Understanding the impacts of space design on local outdoor thermal comfort: An approach combining DepthmapX and XGBoost

Ye Xia [a], Weisheng Lu [a,*], Ziyu Peng [a], Jinfeng Lou [a], Jianxiang Huang [b], Jianlei Niu [c]

[a] Department of Real Estate and Construction, Faculty of Architecture, The University of Hong Kong, Knowles Building, Pokfulam, Hong Kong
[b] Department of Urban Planning and Design, Faculty of Architecture, The University of Hong Kong, Hong Kong
[c] Department of Building Environment and Energy Engineering, Faculty of Construction and Environment, The Hong Kong Polytechnic University, Hong Kong

ARTICLE INFO

ABSTRACT

Researchers and designers alike are keen to explore proper space designs to alleviate the heat waves. A niche research area of this type focuses on the smaller scale, local communities to achieve the so-called 'cool spots' for residents. However, despite the momentum gained so far, it remains unclear what the key design elements are (e. g., space orientation, aspect ratio, or building density) and how they interact with each other in impacting outdoor thermal comfort (OTC), particularly in some complex, vertical 'spots'. This research aims to provide an improved understanding of cool spot reasoning by proposing a new paradigm to engage DepthmapX (an analytic tool for urban spatial configuration), XGBoost (a machine learning tool), and on-site verification. By imple- menting the paradigm on a university campus in Hong Kong for the case study, it was discovered that Percentage of View (PV), Average Height Index (AHI), and Connectivity (CON) are the three most influential factors leading to the formation of a 'cool spot'. DepthmapX can not only help quantify space designs but also help translate the indexes back to real-life space design options. The XGBoost can help better interpret the pathway from different space design indexes to different OTC but more explainable causal relationships are desired. This research advanced our understanding of the impacts of different space designs on OTC and provided references to de- signers in achieving OTC in smaller–scale, local communities. It also opens a new avenue to understand the causal relationships in a more detailed and explainable fashion.

1. Introduction

The rampant heat waves and their induced heat island effects in urban areas have motivated designers and researchers alike to investi- gate strategies to improve Outdoor Thermal Comfort (OTC) (Chen & Ng, 2012b) [18,25]. These include modifying urban geometry, incorpo- rating vegetation, using reflective surfaces, and implementing water bodies in cities [11,15,25]. Amongst these approaches, space design and configuration are found to have a significant impact on microclimate diversity (Chan et al., 2021; [41]; Ouali et al., 2020; [4]. Researchers have examined a range of urban design elements to enhance the local OTC. These elements, which can be categorized as 2D or 3D, may include sky views (Lin et al., 2010), aspect ratio (Jamei et al., 2016), surface density (Ouali et al., 2020), and others. Impressive progress has been achieved so far in this research field.

Yet, it remains ambiguous what the key elements are, and how they interact with each other to impact OTC. Clearly, these design elements

will impose different effects. They are also dynamic and must be further connected with specific climate zones and local weather profiles to take effect [37]. Two factors seem to add to the complexity of this research field. Firstly, researchers and designers tend to focus on broader urban areas (e.g., a district, a city's wind corridor) with scant effort paid to local, smaller communities (e.g., a street corner playground, or a pavilion in a community green park). These smaller, local communities are the places where people, particularly the elderly or children tend to stay [6,42]. These are also the places that could be possibly manipulated by designers among the existing built-up areas. Secondly, it seems that researchers and designers lack handy tools to analyze the impacts of space design elements on the OTC, particularly for the complex, vertical local communities in dense urban areas. Except for the established Computational Fluid Dynamics (CFD) [2] and the Universal Thermal Climate Index (UTCI) [44], researchers are now tapping into the recent advancements in spatial analytics and machine learning (ML).

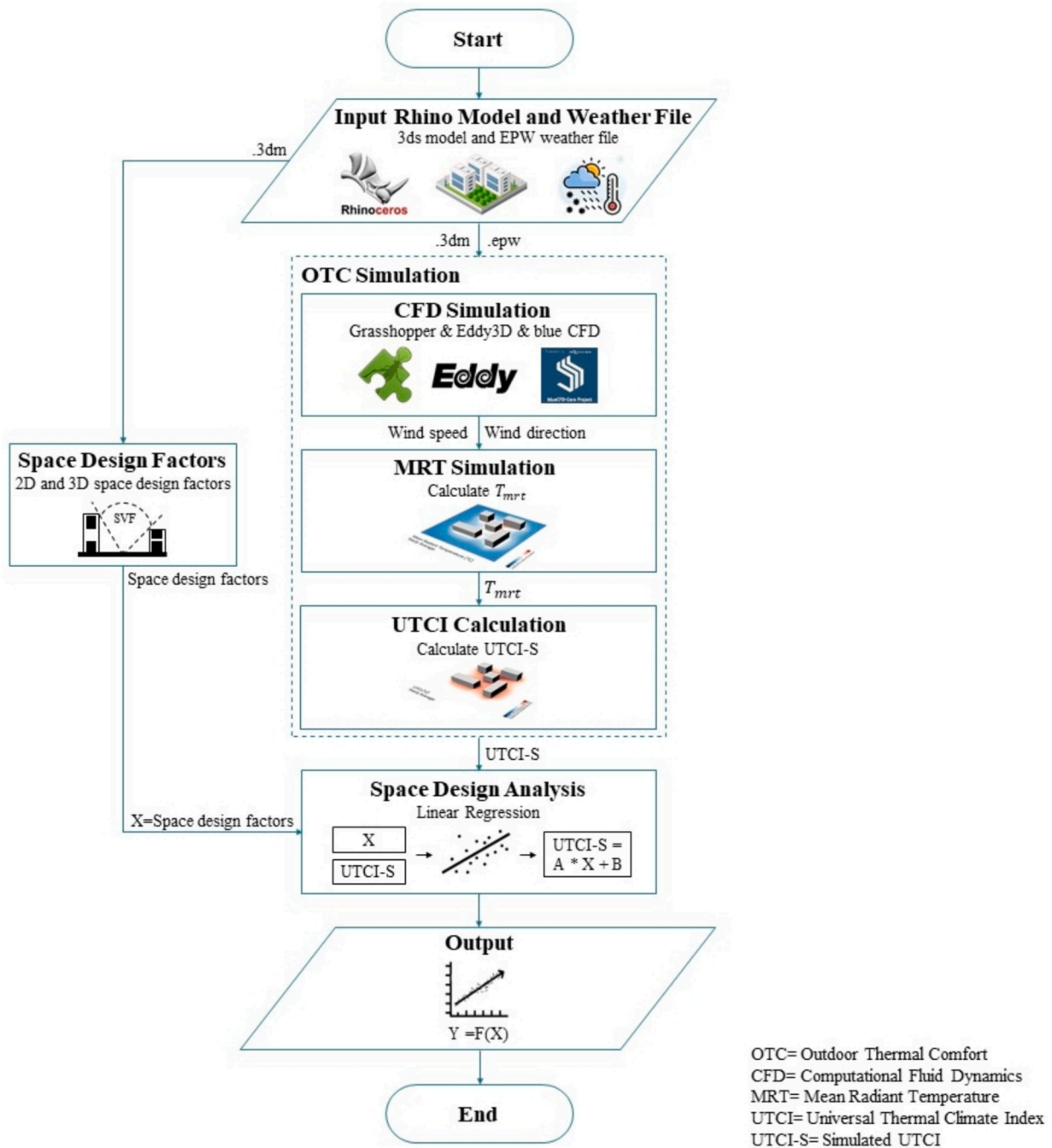This research aims to improve our understanding of the impacts of

Fig. 1. The traditional paradigm of 'cool spot' reasoning.

space design on the OTC of small local communities. These communities with cozy UTCI have a nickname of 'cool spots'. The niche research field can be called cool spot reasoning or auditing. Traditionally, this reasoning work was conducted by adopting a reiterative approach of conducting OTC simulation by positioning the building information models (BIM), simulating mean radiant temperature (MRT), and monitoring UTCI. This research proposes a new paradigm by engaging several analytic tools including DepthmapX, an analytic tool for urban spatial configuration, to improve the static nature of inputting information; and XGBoost, a machine learning (ML) tool, to enhance the simple linear process of reasoning. The new paradigm is then applied in several real-life, observed 'cool spots' on a university's campus in Hong Kong to compare their performance and understand the pathway through which the different space designs can make an impact.

With this vignette, the rest of the paper is organized as follows. Section 2 introduces several space design elements, their impact on OTC, and the traditional paradigm of 'cool spot' reasoning. Section 3 delineates the research methods and a case study. Section 4 describes the results and Section 5 provides a discussion of the results. Section 6 gives the conclusions of this research.

2. Space design elements and OTC

The literature has identified several urban design elements engaged in previous research and their different effects on the urban thermal environment.
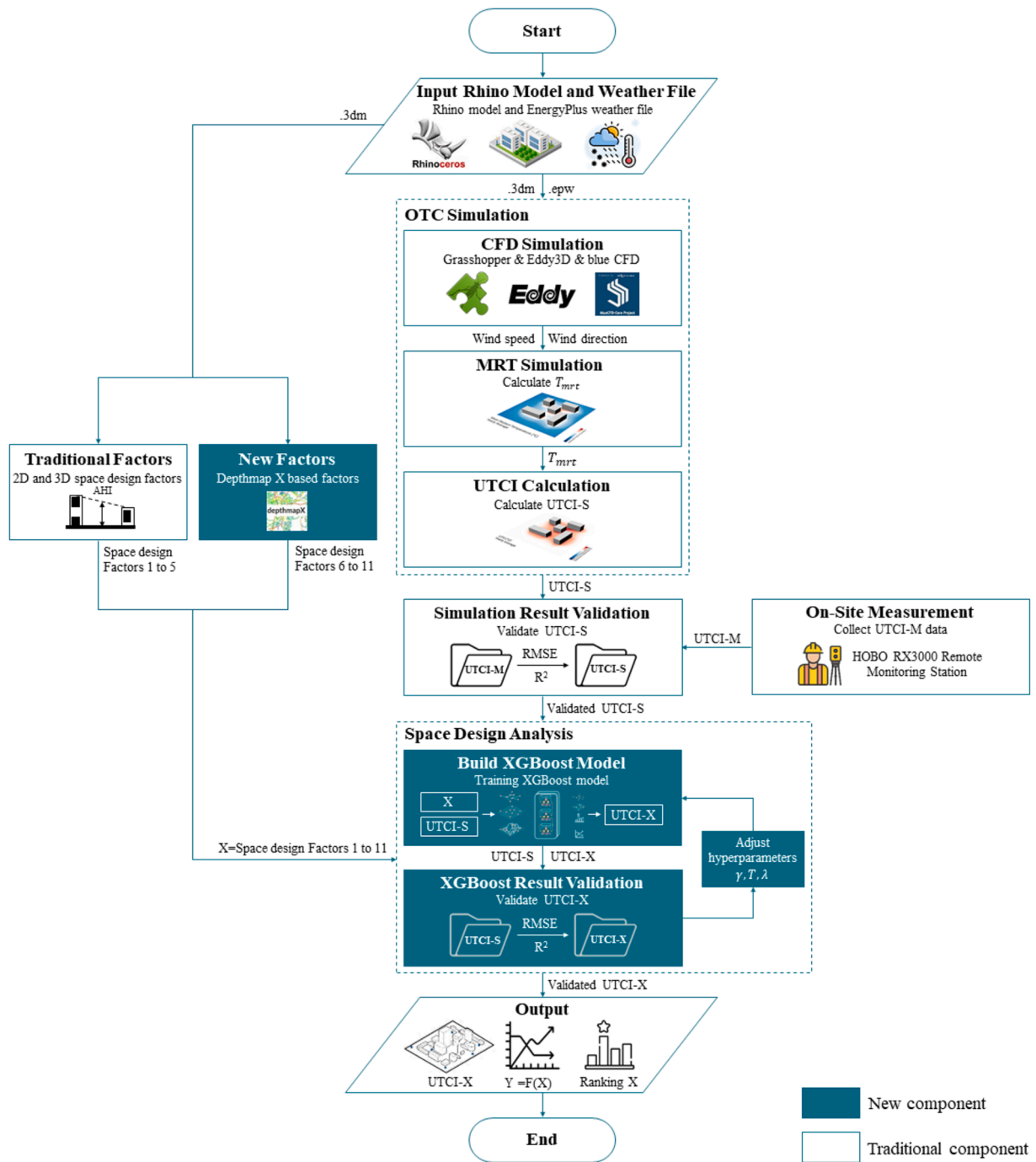
Fig. 2. A new paradigm of 'cool spot' reasoning.

2.1. Space orientation and aspect ratio

The most common factor for evaluating OTC is space orientation [20], which often comes with aspect ratio [35]. They have a significant influence on the thermal environment and consequently, the sensation of people at the same time. For example, Ali-Toudert and Mayer [1] found that canyons with an N-S orientation and high aspect ratio (equal to or greater than H/W = 2) have a much better thermal environment. In a dense central area, Thessaloniki in northern Greece, thermal comfort indices differed a lot due to variations in canyon orientation and differences in aspect ratio [9]. Additionally, in a CFD simulation on OTC in the central business districts of Nanjing, China, Deng & Wong [12] found that E-W streets or setback canyons experience the warmest

thermal conditions and the longest extreme heat durations, regardless of the aspect ratio. However, the duration of the extreme heat ratio decreases as the aspect ratio increases.

2.2. Building density and building height

Building density and height are other important pair of factors that affect OTC. Perini and Magliocco [34] conducted a study in three different locations in Italy (Milan, Genoa, and Rome) on a typical summer day and the hottest summer day to quantify the effects of different atmospheric conditions in a Mediterranean climate. Their study found that higher building density results in higher temperatures, while the height of buildings plays an important role in determining

potential temperatures at 1.6 m from the ground level. Taller buildings provide shading that reduces temperatures, thereby mitigating thermal discomfort. These findings demonstrate the importance of considering building density and height in urban design strategies to improve OTC.

2.3. Sky view factor

Sky view is one of the most studied factors in space design for OTC. The factor measures the openness of a space, which is closely related to solar radiation and wind speed [22]. For example, Lin et al. (2010b) conducted 12 field experiments on a university campus in central Taiwan and found that sky view significantly affects outdoor thermal environments. Their analytical results showed that a high-sky view causes discomfort in summer, while a low-sky view causes discomfort in winter. However, simulation studies conducted in Telheiras by Andrade & Alcoforado [4] showed the coolest conditions occur in open areas such as the center of larger courtyards and marginal areas with high sky views, while the warmest conditions occur in streets with low sky views that are sheltered from the wind. These findings demonstrate the complex relationship between sky view and thermal comfort.

In real-life cases, these factors impose different effects when they have different geometries or features. Moreover, they will interact with each other, e.g., in 2D or 3D ways of connectivity and opening, to influence OTC. There has long been a stream of active research to reason or audit how these different designs can impact the achievement of 'cool spots'.

2.4. 'Cool spot' reasoning

Extant studies to audit cool spots largely follow a paradigm as shown in Fig. 1. It consists of two key components: OTC simulation and space design analyses (SDA).

The OTC simulation takes three types of inputs (1) building information models, (2) terrain and landscape of a complex urban environment, and (3) climate zone and more detailed local weather profiles. It starts with CFD simulation, which will generate the outcomes of airflow distributions, including directions and speeds [5,26]. MRT simulations will take the outcome of the CFD simulation and calculate the mean radiant temperature (T_{mrt}), performs calculations of previous results, and reaches the outcome of OTC [40]. For small local communities, the Universal Thermal Climate Index (UTCI) is suitable for the study of microclimate on the assessment of OTC [8,17,31]. In this case, OTC simulation will generate a simulated UTCI, which is called UTCI-S.

The SDA is mainly for analysts to understand the primary research question, i.e., how different space designs will impact the OTC as measured by UTCI-S. This process can be modeled as a general Eqn (1):

$$X \rightarrow Y \tag{1}$$

where X refers to the space design factors and Y represents the UTCI as caused by X. This process involves complicated analyses but traditionally it is performed with simple linear regression. The space designs they can analyze are simplified without considering their vertical, complex urban environment, etc. The output of this traditional design will be a list of simple design factors and how they cause the UTCI. Hence, the understanding of the nexuses between space designs and local OTC/UTCI is often very limited. The task of cool spot reasoning desires more sophisticated, yet effective approaches to handle the complex dynamics between design elements, configuration, and impacts on OTC [45].

3. Methods

3.1. A new paradigm

The methodology adopted in this study is shown in Fig. 2, which can be read in conjunction with the traditional paradigm as shown in Fig. 1.

One may notice that the new paradigm largely maintains the same structure of components as shown in Fig. 1. However, it introduces four major different components, as shown in different colors in Fig. 2. First, the new paradigm adopts ML tools (e.g., XGBoost) rather than linear regression for SDA to decipher the nexuses between space designs and OTC/UTCI. Second, the research adopts DepthmapX to investigate the more sophisticated urban designs and provides more structured inputs (i.e., X) for the SDA. In addition, by observing that previous studies largely conducted computer simulations without referring to ground truth data, this research adds a new step to verify the output of UTCI simulation (UTCI-S) by conducting onsite measurement (UTCI-M) and comparing the simulated and measured UTCI. Only the verified UTCI can be accepted as the input for the DepthmapX and XGBoost combined analyses. The major components of the new paradigm are introduced in greater detail in the following sections.

3.2. OTC simulation

3.2.1. Digital model and weather input

To conduct an OTC simulation, geometric input is necessary. Boundary representations (e.g., a 3ds format digital model in Rhino) reflecting the physical systems of the real world are used [30]. Weather data is collected as input for OTC simulation. This information is accessible from official weather platforms, such as EnergyPlus, a website with weather data in EPW format for simulation. This weather data should be set up at the same time as the on-site measurement for consistency validation.

3.2.2. CFD simulation

As CFD simulation based on wind tunnel experiments is too expensive and resource-demanding, CFD simulation with Eddy3D, a plugin for Rhino and Grasshopper is adopted in this research. The CFD simulation follows a detailed simulation framework as shown in Fig. 2. Grasshopper, Eddy3D, and blue CFD, which utilizes meshes for OpenFOAM analyses [21].

In detail, Eddy3D, the Grasshopper plugin allows outdoor thermal comfort simulation by calculating wind speed for a given spatial environment and research period [29]. OpenFOAM's *blockMesh* is used for the background environment within the simulation domain, and *snappyHexMesh* is to snap the related background mesh to the study area. This allows to reuse the same simulation mesh for each wind direction, greatly reducing the simulation time and storage volume. This approach allows this study to achieve faster single-node simulations for wind velocity. This research adopts a 2 m grid of points for the community scale OTC research, based on accuracy and calculation speed, and previous research [3,19,27,39,48].

3.2.3. MRT simulation and UTCI calculation

After the airflow distributions are available based on CFD simulation, the methodology turns to MRT simulation. For MRT simulation within Eddy3D, T_{mrt} is calculated based on the building surface temperature, which is calculated with the Surfer algorithm, suggesting surface orientation and shading patterns are the main significant contributors to T_{mrt}. Eddy3D runs sky view analysis with Radiance to account for direct solar gain, accurately simulating diffuse and reflected radiation, and representing geometry and materials. Specifically, this tool uses ray tracing to simulate diffuse and reflected radiation, more accurately representing geometry and material [21]. Besides built environment elements, vegetation is also simulated in Eddy3D with EnergyPlus. Vegetation's translucency is set to be 20 % for the consideration of shading effects. Specifically, T_{mrt} is defined as the result of the view factor weighted building surface temperature, as follows:

$$T_{mrt} = T_{surf} \cdot F_{surf} + T_{sky} \cdot F_{sky} + \Delta T_{mrt,p} \tag{2}$$

in which T_{surf} refers to the surface temperature of all surrounding

Table 1
Technical description of the sensors used in field studies.

Variable	Unit	Resolution	Accuracy
Air Temperature	°C	< 0.03 oC, from 0 oC to 50oC	< ±0.2 °C from 0° to 50 °C
Relative Humidity	–	0.1 % RH	+/- 2.5 % from 10 % to 90 % RH, to a maximum of +/- 3.5 % including hysteresis at 25 °C; below 10 % and above 90 % ±5%
Wind velocity	m/s	0.4 m/s	±0.8 m/s or ± 4 % of reading, whichever is greater
Wind direction	deg	1 deg	0.2 to 3 m/s: ±4 degrees > 3 m/s: ±2 degrees
Black Bulb Temperature	°C	0.1 oC	±0.6°C

environment, including buildings and ground of the study point(°C), F_{surf} represents the surface view factor, T_{sky} corresponds to sky temperature (°C), F_{sky} is the sky view factor, and $\Delta T_{mrt,p}$ refers to the temperature offset based on the outdoor setting of a person (°C). Once the T_{mrt} obtained, UTCI based on OTC simulation (UTCI-S) will be calculated using Eqn. (3):

$$UTCI(T_a, v_a, p_a, T_{mrt}) = T_a + Offset(T_a, v_a, p_a, T_{mrt}) \qquad (3)$$

in which p_a refers to the water vapor pressure (kPa), and the offset refers to, such as the deviation of UTCI from air temperature and other factors.

Admittedly, this simulation does not account for all the responsible causes for the urban heat island effect. This simulation tool does not incorporate the connection between surface temperature and wind speed, which contributes to forced convection. But this effect is considered in the MRT module, and research shows this effect explains less than 1 % variance of the UTCI [21].

3.3. On-site measurement

This research adopts the Onset HOBO RX3000 Remote Monitoring Station to collect on-site meteorological data (UTCI-M). The station is set up at a height of 1.5 m above the ground, which is the standard human sensation height. The measured mean radiant temperature T_{mrt} is calculated using Eqn. (4):

$$T_{mrt} = \left[(T_g + 273.15)^4 + \frac{1.1 \cdot 10^8 \cdot v_a^{0.6}}{\varepsilon \cdot D^{0.4}} (T_g - T_a) \right]^{1/4} - 273.15 \qquad (4)$$

in which T_g refers to the global temperature, which is the black bulb temperature in this research (°C), v_a represents the air velocity at the level of the globe (m/s), ε corresponds to the emissivity of the globe (no dimension), D is the diameter of the globe (m); and T_a refers to the air temperature (°C). Once the T_{mrt} obtained, UTCI based on measurement will be calculated using Eqn. (3). Table 1 refers to a breakdown of these sensors and Fig. 3 shows their features.

3.4. Simulation result validation

Very often, existing studies take the output of UTCI simulation (i.e., UTCI-S) as the input for further analyses. There is a good chance that the simulated results will deviate from the ground truth. Therefore, this research adds an extra layer of verification by introducing the results of on-site measurement. A data set of measured UTCI (called UTCI-M) is obtained as the ground truth data to be compared with UTCI-S.

Two indicators, namely the root mean square error (RMSE) [13,32] and the coefficient of determination (R^2) [43,46], are adopted to measure the resemblance of the two sets of data. The principle is that the smaller the RMSE and the larger the R^2 values (the closer R^2 to 1), the better the resemblance of the simulated data and the ground truth data. RMSE measures the average difference between a statistical model's predicted values and the actual values. RMSE is given as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2}{N}} \qquad (5)$$

where y_i and $\widehat{y}_i$ are the UTCI-M and UTCI-S, respectively, and N is the number of samples.

R^2 represents the proportion of the variation in the dependent variable that is predictable from the independent variable(s). R^2 is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \widehat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \overline{y}_i)^2} \qquad (6)$$
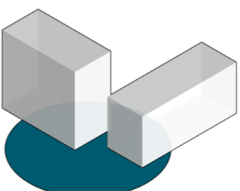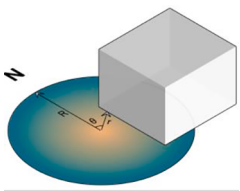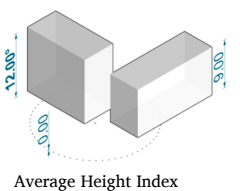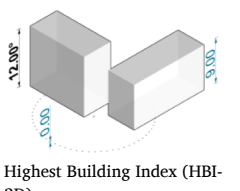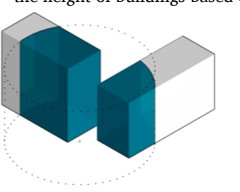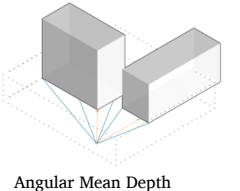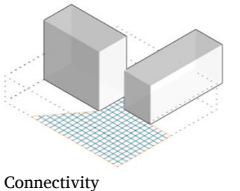
3.5. Space design analyses using DepthmapX and Rhino

Unlike the traditional way of interpreting the nexuses between space designs and OTC, this research adopts DepthmapX to handle the multiple spa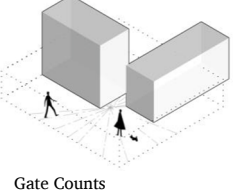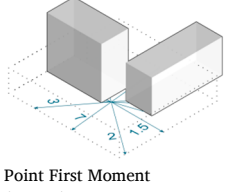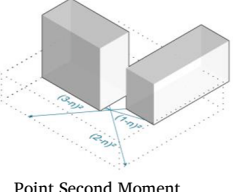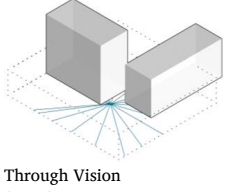ce design elements and their complex 3D, vertical configurations. DepthmapX was developed as a software platform to perform a set

Fig. 3. Onset HOBO RX3000 Remote Monitoring Station.

Table 2

Description of factors considered in the study.

Rhino-based Factors

Indicator	Description	Indicator	Description
Percentage of View (PV-2D) Note: S_{view} refers to the area of view; S_{buffer} refers to the area of the buffer, the radius of the buffer is set to be 10 m.	the area of view divided by the area of buffer $$PV = \frac{S_{view}}{S_{buffer}}$$	Orientation of View (OV-2D) Note: θ_i refers to the angle between a view vector and north; r_i refers to the length of view; R_{buffer_i} refers to the radius of the buffer, which is 10 m; n is the total number of view directions, which is 30.	the sum of angles between north and view directions, which are weighted by view length $$OV = \frac{\sum_{i=1}^{n}\theta_i r_i}{\sum_{i=1}^{n} R_{buffer_i}}$$
Average Height Index (AHI-3D) Note: n is the total number of study points in the buffer, which radius is 10 m ; H_i refers to the height of buildings based on the study points	the average height in the buffer area $$AHI = \frac{1}{n}\sum_{i=1}^{n} H_i$$	Highest Building Index (HBI-3D) Note: H_{max} refers to the maximum height of buildings based on study points; the meaning of other parameters is the same as AHI	the height of the highest building divided by the average height in the buffer area $$HBI = \frac{H_{max}}{\frac{1}{n}\sum_{i=1}^{n} H_i}$$
Building Volume Index (BVI-3D) Note: $V_{buildings}$ is the volume of buildings ; V_{buffer} refers to buffer volume, which is based on the buffer area (radius is 10 m and height is 100 m which is higher than the highest building of the campus).	the volume of buildings divided by the buffer volume $$BVI = \frac{V_{buildings}}{V_{buffer}}$$		

DepthmapX-based Factors

Indicator	Description	Indicator	Description
Angular Mean Depth (AMD-X) Note: U_i is the lowest angular depth calculated from point to successive stochastic global sample locations, n is the total number of directions	the average of all lowest angular depths calculated from point to all successive stochastic global sample locations $$AMD = \frac{1}{n}\sum_{i=1}^{n} U_i$$	Connectivity (CON-X) Note: c_i is the cell visible from a specific cell	the total number of cells visible from a specific cell $$CON = \sum c_i$$
Gate Counts (GC-X) Note: f_{people_i} is the flow of people at sampled locations	establish the flows of people at sampled locations within the environment over the course of a day $$GC = \sum f_{people_i}$$	Point First Moment (PFM-X) Note: $d(v_i, v_j)$ is the steps from one node to other visible nodes; v_i refers to the study node, v_j refers to other visible nodes	the sum of the steps from one node to every other visible node in the system, an indicator of how far it can be seen $$PFM = \sum d(v_i, v_j)$$
Point Second Moment (PSM-X) Note: the meaning of the parameters of PSM is the same as PFM	the standard deviation of the connectivity value, the physical measure is relevant for estimating the perimeter of a space $$PSM = \sum d(v_i, v_j)^2$$	Through Vision (TV-X) Note: the number of times a study point is crossed by a line	the total number of times a study point is crossed by lines drawn between the centroids of all other inter-visible cells $$TV = \sum N_{time}$$

of spatial network analyses designed to understand space configuration within the built environment. It works at a variety of scales from building through small urban to whole cities or states. Here the scale of interest is a small urban community level. This research utilizes eleven factors to quantify outdoor spaces' character, as shown in Table 2. The eleven factors used in this research comprise two 2D factors, three 3D factors, and six DepthmapX-based factors. These factors, grouped in a set of X, will be analyzed further to understand the complex nexuses between space design factors (X) and thermal value (Y).

3.5.1. 2D and 3D factors

Specifically, we select two 2D factors and three 3D factors to understand their impact on the thermal quality of the outdoor spaces. These factors include Percentage of View (PV), Orientation of View (OV), Average Height Index (AHI), Highest Building Index (HBI), and Building Volume Index (BVI). All factors are calculated with Rhino and Grasshopper, and the radius of the standard buffer area is set to be 10 m based on previous research [36]. Following a sensitivity analysis of the radius of the buffer area, 10 m is further verified to be the most proper and representative of the average space design factors for this community-scale research, which can be found in the data analyses section of this research. The detailed calculation formulas are listed in Table 2.

PV measures the openness of the study point, which is greatly related to solar radiation and the space sensation of a certain point. OV is the main direction of a point, and building volume will reduce the direction value of this point, which makes OV a perfect factor for evaluating the impact of space design. AHI represents the average height of the buffer area, and HBI evaluates the highest building height, which showcases the deviation of building height. BVI measures the building volume, because the more volume involved, there will be more shade and less wind, connecting architectural design and thermal environment. AHI, HBI, and BVI factors are based on the height and volume of buildings, which have a significant impact on shadow, solar radiation, and local winds, incorporating 3D elements into this research.

3.5.2. DepthmapX-based factors

DepthmapX, which follows the Space Syntax theory, offers a comprehensive analysis of urban spatial configuration [38]. DepthmapX established a set of techniques for representing, quantifying, and interpreting the spatial configuration and structure of the built environment with visibility, accessibility, and connectivity analysis [7].

These factors include Angular Mean Depth (AMD), Connectivity (CON), Gate Counts (GC), Point First Moment (PFM), Point Second Moment (PSM), and Through Vision (TV). These factors are primarily used to describe visibility and connectivity in the urban environment. However, they can also provide an integrated perspective to evaluate the impact of spatial configuration on thermal comfort. Specifically, AMD examined the space depth, CON values the visual connectivity, GC establishes the total flows of people, PFM and PSM work together to measure the shape of the isovist, and TV refers to the visibility. by incorporating these indexes into this research, we could quantify space design and architectural concepts, examine the connection between space sensation and OTC, connect design with scientific research, and explore accurately about built environment and urban heat island impact.

3.6. Analyzing the nexuses between space design and OTC using XGBoost

3.6.1. The algorithm of XGBoost

The utilization of ML approaches has been increasingly observed in thermal comfort research in recent years [47]. One such tool that has gained popularity in this field is extreme gradient boosting (XGBoost). In this study, we chose XGBoost as the ML method for four reasons:

- This advanced version of classic gradient boosting avoids overfitting issues, making it a highly effective and accurate tool for analyzing and predicting OTC [24,28].
- It is based on partial dependence [16]. Compared with other linear regression, it allows us to explore the true relationship between several inputs and the target [33].
- It implies a gradient-boosting algorithm [14], making it suitable for large datasets.
- XGBoost could provide satisfactory results in ML research with limited time consumption, while other ML methods would need more time to produce satisfying results [10].

The XGBoost assembles several weak tree learners to minimize the following objective function (Eqn. (6)):

$$\mathscr{L}^{(t)} = \sum_{i=1}^{n} l\left(Y_i, \widehat{Y}_i^{(t-1)}\right) + \Omega(f_t) \tag{7}$$

in which $\mathscr{L}^{(t)}$ refers to the objective function at time step t, which consists of two components:

- the loss function l that calculates the error between the input Y_i, c.f. Eqn. (6), the estimated $\widehat{Y}_i^{(t-1)}$ from the last time step $t-1$;
- regularization term $\Omega(f_t)$ at time step t. In this research, the mean squared error (MSE) in Eqn (7) is applied to calculate l and tree complexity in Eqn (8) as the regularization term:

$$l\left(Y_i, \widehat{Y}_i^{(t-1)}\right) = 1 \Big/ n \sum_i \left(Y_i - \widehat{Y}_i^{(t-1)}\right)^2 \tag{8}$$

$$\Omega(f_t) = \gamma T + 1 \Big/ 2\lambda \sum_j^T \Omega_j^2 \tag{9}$$

in which λ is and regularization parameters to be tuned; γ represents the number of leaves and T the number of nodes; and Ω_j^2 calculated the L2-norm complexity score of each node to prevent over-fitting. Through this exercise, the predicted UTCI will be generated, which is UTCI-X.

3.6.2. Xgboost result validation

The effectiveness of UTCI-X will be examined by comparing it with UTCI-S. Similar to the validation of the simulation result in Section 4.4, RMSE and R^2 will be applied in this part, where y_i and $\widehat{y}_i$ are the UTCI-S and UTCI-X. This process will run several times until we get a reliable result. The 'caret' platform in R is applied to optimize Eqn (7) and identify the optimal model with a 10-fold cross-validation technique to find the optimum set of hyperparameters $\{\gamma, T, \lambda\}$ in Eqn (8). After obtaining the optimal model, a validated set of UTCI-X will be generated.

3.6.3. Xgboost analysis

The XGBoost analysis process is shown in Fig. 4, which employs points of interest area as the research subjects. This analysis process can be described as the following steps:

1) It takes the validated UTCI-X (Y_i) and space design factors (X_i) of a set of points as input.
2) After that, the data set $\{(X_i, Y_i)\}_{i=1}^n$ will be introduced to the XGBoost model. Mathematically, this model could be written in Eqn (9):

$$\widehat{Y}_i = \sum_{k=1}^{K} f_k(X_i), f_k \in \mathscr{F} \tag{10}$$

in which K refers to the number of trees, f_k represents the function of functional space $\mathscr{F}$, $\mathscr{F}$ resembles all possible classification and regression trees. The objective function is optimized by Eqn (6), as
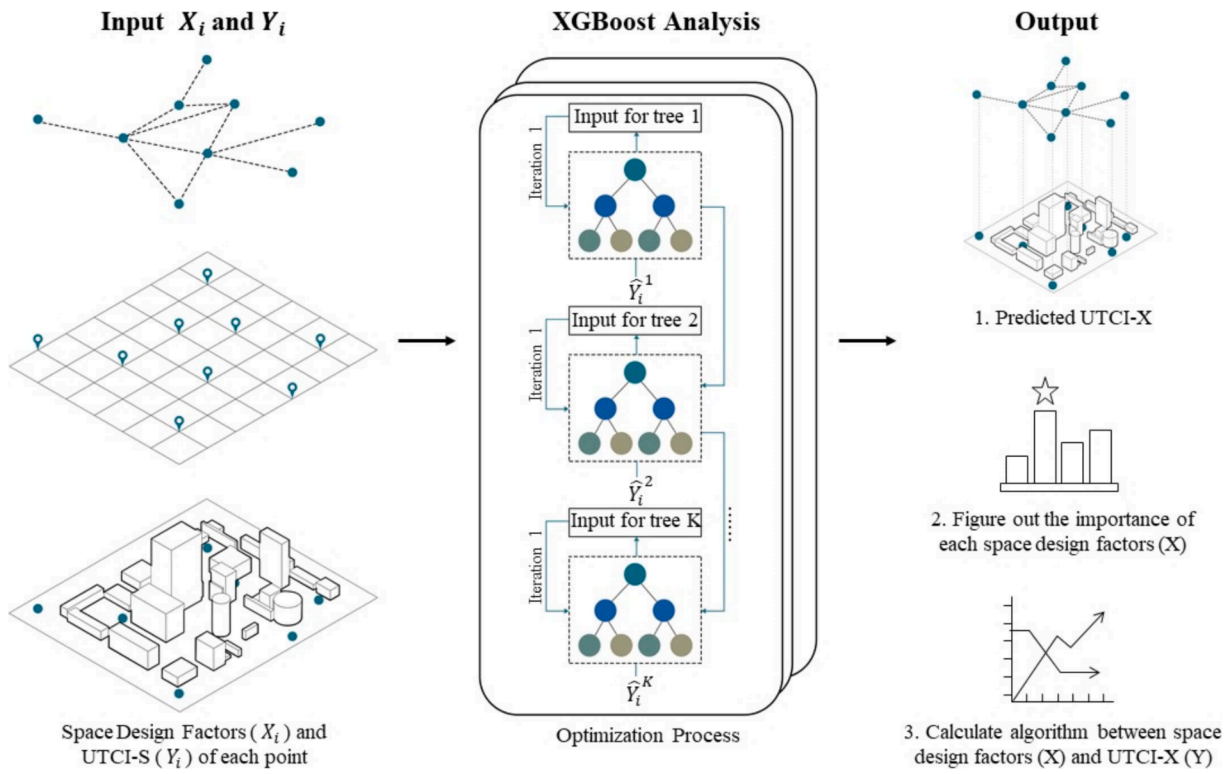
Fig. 4. Graphical scheme of XGBoost analyses.

explained above.

3) Finally, the XGBoost analyses could generate predicted UTCI-X, rank all space design factors according to their importance, and calculate the algorithm between *X* and *Y*. This paradigm could advance the understanding of space design and OTC.

3.7. Case study

3.7.1. Study area

The field study was conducted at the University of Hong Kong, situated in the Central and Western District of Hong Kong Island, Hong Kong. The university's main campus covers an area of 14 ha (720 m x 330 m) (see Fig. 5) and has a latitude of 22°16′58″N, a longitude of 114°8′12″E, and an altitude of 115 m.
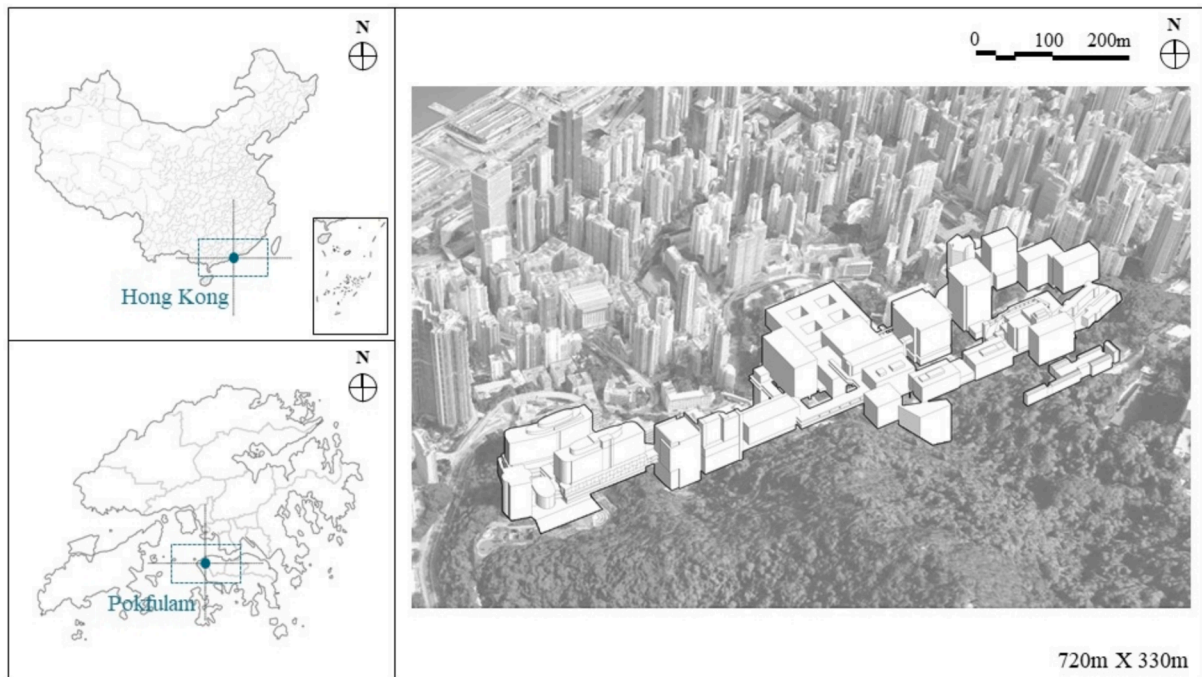

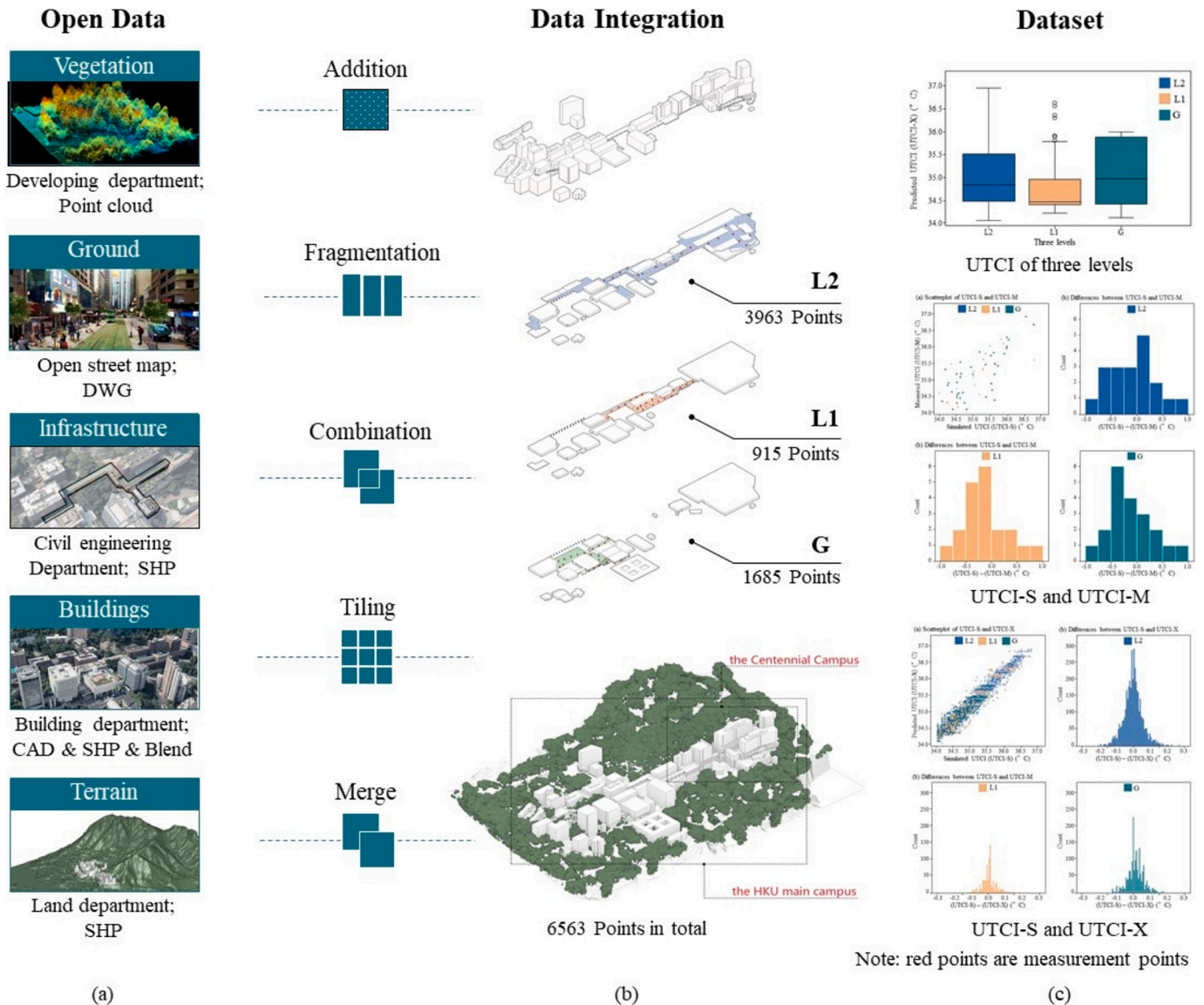
Fig. 5. The case campus as the study area.

Fig. 6. The digital model of the study area.

3.7.2. OTC simulation of HKU

This research develops a digital model of the main campus using diverse official resources, such as 3D digital models from the Land and Civil Engineering Departments, point cloud spatial data from the Development Department, CAD files from the Building Department, and other GIS data from multiple sources. The digital model consists of five layers: building, infrastructure, terrain, ground, and vegetation, which are essential for CFD simulations (see Fig. 6a).

As the campus is multi-layer and vertical, the research focuses on three main levels: Level 2 (L2), Level 1 (L1), and Ground floor (G). These areas are divided into 2 m × 2 m quads to obtain a detailed dataset [3]. This results in 6,563 points calculated in total, including 3,963 points on L2, 915 points on L1, and 1,685 points on G (see Fig. 6b). Features of these points are input data for further XGBoost analyses and the measurement of these features will be discussed later (see Fig. 6c).

This research uses recorded weather data based on Kings Park Meteorological Station, a nearby weather station controlled by Hong Kong Observatory, in EPW formats. In Eddy3D, we take the hottest days in June, July, and August 2024, in Hong Kong based on the record of the Hong Kong observatory. Specifically, the time is set to be June 27th, July 10th, and August 5th, 2024, and the time ranges are all from

13:00–14:00. These data are inputs for CFD simulation.

To conduct CFD simulation on this model, the main campus and its surrounding environment are crucial for assessing air ventilation. Based on the research of Ng (2009b), the radius of the campus is set as R, the radius of the assessment area as 2R, and the radius of the surrounding area as 3R (see Fig. 7a). It adopts a cylindrical domain, with a radius of 2,500 m and a height of 600 m (see Fig. 7b). This ensures that the simulation accounted for the vertical environment of the city and the potential impact on OTC of the study area. Buildings, terrains, and vegetation are three key inputs for the CFD simulation, setting to be mesh, which is crucial for calculating wind speed and direction in a microenvironment.

The material properties of buildings are set up as follows: (1) a 10 cm grey brick for pavement; (2) a 20 cm red concrete tile for some building façade, especially for those located in the Centennial campus; (3) a 10 cm dark red brick tile for the façade of those buildings surrounding the university street; (4) a 5 cm cladding for other building's façade; (5) single glazed window with clear float glass (4 mm); (6) a 5 cm asphalt for roofs. The details about these materials are shown in Table 3.

The wind simulation result of CFD simulation is input for MRT simulation. In addition, the MRT simulation runs three times with cut-
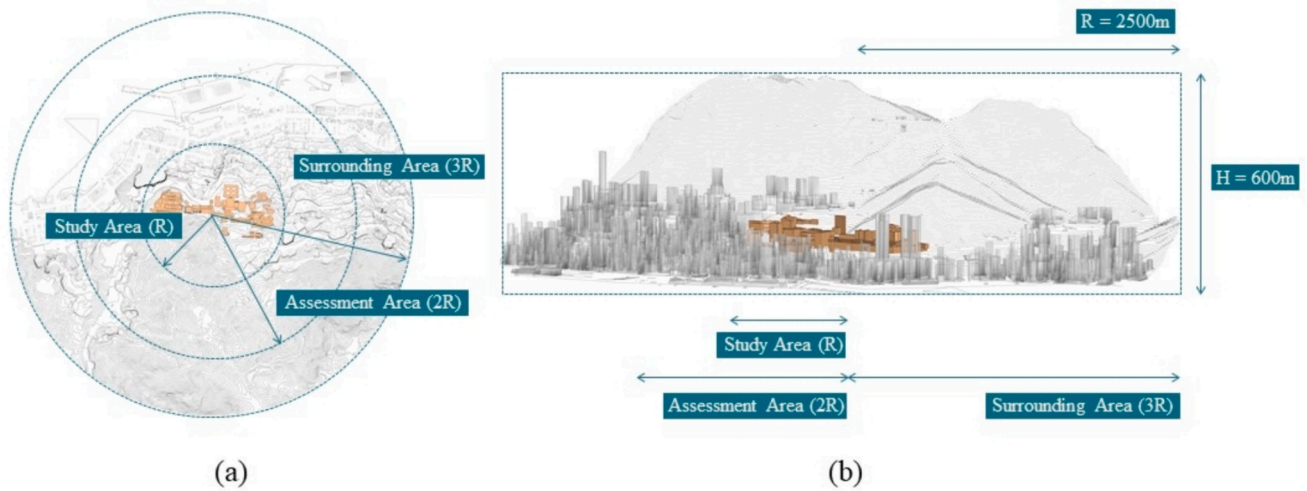
Fig. 7. Total area for CFD simulation.

Table 3
Description of materials settings.

Characters	Grey bricks	Red concrete tile	Dark red brick tile	Cladding	Asphalt	Characters	Clear float glass(4 mm)
Conductance (W/m·K)	0.721	1.100	0.770	0.590	1.200	Reflection	0.07
Density (kg/m³)	1922	1200	1120	1250	1700	Absorption	0.11
Specific Heat (J/kg K)	837	903	840	800	920	Transparency	0.82

ting planes at the elevation of L2, L1, and G. Points of each level get their own T_{mrt} values. Take a standard people as the input, UTCI-S can be calculated based on the formula mentioned above.

3.7.3. On-site measurement of HKU

We conducted on-site measurements on July 10, 2024, from 13:00 to 14:00, the hottest period of the day. The lowest air temperature observed is 28.3 °C and the highest is 32.6 °C. The sky is sunny, a typical summer in Hong Kong (see Table 4). We calculate 20 points for each level, resulting in 60 points in total. The measurement data for each point includes air temperature, black bulb temperature, dew point, wind speed, relative humidity, wind direction, and wind speed. The measurement of each point is the average data of one minute.

The measured globe temperature must adjust for convective heat exchange to convert it into T_{mrt} [23], due to the effect of convective heat loss from the small globe thermometer. Then, UTCI-M of selected points are for further research, with a maximum of 36.1 °C and a minimum of 34.4 °C.

3.7.4. Space design factors

For DepthmapX, the grid is set to be 2 m × 2 m. The intersection points of the grid are the research points. For the first five Rhino-based factors, including PV, OV, AHI, HBI, and BVI, the radius of the buffer area is set to be 10 m for the calculation. After collecting these points, we can continue the calculation of the other seven space design factors in DepthmapX. These data serve as X for further XGBoost analyses.

3.7.5. Xgboost analyses

To generate accurate thermal values with XGBoost, UTCI-S (Y), created by OTC simulation, is input to build an ML model. This process runs several times with adjusted parameters. This ends until the accuracy, which is verified with RMSE, reaches 0.882. Then this model generates accurate UTCI-X of all points. The importance of each factor is calculated and ranked. At last, the algorithm between UTCI-X (Y) and space design factors (X) is also produced.

Table 4
Summary of measured meteorological conditions during the field study in Hong Kong.

Measured Meteorological conditions	July 10, 2024
Weather	Sunny and hot
Mean air temperature °C	30.1
Max air temperature °C	32.6
Min air temperature °C	28.3
Average humidity (%)	71.12
Average wind velocity (m/s)	0.11
Measurement duration	13:00–14:00

4. Data analyses, results, and findings

4.1. Consistent UTCI-S, UTCI-M, and UTCI-X

There is a strong correlation between UTCI-S and UTCI-M of the data collected for July 10th, as well as UTCI-S and UTCI-X for all three days (June 27th, July 10th, and August 5th). A good agreement is observed between the simulated thermal value and measured UTCI.

RMSE of 0.43 and R^2 of 0.74 are observed between UTCI-S and UTCI-M, indicating a great performance of OTC simulation. The scatterplot tends to be a 45-degree line for the study points (see Fig. 8a), indicating strong consistency in this research. The differences between UTCI-S and UTCI-M are displayed in Fig. 8b, which are distributed normally, suggesting an effective simulation.

Based on the above analysis data, we can see that UTCI-S and UTCI-M match with each other, indicating the reliability of conducting thermal simulations with Eddy3D. This combined method (employing onsite measurement, OTC simulation, and machine learning) is fast and efficient, while on-site measurements require more human resources and equipment. The Eddy3D OTC simulation applies an averaging methodology within a time range to ensure convergence. This averaging methodology is different from measurement in this research, for simulation, it is a time range of one hour for all study points; for measurement, it is one minute for one point, 60 points in total.
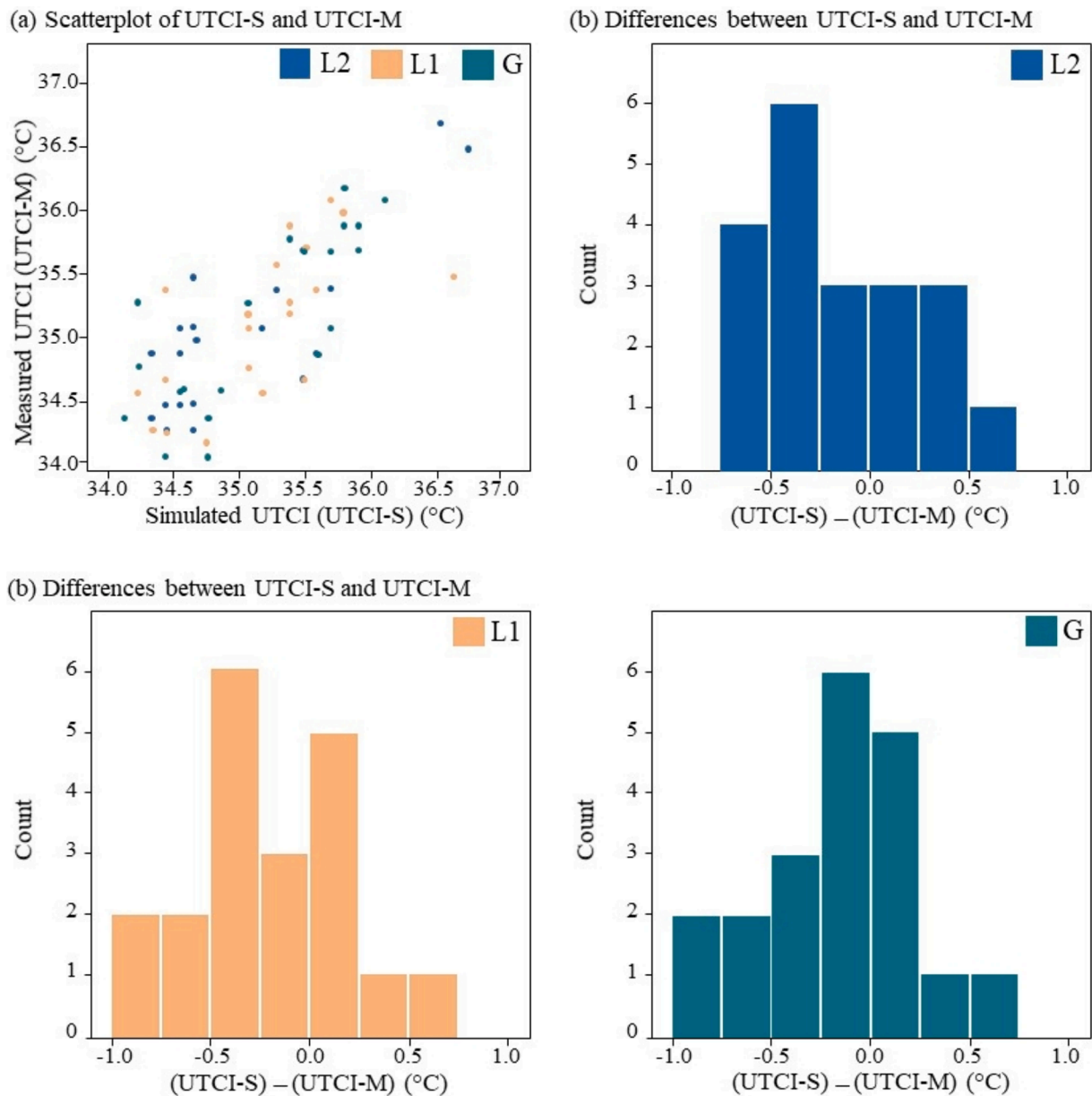
Fig. 8. Comparisons between UTCI-S and UTCI-M, (a) Scatterplot of UTCI-S and UTCI-M and (b) the differences between UTCI-S and UTCI-M.

We conducted Eddy3D simulations on three typical hot days in 2024, gathering the dataset of UTCI-S for all three days (June 27th, July 10th, and August 5th) and plotting the normalized density (see Fig. 9). These three days' thermal data tend to display similar patterns across L2, L1, and G, thus indicating a consistent thermal pattern. This pattern suggests a stable thermal simulation result within the consistent thermal condition and built environment of the campus of HKU.

The UTCI-S and UTCI-X on July 10th are further analyzed in this research. The performance of predicted UTCI (UTCI-X), generated by XGBoost, is provided for comparison with simulated UTCI (UTCI-S). It obtains a RMSE of 0.11 and R^2 of 0.98, suggesting the effectiveness of UTCI-X. Due to the large dataset, the scatterplot presents a clearer 45-degree line (see Fig. 10a). The differences between UTCI-S and UTCI-X display a normal distribution in three levels (see Fig. 10b). Therefore, the result of predicted UTCI is promising. Through these analyses, the simulated results, onsite measured thermal values, and optimized predictions can be treated as a reliable and consistent whole for further discussion. It also shows the robustness of the three measures and proves the effectiveness of this research paradigm on the research of OTC.

4.2. Identifying and verifying cool spots with UTCI-X and site visits

4.2.1. Identifying cool spots with UTCI-X

Take UTCI-X as an example, the microclimate of outdoor spaces exhibits large spatial variability at each level, ranging from 34.0 °C to 37.0 °C. The UTCI of outdoor spaces in L2, L1, and G displays differently (see Fig. 11). The maximum (36.9 °C) and minimum (34.1 °C) thermal values are in L2, but 75 % of the total thermal values are between 34.1 °C to 35.7 °C, indicating several extremely hot places at this level. It is reasonable, because L2 is the highest level, receiving direct sunlight at some spots. As for L1, 75 % of the total thermal values are more concentrated and lower than L2, which ranges from 34.3 °C to 35.0 °C, so L1 is relatively cooler than L2. This is because most points in L1 are shaded by overhanging. The median thermal value in G is the highest among the three levels, which is 35.2 °C, while it is 34.9 °C in L2 and 34.6 °C in L1. However, the 75 % of thermal values in G is similar to L2, and relatively higher than L1. This is because some spaces in G are outdoor open spaces as L2, receiving more direct sunlight than L1.

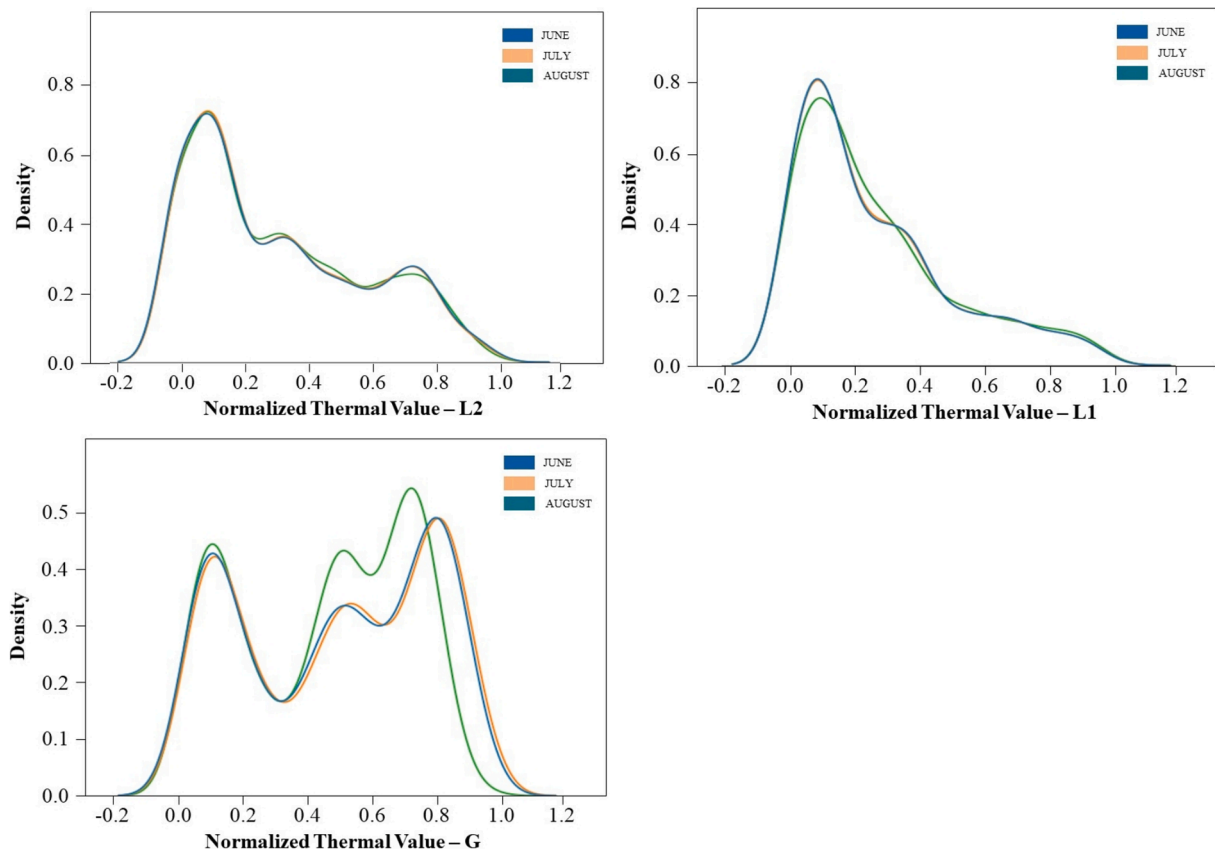Based on UTCI-X, we identify these points with lower values, ranging

Fig. 9. Normalized Density Plots for UTCI-S of June 27th, July 10th, and August 5th.

from 34.0 °C to 35.0 °C, which are relatively cool spots on the research day in the main campus of HKU. These cool spots are mapped out as follows (see Fig. 12).

4.2.2. Site visits to verify the cool spots

Based on these findings, we also conducted on-site research on these cool spots (see Fig. 13). These cool spots are scattered on the campus. They are also very popular, not only for students and for faculties, but also for tourists. The popular cool spots include Centennial Parks, University Street, and Red Wall. Specifically, Red Wall, one of the most popular tourist destinations, is always filled up with tourists because it is such a comfortable cool spot.

4.3. Space design's impact on OTC

4.3.1. The ranking of space design factors

Through XGBoost SHAP analyses, we successfully identify the most important space design factors on thermal value (see Fig. 14). Notably, the most important space design factor differs in three levels. Specifically, the Percentage of View (PV) is the most important factor in Level 2 (L2) of the site; the Average Height Index (AHI) contributes most in OTC for Lever 1 (L1); and the Connectivity (CON) becomes the most influential space design factor to Ground Floor (G) of the campus. Overall, by comparing the SHAP Value analyses of the three levels, we can identify that PV, AHI, and CON list the top 3 most important space design factors for the three levels.

This research conducts correlation analyses between space design factors based on Pearson Correlation (see Fig. 15), examining the effectiveness of these factors. DepthmapX-based space design factors represent the integrated quality of a space, such as visibility and connectivity; while rhino-based space design factors are related to space height and volume and others. The low correlation numbers between

rhino-based space design factors and DepthmapX-based space design factors further prove the difference between these two types of indexes, examining the importance of introducing DepthmapX for a better understanding of space design's impact on OTC.

4.3.2. Analyses of Percentage of view (PV) and UTCI-X

For L2, PV is the most important space design factor on OTC. The 2D factor represents the visibility of a space. Its explanation is shown in Fig. 16a.

We observe that the open spaces in L2 of the campus have higher PV values than other spots in L2 (see Fig. 16b). For points on L2, we find that the closer the points are to the centennial campus buildings (see Fig. 16c), the lower the PV tends to be. Accordingly, we find that the closer the points to the parks, the higher the PV tends to be, which is reasonable. The lower the PV is, the more volume is, and there will be less sunlight, causing lower temperatures.

Based on the above question, we conduct the partial dependence analysis between UTCI and PV and other factors. PV is relatively linear and positively related to UTCI in total (see Fig. 17), which corresponds to the above-mentioned explanation. However, they are not exactly linearly correlated. When PV is more than 0.7, the line is steeper. But if PV is smaller than 0.7, the temperature is more stable. As PV represents the openness of a space of a research point in a standard buffer area, the bigger the PV is, the less volume is involved. Less volume brings more sunlight, contributing to a higher thermal value. Some space design factors in L2 display non-linear trends, such as CON, the third most important index. Therefore, this research sheds light on explaining the relationship between space design and OTC.

4.3.3. Analyses of average height Index (AHI) and UTCI-X

For L1, PV is the second important factor; and AHI, a 3D space design factor, is the most important one. AHI is defined as follows (see
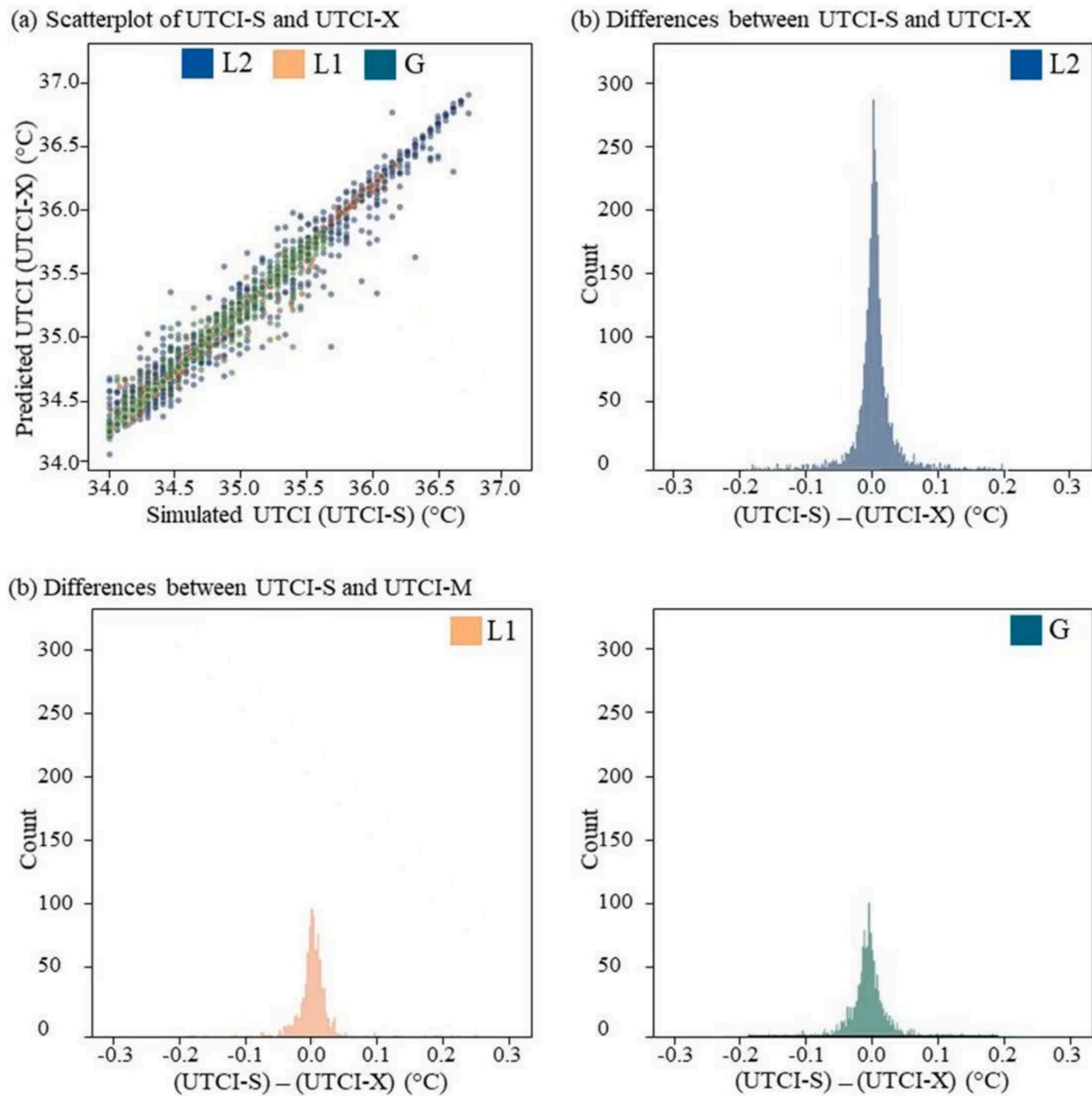
Fig. 10. Comparisons between UTCI-S and UTCI-X: (a) Scatterplot of UTCI-S and UTCI-X and (b) the differences between UTCI-S and UTCI-X.

Fig. 18a). We find that the higher AHI is located near University Street in L1 (see Fig. 18b). As Fig. 18 (c) shows, the closer the point to buildings, the higher AHI, which means the higher buildings are, thus bring more shades.

Contrary to PV, AHI is negative to UTCI in L1 (see Fig. 19). The bigger the AHI, the smaller the thermal value. Importantly, there is a threshold for AHI. When it is larger than 5, the thermal value remains to be 34.6 °C. This finding also applies to other factors. It is necessary to mention that the algorithm between the space design factor and OTC is not the same under different conditions (i.e., AHI displays different trends in L2 and L1). In addition, it is necessary to mention that PV is still relatively linearly related to UTCI in L1, similar to L2.

4.3.4. Analyses of connectivity (CON) and UTCI-X

As to G, CON is the most effective one in predicting UTCI-X. CON represents the total number of cells visible from a specific cell (see Fig. 20a). We observe that the closer the points are to the corner of a space, the lower the value they receive (see Fig. 20b). Take the point near the Main Library as an example. The closer the point to the wall, the

less visible it will be. Then it will be less sunlight. However, at the same time, it will be less wind. Therefore, it is necessary to conduct further analyses on this.

CON is positively related to UTCI as shown in Fig. 21. Because of the impact of both sunlight and wind, the relationship is not linear. The impact of CON is stable within the threshold. It shows that when it is larger than 300 and smaller than 500, the temperature is relatively stable, about 35.1 °C.

Therefore, we can see that linear regression analyses on space design and OTC are not accurate. The relationship between space design and thermal value is complex when considering the counterpart of sunlight and wind. Based on our research, PV, AHI and CON factors are influential on local microclimate. So, it will be effective for architects and urban designers to manipulate these three space factors to achieve outdoor cool spots. Specifically, it is recommended to maintain PV values lower than 0.7 at elevated levels, balancing the impact of shading and visibility, thus reducing heat stress for outdoor spaces. This can be achieved by balancing the openness and volumes and minimizing extremely open spaces. To be more specific, for designers, a space less
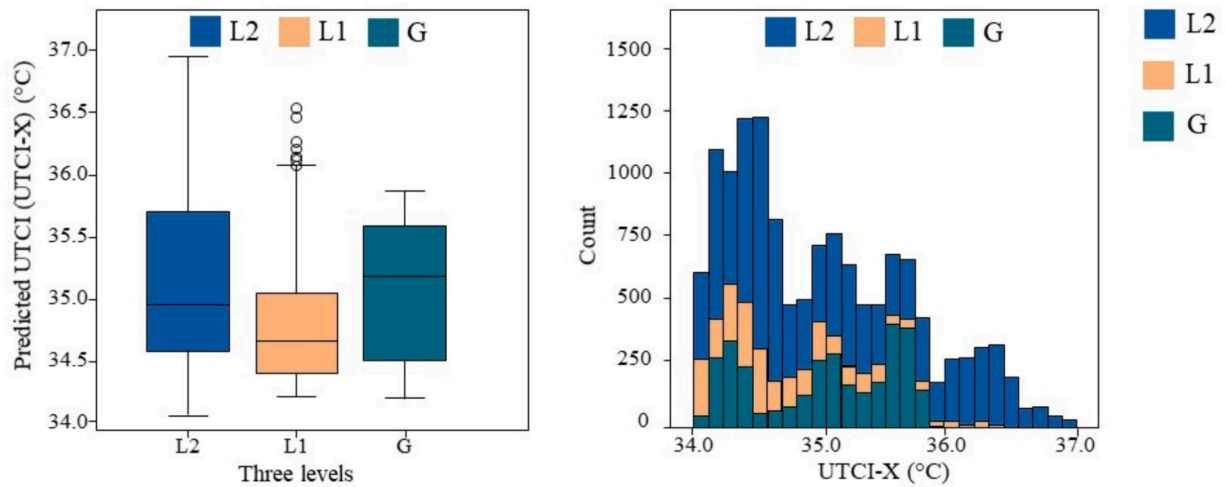
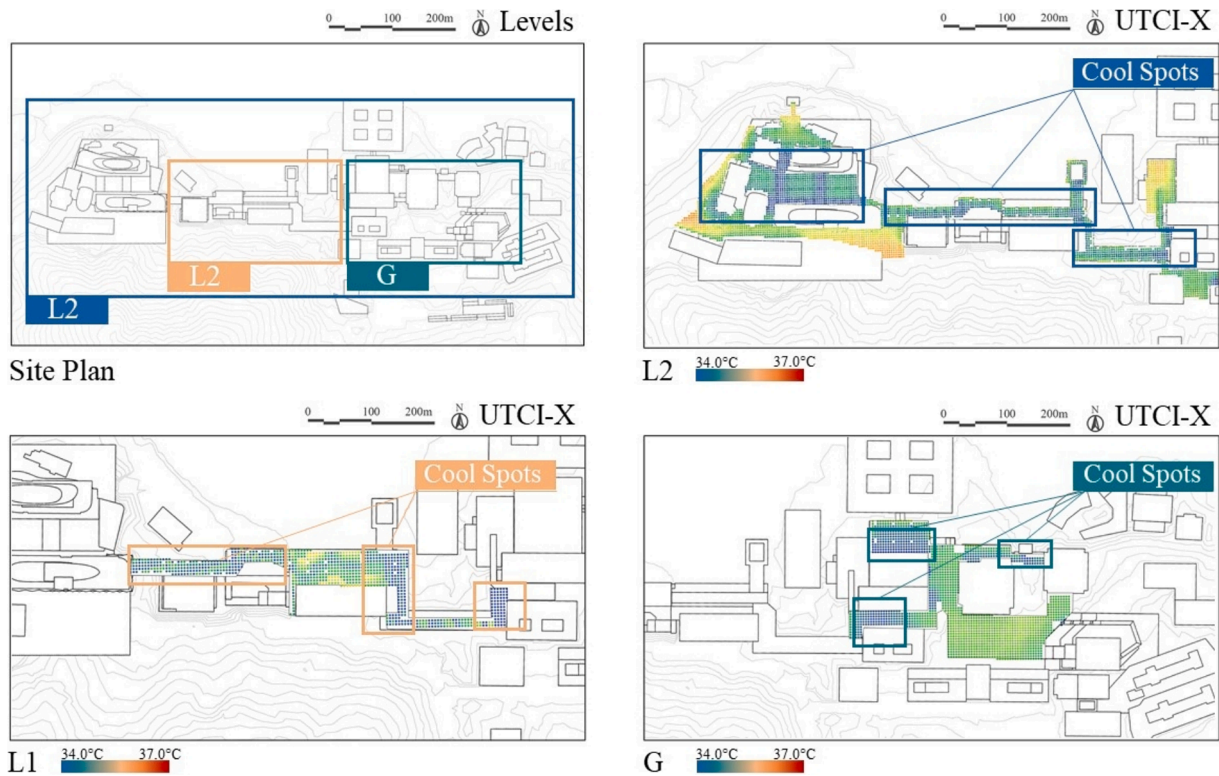Fig. 11. The distribution of UTCI-X in three levels.

Fig. 12. Mapping out the typical cool spots.

open, such as a higher height-to-width aspect ratio (such as more than 4) is beneficial to achieve a cool spot. Besides, ensuring AHI to be greater than 6 at intermediate levels, will provide adequate shading for the research area. In detail, in practice, it is better to keep the height of buildings relatively higher, such as more than four floors of buildings will create a cool spot in a community-scale outdoor space. Also, it is necessary to avoid overly high buildings, which will block the wind, reducing the quality of cool spots. Lastly, for the vertical spaces, such as the campus of HKU, the ground floor is a lower elevated floor, characterized differently from the intermediate level. CON has become the most influential space design factor, and it is suggested to keep CON between 300 and 500 on the ground floor. This indicates it is not beneficial for thermal comfort to be too open or too narrow from the

range of the values of CON. The outdoor spaces around the main library on the campus of HKU could be a good example for the design of cool spots, including one with an overhang on the top of the space and a height-to-width aspect ratio to be 1: 3, and another one without a roof but keeping the height-to-width aspect ratio to be 4: 1. Therefore, this range of CON will ensure adequate connectivity, visibility, and walkability, and balance the impact of sunlight, wind, and shading. This research thus improves the application of OTC research on architects' and urban planers' practice, by giving a detailed and explainable result to maintain cool for community-scale outdoor environments.
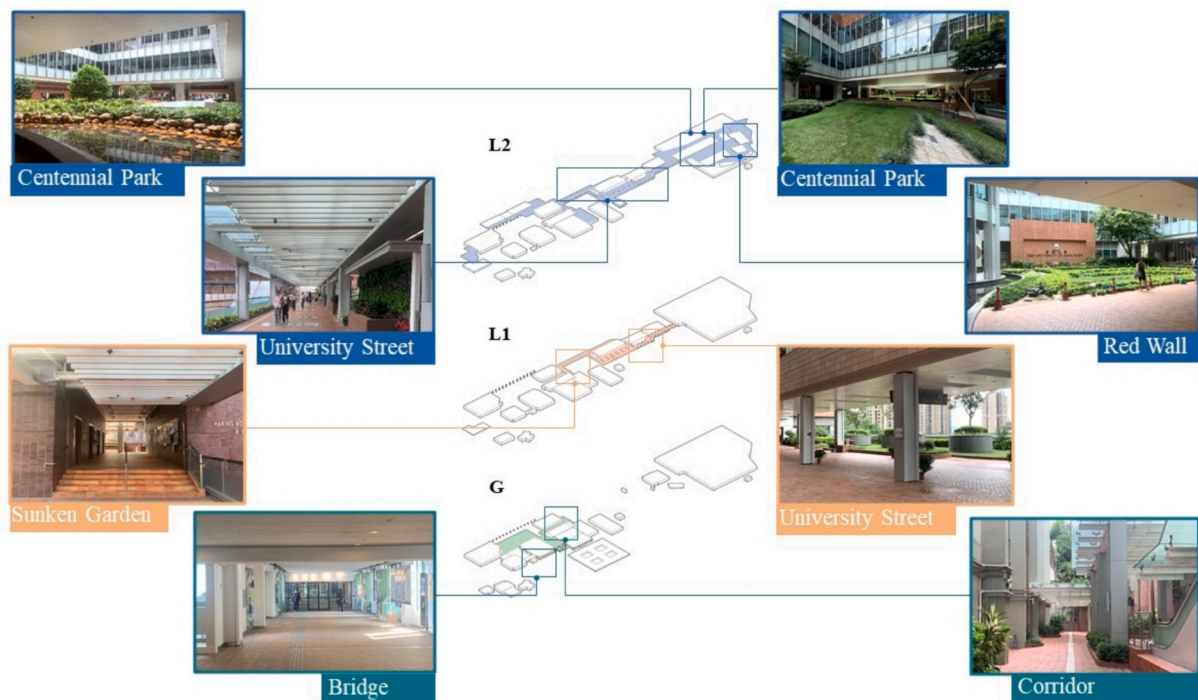
Fig. 13. Mapping cool spots based on on-site visit.

4.4. Sensitivity analysis

A sensitivity analysis is performed on the five rhino-based space design factors to test the impact of the buffer area's radius on the calculation of space design factors, including PV, OV, AHI, HBI, and BVI (see Fig. 22). This analysis is based on comparing the average value of each space design factor at three different levels within different radii of the buffer area. The radius ranges from 2 m to 30 m based on community-scale research.

The results can be found in Fig. 22; we can see that when the radius of the buffer is set to 10 m, the average HBI value of three levels is the biggest, stressing the impact of the highest buildings on space design factors and OTC. For the average AHI value of three levels, when the radius is 10 m, AHI is the minimum. This indicates that the human-scale space is considered which is important for micro-outdoor environment analysis. Besides, both OV and BVI become stable when the radius is 10 m, ensuring reliable and consistent results. The results of PV are negatively linear related to the value of radii, which is less informative for the choice of radii. Apart from the above-mentioned analyses, a 10-meter buffer radius can provide a community-scale, localized perspective on the variation of built environment height. This radius captures the surrounding environment precisely and provides a detailed identification of the outdoor environment. Therefore, it is reasonable to adopt 10 m as the radius of the buffer area for this research.

5. Discussion

5.1. Contributions

To design and build some smaller scale, local communities with better OTC/UTCI is considered a sensible strategy to mitigate the heat weaves that are swamping around the world. Researchers and designers have long been involved in 'cool spot' reasoning/auditing, i.e., to understand what space designs would help better form the places where residents are willing to hang around even during the hot and humid weather. Built upon previous studies, this research proposed and substantiated a new paradigm for 'cool spot' reasoning. It made several non-

trivial contributions to the domain, mainly on the methodological side.

Firstly, traditional methods mainly analyzed five space design factors, including Percentage of View (PV), Orientation of View (OV), Average Height Index (AHI), Highest Building Index (HBI), and Building Volume Index (BVI). This new paradigm, however, by introducing DepthmapX, can analyze more sophisticated space design factors such as Angular Mean Depth (AMD), Connectivity (CON), Gate Counts (GC), Point First Moment (PFM), Point Second Moment (PSM), and Through Vision (TV).

Secondly, traditional methods only adopt linear regression to associate the space design factors with the OTC/UCTI, as measurements of 'cool spots'. This new paradigm, however, by introducing a machine-learning tool, XGBoost, can better grasp the sophisticated, non-linear nexuses between space design and OTC/UTCI.

Thirdly, traditional methods fall short of diving too much into computer simulations (e.g., CFD, MRT, and UTCI simulations) without adequately linking to the real world. This new paradigm, however, by emphasizing triangulation with the ground truth data, can ensure the reliability of input data (e.g., wind speed, direction, and weather) and thus add to the confidence of accepting the simulated and optimized outcomes. The onsite inspection and verification of the 'cool spots' with expert knowledge is another contribution of the research.

Fourthly, with this new paradigm, this research finds that PV, AHI, and CON space design factors are the three most influential factors among the eleven space design factors. According to this case study, to get a relatively cooler outdoor microenvironment, it is recommended to keep a space less open, such as a higher height-to-width aspect ratio at an elevated level. Also, making the average height of a space more than four floors at the intermedia level will be advantageous to achieve a cool spot. Similar to the elevated level, it is beneficial to keep a space not too narrow or too open at ground level, such as by keeping a height-to-width aspect ratio to be 1:3 with an overhang or a height-to-width aspect ratio to be 4:1 with no roof. By limiting the range of these influential indicators, designers, and planners will be more informed about the impact of their designs and regulations on OTC.

Lastly, this research examines the effectiveness of DepthmapX on the study of urban microclimate. By taking the concept of space
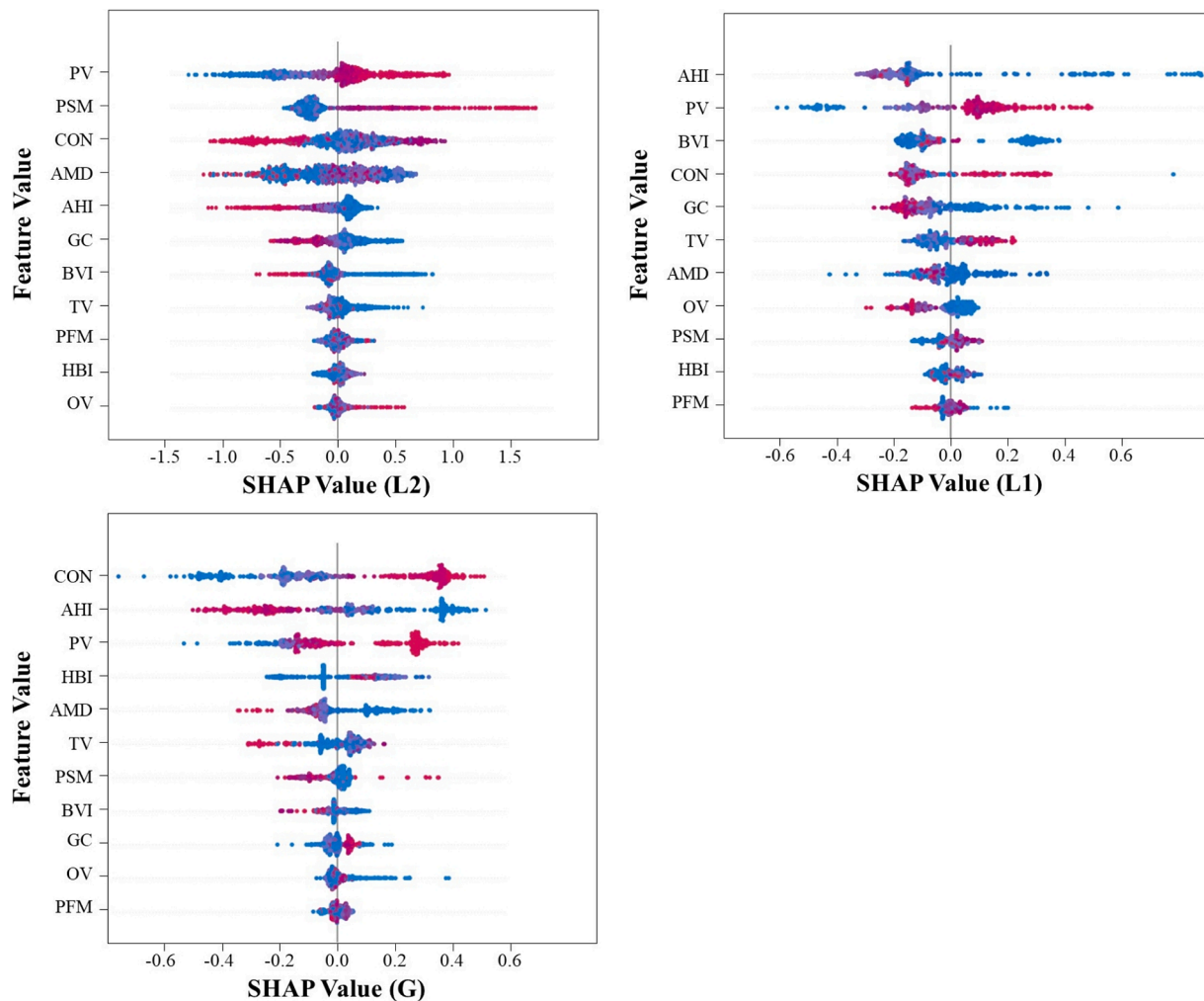
Fig. 14. SHAP analysis on the importance of each factor in L2, L1, and G.

configuration, visibility, and connectivity into consideration, our research provides a set of new DepthmapX-based factors to evaluate the impact of space design on the thermal environment. Future research may incorporate this new paradigm and new factors into the study of urban microclimate.

It is believed that through the improved OTC Simulation and Space Design Analyses, the nexuses between space designs and OTC/UTCI can be better understood. The space design indexes/measures help quantify the design elements and their configurations (e.g., connectivity, visualization) to join the scientific simulations. They can also be translated back to real-life, specific space design options to be considered by designers.

5.2. Weakness and future works

Despite its enormous contributions, the research also shows its weaknesses. Firstly, the study of cool spots is based on a time range, from 13:00 to 14:00 on three typical hot and sunny days in Hong Kong. So, when the time range and data change, the research results will be different. The impact of a sudden wind is also not considered in this research, which will significantly affect human thermal feelings.

Secondly, the identification of the three most influential factors, namely PV, AHI, and CON, does not mean the other nine factors are not important. More case studies related to different 'cool spots' in different microclimates or even different climate zones are desired to find out the influences of other factors.

Additionally, XGBoost helped to analyze the more sophisticated, non-linear dynamics between space design factors and the OTC of the space. However, owing to its 'black box' nature, the pathway from a space design to an OTC/UTCI is not understood in a better way than traditional linear regression methods.

Onsite interpretation of the PV, AHI, and CON with specific, real-life space design factors helps understand the meaning of the simulation and optimization. However, such interpretation is too limited in this study. One can imagine different design options would achieve the same indexes such as connectivity and visibility. Future research with human experts is highly desired to interpret the real-life design options, as they are easy-to-refer materials for designers.

Last but not least, this research adopted an arbitrary grid size (2 m × 2 m) and Rhino-based factor (i.e., 10 m) without trying other combinations, which may significantly impact the results. It only tried one site in a specific season. Overall, the simulation and analytic process in the new 'cool spot' reasoning paradigm is still too complex to implement. Many more studies are desired to further substantiate the paradigm. To this end, the research reported in this paper only opened a promising future to the study of 'cool spots'.

6. Conclusion

This research attempted to better understand what design elements and how to configure them to achieve the so-called 'cool spots', which refer to some local, smaller communities with amenable outdoor
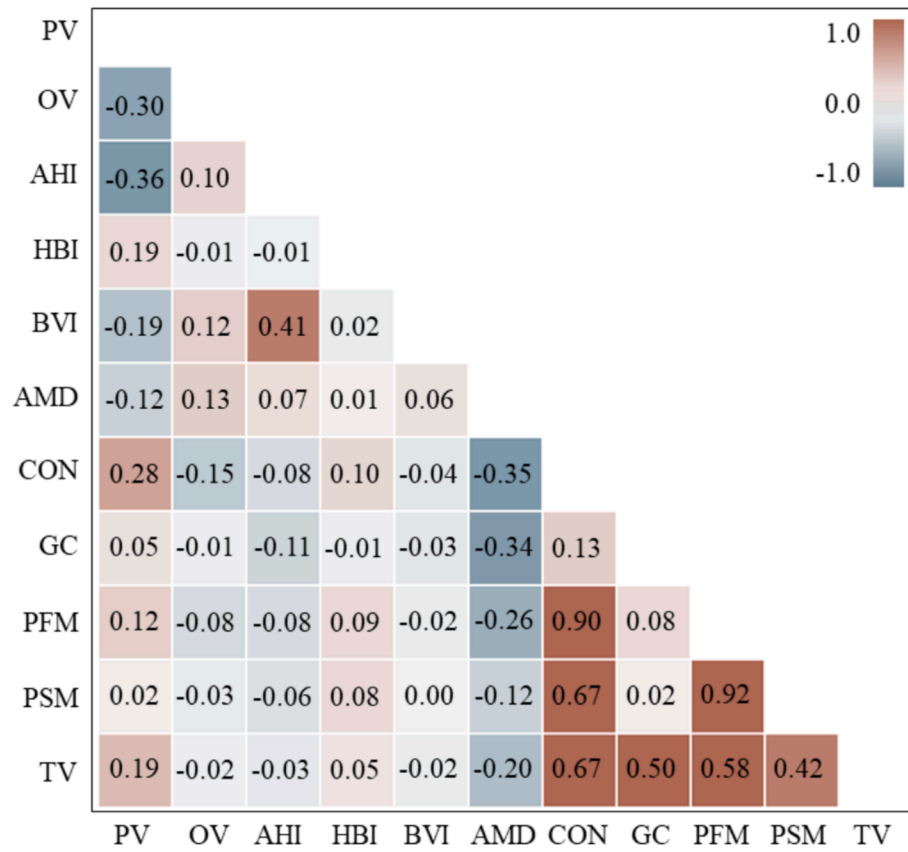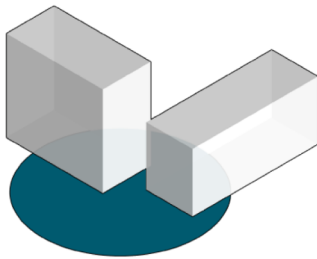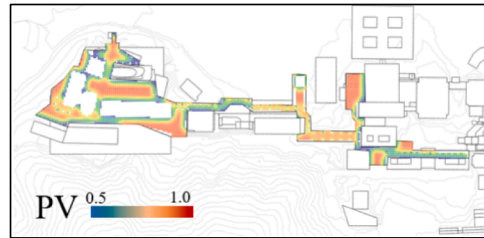
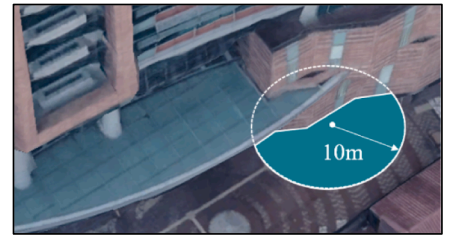**Fig. 15.** Correlation analyses between space design factors.

$$PV = \frac{S_{view}}{S_{buffer}}$$

Note: S_{view} refers to the area of view; S_{buffer} refers to the area of thebuffer, the radius of the buffer is set to be 10 meters.

(a)   (b)   (c)

Fig. 16. The definition, distribution, and example of PV.

thermal comfort (OTC)/ Universal Thermal Climate Index (UTCI). It did so by introducing a new paradigm, which adopts DepthmapX to investigate the sophisticated space designs for the subsequent XGBoost analyses to interpret the nexuses between space designs and OTC/UTCI.

A case study conducted on a university campus in Hong Kong using the new 'cool spot' reasoning paradigm successfully identified the 'cool spots' that are popular on the campus. It found that the Percentage of View (PV), Average Height Index (AHI), and CON (Connectivity) are the three most influential factors leading to the formation of a 'cool spot' under the campus' particular microenvironment. Onsite space design
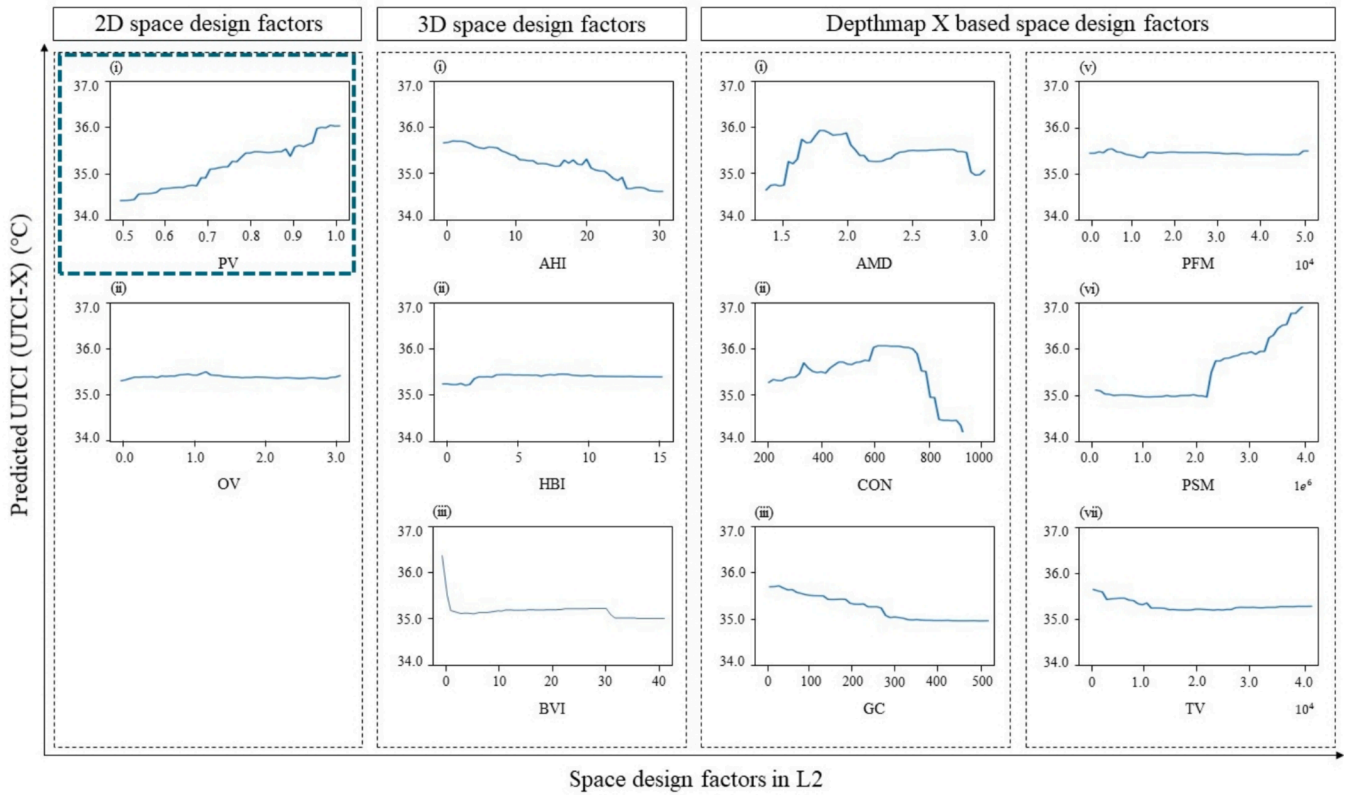
Fig. 17. The algorithm of BVI and other factors in L2.

$$AHI = \frac{1}{n}\sum_{i=1}^{n} H_i$$

Note: Building Volume Index (BVI); $V_{buildings}$ is the volume of buildings; V_{buffer} refers to buffer volume, which is based on the buffer area (radius is 10m and height is 100m).

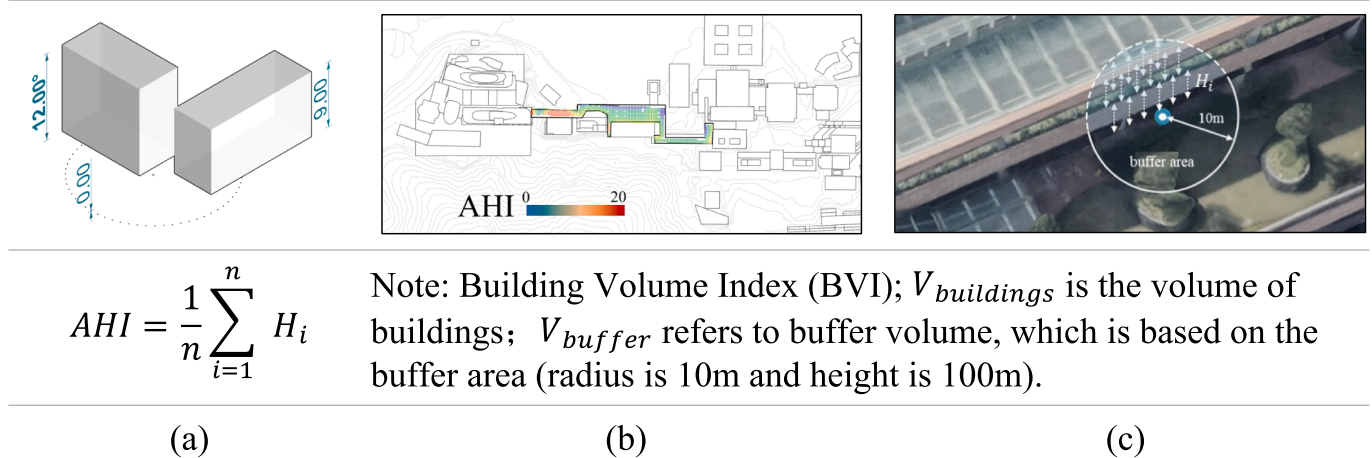(a)   (b)   (c)

Fig. 18. The definition, distribution, and example of AHI in L1.

analyses by expert designers tried to associate the factors with the specific space designs to reason the nexuses. This'epidemiologic knowledge' can be used as easy-to-refer and quantifiable design references for 'cool spots' design in the future.

On the methodology sphere, it is serendipity to find out that the onsite measured, simulated, and optimized UTCIs are of high resemblance and thus add to the confidence to accept them as outputs.

DepthmapX can probe into complex, vertical space designs and provide more structured output for further analyses. XGBoost can better grasp the non-linear dynamics between space designs and local OTC/UTCI. By summarizing all these together, the new paradigm clearly opens a new avenue through which the formation of 'cool spots' can be better reasoned. This paper is of high originality although it deals with an old problem.
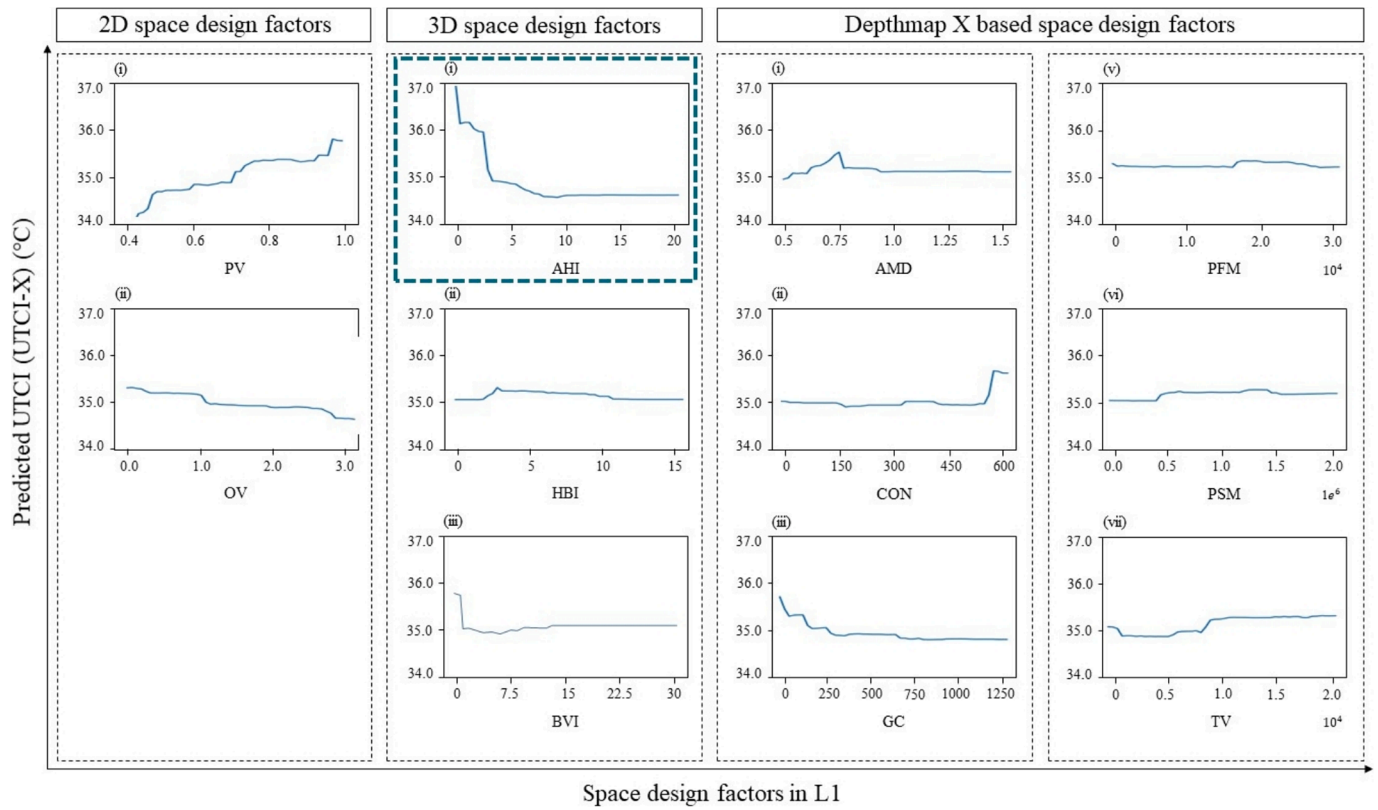
Fig. 19. The algorithm of AHI and other factors in L1.

$$CON = \sum c_i \qquad \text{Note: } c_i \text{ is the cell visible from a specific cell}$$
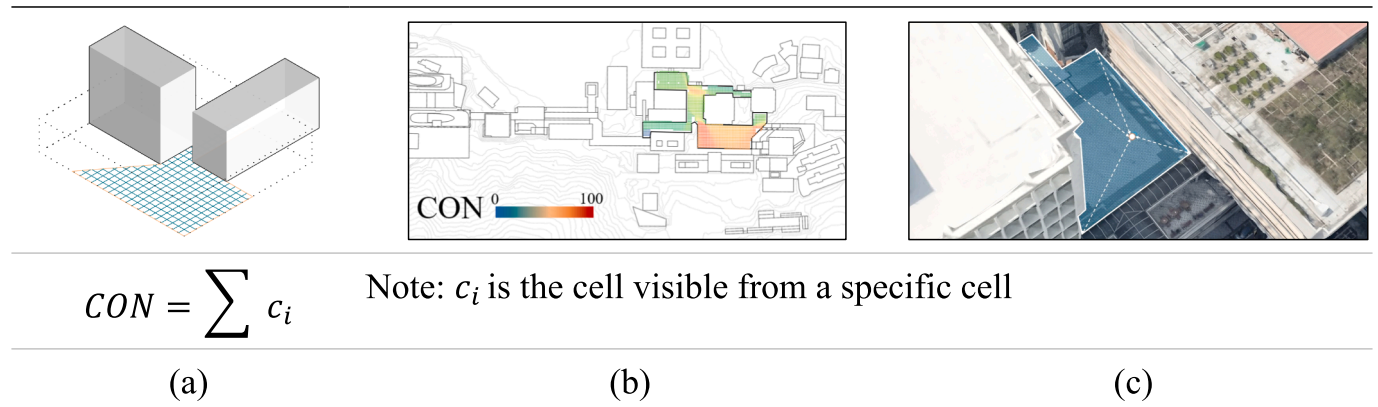
(a)   (b)   (c)

Fig. 20. The definition, distribution, and example of CON in G.

With this new paradigm, we discovered that PV (2D), AHI (3D), and CON (DepthmapX-based) space design factors are the three most influential factors among the eleven space design factors. This research examines the effectiveness of DepthmapX on the study of urban microclimate. By taking the concept of space configuration, visibility, and connectivity into consideration, our research provides a set of new DepthmapX-based factors to evaluate the impact of space design on the thermal environment. Future research may incorporate this new

paradigm and new factors into the study of urban microclimate.

There are some future works to make this 'cool spot' reasoning paradigm truly ready for use. For example, it is unclear why the measured, simulated, and optimization UTCI are so resemble to each other. Whether other space design factors matter to the formation of a cool spot? The real-life examples of space designs to apply PV, AHI, and CON are still limited, waiting for empirical studies to find out more options. The reasoning process of space designs to 'cool spot', even if we
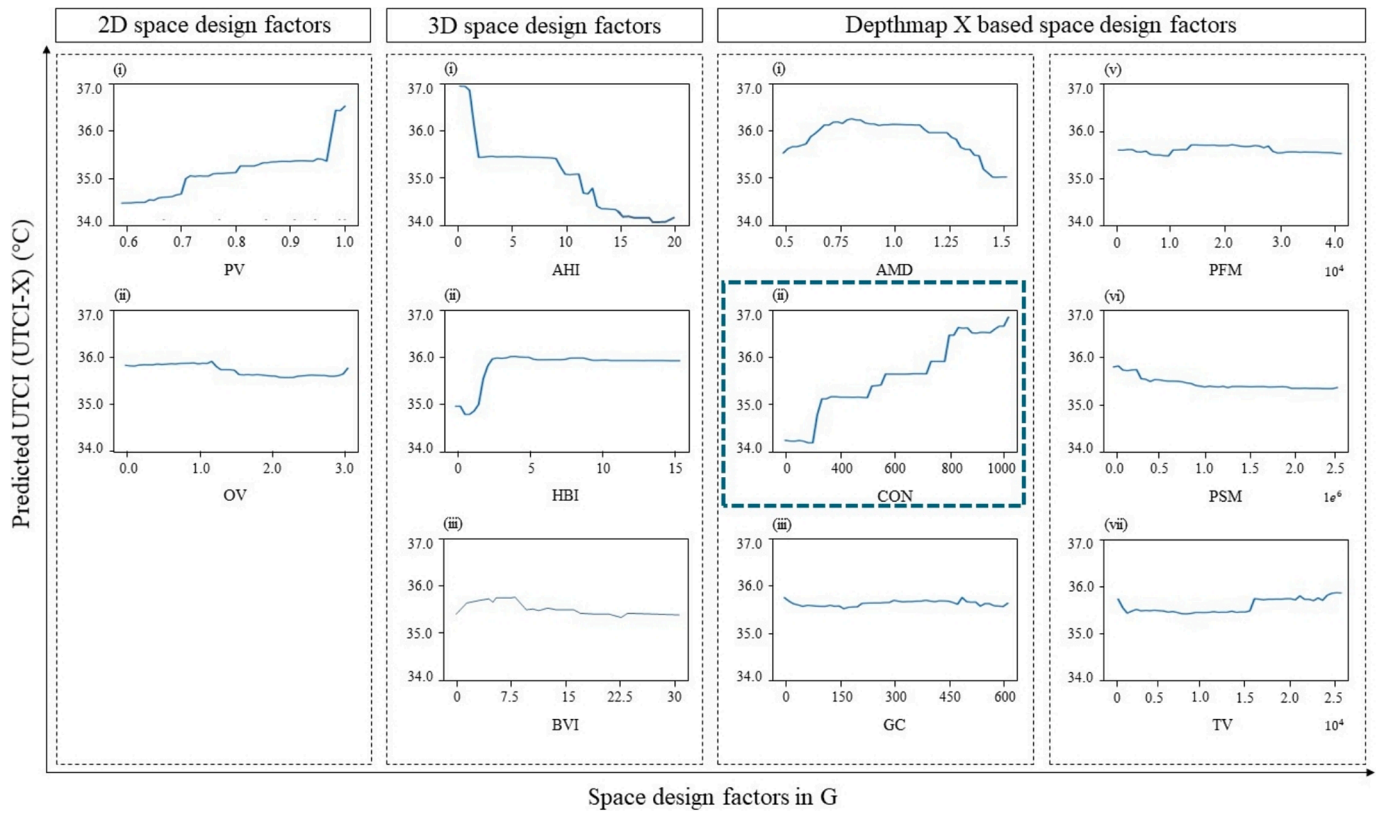
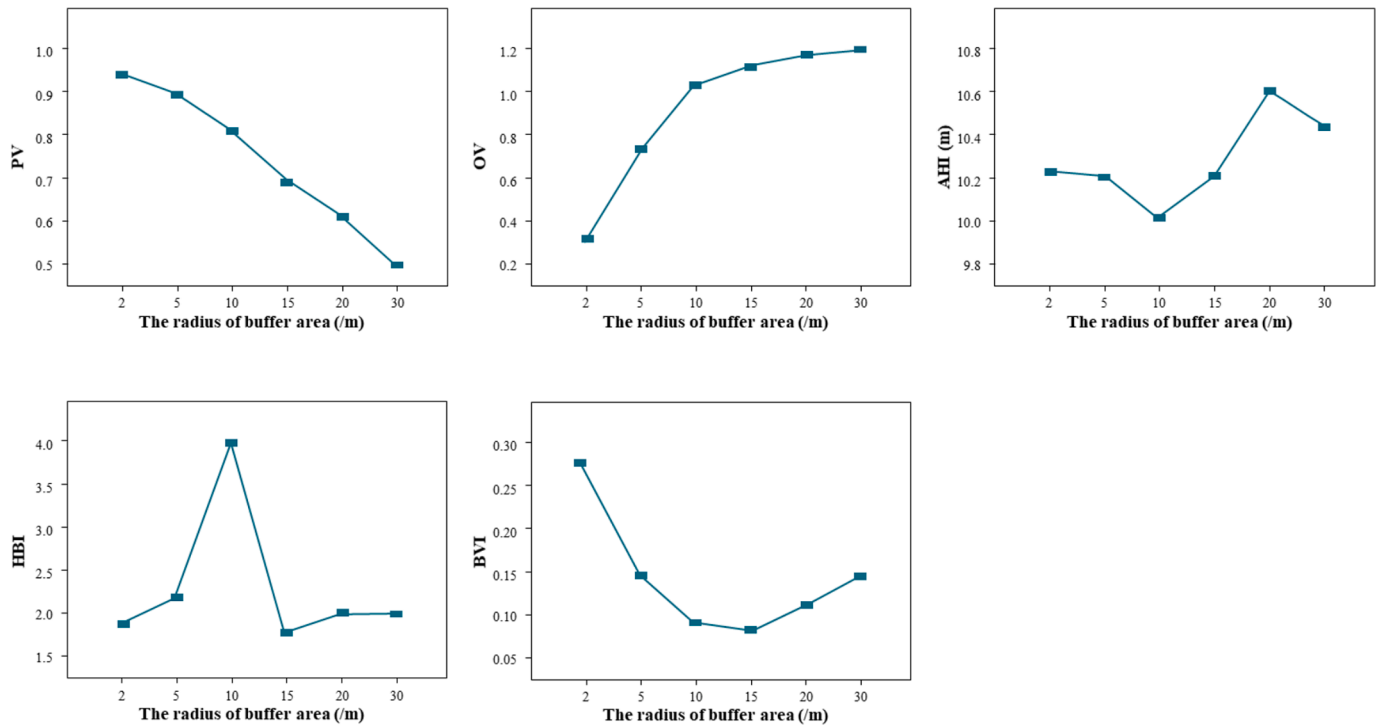**Fig. 21.** The partial dependence analysis between UTCI-X and space design factors in G.

Fig. 22. The sensitivity analysis of space design factors.

adopted the powerful XGBoost, is still far from being fully explainable.

CRediT authorship contribution statement

Ye Xia: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Weisheng Lu:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Ziyu Peng:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Jinfeng Lou:** Writing – review & editing, Visualization, Software, Methodology, Data curation. **Jianxiang Huang:** Writing – review & editing, Validation, Resources, Project administration, Funding acquisition. **Jianlei Niu:** Supervision, Resources, Project administration, Funding acquisition.

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 and funding

Data availability

Data will be made available on request.

References

[1] F. Ali-Toudert, H. Mayer, Numerical study on the effects of aspect ratio and orientation of an urban street canyon on outdoor thermal comfort in hot and dry climate, *Build. Environ.* 41 (2) (2006) 94–108, https://doi.org/10.1016/j.buildenv.2005.01.013.

[2] A. Almhafdy, N. Ibrahim, S.S. Ahmad, J. Yahya, Thermal performance analysis of courtyards in a hot humid climate using computational fluid dynamics CFD method, *Procedia. Soc. Behav. Sci.* 170 (2015) 474–483, https://doi.org/10.1016/j.sbspro.2015.01.012.

[3] C. Altunkasa, C. Uslu, Use of outdoor microclimate simulation maps for a planting design to improve thermal comfort, *Sustain. Cities Soc.* 57 (2020) 102137, https://doi.org/10.1016/j.scs.2020.102137.

[4] H. Andrade, M.-J. Alcoforado, Microclimatic variation of thermal comfort in a district of Lisbon (Telheiras) at night, *Theor. Appl. Climatol.* 92 (3–4) (2007) 225–237, https://doi.org/10.1007/s00704-007-0321-5.

[5] N. Antoniou, H. Montazeri, M. Neophytou, B. Blocken, CFD simulation of urban microclimate: Validation using high-resolution field measurements, *Sci. Total Environ.* 695 (2019) 133743, https://doi.org/10.1016/j.scitotenv.2019.133743.

[6] M.T.B. Larriva, E. Higueras, Health risk for older adults in Madrid, by outdoor thermal and acoustic comfort, *Urban Clim.* 34 (2020) 100724, https://doi.org/10.1016/j.uclim.2020.100724.

[7] M. Barthélemy, Spatial networks, *Phys. Rep.* 499 (1–3) (2011) 1–101, https://doi.org/10.1016/j.physrep.2010.11.002.

[8] K. Blazejczyk, Y. Epstein, G. Jendritzky, H. Staiger, B. Tinz, Comparison of UTCI to selected thermal indices, *Int. J. Biometeorol.* 56 (3) (2011) 515–535, https://doi.org/10.1007/s00484-011-0453-2.

[9] A. Chatzidimitriou, S. Yannas, Street canyon design and improvement potential for urban open spaces; the influence of canyon aspect ratio and orientation on microclimate and outdoor comfort, *Sustain. Cities Soc.* 33 (2017) 85–101, https://doi.org/10.1016/j.scs.2017.05.019.

[10] Chen, T., & Guestrin, C. (2016). XGBoost: a Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, 785–794. https://doi.org/10.1145/2939672.2939785.

[11] A.M. Coutts, N.J. Tapper, J. Beringer, M. Loughnan, M. Demuzere, Watering our cities, *Prog. Phys. Geogr.: Earth Environ.* 37 (1) (2012) 2–28, https://doi.org/10.1177/0309133312461032.

[12] J.-Y. Deng, N.H. Wong, Impact of urban canyon geometries on outdoor thermal comfort in central business districts, *Sustain. Cities Soc.* 53 (2020) 101966, https://doi.org/10.1016/j.scs.2019.101966.

[13] C. Ding, K.P. Lam, Data-driven model for cross ventilation potential in high-density cities based on coupled CFD simulation and machine learning, *Build. Environ.* 165 (2019) 106394, https://doi.org/10.1016/j.buildenv.2019.106394.

[14] J.H. Friedman, Greedy function approximation: A gradient boosting machine, *Ann. Stat.* 29 (5) (2001) 1189–1232, https://doi.org/10.1214/aos/1013203451.

[15] E.J. Gago, J. Roldan, R. Pacheco-Torres, J. Ordóñez, The city and urban heat islands: A review of strategies to mitigate adverse effects, *Renew. Sustain. Energy Rev.* 25 (1) (2013) 749–758, https://doi.org/10.1016/j.rser.2013.05.057.

[16] M.B. Greenwell, pdp: an R package for constructing partial dependence plots, *The R Journal* 9 (1) (2017) 421, https://doi.org/10.32614/rj-2017-016.

[17] T. Hao, H. Chang, S. Liang, P. Jones, P.W. Chan, L. Li, J. Huang, Heat and park attendance: Evidence from "small data" and "big data" in Hong Kong, *Build. Environ.* 234 (2023) 110123, https://doi.org/10.1016/j.buildenv.2023.110123.

[18] C. Heaviside, H. Macintyre, S. Vardoulakis, The urban heat island: implications for health in a changing environment, *Curr. Environ. Health Rep.* 4 (3) (2017) 296–305, https://doi.org/10.1007/s40572-017-0150-3.

[19] Y. Ibrahim, T. Kershaw, P. Shepherd, I. Elwy, A parametric optimisation study of urban geometry design to assess outdoor thermal comfort, *Sustain. Cities Soc.* 75 (2021) 103352, https://doi.org/10.1016/j.scs.2021.103352.

[20] M. Karimimoshaver, M.S. Shahrak, The effect of height and orientation of buildings on thermal comfort, *Sustain. Cities Soc.* 79 (2022) 103720, https://doi.org/10.1016/j.scs.2022.103720.

[21] P. Kastner, T. Dogan, Eddy3D: A toolkit for decoupled outdoor thermal comfort simulations in urban areas, *Build. Environ.* 108639 (2021), https://doi.org/10.1016/j.buildenv.2021.108639.

[22] E.L. Krüger, F.O. Minella, F. Rasia, Impact of urban geometry on outdoor thermal comfort and air quality from field measurements in Curitiba, Brazil, *Building and Environment* 46 (3) (2011) 621–634, https://doi.org/10.1016/j.buildenv.2010.09.006.

[23] L.A. Kuehn, R.A. Stubbs, R.S. Weaver, Theory of the globe thermometer, *J. Appl. Physiol.* 29 (5) (1970) 750–757, https://doi.org/10.1152/jappl.1970.29.5.750.

[24] P. Kumari, D. Toshniwal, Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance, *J. Clean. Prod.* (2020) 123285, https://doi.org/10.1016/j.jclepro.2020.123285.

[25] D. Lai, W. Liu, T. Gan, K. Liu, Q. Chen, A review of mitigating strategies to improve the thermal environment and thermal comfort in urban outdoor spaces, *Sci. Total Environ.* 661 (2019) 337–353, https://doi.org/10.1016/j.scitotenv.2019.01.062.

[26] S. Liu, W. Pan, X. Zhao, H. Zhang, X. Cheng, Z. Long, Q. Chen, Influence of surrounding buildings on wind flow around a building predicted by CFD simulations, *Build. Environ.* 140 (2018) 1–10, https://doi.org/10.1016/j.buildenv.2018.05.011.

[27] C. Mackey, T. Galanos, L. Norford, M. Sadeghipour Roudsari, Wind, Sun, Surface Temperature, and Heat Island: Critical Variables for High-Resolution Outdoor Thermal Comfort, *Building Simulation Conference Proceedings* (2017), https://doi.org/10.26868/25222708.2017.260.

[28] M. Martínez-Comesaña, P. Eguía-Oller, J. Martínez-Torres, L. Febrero-Garrido, E. Granada-Álvarez, Optimisation of thermal comfort and indoor air quality estimations applied to in-use buildings combining NSGA-III and XGBoost, *Sustain. Cities Soc.* 80 (2022) 103723, https://doi.org/10.1016/j.scs.2022.103723.

[29] J. Natanian, P. Kastner, T. Dogan, T. Auer, From energy performative to livable Mediterranean cities: An annual outdoor thermal comfort and energy balance cross-climatic typological study, *Energ. Buildings* 224 (2020) 110283, https://doi.org/10.1016/j.enbuild.2020.110283.

[30] W.L. Oberkampf, T.G. Trucano, Verification and validation in computational fluid dynamics, *Prog. Aerosp. Sci.* 38 (3) (2002) 209–272, https://doi.org/10.1016/s0376-0421(02)00005-2.

[31] W. Ouyang, G. Ren, Z. Tan, Y. Li, C. Ren, Natural shading vs. artificial shading: A comparative analysis of their cooling efficacy in extreme hot weather, *Urban Clim.* 55 (2024) 101870, https://doi.org/10.1016/j.uclim.2024.101870.

[32] Z. Peng, W. Lu, T. Hao, X. Tang, J. Huang, C. Webster, Cost-aware generative design for urban "cool spots": A random forest-principal component analysis-augmented combinatorial optimization approach, *Energ. Buildings* 295 (2023) 113317, https://doi.org/10.1016/j.enbuild.2023.113317.

[33] Z. Peng, W. Lu, L. Yuan, Y. Zhang, The rise of syndicates: Social network analyses of construction waste haulers in Hong Kong using a novel InfoMap-XGBoost method, *Resour. Conserv. Recycl.* 206 (2024) 107663, https://doi.org/10.1016/j.resconrec.2024.107663.

[34] K. Perini, A. Magliocco, Effects of vegetation, urban density, building height, and atmospheric conditions on local temperatures and thermal comfort, *Urban For. Urban Green.* 13 (3) (2014) 495–506, https://doi.org/10.1016/j.ufug.2014.03.003.

[35] R. Priyadarsini, W.N. Hien, C.K. Wai David, Microclimatic modeling of the urban thermal environment of Singapore to mitigate urban heat island, *Sol. Energy* 82 (8) (2008) 727–745, https://doi.org/10.1016/j.solener.2008.02.008.

[36] Y. Shi, L. Katzschner, E. Ng, Modelling the fine-scale spatiotemporal pattern of urban heat island effect using land use regression approach in a megacity, *Sci. Total Environ.* 618 (2018) 891–904, https://doi.org/10.1016/j.scitotenv.2017.08.252.

[37] N. Skarbit, I.D. Stewart, J. Unger, T. Gál, Employing an urban meteorological network to monitor air temperature conditions in the "local climate zones" of Szeged, Hungary, *Int. J. Climatol.* 37 (2017) 582–596, https://doi.org/10.1002/joc.5023.

[38] H. Soltanifard, K. Aliabadi, Impact of urban spatial configuration on land surface temperature and urban heat islands: a case study of Mashhad, Iran*Theoretical and Applied Climatology* 137 (3–4) (2019) 2889–2903, https://doi.org/10.1007/s00704-018-2738-4.

[39] M. Srivanit, D. Jareemit, Modeling the influences of layouts of residential townhouses and tree-planting patterns on outdoor thermal comfort in Bangkok

suburb, *Journal of Building Engineering* 30 (2020) 101262, https://doi.org/10.1016/j.jobe.2020.101262.

[40] R. Sun, D. Lai, W. Liu, A computationally affordable and reasonably accurate approach for annual outdoor thermal comfort assessment on an hourly basis, *Energ. Buildings* 316 (2024) 114323, https://doi.org/10.1016/j.enbuild.2024.114323.

[41] M. Taleghani, L. Kleerekoper, M. Tenpierik, A. van den Dobbelsteen, Outdoor thermal comfort within five different urban forms in the Netherlands, *Build. Environ.* 83 (2015) 65–78, https://doi.org/10.1016/j.buildenv.2014.03.014.

[42] B. Yang, T. Olofsson, G. Nair, A. Kabanshi, Outdoor thermal comfort under subarctic climate of north Sweden – A pilot study in Umeå, *Sustain. Cities Soc.* 28 (2017) 387–397, https://doi.org/10.1016/j.scs.2016.10.011.

[43] J. Yang, Y. Yang, D. Sun, C. Jin, X. Xiao, Influence of urban morphological characteristics on thermal environment, *Sustain. Cities Soc.* 72 (2021) 103045, https://doi.org/10.1016/j.scs.2021.103045.

[44] S. Zare, N. Hasheminejad, H.E. Shirvan, R. Hemmatjo, K. Sarebanzadeh, S. Ahmadi, Comparing Universal Thermal Climate Index (UTCI) with selected thermal indices/environmental parameters during 12 months of the year, *Weather Clim. Extremes* 19 (2018) 49–57, https://doi.org/10.1016/j.wace.2018.01.004.

[45] L. Zheng, W. Lu, Urban micro-scale street thermal comfort prediction using a "graph attention network" model, 111780–111780, *Build. Environ.* (2024), https://doi.org/10.1016/j.buildenv.2024.111780.

[46] L. Zheng, W. Lu, Q. Zhou, Weather image-based short-term dense wind speed forecast with a ConvLSTM-LSTM deep learning model, *Build. Environ.* 239 (2023) 110446, https://doi.org/10.1016/j.buildenv.2023.110446.

[47] X. Zhou, L. Xu, J. Zhang, B. Niu, M. Luo, G. Zhou, X. Zhang, Data-driven thermal comfort model via support vector machine algorithms: Insights from ASHRAE RP-884 database, *Energ. Buildings* 211 (2020) 109795, https://doi.org/10.1016/j.enbuild.2020.109795.

[48] T. Zölch, M.A. Rahman, E. Pfleiderer, G. Wagner, S. Pauleit, Designing public squares with green infrastructure to optimize human thermal comfort, *Build. Environ.* 149 (2019) 640–654, https://doi.org/10.1016/j.buildenv.2018.12.051.