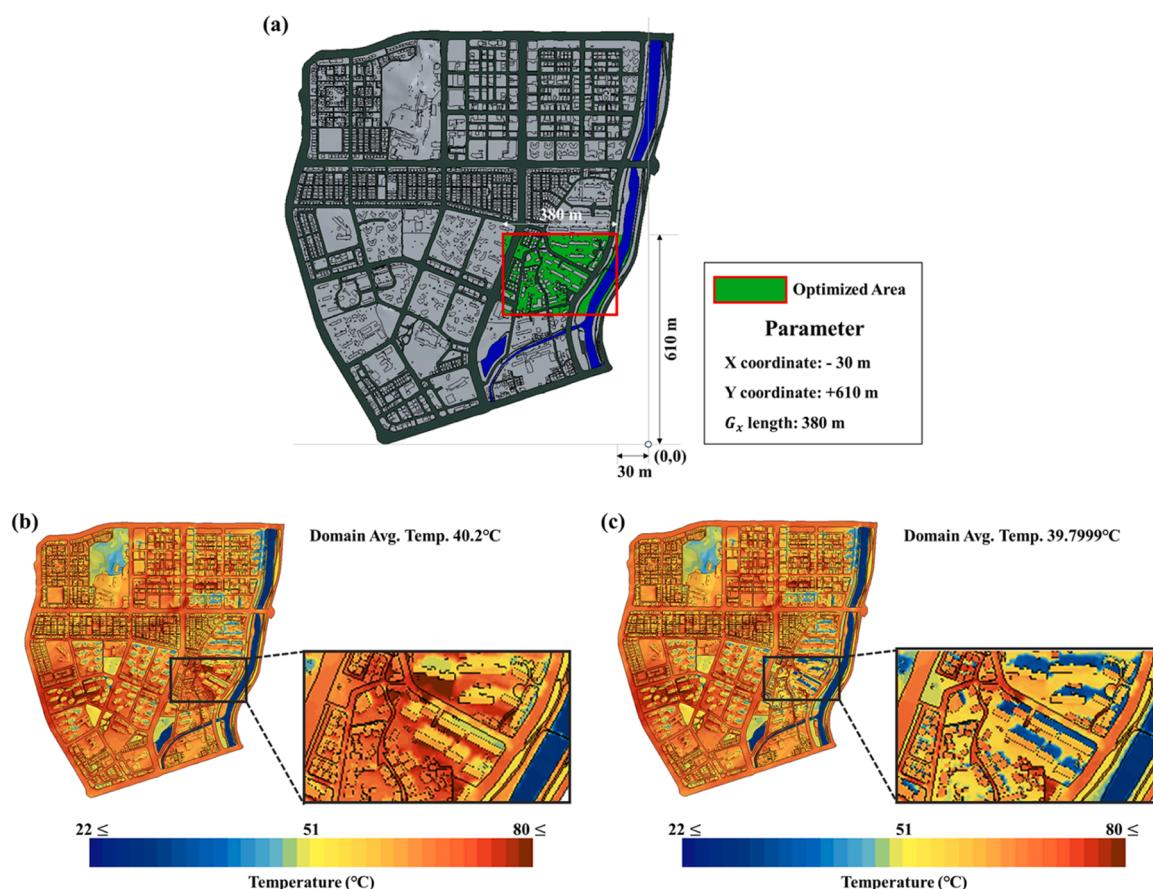


Fig. 10. Impact of green space optimization and urban temperature distribution: (a) Location of the optimized space and temperature distribution (b) before and (c) after applying green space optimization.

and construction materials. The developed model is site-, day-, and season-specific, which limits its broader applicability. While conducting 100 cases requires a significant amount of time, future research should focus on developing methods to further reduce the computational time. Additionally, the model should be enhanced to account for seasonal variations and day-night effects to improve its accuracy and applicability across different scenarios. To enhance the generalizability of the methodology, future research should focus on incorporating comprehensive material characteristics, including detailed building envelope properties, and consider a wider range of environmental variables. Additionally, developing a more universal modeling framework that includes parameters such as terrain elevation, urban density, and building forms would allow the methodology to be adapted to various urban settings. This would improve the accuracy of simulations and support the creation of an accessible UHI modeling program that can be utilized by practitioners and policymakers in diverse contexts.

Furthermore, there are some limitations regarding the omission of night-time evaluation in both measurements and CFD-based validation. This study specifically focused on addressing the issue of peak day time temperatures, with simulations conducted at 2 P.m., when the UHI is most pronounced. Night-time evaluations were beyond the scope of this research and should be explored in future studies. For instance, [Zhu et al. \(2024\)](#) conducted high resolution CFD simulations considering various spatiotemporal boundary impacts, including different times of the day and night. Incorporating such methods in future research would help address the known tendency of CFD simulations to over-predict the cooling effects of urban green infrastructure at night.

Despite these limitations, the machine learning model developed in this study can serve as an accessible tool to other developers and citizens. This model can be a valuable policy resource for decision makers. This represents a cost-effective alternative to outsourcing urban planning and green space planning simulations, potentially leading to significant budget savings for municipalities and government bodies. Using this model, policymakers can make reliable and data-driven decisions in urban planning. For example, a city planner could use the model to simulate different scenarios of green space distribution and their impact on urban temperature. This allows different planning strategies to be evaluated before implementation, ensuring that the most effective approach is chosen. By reducing the need for external simulation services, the model not only saves financial resources but also accelerates the planning process. Overall, the availability of this machine learning model as an accessible tool is a step forward in making urban planning and environmental management more accessible, empowering a wide range of stakeholders to participate in shaping sustainable and livable urban environments.

5.2. Discussion on the CFD Analysis

The research findings revealed that factors such as the density and arrangement of buildings, as well as wind paths, are crucial when selecting locations for green spaces. Reliable modeling was carried out using simulation and an evidence-based high-resolution model (CFD), providing a green space planning method applicable to both urban regeneration plans in old downtown areas and new urban planning. The research findings revealed that factors such as the density and arrangement of buildings in a city as well as wind paths are important elements when selecting locations for green spaces. However, it is important to note that various confounding factors-such as building materials, traffic patterns, and microclimate variations- could influence the results.

These results support the importance of green space planning, that emphasizes the connectivity of green networks in UHIs and are consistent with previous studies that have shown the contribution of green connectivity to the improving thermal comfort ([Juorio and Harbich, 2019](#); [Ornelas et al., 2023](#); [Teoh et al., 2022](#)). Nonetheless, the study acknowledges that the form and function of UHI, as well as other

external factors, need to be addressed to better understand the overall impact of urban green infrastructure.

The integration of CFD and AI to achieve faster and more accurate predictions of urban environmental conditions, providing substantial benefits for urban climatology and smart city operations. This approach is particularly valuable in smart cities, where swift CFD analyses can bolster UHI mitigation efforts and improve urban planning initiatives. By providing quick and precise assessments, this research supports the deployment of urban cooling services and helps develop climate-resilient urban areas. It highlights the practical advantages of combining CFD and AI for superior urban management and planning.

5.3. Limitations and Future Directions

While the study provides significant insights, some limitations remain. The exclusion of detailed material characteristics, such as the omission of glazing layers, could affect the accuracy of thermal behavior simulations, particularly for interior air temperature and diurnal surface temperature variations. Additionally, the methodology used may have limited generalizability to other regions with different climatic conditions and urban morphologies. To address these limitations, future research should focus on incorporating a broader range of material properties and environmental variables. Developing a more universal modeling framework that includes parameters such as terrain Elevation, urban density, and building forms would enhance the applicability of the methodology to various urban settings. This would improve the accuracy of simulations and support the creation of a more accessible UHI modeling tool for practitioners and policymakers.

Although the integration of CFD and AI offers significant insights and improvements for urban environmental analysis, there are some limitations to this study. Most notably, the types and configurations of green space layouts were not explicitly considered in this research. The Design Manager tool focused on simulating the impact of green space coverage rather than the specific shapes, arrangements, or layouts of green spaces. This omission limits the ability to draw conclusions on how different green space layouts might affect surrounding temperatures or thermal comfort in urban environments. Future research should explore different green space layout types and their influence on urban cooling to further enhance the model's utility for urban planning.

6. Conclusions

The present study represents a significant advance in urban climate resilience through the use of state-of-the-art machine learning techniques. The primary aim of this study was to develop a robust methodology that integrates data augmentation and green space optimization to counteract the UHI effect. The followings are the main conclusions:

- High-accuracy data augmentation and green space optimization modeling

A machine learning model was developed using ensemble methods that combine Random Forest and XGBoost algorithms. This model was meticulously trained and tested on a dataset obtained from CFD simulations. During the validation phase, where the machine learning predictions with actual CFD simulation results, the model demonstrated a high level of accuracy, as evidenced by an R^2 value of over 0.8. The impressive accuracy of data augmentation and green space optimization modeling represents a significant advance in urban temperature regulation strategies.

- Enhanced efficiency in CFD simulation analysis via machine learning

The machine learning model developed showed a significant leap in efficiency compared to traditional, labor-intensive CFD simulation methods. While the conventional approach requires over 400,000 hours for exhaustive data analysis, the machine learning model delivers comparable results within an hour. This significant increase in analysis speed represents a critical improvement in urban

climate studies, enabling rapid and accurate environmental assessments.

- Effectiveness of green space optimization in urban temperature reduction

The study aimed to evaluate the influence of strategically placed green spaces on temperature reduction in urban environments. The optimized green space configuration resulted in a temperature reduction of approximately 1 %, highlighting the importance of strategic green space placement in densely populated urban areas. Although the overall temperature reduction appears modest, it underscores the necessity for deliberate placement of green spaces to mitigate UHI effects.

The study's method and findings make valuable contributions to urban planning and environmental policy. The integration of advanced machine learning models with environmental data has ushered in a new era of urban development strategies, leading to more informed, efficient, and sustainable approaches. This aligns with the overarching goal of creating livable and climate-resilient urban spaces. However, the present study focused solely on modeling Jeonju, and did not investigate the applicability of the machine learning model to other locations. Future research should consider these aspects and conduct studies that cover a broader range of target areas. Furthermore, this study primarily focused on the thermal characteristics of urban environments to assess the impact of land cover changes on temperature reduction. Consequently, we did not measure or validate other meteorological data such as solar radiation and relative humidity. These factors can also significantly influence urban microclimates and heat transfer processes. Future

research should consider incorporating these variables to provide more comprehensive analysis of urban climate dynamics and to enhance the accuracy of simulations.

CRediT authorship contribution statement

Junsuk Kang: Supervision, Investigation, Conceptualization. **Junghyeon Ahn:** Writing – original draft, Software, Methodology. **Jaekeyoung Kim:** Writing – review & editing, Validation, Data curation.

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.

Acknowledgements

This work was financially supported by the Korea Ministry of Land, Infrastructure and Transport (MOLIT) through the Innovative Talent Education Program for Smart City and the Knowledge-based Environmental Service Program and Korea Environment Industry & Technology Institute (KEITI) through the Climate Change R&D Project for New Climate Regime, and funded by the Korea Ministry of Environment (MOE) (Grant no. 2022003570004) and the National Research Foundation of Korea Grant through the Korean Government (Grant no. NRF-RS-2023-00259403).

Appendix A. Training and test code for machine learning import pandas as pd

```
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.ensemble import RandomForestRegressor
import xgboost as xgb
f = pd.read_csv('/content/drive/MyDrive/Paper/ML_dataset_gr.csv')
X = f.iloc[:,2:].to_numpy()
Y = f.iloc[:,1].to_numpy()
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.1)
print(len(X_train), len(X_test))
rf_model = RandomForestRegressor(n_estimators=100)
xgb_model = xgb.XGBRegressor()
rf_model.fit(X_train, y_train)
xgb_model.fit(X_train, y_train)
rf_predictions = rf_model.predict(X_test)
xgb_predictions = xgb_model.predict(X_test)
ensemble_predictions = (rf_predictions + xgb_predictions) / 2
mse = mean_squared_error(y_test, ensemble_predictions)
print("Ensemble mean_squared_error =", mse)
min_mse_index = np.argmin(np.abs(ensemble_predictions - y_test))
print("Best ensemble prediction =", ensemble_predictions[min_mse_index], ", test result =", y_test[min_mse_index])
best_parameters = X_test[min_mse_index]
print("Parameters for the best prediction:")
for i, param in enumerate(best_parameters):
    print(f"Parameter {i + 1}:", param)
```

Appendix B. Data augmentation and optimization import itertools

```
import numpy as np
import random
feature_ranges = [
    np.arange(200, 500, 10),
    np.arange(0, 1523, 10),
```

```

np.arange(0, 1101, 10),
all_combinations = list(itertools.product(*feature_ranges))
random_combinations = random.sample(all_combinations, 100)
min_prediction = float('inf')
best_parameters_for_min_pred = None
random_combination_results =
for combination in all_combinations:
combination_array = np.array(combination).reshape(1, -1)
rf_pred = rf_model.predict(combination_array)
xgb_pred = xgb_model.predict(combination_array)
ensemble_pred = (rf_pred + xgb_pred) / 2
if ensemble_pred < min_prediction:
min_prediction = ensemble_pred
best_parameters_for_min_pred = combination_array
if combination in random_combinations:
random_combination_results.append((combination, ensemble_pred[0]))
print("Minimum prediction using ensemble grid search:", min_prediction[0])
print("Parameters for the minimum prediction:")
for i, param in enumerate(best_parameters_for_min_pred[0]):
print(f"Parameter {i + 1}: {param}")
print("\nRandomly selected 100 combinations:")
print('*'*80)
print(f'{Parameter 1:<15} {Parameter 2:<15} {Parameter 3:<15} {Prediction:<15}')
print('*'*80)
for combo, pred in random_combination_results:
print(f'{combo[0]:<15} {combo[1]:<15} {combo[2]:<15} {pred:<15.8 f}'")

```

Appendix C. Importance factor

```

import matplotlib.pyplot as plt
rf_mse = mean_squared_error(y_test, rf_model.predict(X_test))
xgb_mse = mean_squared_error(y_test, xgb_model.predict(X_test))
rf_weight = 1 / rf_mse
xgb_weight = 1 / xgb_mse
total_weight = rf_weight + xgb_weight
rf_weight /= total_weight
xgb_weight /= total_weight
ensemble_importances = rf_weight * rf_model.feature_importances_ + xgb_weight * xgb_model.feature_importances_
print("RandomForest Feature Importances:")
for i, imp in enumerate(rf_model.feature_importances_):
print(f"Parameter {i + 1}: {imp}")
print("\nXGBoost Feature Importances:")
for i, imp in enumerate(xgb_model.feature_importances_):
print(f"Parameter {i + 1}: {imp}")
print("\nEnsemble Feature Importances (Weighted Average):")
for i, imp in enumerate(ensemble_importances):
print(f"Parameter {i + 1}: {imp}")
indices = np.arange(len(rf_model.feature_importances_))
plt.figure(figsize=(18, 6))
plt.subplot(1, 3, 1)
plt.title("RandomForest Feature Importances")
plt.bar(indices, rf_model.feature_importances_, color='b')
plt.xlabel('Parameter Index')
plt.ylabel('Importance')
plt.xticks(indices, indices+1)
plt.subplot(1, 3, 2)
plt.title("XGBoost Feature Importances")
plt.bar(indices, xgb_model.feature_importances_, color='r')
plt.xlabel('Parameter Index')
plt.ylabel('Importance')
plt.xticks(indices, indices+1)
plt.subplot(1, 3, 3)
plt.title("Ensemble Feature Importances (Weighted Average)")
plt.bar(indices, ensemble_importances, color='g')
plt.xlabel('Parameter Index')

```

```
plt.ylabel('Importance')
plt.xticks(indices, indices+1)
plt.show()
```

Appendix D. Input Data for AI Model Training Generated via Design Manager

Design	Temperature	width	X	Y
1	39.828	200	1523	0
2	39.8434	200	0	340
3	39.8612	389.474	1523	1100
4	39.8259	800	481	750
5	39.8679	200	1523	750
6	39.8107	800	481	0
7	39.8849	200	1523	920
8	39.8131	547.368	1042	630
9	39.8561	263.158	1523	750
10	39.8406	578.947	0	0
11	39.8065	326.316	481	0
12	39.8193	547.368	641	630
13	39.8107	800	481	0
14	39.8299	800	1523	0
15	39.8153	326.316	1042	630
16	39.8303	800	481	980
17	39.8098	326.316	401	50
18	39.8118	768.421	962	230
19	39.8098	326.316	401	50
20	39.8098	294.737	721	110
21	39.821	515.789	1283	230
22	39.8098	326.316	401	0
23	39.8107	515.789	962	0
24	39.8126	389.474	962	0
25	39.8046	326.316	802	0
26	39.8041	357.895	721	0
27	39.8039	452.632	802	0
28	39.8155	768.421	1042	230
29	39.8164	800	641	340
30	39.8268	263.158	160	0
31	39.8098	357.895	401	0
32	39.8024	515.789	802	0
33	39.8073	547.368	802	400
34	39.8005	452.632	641	0
35	39.8027	578.947	641	50
36	39.8021	452.632	641	110
37	39.8034	610.526	641	0
38	39.8083	484.211	882	0
39	39.8034	610.526	641	0
40	39.8266	610.526	641	1040
41	39.8018	515.789	641	0
42	39.8012	421.053	561	0
43	39.8073	547.368	802	400
44	39.7999	357.895	641	0
45	39.8182	452.632	321	0
46	39.8069	578.947	802	400
47	39.8017	326.316	641	0
48	39.8017	326.316	641	0
49	39.8259	547.368	1122	110
50	39.7999	357.895	641	0
51	39.8114	357.895	641	460
52	39.7998	357.895	641	0
53	39.8296	200	160	340
54	39.8051	357.895	802	0
55	39.7997	389.474	641	0
56	39.8046	389.474	481	0
57	39.8022	484.211	561	110
58	39.8205	673.684	561	870
59	39.7997	389.474	641	0
60	39.7997	389.474	641	0
61	39.8073	547.368	802	400
62	39.7998	357.895	641	0
63	39.7997	389.474	641	0
64	39.8015	484.211	641	0
65	39.8019	547.368	802	0
66	39.8001	421.053	641	0
67	39.8003	421.053	641	50
68	39.7999	357.895	641	0
69	39.8076	421.053	721	290

(continued on next page)

(continued)

Design	Temperature	width	X	Y
70	39.8044	484.211	721	110
71	39.8406	800	0	0
72	39.8001	421.053	641	0
73	39.7996	389.474	641	50
74	39.7999	357.895	641	0
75	39.7999	357.895	641	0
76	39.8001	421.053	641	0
77	39.8018	547.368	641	50
78	39.8024	389.474	641	170
79	39.7997	389.474	641	0
80	39.8001	421.053	641	0
81	39.7997	389.474	641	0
82	39.8001	357.895	641	50
83	39.8036	547.368	641	110
84	39.7997	389.474	641	0
85	39.7996	389.474	641	50
86	39.802	326.316	641	50
87	39.7996	389.474	641	50
88	39.8001	421.053	641	0
89	39.8284	357.895	641	1100
90	39.8017	421.053	641	110
91	39.8071	736.842	802	400
92	39.8001	357.895	641	50
93	39.8001	357.895	641	50
94	39.8187	389.474	641	870
95	39.8001	421.053	641	0
96	39.8002	389.474	561	50
97	39.8308	389.474	160	50
98	39.8001	421.053	641	0
99	39.8064	736.842	641	50
100	39.7996	389.474	641	50

References

- Acero, J.A., Herranz-Pascual, K., 2015. A comparison of thermal comfort conditions in four urban spaces by means of measurements and modelling techniques. *Build. Environ.* 93, 245–257. <https://doi.org/10.1016/j.buildenv.2015.06.028>.
- Allegri, J., Carmeliet, J., 2017. Coupled CFD and building energy simulations for studying the impacts of building height topology and buoyancy on local urban microclimates. *Urban Clim.* 21, 278–305. <https://doi.org/10.1016/j.ulclim.2017.07.005>.
- Allegri, J., Carmeliet, J., 2018. Simulations of local heat islands in Zürich with coupled CFD and building energy models. *Urban Clim.* 24, 340–359. <https://doi.org/10.1016/j.ulclim.2017.02.003>.
- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements – FAO Irrigation and Drainage Paper 56, 300. Food and Agriculture Organization, Rome, p. D05109.
- Antoniou, N., Montazeri, H., Neophytou, M., Blocken, B., 2019. CFD simulation of urban microclimate: Validation using high-resolution field measurements. *Sci. Total Environ.* 695, 133743. <https://doi.org/10.1016/j.scitotenv.2019.133743>.
- Arias, P., Bellouin, N., Coppola, E., Jones, R., Krinner, G., Marotzke, J., Naik, V., Palmer, M., Plattner, G.-K., Rogelj, J., 2021. The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; technical summary. *Climate Change*.
- Augusto, B., Roebelein, P., Rafael, S., Ferreira, J., Ascenso, A., Bodilis, C., 2020. Short and medium- to long-term impacts of nature-based solutions on urban heat. *Sustain. Cities Soc.* 57, 102122. <https://doi.org/10.1016/j.scs.2020.102122>.
- Bae, C., Lee, D.K., 2020. Effects of low-impact development practices for flood events at the catchment scale in a highly developed urban area. *Int. J. Disaster Risk Reduc.* 44, 101412. <https://doi.org/10.1016/j.ijdrr.2019.101412>.
- Batty, M., 2018. Digital twins. *Environ. Plan. B: Urban Anal. City Sci.* 45 (5), 817–820. <https://doi.org/10.1177/2399808318796416>.
- Bayulken, B., Huisingsh, D., Fisher, P.M.J., 2021. How are nature based solutions helping in the greening of cities in the context of crises such as climate change and pandemics? A comprehensive review. *J. Clean. Prod.* 288, 125569. <https://doi.org/10.1016/j.jclepro.2020.125569>.
- Beckett, S., 2022. Smart city digital twins, 3D modeling and visualization tools, and spatial cognition algorithms in artificial intelligence-based urban design and planning. *Geopol. Hist. Int. Relat.* 14, 123–138.
- Bosello, F., Roson, R., Tol, R.S.J., 2007. Economy-wide estimates of the implications of climate change: Sea level rise. *Environ. Resour. Econ.* 37, 549–571. <https://doi.org/10.1007/s10640-006-9048-5>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Buck, A.L., 1981. New equations for computing vapor pressure and enhancement factor. *J. Appl. Meteor.* 20, 1527–1532. [https://doi.org/10.1175/1520-0450\(1981\)020<1527:NEFCVP>2.0.CO;2](https://doi.org/10.1175/1520-0450(1981)020<1527:NEFCVP>2.0.CO;2).
- Calautit, J.K., Hughes, B.R., Chaudhry, H.N., Ghani, S.A., 2013. CFD analysis of a heat transfer device integrated wind tower system for hot and dry climate. *Appl. Energy* 112, 576–591. <https://doi.org/10.1016/j.apenergy.2013.01.021>.
- Clarke, B., Otto, F., Stuart-Smith, R., Harrington, L., 2022. Extreme weather impacts of climate change: An attribution perspective. *Environ. Res. Clim.* 1, 012001.
- Cui, Y., Cai, M., Stanley, H.E., 2017. Comparative analysis and classification of cassette exons and constitutive exons. *BioMed. Res. Int.* 2017 (1), 7323508. <https://doi.org/10.1155/2017/7323508>.
- Data Driven Lab. (2015). *Measuring urban sustainability in Seoul, South Korea*. Retrieved from (<https://datadrivendata.org/climate/measuring-urban-sustainability-in-seoul-south-korea/>).
- Depietri, Y., Renaud, F.G., Kallis, G., 2012. Heat waves and floods in urban areas: A policy-oriented review of ecosystem services. *Sustain. Sci.* 7, 95–107. <https://doi.org/10.1007/s11625-011-0142-4>.
- Dhyani, S., Karki, M., Gupta, A.K., 2020. Opportunities and advances to mainstream nature-based solutions in disaster risk management and climate strategy. In: Dhyani, S., Gupta, A., Karki, M. (Eds.), *Nature-Based Solutions for Resilient Ecosystems and Societies*. Springer, Singapore, pp. 1–26. https://doi.org/10.1007/978-981-15-4712-6_1.
- Erell, E., Williamson, T., 2006. Comments on the correct specification of the analytical CTTC model for predicting the urban canopy layer temperature. *Energy Build.* 38, 1015–1021. <https://doi.org/10.1016/j.enbuild.2005.11.013>.
- Fang, X., Li, J., Ma, Q., 2023. Integrating green infrastructure, ecosystem services and nature-based solutions for urban sustainability: A comprehensive literature review. *Sustain. Cities Soc.* 98, 104843. <https://doi.org/10.1016/j.scs.2023.104843>.
- Ford, J.D., Pearce, T., McDowell, G., Berrang-Ford, L., Sayles, J.S., Belfer, E., 2018. Vulnerability and its discontents: The past, present, and future of climate change vulnerability research. *Clim. Change* 151, 189–203. <https://doi.org/10.1007/s10584-018-2304-1>.
- Franke, J., Hirsch, C., Jensen, A., Krüs, H., Schatzmann, M., Westbury, P., Miles, S., Wisse, J., Wright, N., 2004. Recommendations on the use of CFD in predicting pedestrian wind environment, in: *Proceedings of the International Conference on Urban Wind Engineering and Building Aerodynamics* pp. C. 1.1-C1.11.
- Fu, G., Jin, Y., Sun, S., Yuan, Z., Butler, D., 2022. The role of deep learning in urban water management: A critical review. *Water Res.* 223, 118973. <https://doi.org/10.1016/j.watres.2022.118973>.
- Gasper, R., Blohm, A., Ruth, M., 2011. Social and economic impacts of climate change on the urban environment. *Curr. Opin. Environ. Sustain.* 3, 150–157. <https://doi.org/10.1016/j.cosust.2010.12.009>.
- Ghaffarianhoseini, A., Berardi, U., Ghaffarianhoseini, A., 2015. Thermal performance characteristics of unshaded courtyards in hot and humid climates. *Build. Environ.* 87, 154–168. <https://doi.org/10.1016/j.buildenv.2015.02.001>.

- Grimm, N.B., et al., 2013. The impacts of climate change on ecosystem structure and function. *Front. Ecol. Environ.* 11, 474–482.
- Hambrecht, E., Tolhurst, R., Whittaker, L., 2022. Climate change and health in informal settlements: A narrative review of the health impacts of extreme weather events. *Environ. Urb.* 34, 122–150. <https://doi.org/10.1177/09562478221083896>.
- Hayes, A.T., Jandaghian, Z., Lacasse, M.A., Gaur, A., Lu, H., Laouadi, A., Ge, H., Wang, L., 2022. Nature-based solutions (nbs) to mitigate urban heat island (UHI) effects in Canadian cities. *Buildings* 12, 925. <https://doi.org/10.3390/buildings12070925>.
- Henriksen, H.J., Schneider, R., Koch, J., Ondracek, M., Troldborg, L., Seidenfaden, I.K., Kragh, S.J., Bøgh, E., Stisen, S., 2022. A new digital twin for climate change adaptation, water management, and disaster risk reduction (HIP Digital Twin). *Water* 15, 25. <https://doi.org/10.3390/w15010025>.
- Hong, I.P., Ki, H.S., Jung, E.Y., Song, D.S., 2012. A study on inlet flow profile in CFD simulation for predicting wind environment in urban area. *J. Archit. Inst. Korea* 28, 311–318.
- Hsu, A., Sheriff, G., Chakraborty, T., Manya, D., 2021. Disproportionate exposure to urban heat island intensity across major US cities. *Nat. Commun.* 12, 2721. <https://doi.org/10.1038/s41467-021-22799-5>.
- Hua, W., Dong, X., Liu, Q., Zhou, L., Chen, H., Sun, S., 2021. High-resolution WRF simulation of extreme heat events in Eastern China: Large sensitivity to land surface schemes. *Front. Earth Sci.* 9, 770826. <https://doi.org/10.3389/feart.2021.770826>.
- Iuorio, O., Harbich, L.A., 2019. Green networks as a Key of urban planning with thermal comfort, in: de Oliveira, L., In: Mell, I. (Ed.), *Planning Cities with Nature: Theories, Strategies and Methods*. Springer, Cham, pp. 97–109.
- Jang, G., Kim, S., 2021. Are decline-oriented strategies thermally sustainable in shrinking cities? *Urban Clim.* 39, 100924. <https://doi.org/10.1016/j.ulclim.2021.100924>.
- Jang, G., Kim, S., Lee, J.S., 2022. Planning scenarios and microclimatic effects: The case of high-density riverside residential districts in Seoul, South Korea. *Build. Environ.* 223, 109517. <https://doi.org/10.1016/j.buildenv.2022.109517>.
- Jones, D., Snider, C., Nashehi, A., Yon, J., Hicks, B., 2020. Characterising the Digital Twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* 29, 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>.
- Kafy, A.-A., et al., 2021. Prediction of seasonal urban thermal field variance index using machine learning algorithms in Cumilla, Bangladesh. *Sustain. Cities Soc.* 64, 102542. <https://doi.org/10.1016/j.scs.2020.102542>.
- Kafy, A.-A., et al., 2022. Predicting the impacts of land use/land cover changes on seasonal urban thermal characteristics using machine learning algorithms. *Build. Environ.* 217, 109066. <https://doi.org/10.1016/j.buildenv.2022.109066>.
- Kazidenev, D., Omirbekov, S., Amanbek, Y., 2023. Optimal time-step for coupled CFD-DEM model in sand production. *Lect. Notes Comput. Sci. International Conference on Computational Science and Its Applications – ICCSA, 2023 Workshops*. Springer, Cham.
- Kim, G., Lee, G., 2024. Validation of CFD models of urban microclimates under high temperature and humidity conditions during daytime heatwaves in dense low-rise areas. *Build. Environ.* 266, 112087.
- Kim, H., Jung, Y., Oh, J.I., 2019. Transformation of urban heat island in the three-center city of Seoul, South Korea: The role of master plans. *Land Use Policy* 86, 328–338. <https://doi.org/10.1016/j.landusepol.2019.05.016>.
- Kim, J., Kang, J., 2020. Analysis of flood damage in the Seoul Metropolitan government using climate change scenarios and mitigation technologies. *Sustainability* 13, 105. <https://doi.org/10.3390/su13010105>.
- Kim, J., Kang, J., 2022. Evaluating the efficiency of fog cooling for climate change adaptation in vulnerable groups: A case study of Daegu Metropolitan City. *Build. Environ.* 217, 109120. <https://doi.org/10.1016/j.buildenv.2022.109120>.
- Kim, J., Kang, J., 2023a. AI based temperature reduction effect model of fog cooling for human thermal comfort: Climate adaptation technology. *Sustain. Cities Soc.* 95, 104574. <https://doi.org/10.1016/j.scs.2023.104574>.
- Kim, J., Kang, J., 2023b. Development of hazard capacity factor design model for net-zero: Evaluation of the flood adaptation effects considering green – gray infrastructure interaction. *Sustain. Cities Soc.* 96, 104625. <https://doi.org/10.1016/j.scs.2023.104625>.
- Kim, J., Lee, J., Hwang, S., Kang, J., 2022. Urban flood adaptation and optimization for net-zero: Case study of Dongjak-gu, Seoul. *J. Hydrol. Reg. Stud.* 41, 101110. <https://doi.org/10.1016/j.ejrh.2022.101110>.
- Kim, J., Lee, S.Y., Kang, J., 2020. Temperature reduction effects of rooftop garden arrangements: A case study of Seoul National University. *Sustainability* 12, 6032. <https://doi.org/10.3390/su12156032>.
- Lee, H., Mayer, H., Chen, L., 2016. Contribution of trees and grasslands to the mitigation of human heat stress in a residential district of Freiburg, Southwest Germany. *Landsc. Urban Plan.* 148, 37–50. <https://doi.org/10.1016/j.landurbplan.2015.12.004>.
- Lee, K., Kim, Y., Sung, H.C., Kim, S.H., Jeon, S.W., 2022a. Surface urban heat island in South Korea's new towns with different urban planning. *Environ. Monit. Assess.* 194, 360. <https://doi.org/10.1007/s10661-022-09967-w>.
- Lee, S., Yoo, C., Im, J., Cho, D., Lee, Y., Bae, D., 2023. A hybrid machine learning approach to investigate the changing urban thermal environment by dynamic land cover transformation: A case study of Suwon, Republic of Korea. *Int. J. Appl. Earth Obs. Geoinf.* 122, 103408. <https://doi.org/10.1016/j.jag.2023.103408>.
- Lee, W., Choi, M., Bell, M.L., Kang, C., Jang, J., Song, I., Kim, Y.O., Ebi, K., Kim, H., 2022b. Effects of urbanization on vulnerability to heat-related mortality in urban and rural areas in South Korea: A nationwide district-level time-series study. *Int. J. Epidemiol.* 51, 111–121. <https://doi.org/10.1093/ije/dyab148>.
- Li, Y., Schubert, S., Kropp, J.P., Rybski, D., 2020. On the influence of density and morphology on the Urban Heat Island intensity. *Nat. Commun.* 11, 2647. <https://doi.org/10.1038/s41467-020-16461-9>.
- Lowe, D., Ebi, K.L., Forsberg, B., 2011. Heatwave early warning systems and adaptation advice to reduce human health consequences of heatwaves. *Int. J. Environ. Res. Public Health* 8, 4623–4648. <https://doi.org/10.3390/ijerph8124623>.
- Masson-Delmotte, V., Zhai, P., Pirani, S., Connors, C., Péan, S., Berger, N., Caud, Y., Chen, L., Goldfarb, M., Scheel Monteiro, P.M., 2021. The physical science basis. Contribution of Working Group I to the sixth assessment report of the intergovernmental panel on climate change. *IPCC, 2021: Summary for policymakers. Clim. Change*, Cambridge, United Kingdom and New York.
- McNeel, R. et al., 2023. Rhino – New in rhino 7. <https://www.rhino3d.com/7/new/> (Accessed Nov 19, 2023).
- Mimura, N., 2013. Sea-level rise caused by climate change and its implications for society. *Proc. Jpn Acad. Ser. B Phys. Biol. Sci.* 89, 281–301. <https://doi.org/10.2183/pjab.89.281>.
- Ministry of the Environment, 2018. Heatwave response, adaptation to climate change requires consideration of local conditions and capabilities. <https://me.go.kr/home/web/board/read.do?boardMasterId=1&boardId=888690&menuId=286>. (Accessed Nov 19, 2023).
- Nasrollahi, N., Hatami, M., Khastar, S.R., Taleghani, M., 2017a. Numerical evaluation of thermal comfort in traditional courtyards to develop new microclimate design in a hot and dry climate. *Sustain. Cities Soc.* 35, 449–467. <https://doi.org/10.1016/j.scs.2017.08.017>.
- Nazarian, N., Kleissl, J., 2015. CFD simulation of an idealized urban environment: Thermal effects of geometrical characteristics and surface materials. *Urban Clim.* 12, 141–159. <https://doi.org/10.1016/j.ulclim.2015.03.002>.
- Oliveira, A., Lopes, A., Niza, S., Soares, A., 2022. An urban energy balance-guided machine learning approach for synthetic nocturnal surface Urban Heat Island prediction: A heatwave event in Naples. *Sci. Total Environ.* 805, 150130. <https://doi.org/10.1016/j.scitotenv.2021.150130>.
- Ornelas, A., Cordeiro, A., Lameiras, J.M., 2023. Thermal comfort assessment in urban green spaces: Contribution of thermography to the study of thermal variation between tree canopies and air temperature. *Land* 12, 1568. <https://doi.org/10.3390/land12081568>.
- Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., Dasgupta, P., 2014. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. *Synth. Rep. 2014*.
- Parmesan, C., Morecroft, M.D., Trisurat, Y., 2022. Impacts, adaptation and vulnerability. *Clim. Change GIEC*.
- Park, S., Kim, J., Kim, Y., Kang, J., 2024. Participatory Framework for Urban Pluvial Flood Modeling in the Digital Twin Era. *Sustain. Cities Soc.* 108, 105496. <https://doi.org/10.1016/j.scs.2024.105496>.
- Pearce-Higgins, J.W., et al., 2022. A framework for climate change adaptation indicators for the natural environment. *Ecol. Indic.* 136, 108690.
- Pirkler, S., Lichtenegger, T., 2018. Efficient time-extrapolation of single- and multiphase simulations by transport based recurrence CFD (rCFD). *Chem. Eng. Sci.* 188, 65–83. <https://doi.org/10.1016/j.ces.2018.04.059>.
- Rentschler, J., Avner, P., Marconcini, M., Su, R., Strano, E., Vousdoukas, M., Hallegatte, S., 2023. Global evidence of rapid urban growth in flood zones since 1985. *Nature* 622, 87–92. <https://doi.org/10.1038/s41586-023-06468-9>.
- Rodriguez, C.E.A., Velazquez, J.F., 2019. CFD simulation of heat and mass transfer for climate control in greenhouses. In: Irazoqui, A. (Ed.), *Heat and Mass Transfer – Advances in Science and Technology Applications*. IntechOpen.
- Rowland, Z., Cug, J., Nica, E., 2022. The geopolitics of smart city digital twins: Urban sensing and immersive virtual technologies, Spatio-temporal Fusion Algorithms, and Visualization Modeling Tools. *Geopolit. Hist. Int. Relat.* 14 (2), 56–71.
- Salata, F., Golasi, I., de Lieto Vollaro, R., de Lieto Vollaro, A., 2016. Urban microclimate and outdoor thermal comfort. A proper procedure to fit ENVI-met simulation outputs to experimental data. *Sustain. Cities Soc.* 26, 318–343. <https://doi.org/10.1016/j.scs.2016.07.005>.
- Seddon, N., Smith, A., Smith, P., Key, I., Chausson, A., Girardin, C., House, J., Srivastava, S., Turner, B., 2021. Getting the message right on nature-based solutions to climate change. *Glob. Change Biol.* 27, 1518–1546. <https://doi.org/10.1111/gcb.15513>.
- Shao, G., Kibria, D., 2018, December. Digital manufacturing: Requirements and challenges for implementing digital surrogates. In *2018 Winter Simulation Conference (WSC)*. IEEE, pp. 1226–1237. <https://doi.org/10.1109/WSC.2018.8632242>.
- Shih, T.-H., Liou, W.W., Shabbir, A., Yang, Z., Zhu, J., 1995. A new k-e eddy viscosity model for high Reynolds number turbulent flows. *Comput. Fluids* 24, 227–238. [https://doi.org/10.1016/0045-7930\(94\)00032-T](https://doi.org/10.1016/0045-7930(94)00032-T).
- Shukla, P.R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H.O., Roberts, D., Zhai, P., Slade, R., Connors, S., Van Diemen, R., 2019. *IPCC, 2019: Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*.
- Taleb, D., Abu-Hijleh, B., 2013. Urban heat islands: Potential effect of organic and structured urban configurations on temperature variations in Dubai, UAE. *Renew. Energy* 50, 747–762. <https://doi.org/10.1016/j.renene.2012.07.030>.
- Teoh, M.-Y., Shinozaki, M., Saito, K., Said, I., 2022. Developing climate-led landscapes and greenery in urban design: A case study at Ipoh, Malaysia. *J. Asian Archit. Build. Eng.* 21, 1640–1656. <https://doi.org/10.1080/13467581.2021.1942881>.
- Thompson, V., Mitchell, D., Hegerl, G.C., Collins, M., Leach, N.J., Slingo, J.M., 2023. The most at-risk regions in the world for high-impact heatwaves. *Nat. Commun.* 14, 2152. <https://doi.org/10.1038/s41467-023-3754-1>.
- Tominaga, Y., Mochida, A., Yoshie, R., Kataoka, H., Nozu, T., Yoshikawa, M., Shirasawa, T., 2008. AJI guidelines for practical applications of CFD to pedestrian

- wind environment around buildings. *J. Wind Eng. Ind. Aerodyn.* 96, 1749–1761. <https://doi.org/10.1016/j.jweia.2008.02.058>.
- Toparlar, Y., Blocken, B., Maiheu, B., Van Heijst, G.J.F., 2018a. Impact of urban microclimate on summertime building cooling demand: A parametric analysis for Antwerp, Belgium. *Appl. Energy* 228, 852–872. <https://doi.org/10.1016/j.apenergy.2018.06.110>.
- Toparlar, Y., Blocken, B., v, Maiheu, B., Van Heijst, G.J.F., 2018b. The effect of an urban park on the microclimate in its vicinity: A case study for Antwerp, Belgium. *Int. J. Climatol.* 38, e303–e322. <https://doi.org/10.1002/joc.5371>.
- Toparlar, Y., Blocken, B., v, Vos, P., Van Heijst, G.J.F., Janssen, W.D., van Hooff, T., Montazeri, H., Timmermans, H.J.P., 2015. CFD simulation and validation of urban microclimate: A case study for Bergpolder Zuid, Rotterdam. *Build. Environ.* 83, 79–90. <https://doi.org/10.1016/j.buildenv.2014.08.004>.
- Vu, A., Rutherford, S., Phung, D., 2019. Heat health prevention measures and adaptation in older populations-A systematic review. *Int. J. Environ. Res. Public Health* 16, 4370. <https://doi.org/10.3390/ijerph16224370>.
- Ye, X., Du, J., Han, Y., Newman, G., Retchless, D., Zou, L., Ham, Y., Cai, Z., 2023. Developing Human-Centered Urban Digital Twins for Community Infrastructure Resilience: A Research Agenda. *J. Plan. Lit.* 38, 187–199. <https://doi.org/10.1177/08854122221137861>.
- Zhang, M., Zhang, C., Kafy, A.-A., Tan, S., 2021. Simulating the relationship between land use/cover change and urban thermal environment using machine learning algorithms in Wuhan City, China. *Land* 11, 14. <https://doi.org/10.3390/land11010014>.
- Zhang, X., 2009. Danmarks Tekniske Universitet, Risø Nationallaboratoriet for Bæredygtig Energi. CFD Simul. Neutral ABL Flows.
- Zhu, H.C., Xi, C., Ren, C., Wang, J., Cao, S.J., 2024. High-resolution urban temperature simulation method considering various spatiotemporal boundary impacts. *Phys. Fluids* 36 (7). <https://doi.org/10.1063/5.0215759>.