Leveraging machine learning to explore nonlinear associations between urban heat vulnerability and morbidity risk

Jiaming Yang [a], Zhaomin Tong [a], Jiwei Xu [a], Rui An [a], Yanfang Liu [a,b,c], Yaolin Liu [a,b,c,d,*]

[a] *School of Resource and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, PR China*
[b] *Key Laboratory of Geographic Information System of Ministry of Education, 129 Luoyu Road, Wuhan University, Wuhan 430079, PR China*
[c] *Collaborative Innovation Center of Geospatial Technology, 129 Luoyu Road, Wuhan University, Wuhan 430079, PR China*
[d] *Duke Kunshan University, No. 8 Duke Avenue, Kunshan 215316, Jiangsu Province, PR China*

ARTICLE INFO

ABSTRACT

Urban heat vulnerability (UHV) caused by anthropogenic activities and climate changes has given rise to heat health issues in urban areas worldwide. Previous studies have extensively revealed a simple linear relationship between heat vulnerability indices (HVIs) and morbidity or mortality of heat-related illnesses, but the nonlinear relationship and interactions between main HVIs have not yet been fully explored. Based on vulnerability assessment framework, this paper selected fifteen indicators from built environment, sociodemographic and socioeconomic attributes, resource accessibility and residential thermal comfort, obtained from multisource data. Through the evaluation and analysis of composite HVI and its dimensions, we found that Qingshan district and East Lake scenic area contain more high to very high heat vulnerability communities. The performances of the ordinary least squares (OLS) and gradient boosting decision trees (GBDT) were compared, and results indicate GBDT outperforms the OLS model and captures the nonlinear relationship more efficiently in study areas with higher accuracy. When analyzing HVIs' contributions and interactions with the GBDT model and the SHAP algorithm, nighttime light (NTL), building year (BY), PM2.5, floor area ratio (FAR), number of elderly (≥65 years) (NE) and urban surface roughness (USR) are six key indicators of morbidity of heat-related diseases (mean SHAP value>2.5), and they have an evident nonlinear relationship with the threshold effect and spatially heterogeneous contributions for the morbidity variation of heat-related diseases. Our study provides insights into machine learning (ML) model for the effect of heat vulnerability on city residential health and mitigation and adaptation strategies for governments and urban planners to develop heat resilience cities.

1. Introduction

Against the backdrop of rapid urbanization, the urban heat island (UHI) affecting the urban environment and human health has been an overly concerning issue for much more than 40 years (Chen et al., 2006; Ellena et al., 2023; Tesfamariam et al., 2023). As a result of human-caused climate changes, extreme high-temperature events greatly disordering human physiological systems have

---

* Corresponding author at: School of Resource and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, PR China.
E-mail address: yaolin610@163.com (Y. Liu).

posed serious threats to human health, giving rise to increased morbidity and mortality of heat-related illnesses (Elizabeth Loughnan et al., 2014; Hu et al., 2017; Sherwood and Huber, 2010). More specifically, in the summer of 2003, more than 70,000 deaths resulted from the heatwave in Europe (Robine et al., 2008), and by the end of the 21st century, 1.3 million deaths can be caused by extreme high-temperature event in the USA (Schiermeier, 2019). In the summer of 2013, the strongest intensity of heat waves since 1951 hit China, and 5758 heat-related diseases cases were reported by Chinese Center for Disease Control and Prevention (China CDC) (Gu et al., 2015). Approximately 60 % of the global population of 8.2 billion is expected to reside in cities by 2030 and the global average surface temperature may increase by 1.5 °C from 2030 to 2052 (Delic and Kincl, 2008; IPCC, 2018). Cities experiencing high-intensity, high-frequency heatwaves with limited heat resistance, especially urban centers with dense built-up areas, are more likely to bring adverse heat-induced effects that threaten the residents' physical and mental health, such as fatal cases of cardiovascular and respiratory diseases and hospital admissions for anxiety or depressive disorders (Li et al., 2021; Liu et al., 2021a). Cities find responding to heat disasters challenging given numerous vulnerable groups (e.g., children, elderly individuals and ethnic minority) who have a lower ability to handle heat events compared with other social groups (Zeng et al., 2022). Therefore, cities become more vulnerable to extreme heat, and comprehensively understanding the urban heat vulnerability (UHV) exacerbated by UHI and developing mitigation and adaptation strategies to strengthen urban heat resilience are imperative (Li et al., 2019).

The conceptual definition of vulnerability is extensively used according to the Intergovernmental Panel on Climate Change (IPCC), and vulnerability is a function of climate changes to which a system is exposed, its sensitivity, and its adaptive capacity (IPCC, 2007). UHV is also a consequence of the interaction of exposure, sensitivity and adaptive capacity with evaluation indicators for all the three. Heat exposure demonstrates the degree of heat-induced disaster proximity and perception for people. Temperature is commonly used to represent the spatial distribution and temporal changes of the extent of heat exposure, including land surface temperature (LST) and air temperature (Aubrecht and Özceylan, 2013; Estoque et al., 2020). The outdoor thermal sensation of people also provide researchers a great research aspect, such as Physiological Equivalent Temperature (PET) (Zemtsov et al., 2020), universal thermal climate index (UTCI)(Liu et al., 2021b) and the COMFA (Comfort Formula) model (Kim et al., 2023). In addition, normalized difference vegetation index (NDVI), and the proportion of impervious surfaces that characterize land cover are used for the quantitative expression of high-temperature exposure (Johnson et al., 2012; Uejio et al., 2011). However, few studies focus on urban ventilation and air pollution conditions. The size of urban ventilation potential plays an irreplaceable role in reducing UHI, thereby improving the wind, thermal environments, and thermal comfort (Liu et al., 2021b; Wang et al., 2022). The interactions of high temperature and air pollution can be harmful to the health conditions of urban residents (Harlan and Ruddell, 2011).

In terms of sensitivity, heat sensitivity reflects adverse conditions that heighten individuals' exposure to harmful heat effects. Previous studies have mainly considered the physiological state of people and their socio-economic state, such as the percentage of population ≥ 65 years of age, percentage of the population below high school education and the proportion of people in low-income households. (Chen et al., 2022; Guo et al., 2019; Palinkas et al., 2022). People who are older and have lower economic status are more sensitive to heat, making them more susceptible to heat-related diseases. However, the built environment considerably affects the urban thermal environment, such as floor area ratio (FAR), road density (RD) and urban surface roughness (USR), but is rarely included in UHV (He et al., 2019; Lu et al., 2021). Previous studies have shown the variation of high temperature mediated by built environment has a considerable influence on the rate of chronic diseases (e.g., kidney disease) (Shen, 2022). The urban built environment, social demography, and socioeconomic attributes jointly contribute to UHV (Macintyre et al., 2018; Song et al., 2020; Xiang et al., 2022). Heat adaptive capacity can be expressed as the ability of people to resist adverse heat effects. Many scholars adopted various indicators to measure adaptability, such as economic conditions, accessibility of blue and green spaces, and available medical resources (Eisenman et al., 2016; Sun and Chen, 2012; Wu et al., 2022b). The high availability and accessibility of cooling facilities and medical resources can enhance people's adaptability to high temperature, thereby reducing incidence of heat-related diseases.

Recently, the urban heat vulnerability has become a research focus, a great number of scholars have attempted to use the conceptual framework proposed by the IPCC to assess and map the intra-city heat-related health risk and vulnerability (Chen et al., 2022; Wu et al., 2022a; Zhu and Yuan, 2023). Appreciating the spatial heterogeneity of urban heat health risk and where to mitigate adverse heat-induced effects is helpful. Meanwhile, exploring the profound adverse health effects of UHV is indispensable, and further research on the relationship between heat-related morbidity or mortality and heat vulnerability indices (HVIs) is needed. However, no unified standard exists in the selection of evaluation indicators. For instance, human thermal comfort, air pollution and outdoor ventilation conditions in exposure component, built environment variables in sensitivity component and nighttime light (NTL) in adaptive capacity component are rarely or not considered by scholars. Moreover, different statistical methods, such as Poisson regression, multiscale geographically weighted regression (MGWR) and other linear regression methods, are commonly used to detect the effects of HVIs on heat-related morbidity or mortality (Maier et al., 2014; Song et al., 2021). However, due to the assumption of independent variables, the interaction effects between indicators on the morbidity or mortality of heat-related diseases have not been fully explored. This phenomenon weakens the comprehensive understanding of explanatory variables and reduces the reliability of explanatory models when mining the relationship between UHV and morbidity of heat-related diseases. Meanwhile, the complex nonlinear effect of HVIs on heat-induced adverse health effects have not been sufficiently explored either. Although some scholars utilized the distributed lag nonlinear model (DLNM) to analyze the associations of different hot extremes with mortality risk at the street level, uncertainty estimates (e.g., confidence intervals, hypothesis testing) and all necessary assumptions make the selection of indicators not comprehensive enough, and the results of DLNM are difficult to summarize due to complex parameterization (Gasparrini et al., 2010; Li et al., 2021).

Given these circumstances, machine learning (ML) techniques (e.g., random forest, XGBoost, deep learning neural networks) are gradually prevalent but rarely used in research related to UHV. ML techniques, especially tree-based models, can provide convincing prediction results with minimal assumptions about data-generating processes and also be more interpretable by using TreeExplainer

than linear models owing to model-mismatch effects (Bzdok et al., 2018; Lundberg et al., 2020). A series of recent studies suggest gradient boosting decision trees (GBDT) is an effective approach to overcome the limitations of traditional linear regression methods with higher prediction accuracy and visual effect process of explanatory variables (Tong et al., 2022; Xu et al., 2021; Yang et al., 2021). The Shapley additive explanation (SHAP) method used for interpreting ML black-box models can estimate the total, main, and interaction effects of influence factors (Dong et al., 2022). As a method for post hoc interpretation, the Tree SHAP has many advantages (Lundberg et al., 2020; Molnar, 2022): Tree SHAP not only can sharply reduce time complexity of interpretation but also global interpretation based on numerous local interpretations can be obtained by plotting the beeswarm-style SHAP summary plot. However, traditional feature importance plot can only represent the magnitude and order of relative importance for each variable. In addition, SHAP dependence plots can simultaneously reveal influences on the prediction and interaction effects of every sample and can be an alternative to traditional partial dependence plots (PDPs) and accumulated local effect (ALE) Plots.

To address the above knowledge gap, this paper built a comprehensive UHV evaluation index system consisting of fifteen distinct factors by using multi-source data (e.g., POI, remote sensing data, building data, etc.). Next, we utilize the relative importance of SHAP as the indicator weight to assess composite heat vulnerability and its three dimensions. Then, the non-linear relationship between the key factors of HVIs and the community-level morbidity of heat-related diseases, and the threshold effects among key indicators in the highly urbanized city of Wuhan in China were investigated by employing GBDT and SHAP algorithm. The interaction effects among the key factors affecting the morbidity of heat-related diseases were explored by plotting the SHAP dependence plot, and the spatially heterogeneous contributions of the key factors of HVIs to the morbidity of heat-related diseases were determined by mapping the SHAP values. The results from this paper can help researchers better understand how UHV affects human health, and targeted mitigation and adaptation strategies and policies should be implemented to improve individual health. This study proposes to answer the following key questions:

(1) How do the key indicators of HVIs influence the morbidity of heat-related diseases (including the nonlinear relationship and threshold effects) at a community scale? What are the specific contributions of the three types of indicators to the prediction of morbidity and what is their importance ranking? What are the key variables?
(2) What are the influences of the interaction effects between the key factors of HVIs on the morbidity of heat-related diseases at a community scale?
(3) What are the spatial variations of the contribution of the key factors of HVIs to morbidity prediction of heat-related diseases?

Compared with previous studies on the heat health effects, the main novel contributions for this paper are as follows: First, the combination of medical emergency investigation data and census data spatially quantifies the morbidity of heat-related diseases, which provides a basis for further spatial analysis. Second, a nonlinear relationship exists between HVIs and the morbidity of heat-related diseases, and the interaction effects of the key factors of HVIs substantially affect this relationship. Third, without too many assumptions about data, the contributions of key factors to the morbidity of heat-related diseases are spatially heterogeneous.
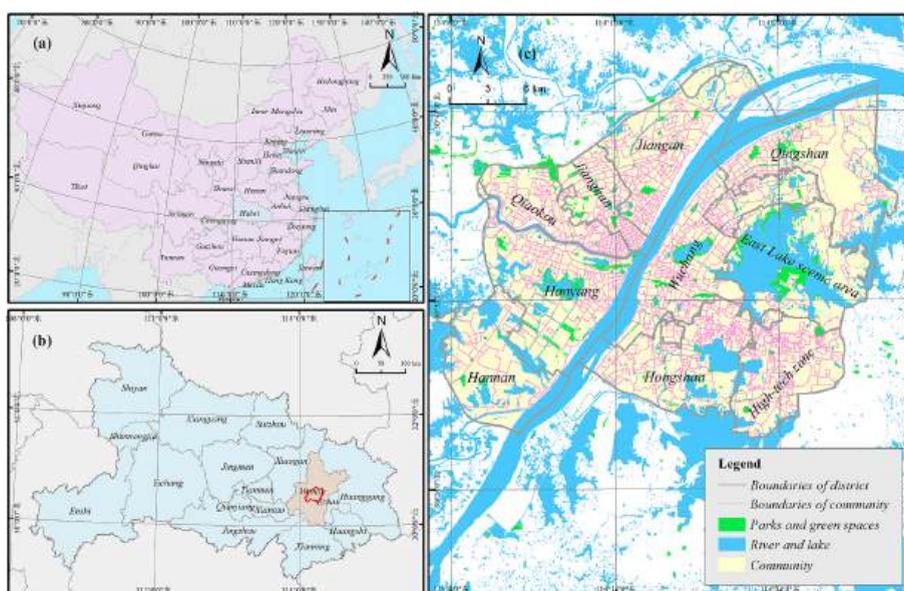
Fig. 1. Study area.

2. Study area and dataset

2.1. Study area

Wuhan is the capital city of Hubei Province in Central China. It is located in the east of Jianghan Plain and the middle reaches of the Yangtze River (Fig. 1). The city has crisscross rivers and intertwined lakes and ports. The water area accounts for a quarter of the total area of the city, each importantly helping regulate the urban thermal environment. Wuhan has a humid subtropical monsoon climate, with rainfall and hot summer mostly from June to August, and as such is called one of China's four furnaces. Due to the rapid expansion of the city in recent years, by the end of 2020, the permanent population of the city has reached about 12.40 million, according to the results of The Seventh National Population Census of the People's Republic of China, and the number of people over 65 in Wuhan has reached 1.1 million, accounting for 17.1 % of the total population, proving that the problem of aging population is becoming more serious. The main urban area of Wuhan has a concentrated population and dense buildings, and forming heat island effect is easy in summer, so it was selected as the study area. Within the study area, 1031 communities were taken as the research units, with an average area of 0.569 km^2.

2.2. Data collection and preprocessing

2.2.1. Heat-related diseases
In this paper, as shown in Table 1, heat-related diseases include heat-related chronic diseases and heat-related acute diseases. According to previous studies, heat-related chronic diseases can be divided into four types: respiratory disease, cardiovascular disease, digestive system disease and diabetes (Basagaña et al., 2011), and heat-related acute diseases including heatstroke (Liu et al., 2020). A total of 6996 data on heat-related diseases morbidity from June to August 2017 to 2019 and 55 medical emergency stations were collected from Wuhan Municipal Health Commission (WMHC). The dataset consists of the patient's demographic characteristics (e.g., sex, age and so on), disease type, hospitals treated, and address coordinates. Heat-related morbidity can be expressed as the density of heat-related disease by measuring the amount of heat-related diseases per unit of residential land area in each community (Fig. 2).

2.2.2. Spatial variables for HVIs
This paper collected multisource data to access HV in the main area of Wuhan, including building data, remote sensing data, air quality data, demographic and socioeconomic data, road network data, land use data and point of interest data. The time nodes of all data related to HVIs were unified to 2018. Table 2 visually represents the indicators selected in this paper, and lists the detailed information, including data sources and formula (or data processing).

(1) Urban heat exposure

In this paper, temperature-humidity index (TI), normalized difference water index (NDWI), ventilation index (VI) and PM2.5 were used to indicate urban heat exposure. First, TI was calculated at the community scale. TI is the normalized modified temperature–humidity index (MTHI), which commonly represents urban thermal comfort (Roshan et al., 2022). In summer, a smaller TI means more comfortable, thermal perception, and the reverse holds in winter (Feng et al., 2020). TI is extremely useful because it only requires easily accessible remote sensing data and does not require complex models to represent outdoor thermal comfort.

Water bodies have a stronger cooling effect than green spaces under the same weather conditions (Yao et al., 2022). Areas lacking surface cover such as water bodies are more susceptible to adverse heat-induced effects. Given the substantial proportion of water area in Wuhan, NDWI is used to indicate a thermal exposure indicator.

Urban canopy airflow is a key factor in relieving the UHIs and improving the urban thermal environment. The airflow in the urban canopy is forced to branch, which is highly similar to the branch current in circuit networks composed of nodes and resistors (Xie et al., 2020). Based on building data, the circuit theory method was used to simulate the urban ventilation and calculate the VI at the community level, which can effectively reflect the extent of exposure to adverse heat.

Table 1
Heat-related diseases.

Categories	Types	Content
Chronic diseases	Respiratory	Pneumonia
		Chronic lower respiratory disorder.
		Other respiratory disorder.
	Cardiovascular	Heart: Ischemic
		Cerebrovascular
		Atherosclerosis
		Hypertensive
	Digestive system	Ulcers
		Liver diseases
	Endocrine system	Diabetes
Acute diseases	Heatstroke	
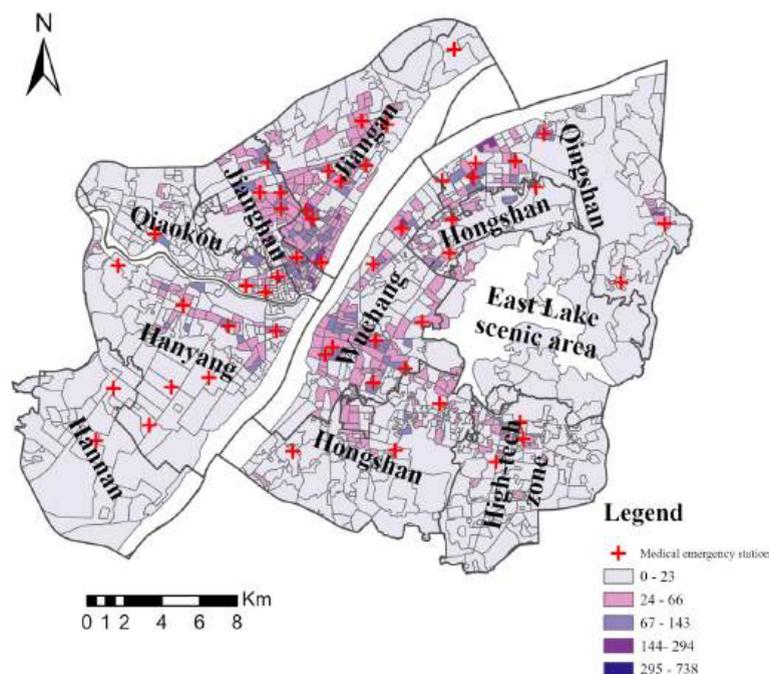
Fig. 2. Spatial distribution of morbidity of heat-related diseases and medical emergency stations.

Air pollution has always been a concerning issue, heat and air pollution have a synergistic effect, and this effect may have a more severe influence on human health than a single factor among them (Wang et al., 2023), which can increase the risk of cardiovascular disease and respiratory disease mortality, especially in summer (Rai et al., 2023). Thus, PM2.5 was used to represent particulate matter in the air, the PM2.5 concentration ($\mu g/m^3$) was calculated at the community scale, and the relationship between PM2.5 concentration and morbidity of heat-related diseases was explored.

(2) Urban heat sensitivity

Children under 5 years old and elderly people over 65 years old are heat sensitive due to their susceptible physical constitution to heat disaster. The number of these two sensitive populations, which was used to characterize urban thermal sensitivity, were collected.

The compact built environment in cities increases the urban thermal sensitivity and the risk of adverse health effects, and four variables were calculated to indicate heat sensitivity, namely, USR, FAR, RD and LUM in each community. USR and FAR are highly related to the building height, which could significantly affect urban thermal environment (Hou et al., 2023; Zhou and Chen, 2018). RD is associated with traffic intensity, more dense roads represent more anthropogenic heat emissions (He et al., 2019), which may exacerbate UHV. In addition, more diverse land use may promote increasing the attractiveness of a land plot and active travel in a community, especially for the older population (Xu et al., 2023), which increase the heat sensitivity in community. Thus, five types of land related to urban thermal environment were chosen to calculate LUM: commercial, industry, residential, public, and transportation land.

Building year(BY) information were also collected to indicate urban heat sensitivity, considering its relationship with age and cooling facilities (Cai et al., 2019).

(3) Urban heat adaptive capacity

First, HV information which reflect the economic conditions of dwellers in communities was collected for each community (Atalay et al., 2017). NTL intensity can represent the regional economic development status and energy consumption deteriorating the urban heat island effect (Keola et al., 2015). NTL was used to represent socioeconomic status in each community. The accessibility of each patient to medical emergency stations and cooling places was also determined. Cooling places include parks and green spaces, supermarkets and libraries. The richness and availability of medical resources and cooling facilities reduce the likelihood of outdoor morbidity and enhance social heat adaptation (He et al., 2019; Zhang et al., 2019).

2.2.3. Heat vulnerability index (HVI)

Similar to current studies on heat risk or heat vulnerability, we weighted all HVIs to comprehensively evaluate the degree of heat impact on local areas, after normalizing all indicators using the min-max normalization method (Guo et al., 2019; Lopes et al., 2023). In order to avoid subjectively assigning weights to indicators, we use the normalized SHAP feature importance as the weight for each

Table 2
Data sources and processing of each indicator.

Target	Indicator	Data sources	Formula (or data processing)
Urban Heat Exposure	Temperature-humidity Index (TI)	MOD11A1 and MOD09GA datasets from June to August, obtained from National Aeronautics and Space Administration (NASA)	$TI = \dfrac{MTHI - MTHI_{mean}}{MTHI_{mean}}$ $MTHI = 1.8 \times LST + 32 - 0.55 \times (1 - NDMI) \times (1.8 \times LST - 26)$ $NDMI = \dfrac{NIR - SWIR}{NIR + SWIR}$ MTHI is the normalized modified temperature-humidity index, LST is the land surface temperature in °C, NDMI is normalized difference moisture index, NIR is the near-infrared band and SWIR is the short wavelength infrared band
	NDWI	MOD09GA datasets from June to August, obtained from NASA	$NDWI = \dfrac{Green - NIR}{Green + NIR}$ Green is the short wavelength infrared band and NIR is near-infrared band
	Ventilation Index (VI)	Building data obtained from Baidu Map, Meteorological Station Data obtained from National Climatic Data Center (NCDC)	$VI = \dfrac{A_h - A_f}{A_P}, \lambda_f = \dfrac{A_f}{A_p}$ A_h is the projected background area of the windward side of the building with height h, λ_f is the building frontal area index, A_f is the projected building area of windward side of the building, and A_P is planar area.
	PM2.5	ChinaHighPM2.5 dataset obtained from National Earth System Science Data Center, National Science & Technology Infrastructure of China (Wei et al., 2020; Wei et al., 2021)	–
	Number of elderly (≥65 years) (NE)	Data provided by Wuhan Geomatics Institute (WGI)	–
	Number of infants (≤5 years) (NI)	Data provided by WGI	–
Urban Heat Sensitivity	Urban Surface Roughness (USR)	Building data obtained from Baidu Map	$\dfrac{Z_d}{Z_h} = 1 - \dfrac{1 - exp\left[-\left(7.5 \times 2 \times \lambda_f\right)^{0.5}\right]}{\left(7.5 \times 2 \times \lambda_f\right)^{0.5}}$ $\dfrac{Z_0}{Z_h} = \left(1 - \dfrac{Z_d}{Z_h}\right) exp\left(-0.4 \times \dfrac{U_h}{u_*} + 0.193\right)$ $\dfrac{u_*}{U_h} = min\left[\left(0.003 + 0.3 \times \lambda_f\right)^{0.5}, 0.3\right]$ $\lambda_f = \dfrac{A_f}{A_p}$ Z_0 is the surface roughness, Z_d is the height of the zero plane displacement, Z_h is the building height, U_h is the wind speed, u_* is the shear velocity, λ_f is the building frontal area index.
	Floor Area Ratio (FAR)	Building data obtained from Baidu Map	$FAR = \dfrac{\sum_{i=1}^{N} A_i \times f_i}{A_l}$ A_i is the base area of building in a land plot, f_i is the number of building floors, and A_l is land plot area.
	Road Density (RD)	Data provided by WGI	Density of Road is the ratio of total length of all roads to the area in each community(km/km²).
	Land Use Mix (LUM)	Data provided by WGI	$EI = -\dfrac{\sum_{k=1}^{N} p_i ln p_i}{ln k}$ EI is the entropy index, p_i is the proportion of the ith category of land use, and k is the number of land use categories.
	Building Year (BY)	Real estate network platform	–
	House Value (HV)	Real estate network platform	–
	Nighttime Light (NTL)	VIIRS-DNB dataset, obtained from Earth Observation Group (EOG) (Elvidge et al., 2017)	–
Urban Heat Adaptive Capacity	Access to medical services (AMSs)	The medical emergency station points obtained from WMHC, road networks data obtained from WGI	The OD cost matrix and ordinary kriging interpolation method in ArcGIS were used to calculate the accessibility from medical emergency points to patient points.
	Access to cooling places (ACPs)	Point of interest (POI) data obtained from Baidu Map	The OD cost matrix and ordinary kriging interpolation method in ArcGIS were used to calculate the accessibility from cooling places (e.g., supermarket, library, etc) points to patient points.

indicator. The composite HVI is defined as:

$$HVI = EI + SI - ACI \tag{1}$$

Where EI is the Exposure Index, weighted by exposure level indicators; SI is the Sensitivity Index, weighted by sensitivity level indicators; ACI is the Adaptive Capacity Index, weighted by adaptive capacity level indicators.

3. Methods

3.1. Ordinary least squares

The ordinary least squares (OLS) model is a traditional global linear regression model, which is often used to quantify the relationship between dependent variables and independent variables. The model assumes no multicollinearity between variables and ignores the mutual influence of variables in adjacent regions in space (Gao et al., 2022). The OLS is defined as follows:

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \varepsilon \tag{2}$$

where y is the dependent variable, β_0 represents the intercept, β_i is the coefficient of the independent variable x_i, and n represents the number of independent variables, and is the random error.

3.2. Gradient boosting decision trees

This paper employed GBDT and SHAP explanation to explore the relationships between the morbidity of heat-related diseases and HVIs, and the interaction effects of influential factors on morbidity of heat-related diseases at the community level.

The GBDT is a ML method consisting of gradient boosting and decisions trees. To minimize the prediction error, the algorithm builds a set of decision trees at each stage of the iteration and selects the optimal number of the trees. Gradient boosting integrates multiple weak models into one strong model to minimize the prediction error along the gradient direction. In addition, the algorithm could reflect the potential nonlinear relationships between independent variables and dependent variables and acquire the relative importance of variables with higher accuracy than other algorithms.

$f(x)$ is an approximation function of the independent variable y (i.e., density of heat-related diseases) explained as an additive expansion of basis function $h(x; a_m)$, which represents the results of multiple decision trees:

$$f(x) = \sum_{m=1}^{M} f_m(x) = \sum_{m=1}^{M} \beta_m h(x; a_m) \tag{3}$$

where a_m is the average values of the split positions and the terminal node in an individual decision tree $h(x; a_m)$, β_m represents the weights assigned to each tree nodes and is calculated by minimizing the loss function, and M is the total number of the regression trees.

To estimate the parameters, the gradient boosting method is applied (Friedman, 2001). The procedure includes several iterative steps.

First, regression tree initialization:

$$f_0(x) = argmin_\beta \sum_{i=1}^{N} L(y_i, \beta) \tag{4}$$

For all samples N, the total loss can be represented as follows:

$$L_{all} = \sum_{i=1}^{N} L(y_i, f_m(x_i)) \tag{5}$$

where y_i and $f_m(x_i)$ are the true values of sample x_i and the predicted values of the mth model, respectively. For each iteration, the residual can be expressed by a negative gradient as follows:

$$\widetilde{y}_{im} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f=f_{m-1}} \tag{6}$$

Then, an optimal gradient descent step size for the seeking direction of the gradient update can be presented as follows:

$$\beta_m = argmin_\beta \sum_{i=1}^{N} L(y_i, f_{m-1}(x_i) + \beta t(x_i; a_m)) \tag{7}$$

To solve the overfitting problem (Friedman, 2001), the algorithm uses shrinkage (i.e., learning rate) to control the contribution of each base tree model by introducing a factor of ξ ($0 < \xi \leq 1$) as shown below:

$$f_m(x) = f_{m-1}(x) + \xi \bullet \beta_m h(x; a_m), where \ 0 < \xi \leq 1 \tag{8}$$

In addition, the fivefold cross-validation method was used to train the model for solving overfitting problems. Moreover, to obtain better model results, the optimal parameters in the model were adjusted. In this paper, 5000 decision trees with the depth at 5 were adjusted, and the shrinkage was set to 0.001. Furthermore, the result with a small root mean square error (RMSE) was obtained, and the pseudo-R^2 was about 0.33 in our model.

3.3. Nonlinear interpretation of HVIs

In the field of ML, the famous Black Box Problem is ubiquitous, and thus even the most advanced ML models, such as XGBoost and GBDT, lose the same interpretability as the linear models. Considering the above problem, the SHAP algorithm is introduced to solve the problem and offer a more reasonable nonlinear effect between HVIs and the morbidity of heat-related diseases. SHAP is an approach for estimating Shapley values and the main function of the SHAP based on cooperative game theory Shapley value is to obtain the marginal contribution of the sample point added to the model. Shapley values can be used in machine learning to quantify the contribution of each feature in the model that collectively delivers the prediction (Štrumbelj and Kononenko, 2013). The Shapley value for X_i feature in a model can be expressed by:

$$Shapley(X_i) = \sum_{S \subseteq N \setminus \{i\}} \frac{k!(p-k-1)!}{p!} (f(S \cup \{i\}) - f(S)) \tag{9}$$

where p refers to the total number of features; S represents the permutation subset of features used in the model, the dimension is k; $N \setminus \{i\}$ is a set of all possible combinations of features excluding X_i; $\frac{k!(p-k-1)!}{p!}$ is the weight of the difference between samples with feature i and without feature i using the feature set; $f(S)$ is the average predicted value of samples using the feature set S; $f(S \cup \{i\})$ is the average predicted value of samples with feature in S plus feature X_i. The interpretation of Eq. (8) is that the Shapley value is its marginal contribution to model prediction averaged over all possible models with different combinations of features (Lipovetsky and Conklin, 2001). Moreover, the positive and negative values of $Shapley(X_i)$ represent the degree of influence of different heat vulnerability features at per-sample points.

The absolute Shapley values can be used to measure the global importance of each feature:

$$I_j = \frac{1}{n} \sum_{i=1}^{n} |Shapley(X_i)| \tag{10}$$

where I_j represents the SHAP feature importance measured as the mean absolute Shapley values. Not only is the SHAP feature importance a substitute for the importance of permutation feature importance, but also the beeswarm-style SHAP summary plot based on it contain more information than traditional feature importance plot.

SHAP measures the pairwise interaction effect $Shapley(X_i)$ based on the Shapley value for each prediction. The calculation is as follows:

$$Shapley(X_{i,j}) = \sum_{S \subseteq N \setminus \{i,j\}} \frac{k!(p-k-2)!}{2(p-1)!} \nabla ij(S) \tag{11}$$

here, for i $\neq$ j, $\nabla ij(S) = f(S \cup \{i,j\}) - f(S \cup \{i\}) - f(S \cup \{j\}) + f(S)$. And this formula subtracts the main effect of features, so that after considering the individual effect, we get the pure interaction effect. Compared with traditional partial dependence plots, the SHAP feature dependence plot colored with the strongest interaction effect.

SHAP's main contribution is to generate locally additive feature attribution, as described in Eq. (11):

$$\widehat{y_i} = shap_0 + shap(X_{1i}) + shap(X_{2i}) + \ldots + shap(X_{pi}) \tag{12}$$

where $\widehat{y_i}$ is the model prediction value for the observation i, $shap_0 = E(\widehat{y_i})$ is the mean prediction across all observation, and $shap(X_{ji})$ represents the SHAP value of jth feature for observation i which represents the marginal contribution of the feature to the prediction. In this way, the sum of all SHAP values is equal to the difference between the actual prediction and the average prediction. As for the morbidity of heat-related diseases prediction model. The SHAP algorithm calculates the marginal effect of each feature to the model's prediction results, thus solving the black-box problem of empirical model.

Table 3
Comparison of model performance for the OLS and GBDT results.

Model Type	R^2	MSE	MAE
OLS	0.198	1617.731	20.621
GBDT	0.334	1345.937	18.970

4. Results

4.1. Comparison of the OLS and GBDT models

The ML regression method of building models and making predictions through training data differs from traditionally statistical regression methods. Here, the OLS model and the GBDT model were used to predict the effect of HVI on the density of heat-related diseases occurrence, and we compared the R^2, mean squared error (MSE), and mean absolute error (MAE) of two models were compared, as shown in Table 3.

As expected, the GBDT approach performs better for HVIs with higher R^2 and lower MSE and MAE. More specifically, the GBDT model fitting represented by R^2 improved by 13.6 % compared with the OLS model. MSE and MAE are 271.794 and 1.651 lower than those of the OLS model, respectively. Therefore, the GBDT model has a better predictive performance than the OLS model.

4.2. Relative importance of HVI in predicting heat-related diseases

Fig. 3 shows a summary plot of estimated SHAP values colored by feature values of each point. The density distribution of SHAP value per feature has overlapping points jittered in the y-axis direction (Guliyev and Mustafayev, 2022). Overall, the first six most important variables are NTL, BY, PM2.5, FAR, NE and USR.

For the exposure variables, PM2.5 is the most highly related with the density of heat-related diseases (mean SHAP value = 3.89), and a heat-related diseases' negative feedback and insignificant positive feedback from its distribution of feature values are observed. However, TI ranks 12 and had a low importance in predicting heat-related diseases. For sensitivity variables, BY has a more important marginal negative effect on the prediction of morbidity of heat-related diseases (mean SHAP value = 3.95), followed by FAR, NE and USR with mean SHAP values of 3.35, 3.02, and 2.95, respectively. However, all the variables including FAR, NE and USR have a marginal positive influence on the prediction of heat-related diseases. For adaptive capacity variables, NTL is the most important variable in predicting heat-related diseases (mean SHAP value = 4.14) with the most significant variability in SHAP values, followed by access to medical services (AMSs) (mean SHAP value = 2.82). NTL has a marginal positive influence on the heat-related disease prediction, whereas AMSs has a negative correlation. Therefore, in terms of the distribution of feature value, not all variables have red and blue dots separated by SHAP values equal to 0, which means the relationships between HVIs and the morbidity of heat-related diseases are not single positive or negative trends.

4.3. Spatial representation of composite vulnerability assessment and vulnerability dimensions

Mapping composite UHV and its three dimensions is essential for identifying potential high-vulnerable areas. The exposure of the main urban area demonstrated an increasing trend from center areas to the periphery. Qingshan district has the largest number of community high heat exposure, followed by Hongshan district (Fig. 4a). Qingshan district experiences a certain level of persistent industrial waste heat due to its industrial areas. The areas surrounding the Yezhi Lake and the South Lake experience high heat exposure due to strong evaporation during the summer. A large number of the high to very high heat sensitivity communities can be
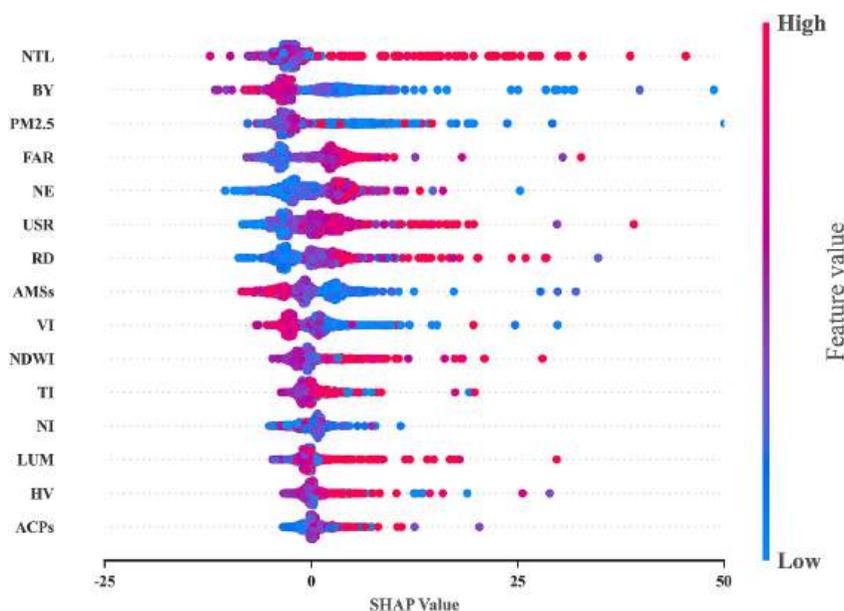
Fig. 3. SHAP summary plot for each indicators.
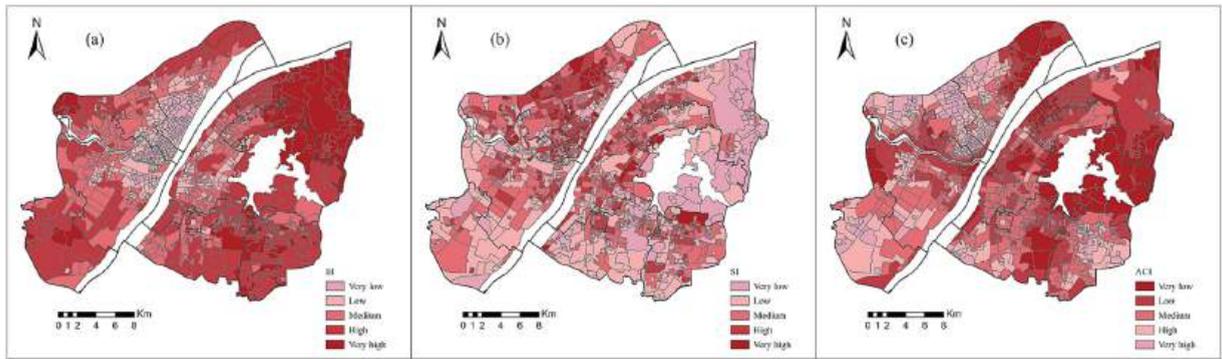
Fig. 4. Spatial representation of heat vulnerability dimensions.

found in Wuchang district, Qiaokou district, Jianghan district and Jiangan district (Fig. 4b). This phenomenon is the result of the combined influence of sensitive populations and the built environment. Furthermore, low to very low heat adaptive capacity areas are concentrated along the Yangtze River and Hanjiang River, as well as the East Lake scenic area and surrounding areas of the South Lake (Fig. 4c). This indicates that the uneven distribution of public health resources and disparities in socioeconomic levels among the population.

37 % of the main urban area exhibits high thermal vulnerability, with regions of high composite heat vulnerability concentrated more in the northern part of the main urban area (Fig. 5a). In Qingshan District, 45 % of communities exhibit a very high composite HVI, meriting focused attention, high vulnerability to heat may be driven by prolonged air pollution in the industrial area, the concentration of retired workers in nearby communities, and relatively low socioeconomic status (Fig. 5b).

4.4. Direction of association, nonlinear relationships, and interaction effects

The marginal effects of HVIs on the predicted morbidity of heat-related diseases can be observed by drawing SHAP dependence plots, and analyzing the interaction effects among the HVIs (Fig. 6). The color variance expresses the value of another variable with the strongest interaction with this variable in each community. In addition, available information from the nonlinear pattern and threshold effects from the SHAP dependence plots can be obtained.

Variables that are positively correlated with the morbidity of heat-related diseases include NE, USR, FAR, and NTL intensity. People over 65 years old are extremely sensitive to temperature. When NE (Fig. 6e) is less than 600, the increase in the FAR weakens its positive effect. However, when NE is between 600 and 1000, the situation is converse. After the threshold of 1000, its influence on the density of the heat-related diseases is stable. USR and FAR are affected by the interaction of PM2.5. For USR (Fig. 6f), the prediction effect has a small amplitude with 21 person/km^2, and as the USR increases from 0 to 4, the increase in the PM2.5 slightly suppresses its positive effects. When it is less than 2.5, the increase in the PM2.5 weakens its positive effect. However, as the FAR (Fig. 6d) increases from 2.5 to 3.5, the predicted morbidity of heat-related diseases increases rapidly then it tends to stabilize. As for NTL (Fig. 6a), its predictive ability has not changed significantly until 50 nW/cm^2/sr, when it increases from 40 nW/cm^2/sr, NE promotes its positive effect.

The following variables are negatively associated with the morbidity of heat-related diseases: VI, BY and AMSs. The urban
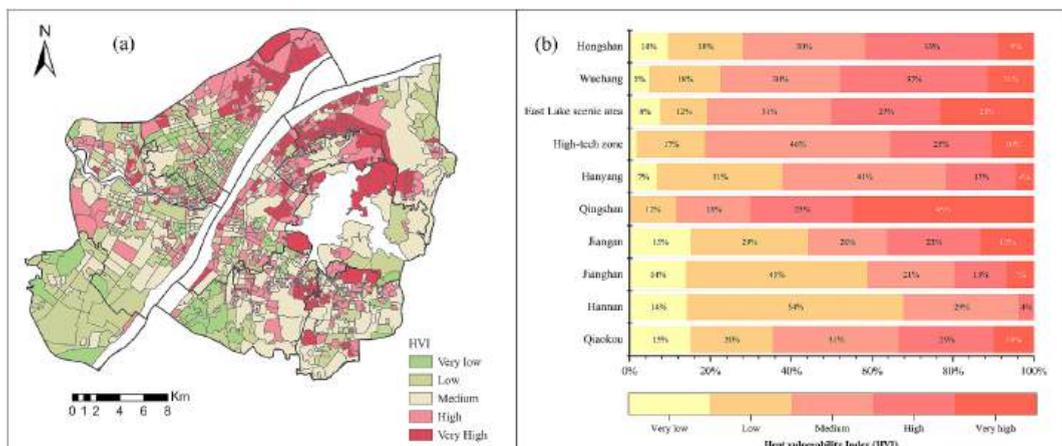
Fig. 5. Spatial representation and summary of composite HVI.
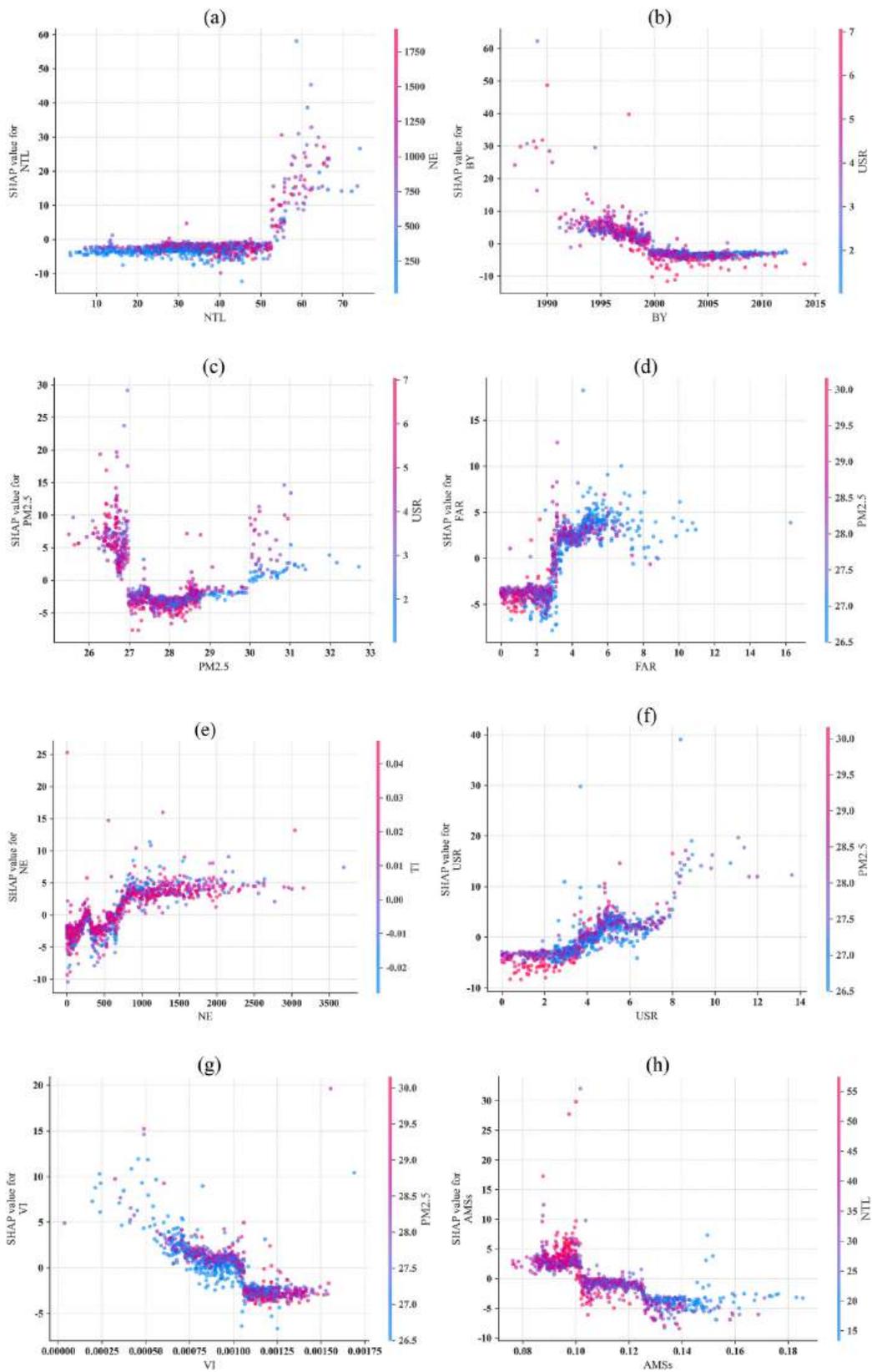
Fig. 6. SHAP dependency graphs of eight indicators of HVIs with SHAP values.
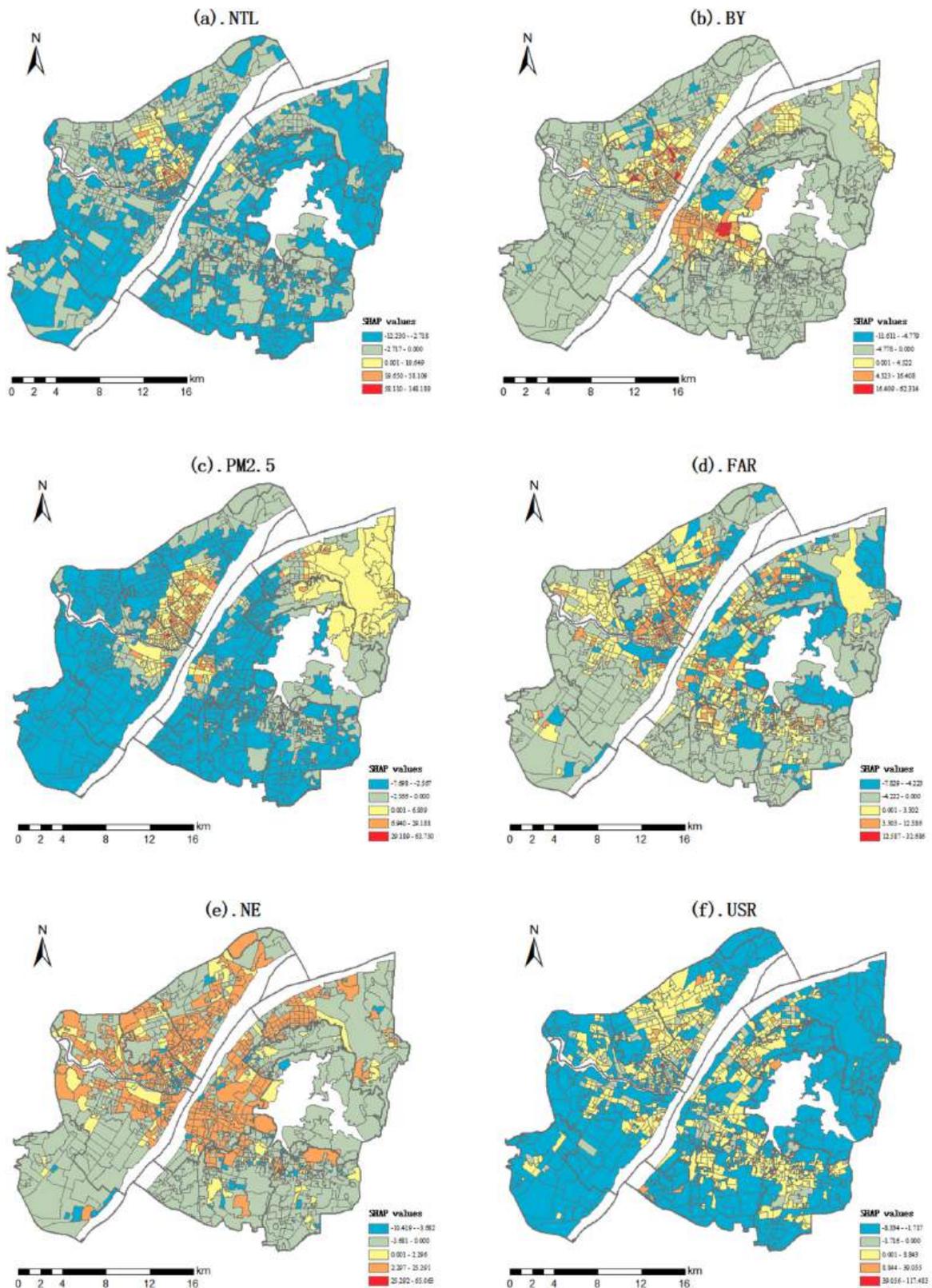
Fig. 7. The spatial distribution of SHAP values for the key factors of HVIs.

ventilation potential is closely related to thermal environment. As the VI (Fig. 6g) increases to the threshold of 0.106 %, the negative effect that the increase in PM2.5 promotes tends to be stable. When BY (Fig. 6b) is less than 2000, the increase in the USR suppresses its negative impact, after the threshold of 2000, the negative effect that the increase in USR promotes stabilizes. The correlation between AMSs and heat-related diseases shows a stepped shape (Fig. 6h). When the AMSs is less than 0.1 h, the increase in the NTL slightly suppresses its negative effects. However, when it is over 0.1 or increases to the threshold of 0.125, the predicted density of heat-related diseases decreases rapidly and then stabilizes, and the increase in NTL promotes its negative effect.

The morbidity of heat-related diseases decreases significantly when PM2.5 is less than 27, and the increase in USR suppresses its negative effect. However, when PM2.5 (Fig. 6c) is between 27 and 28.5, the increase in USR promotes its negative effect. After the value of 28.5, the increase PM2.5 has a positive effect on predicting the density of heat-related diseases.

4.5. Spatial patterns of predictive contribution of the key indicators of HVIs

Previous studies have often used spatial statistical model to evaluate the spatial effects of independent variables, such as such as spatial econometrics models (e.g. spatial lag and error models) and spatially varying coefficients models (e.g. geographically weighted regression). The local SHAP values can be mapped to the spatial background for visual analysis to explore the spatial distributions of the influence of HVIs on the morbidity of heat-related diseases. Fig. 7 shows the SHAP values of the key indicators of HVIs have an explicit spatial pattern of predictive contributions to the morbidity of heat-related diseases. The range of SHAP values was divided into five levels, where positive contributions can be expressed by areas closer to red, whereas negative contributions can be represented by areas closer to blue. The top six HVIs are important to the morbidity of heat-related diseases. NTL (Fig. 7a) has a negative contribution in most regions, and only has a positive contribution in Jianghan and southern Jiangan. In terms of BY (Fig. 7b), the negative contributions can be found in the periphery of the main urban area, such as Hannan and the high-tech zone. However, the central region has the largest negative contribution, such as northern Jianghan, central Wuchang and northern Jiangan. The positive contributions are concentrated in the central region, including southern Wuchang, Qiaokou, Jianghan, and Jiangan. The east and west sides of Qingshan can also exhibit positive contribution. PM2.5 (Fig. 7c) has positive contributions in southern Jiangan, southern Jianghan, southeastern Qiaokou, northeastern Hanyang, western Wuchang, and Qingshan, and the rest of regions are all negative contributions.

In terms of FAR (Fig. 7d), positive and negative contribution areas are alternately distributed, the degree of negative contribution gradually increases from the periphery to the central region. The positive contributions are concentrated in the central region including Jiangan, Jianghan, Qiaokou, northern Hanyang, northern Hongshan, Wuchang, western Qingshan and western High-tech zone. As for NE (Fig. 7e), the positive contributions can be found in northern Hanyang, Qiaokou, Jianghan, Jiangan, Wuchang, northern Hongshan, western Qingshan and northern high-tech zone. The negative contributions are concentrated in the periphery, such as Hannan, southern Hanyang, southern Hongshan, southern High-tech zone and eastern Qingshan. As for USR (Fig. 7f), the positive contributions can be found in Jiangan, northern Jianghan, Qiaokou, eastern Hanyang, northern Hongshan, western high-tech zone, Wuchang, and western Qingshan. The negative contributions are concentrated in the periphery, such as eastern Qingshan, Hannan and northern Jiangan.

5. Discussion

5.1. Spatial heterogeneity of HVI and vulnerability dimensions

In our paper, we built a comprehensive UHV index system, and mapped composite HVI and vulnerability dimensions. To avoid subjectivity, we use the normalized SHAP feature importance as the weight of each indicator. Consistent with prior results, urban centers exhibit a greater vulnerability to heat, characterized by distinct spatial heterogeneity across different areas (Li et al., 2021). Additionally, the spatial patterns of values show distinct variations among the three dimensions of EI, SI, and ACI (Fig. 4), which coincides with the previous findings (Xiang et al., 2022).

In urban areas, higher heat exposure, sensitivity and lower adaptive capacity demonstrate lower HVI. Wuchang district, Qingshan district and East Lake scenic area contain more high to very high heat vulnerability communities, targeted policies and mitigation strategies should be implemented in these areas.

5.2. Nonlinear effect of HVIs on community-level morbidity of heat-related diseases

The paper explored the non-linear influence of HVIs on the morbidity of heat-related diseases using ML methods that differ from traditional linear regression methods. The results indicate the superior predictive performance of the GBDT model. Furthermore, by analyzing the relative importance and SHAP partial dependence, the key variables that affect the incidence of heat-related diseases were identified and the threshold effects and interaction effects were explored.

In this paper, we define the top six HVIs important to the prediction as key indicators of HVIs. Fig. 3 shows that comparing the performances of different HVIs, NTL, BY and PM2.5 are the three most important factors influencing the incidence of heat-related diseases. Among them, NTL plays the dominant role in influencing the heat-related disease prediction, which demonstrates that socio-economic conditions are strongly associated with the morbidity of heat-related diseases. Similar to prior results (Li et al., 2020; Yuan et al., 2022), as NTL increases, the slope of the association flattens out with low SHAP values, which suggests that the boost of economic development may contribute to mitigating the UHI effect and reducing the incidence of heat-related diseases. Nevertheless, the urban polycentric areas formed by economic development also have a dense population (Jung et al., 2022), which increases the

likelihood of the incidence of heat-related diseases, especially among sensitive populations.

Housing conditions with good cooling facilities alleviate HV. When the year of construction is before 2000, the incidence of heat-related diseases increases. Some studies have shown that old buildings are always have a higher risk of overheating because of poor insulation performance and insufficient cooling facilities (Cai et al., 2019; Samuelson et al., 2020). However, in some old city districts, higher USR means stronger obstruction on urban ventilation (Fang et al., 2021), which may exacerbate the UHI effects leading to an increase in the incidence of heat-related diseases.

The effect of air pollution on human health cannot be ignored. Most of previous studies have shown that long-term exposure to high concentrations of PM2.5 is positively associated with deaths due to heat-related diseases. According to Fig. 4c, when it is less than 28, it has an abnormal negative effect. The PM2.5 concentration in the entire study area is less than 35, the industrial area in Qingshan district has a high concentration of PM2.5. However, the concentration of PM2.5 is less than 27 in the south of Jiangan district, but the density of heat-related diseases is high. A possible explanation is that PM2.5 is not the most crucial factor contributing to the incidence of the heat-related disease.

Furthermore, FAR, NE, and USR also presented substantial contribution to the prediction of incidence of the heat-related diseases. This outcome indicates that sensitive population and building environment are greatly affecting the probability of heat-related diseases incidence. The 3D urban morphology has a profound influence on the SUHIs (Shao et al., 2023). Some studies have shown a negative correlation between building height and LST in summer due to the shadow effects (Huang and Wang, 2019; Yu et al., 2020), while some have come to the opposite conclusion (Berger et al., 2017). In this paper, a positive correlation between FAR and the morbidity of heat-related diseases as a whole is found. This phenomenon may be attributed to the fact that high, dense building areas absorb short-wave radiation and cause airflow obstruction effects and thus, form an adverse thermal environment (Ward et al., 2016; Zhou and Chen, 2018).

In addition, growing evidence suggests USR is one of the most critical determinants of the UHI effects (Hou et al., 2023). The higher USR is, the greater the likelihood of heat-related diseases occurring, which could be attributed to that the higher USR, the more chaotic airflow, and the increased difficulty in effectively alleviating the adverse heat effects leading to the occurrence of diseases (Lopes et al., 2011).

Elderly people are more vulnerable than younger groups in dealing with urban heat risks. Many studies have reported that elderly people belong to heat sensitive population, and they have the possibility of living alone and being disabled (Zhu and Yuan, 2023). As expected, NE is positively correlated with the incidence of heat-related diseases, which coincides with the study that the heat-related EMS incidents in neighborhoods increases with higher percentage of the elderly population (Seong et al., 2022).

5.3. Spatial effects in the influence of HVIs on community-level morbidity of heat-related diseases

For various HVIs, the SHAP values vary over space, thereby the effect of HVIs on morbidity of diseases may differ from locations of communities. Previous studies revealed spatial variability in the influence of heat health risk indicators on heat-related illnesses by using linear regression models, such as the MGWR and spatial error models (Kovach et al., 2015; Song et al., 2021). In terms of socio-economic status, the intensity of NTL in most areas of the main urban area negatively contribute to the morbidity of heat-related diseases. However, previous studies indicated socio-economic status is anticipated to contribute positively to heat-related health risk at the global scale (Song et al., 2021). This phenomenon may be attributed to the adoption of different indicators to represent socio-economic status. As for demographic attributes, NE positively contributed to the morbidity of heat-related diseases in the central and western regions of the main urban area. This finding is similar to previous research, which suggests the number of the elderly is positively associated with the emergency department (ED) visits at the area level (Kovach et al., 2015). In terms of the built environment, the positive contributions of BY, FAR, and USR to the morbidity of heat-related diseases are concentrated near both sides of the Yangtze River. These areas with high population and building density easily form a thermal environment adverse to health. Therefore, compared with some linear regression models, SHAP can be used to identify spatial heterogeneity of the contribution of each feature to the dependent variable at different scales without many assumptions.

5.4. Implications and limitations

Methodologically, the organic integration of ML algorithm and SHAP can be a fantastic tool for identifying significant features and understanding relationships between accessible geospatial data (Li, 2022). More specifically, the totality of the results confirms the nonlinear association of HVIs on the morbidity of heat-related diseases and disputes the widely held assumption of the simple linear relationship in previous studies. Relatively more accurate prediction could suggest that relevant urban planners could reduce the morbidity of heat-related diseases by managing several vital drivers.

The trade-off between socioeconomic development and UHI effects should be paid much attention by urban planners, socioeconomic factors provide possibilities to alleviate the UHI, but in the face of population growth, sustainable economic development becomes increasingly crucial. For building environment, improving unified building standards, such as height, width, and layout, is conducive to creating a favorable urban ventilation environment, and thus enhancing people's outdoor thermal comfort. Meanwhile, building soft structures and green measures can also be improved to upgrade building design (Xiang et al., 2022).

In the post COVID-19 era, reducing human thermal exposure is a non-negligible factor to alleviate heatwave-induced human health risk, and strategies that may ameliorate the heat-related morbidity burden associated with heat exposure are mainly concerned with modifying behavior. One strategy is for individuals to improve self-awareness on how to manage heat (Zhao et al., 2023). More specifically, either staying indoors in an air-conditioned environment for reducing outdoor activities or moving away from the city's

hot center seem to be a feasible solution (Estoque et al., 2020; Xu et al., 2019). Social vulnerability reduction is the core of urban planning, and more deserved attention should be paid to empty nesters, left-behind children and other sensitive, disadvantaged groups who urgently need assistance from public policies. The local government should improve health intervention services and provide well-equipped high-temperature emergency supplies to reduce the morbidity of heat-related diseases at the community level, such as providing regular physical examination, health assessment and cooling equipment (i.e. fans or air conditioners).

For megacity planning, strategies enhancing urban heat health resilience for local residents need further research and application. On the one hand, given the balance between the supply of public resources and people's demand, how to achieve reasonable allocation of resources is worth pondering. For example, the optimization of location for the EMS station determining the speed of service is vital to a patient (Hashemi et al., 2022). On the other hand, with respect to old neighborhoods, the adoption of the mitigation strategies including cool pavement, green roof, and pocket park can be a good way to improve thermal comfort if residents need (Evola et al., 2017).

From the perspective of territory and management, different districts should have different heat mitigation policies. For example, Qingshan district government should adopt heat recovery technology, heat generated during production can be redirected for heating, power generation, or other processes, reducing heat loss and minimizing waste. Retired workers from nearby factories should receive corresponding high-temperature subsidies. The buildings in Wuchang district are dense, however, there are water areas in the East and west of Wuchang District, so the building layout should be reasonable to enhance the ventilation corridor. The East Lake scenic area is rich in tourism resources, and the government should strengthen the protection of blue and green space.

Admittedly, some limitations are noted. First, some variables associated with sociodemographic and socioeconomic aspects are not included in HVIs (eg., family size, educational and income level, sexual structure and so on) because of limited data availability. Second, the research unit selected as community has certain feasibility, but faces modified areal unit problem (MAUP). The measurement of HVIs and morbidity of heat-related diseases will be influenced by the unit division scheme, which affects the final result. In the future, comparative studies at multiple scales can be conducted. Third, overfitting and sparse data problems cannot be perfectly solved in the GBDT model, and the nonlinear results demonstrate more correlation instead of causality. Finally, this paper focuses more on heat-related diseases. Various health information data (eg., past medical history, various physical function data and so on) need to be included to understand better the relationships between HVIs and physical indicators for urban dwellers in the future.

6. Conclusions

The urban heat-related issue has placed residential health and socioeconomic system into highly vulnerable position (Sun et al., 2022). Taking the highly urbanized region of Wuhan as a sample area, this paper aims to explore the relationship between HVIs and the morbidity of heat-related diseases. Different from previous research, this paper is the first to take advantage of advanced interpretable ML method–GBDT and SHAP algorithm to map composite UHV and its three dimensions, and identify the contribution of HVIs and the interaction effects of the morbidity of heat-related diseases at the community level by using multi-source data. First, by comparing the OLS model with the GBDT model, that the results show the GBDT model has higher R^2, lower MSE and MAE, and demonstrates the better potential of using ML methods to investigate the nonlinear relationship. Next, by using the relative importance of SHAP as the indicator weight for composite heat vulnerability assessment, we found that Qingshan district has the highest number of high to very high heat-vulnerable communities (Fig. 5). Then, by further analyzing the relative importance and nonlinear relationships, it can be concluded that the top six HVIs important to the morbidity of heat-related diseases are NTL, BY, PM2.5, FAR, NE and USR. Among HVIs, threshold characteristics are ubiquitous, and in general, NE, USR, FAR, and NTL have a positive effect on the morbidity of heat-related diseases, whereas VI, BY, and AMSs show opposite correlation patterns, and PM2.5 has a slight effect on heat-related diseases at the community level. In addition, evident spatial heterogeneity exists in the contributions of the built environment, socioeconomic and demographic attributes to the morbidity of heat-related diseases. The built environment, elderly population and socio-economic indicators should be paid more attention to alleviate UHV. These importance indicators with their marginal effects on morbidity of heat-related diseases may be worthy of reference for the government to prioritize health intervention strategies and assist urban planners and decision-makers in the strategic planning stage of a heat resilience city in Wuhan.

In the future, we can adopt more reasonable and applicable indicators to meet the needs of different scenarios and urban thermal environment. From natural environment perspective, LST can directly reflect changes in heat hazard or exposure dimension, and the accumulation of surface heat is a direct factor contributing to UHI effects. Air temperature is also the primary measure of heat directly perceived by the human body. NDVI represents vegetation coverage density, which is related to urban cooling, and more vegetation means a larger heat adaptation area or urban cold island (UCI). In terms of built environment, transportation convenience is also a factor worth considering, because traffic heat emissions also have a certain impact on the urban thermal environment. As for demographic and socioeconomic characteristics, we should consider the impact of gender difference, because heat perception tolerance may vary by gender. Personal income and education level are also important indicators, because they are related to household air conditioning ownership rates and more opportunities for medical services. Considering these indicators enables the establishment of a more comprehensive UHV index system tailored to diverse urban environments, facilitating a more precise assessment of vulnerability and the development of targeted strategies(Li et al., 2022).

CRediT authorship contribution statement

Jiaming Yang: Writing – original draft, Data curation. **Zhaomin Tong:** Software, Methodology, Conceptualization. **Jiwei Xu:** Data curation. **Rui An:** Visualization, Investigation. **Yanfang Liu:** Supervision. **Yaolin Liu:** Supervision.

Declaration of competing interest

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

Acknowledgements

Data availability

Data will be made available on request.

References

Atalay, K., Edwards, R., Liu, B.Y.J., 2017. Effects of house prices on health: new evidence from Australia. Soc. Sci. Med. 192, 36–48. https://doi.org/10.1016/j.socscimed.2017.09.008.

Aubrecht, C., Özceylan, D., 2013. Identification of heat risk patterns in the U.S. National Capital Region by integrating heat stress and related vulnerability. Environ. Int. 56, 65–77. https://doi.org/10.1016/j.envint.2013.03.005.

Basagaña, X., Sartini, C., Barrera-Gómez, J., Dadvand, P., Cunillera, J., Ostro, B., Sunyer, J., Medina-Ramón, M., 2011. Heat waves and cause-specific mortality at all ages. Epidemiology 22, 765–772. https://doi.org/10.1097/EDE.0b013e31823031c5.

Berger, C., Rosentreter, J., Voltersen, M., Baumgart, C., Schmullius, C., Hese, S., 2017. Spatio-temporal analysis of the relationship between 2D/3D urban site characteristics and land surface temperature. Remote Sens. Environ. 193, 225–243. https://doi.org/10.1016/j.rse.2017.02.020.

Bzdok, D., Altman, N., Krzywinski, M., 2018. Statistics versus machine learning. Nat. Methods 15, 233–234. https://doi.org/10.1038/nmeth.4642.

Cai, Z., Tang, Y., Chen, K., Han, G., 2019. Assessing the heat vulnerability of different local climate zones in the old areas of a Chinese megacity. Sustainability 11. https://doi.org/10.3390/su11072032.

Chen, X.-L., Zhao, H.-M., Li, P.-X., Yin, Z.-Y., 2006. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sens. Environ. 104, 133–146. https://doi.org/10.1016/j.rse.2005.11.016.

Chen, T.-L., Lin, H., Chiu, Y.-H., 2022. Heat vulnerability and extreme heat risk at the metropolitan scale: a case study of Taipei metropolitan area, Taiwan. Urban Clim. 41. https://doi.org/10.1016/j.uclim.2021.101054.

Delic, A., Kincl, B., 2008. World urbanization – new visions. Sustain. City V, 399–409. https://doi.org/10.2495/sc080381.

Dong, S., Khattak, A., Ullah, I., Zhou, J., Hussain, A., 2022. Predicting and analyzing road traffic injury severity using boosting-based ensemble learning models with SHAPley additive exPlanations. Int. J. Environ. Res. Public Health 19. https://doi.org/10.3390/ijerph19052925.

Eisenman, D.P., Wilhalme, H., Tseng, C.-H., Chester, M., English, P., Pincetl, S., Fraser, A., Vangala, S., Dhaliwal, S.K., 2016. Heat death associations with the built environment, social vulnerability and their interactions with rising temperature. Health Place 41, 89–99. https://doi.org/10.1016/j.healthplace.2016.08.007.

Elizabeth Loughnan, M., Tapper, J.N., Phan, T.A., McInnes, J., 2014. Can a spatial index of heat-related vulnerability predict emergency service demand in Australian capital cities? Int. J. Emerg. Serv. 3, 6–33. https://doi.org/10.1108/ijes-10-2012-0044.

Ellena, M., Melis, G., Zengarini, N., Di Gangi, E., Ricciardi, G., Mercogliano, P., Costa, G., 2023. Micro-scale UHI risk assessment on the heat-health nexus within cities by looking at socio-economic factors and built environment characteristics: the Turin case study (Italy). Urban Clim. 49. https://doi.org/10.1016/j.uclim.2023.101514.

Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C., Ghosh, T., 2017. VIIRS night-time lights. Int. J. Remote Sens. 38, 5860–5879. https://doi.org/10.1080/01431161.2017.1342050.

Estoque, R.C., Ooba, M., Seposo, X.T., Togawa, T., Hijioka, Y., Takahashi, K., Nakamura, S., 2020. Heat health risk assessment in Philippine cities using remotely sensed data and social-ecological indicators. Nat. Commun. 11. https://doi.org/10.1038/s41467-020-15218-8.

Evola, G., Gagliano, A., Fichera, A., Marletta, L., Martinico, F., Nocera, F., Pagano, A., 2017. UHI effects and strategies to improve outdoor thermal comfort in dense and old neighbourhoods. Sustain. Energy Build. 2017 (134), 692–701. https://doi.org/10.1016/j.egypro.2017.09.589.

Fang, Y., Gu, K., Qian, Z., Sun, Z., Wang, Y., Wang, A., 2021. Performance evaluation on multi-scenario urban ventilation corridors based on least cost path. J. Urban Manag. 10, 3–15. https://doi.org/10.1016/j.jum.2020.06.006.

Feng, L., Zhao, M., Zhou, Y., Zhu, L., Tian, H., 2020. The seasonal and annual impacts of landscape patterns on the urban thermal comfort using Landsat. Ecol. Indic. 110. https://doi.org/10.1016/j.ecolind.2019.105798.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29. https://doi.org/10.1214/aos/1013203451.

Gao, Y., Zhao, J., Yu, K., 2022. Effects of block morphology on the surface thermal environment and the corresponding planning strategy using the geographically weighted regression model. Build. Environ. 216. https://doi.org/10.1016/j.buildenv.2022.109037.

Gasparrini, A., Armstrong, B., Kenward, M.G., 2010. Distributed lag non-linear models. Stat. Med. 29, 2224–2234. https://doi.org/10.1002/sim.3940.

Gu, S., Huang, C., Bai, L., Chu, C., Liu, Q., 2015. Heat-related illness in China, summer of 2013. Int. J. Biometeorol. 60, 131–137. https://doi.org/10.1007/s00484-015-1011-0.

Guliyev, H., Mustafayev, E., 2022. Predicting the changes in the WTI crude oil price dynamics using machine learning models. Res. Policy 77. https://doi.org/10.1016/j.resourpol.2022.102664.

Guo, X., Huang, G., Jia, P., Wu, J., 2019. Estimating fine-scale heat vulnerability in Beijing through two approaches: spatial patterns, similarities, and divergence. Remote Sens. 11. https://doi.org/10.3390/rs11202358.

Harlan, S.L., Ruddell, D.M., 2011. Climate change and health in cities: impacts of heat and air pollution and potential co-benefits from mitigation and adaptation. Curr. Opin. Environ. Sustain. 3, 126–134. https://doi.org/10.1016/j.cosust.2011.01.001.

Hashemi, S.E., Jabbari, M., Yaghoubi, P., 2022. A mathematical optimization model for location emergency medical service (EMS) centers using contour lines. Healthcare Analyt. 2. https://doi.org/10.1016/j.health.2022.100026.

He, C., Ma, L., Zhou, L., Kan, H., Zhang, Y., Ma, W., Chen, B., 2019. Exploring the mechanisms of heat wave vulnerability at the urban scale based on the application of big data and artificial societies. Environ. Int. 127, 573–583. https://doi.org/10.1016/j.envint.2019.01.057.

Hou, H., Su, H., Yao, C., Wang, Z.-H., 2023. Spatiotemporal patterns of the impact of surface roughness and morphology on urban heat island. Sustain. Cities Soc. 92. https://doi.org/10.1016/j.scs.2023.104513.

Hu, K., Yang, X., Zhong, J., Fei, F., Qi, J., 2017. Spatially explicit mapping of heat health risk utilizing environmental and socioeconomic data. Environ. Sci. Technol. 51, 1498–1507. https://doi.org/10.1021/acs.est.6b04355.

Huang, X., Wang, Y., 2019. Investigating the effects of 3D urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: a case study of Wuhan, Central China. ISPRS J. Photogramm. Remote Sens. 152, 119–131. https://doi.org/10.1016/j.isprsjprs.2019.04.010.

IPCC, 2007. Climate change 2007: Impacts, adaptation and vulnerability. In: Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK. https://www.ipcc.ch/site/assets/uploads/2018/03/ar4_wg2_full_report.pdf.

IPCC, Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P.R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B.R., Chen, Y., Zhou, X., Gomis, M.I., Lonnoy, E., Maycock, T., Tignor, M., Waterfield, T., 2018. Summary for policymakers. In: Global Warming of 1.5°C.An IPCC Special Report on the Impacts of Global Warming of 1.5°C above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 1–24. https://doi.org/10.1017/9781009157940.001.

Johnson, D.P., Stanforth, A., Lulla, V., Luber, G., 2012. Developing an applied extreme heat vulnerability index utilizing socioeconomic and environmental data. Appl. Geogr. 35, 23–31. https://doi.org/10.1016/j.apgeog.2012.04.006.

Jung, M.C., Kang, M., Kim, S., 2022. Does polycentric development produce less transportation carbon emissions? Evidence from urban form identified by night-time lights across US metropolitan areas. Urban Clim. 44. https://doi.org/10.1016/j.uclim.2022.101223.

Keola, S., Andersson, M., Hall, O., 2015. Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. World Dev. 66, 322–334. https://doi.org/10.1016/j.worlddev.2014.08.017.

Kim, Y., Li, D., Xu, Y., Zhang, Y., Li, X., Muhlenforth, L., Xue, S., Brown, R., 2023. Heat vulnerability and street-level outdoor thermal comfort in the city of Houston: application of google street view image derived SVFs. Urban Clim. 51. https://doi.org/10.1016/j.uclim.2023.101617.

Kovach, M.M., Konrad, C.E., Fuhrmann, C.M., 2015. Area-level risk factors for heat-related illness in rural and urban locations across North Carolina, USA. Appl. Geogr. 60, 175–183. https://doi.org/10.1016/j.apgeog.2015.03.012.

Li, Z., 2022. Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. Comput. Environ. Urban. Syst. 96. https://doi.org/10.1016/j.compenvurbsys.2022.101845.

Li, Z., Guo, X., Yang, Y., Hong, Y., Wang, Z., You, L., 2019. Heatwave trends and the population exposure over China in the 21st century as well as under 1.5 °C and 2.0 °C global warmer future scenarios. Sustainability 11. https://doi.org/10.3390/su11123318.

Li, Y., Sun, Y., Li, J., Gao, C., 2020. Socioeconomic drivers of urban heat island effect: empirical evidence from major Chinese cities. Sustain. Cities Soc. 63. https://doi.org/10.1016/j.scs.2020.102425.

Li, Z., Hu, J., Meng, R., He, G., Xu, X., Liu, T., Zeng, W., Li, X., Xiao, J., Huang, C., Du, Y., Ma, W., 2021. The association of compound hot extreme with mortality risk and vulnerability assessment at fine-spatial scale. Environ. Res. 198. https://doi.org/10.1016/j.envres.2021.111213.

Li, F., Yigitcanlar, T., Nepal, M., Thanh, K., Dur, F., 2022. Understanding urban heat vulnerability assessment methods: a PRISMA review. Energies 15. https://doi.org/10.3390/en15196998.

Lipovetsky, S., Conklin, M., 2001. Analysis of regression in game theory approach. Appl. Stoch. Model. Bus. Ind. 17, 319–330. https://doi.org/10.1002/asmb.446.

Liu, X., Yue, W., Yang, X., Hu, K., Zhang, W., Huang, M., 2020. Mapping urban heat vulnerability of extreme heat in Hangzhou via comparing two approaches. Complexity 2020, 1–16. https://doi.org/10.1155/2020/9717658.

Liu, J., Varghese, B.M., Hansen, A., Xiang, J., Zhang, Y., Dear, K., Gourley, M., Driscoll, T., Morgan, G., Capon, A., Bi, P., 2021a. Is there an association between hot weather and poor mental health outcomes? A systematic review and meta-analysis. Environ. Int. 153. https://doi.org/10.1016/j.envint.2021.106533.

Liu, W., Zhang, G., Jiang, Y., Wang, J., 2021b. Effective range and driving factors of the urban ventilation corridor effect on urban thermal comfort at unified scale with multisource data. Remote Sens. 13. https://doi.org/10.3390/rs13091783.

Lopes, A., Saraiva, J., Alcoforado, M.J., 2011. Urban boundary layer wind speed reduction in summer due to urban growth and environmental consequences in Lisbon. Environ. Model Softw. 26, 241–243. https://doi.org/10.1016/j.envsoft.2010.05.015.

Lopes, H., Remoaldo, P., Ribeiro, V., Martín-Vide, J., 2023. The Impacts of Climate Change on Human Wellbeing in the Municipality of Porto—An Analysis Based on Remote Sensing, Climate Change and Health Hazards, pp. 135–172. https://doi.org/10.1007/978-3-031-26592-1_8.

Lu, Y., Yue, W., Liu, Y., Huang, Y., 2021. Investigating the spatiotemporal non-stationary relationships between urban spatial form and land surface temperature: a case study of Wuhan, China. Sustain. Cities Soc. 72. https://doi.org/10.1016/j.scs.2021.103070.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intellig. 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Macintyre, H.L., Heaviside, C., Taylor, J., Picetti, R., Symonds, P., Cai, X.M., Vardoulakis, S., 2018. Assessing urban population vulnerability and environmental risks across an urban area during heatwaves – implications for health protection. Sci. Total Environ. 610-611, 678–690. https://doi.org/10.1016/j.scitotenv.2017.08.062.

Maier, G., Grundstein, A., Jang, W., Li, C., Naeher, L.P., Shepherd, M., 2014. Assessing the performance of a vulnerability index during oppressive heat across Georgia, United States. Weather Climate Soc. 6, 253–263. https://doi.org/10.1175/wcas-d-13-00037.1.

Molnar, C., 2022. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable.

Palinkas, L.A., Hurlburt, M.S., Fernandez, C., De Leon, J., Yu, K., Salinas, E., Garcia, E., Johnston, J., Rahman, M.M., Silva, S.J., McConnell, R.S., 2022. Vulnerable, resilient, or both? A qualitative study of adaptation resources and behaviors to heat waves and health outcomes of low-income residents of urban Heat Islands. Int. J. Environ. Res. Public Health 19. https://doi.org/10.3390/ijerph191711090.

Rai, M., Stafoggia, M., de Donato, F., Scortichini, M., Zafeiratou, S., Vazquez Fernandez, L., Zhang, S., Katsouyanni, K., Samoli, E., Rao, S., Lavigne, E., Guo, Y., Kan, H., Osorio, S., Kyselý, J., Urban, A., Orru, H., Maasikmets, M., Jaakkola, J.J.K., Ryti, N., Pascal, M., Hashizume, M., Fook Sheng Ng, C., Alahmad, B., Hurtado Diaz, M., De la Cruz Valencia, C., Nunes, B., Madureira, J., Scovronick, N., Garland, R.M., Kim, H., Lee, W., Tobias, A., Íñiguez, C., Forsberg, B., Åström, C., Maria Vicedo-Cabrera, A., Ragettli, M.S., Leon Guo, Y.-L., Pan, S.-C., Li, S., Gasparrini, A., Sera, F., Masselot, P., Schwartz, J., Zanobetti, A., Bell, M.L., Schneider, A., Breitner, S., 2023. Heat-related cardiorespiratory mortality: effect modification by air pollution across 482 cities from 24 countries. Environ. Int. 174. https://doi.org/10.1016/j.envint.2023.107825.

Robine, J.-M., Cheung, S.L.K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.-P., Herrmann, F.R., 2008. Death toll exceeded 70,000 in Europe during the summer of 2003. C. R. Biol. 331, 171–178. https://doi.org/10.1016/j.crvi.2007.12.001.

Roshan, G., Sarli, R., Fitchett, J.M., 2022. Urban heat island and thermal comfort of Esfahan City (Iran) during COVID-19 lockdown. J. Clean. Prod. 352. https://doi.org/10.1016/j.jclepro.2022.131498.

Samuelson, H., Baniassadi, A., Lin, A., Izaga González, P., Brawley, T., Narula, T., 2020. Housing as a critical determinant of heat vulnerability and health. Sci. Total Environ. 720. https://doi.org/10.1016/j.scitotenv.2020.137296.

Schiermeier, Q., 2019. Europe's mega-heatwave boosted by climate change. Nature 571, 155. https://doi.org/10.1038/d41586-019-02071-z.

Seong, K., Jiao, J., Mandalapu, A., 2022. Evaluating the effects of heat vulnerability on heat-related emergency medical service incidents: lessons from Austin, Texas. Environ. Plan. B Urban Analyt. City Sci. 50, 776–795. https://doi.org/10.1177/23998083221129618.

Shao, L., Liao, W., Li, P., Luo, M., Xiong, X., Liu, X., 2023. Drivers of global surface urban heat islands: surface property, climate background, and 2D/3D urban morphologies. Build. Environ. 242. https://doi.org/10.1016/j.buildenv.2023.110581.

Shen, Y.-S., 2022. Multiple pathways and mediation effects of built environment on kidney disease rate via mitigation of atmospheric threats. Sci. Total Environ. 833. https://doi.org/10.1016/j.scitotenv.2022.155177.

Sherwood, S.C., Huber, M., 2010. An adaptability limit to climate change due to heat stress. Proc. Natl. Acad. Sci. 107, 9552–9555. https://doi.org/10.1073/pnas.0913352107.

Song, J., Huang, B., Kim, J.S., Wen, J., Li, R., 2020. Fine-scale mapping of an evidence-based heat health risk index for high-density cities: Hong Kong as a case study. Sci. Total Environ. 718. https://doi.org/10.1016/j.scitotenv.2020.137226.

Song, J., Yu, H., Lu, Y., 2021. Spatial-scale dependent risk factors of heat-related mortality: a multiscale geographically weighted regression analysis. Sustain. Cities Soc. 74. https://doi.org/10.1016/j.scs.2021.103159.

Štrumbelj, E., Kononenko, I., 2013. Explaining prediction models and individual predictions with feature contributions. Knowl. Inf. Syst. 41, 647–665. https://doi.org/10.1007/s10115-013-0679-x.

Sun, R., Chen, L., 2012. How can urban water bodies be designed for climate adaptation? Landsc. Urban Plan. 105, 27–33. https://doi.org/10.1016/j.landurbplan.2011.11.018.

Sun, Y., Li, Y., Ma, R., Gao, C., Wu, Y., 2022. Mapping urban socio-economic vulnerability related to heat risk: a grid-based assessment framework by combing the geospatial big data. Urban Clim. 43. https://doi.org/10.1016/j.uclim.2022.101169.

Tesfamariam, S., Govindu, V., Uncha, A., 2023. Spatio-temporal analysis of urban heat island (UHI) and its effect on urban ecology: the case of Mekelle city, Northern Ethiopia. Heliyon 9. https://doi.org/10.1016/j.heliyon.2023.e13098.

Tong, Z., An, R., Zhang, Z., Liu, Y., Luo, M., 2022. Exploring non-linear and spatially non-stationary relationships between commuting burden and built environment correlates. J. Transp. Geogr. 104. https://doi.org/10.1016/j.jtrangeo.2022.103413.

Uejio, C.K., Wilhelmi, O.V., Golden, J.S., Mills, D.M., Gulino, S.P., Samenow, J.P., 2011. Intra-urban societal vulnerability to extreme heat: the role of heat exposure and the built environment, socioeconomics, and neighborhood stability. Health Place 17, 498–507. https://doi.org/10.1016/j.healthplace.2010.12.005.

Wang, W., Wang, D., Chen, H., Wang, B., Chen, X., 2022. Identifying urban ventilation corridors through quantitative analysis of ventilation potential and wind characteristics. Build. Environ. 214. https://doi.org/10.1016/j.buildenv.2022.108943.

Wang, S., Cai, W., Tao, Y., Sun, Q.C., Wong, P.P.Y., Huang, X., Liu, Y., 2023. Unpacking the inter- and intra-urban differences of the association between health and exposure to heat and air quality in Australia using global and local machine learning models. Sci. Total Environ. 871. https://doi.org/10.1016/j.scitotenv.2023.162005.

Ward, K., Lauf, S., Kleinschmit, B., Endlicher, W., 2016. Heat waves and urban heat islands in Europe: a review of relevant drivers. Sci. Total Environ. 569-570, 527–539. https://doi.org/10.1016/j.scitotenv.2016.06.119.

Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., Song, Y., 2020. Improved 1 km resolution $PM_{2.5}$ estimates across China using enhanced space–time extremely randomized trees. Atmos. Chem. Phys. 20, 3273–3289. https://doi.org/10.5194/acp-20-3273-2020.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., Cribb, M., 2021. Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. Remote Sens. Environ. 252. https://doi.org/10.1016/j.rse.2020.112136.

Wu, C., Shui, W., Huang, Z., Wang, C., Wu, Y., Wu, Y., Xue, C., Huang, Y., Zhang, Y., Zheng, D., 2022a. Urban heat vulnerability: a dynamic assessment using multi-source data in coastal metropolis of Southeast China. Front. Public Health 10. https://doi.org/10.3389/fpubh.2022.989963.

Wu, X., Liu, Q., Huang, C., Li, H., 2022b. Mapping heat-health vulnerability based on remote sensing: a case study in Karachi. Remote Sens. 14. https://doi.org/10.3390/rs14071590.

Xiang, Z., Qin, H., He, B.-J., Han, G., Chen, M., 2022. Heat vulnerability caused by physical and social conditions in a mountainous megacity of Chongqing, China. Sustain. Cities Soc. 80. https://doi.org/10.1016/j.scs.2022.103792.

Xie, P., Yang, J., Wang, H., Liu, Y., Liu, Y., 2020. A new method of simulating urban ventilation corridors using circuit theory. Sustain. Cities Soc. 59. https://doi.org/10.1016/j.scs.2020.102162.

Xu, R., Zhao, Q., Coelho, M.S.Z.S., Saldiva, P.H.N., Zoungas, S., Huxley, R.R., Abramson, M.J., Guo, Y., Li, S., 2019. Association between heat exposure and hospitalization for diabetes in Brazil during 2000–2015: a Nationwide case-crossover study. Environ. Health Perspect. 127. https://doi.org/10.1289/ehp5688.

Xu, Y., Yan, X., Liu, X., Zhao, X., 2021. Identifying key factors associated with ridesplitting adoption rate and modeling their nonlinear relationships. Transp. Res. A Policy Pract. 144, 170–188. https://doi.org/10.1016/j.tra.2020.12.005.

Xu, J., Jing, Y., Xu, X., Zhang, X., Liu, Y., He, H., Chen, F., Liu, Y., 2023. Spatial scale analysis for the relationships between the built environment and cardiovascular disease based on multi-source data. Health Place 83. https://doi.org/10.1016/j.healthplace.2023.103048.

Yang, J., Cao, J., Zhou, Y., 2021. Elaborating non-linear associations and synergies of subway access and land uses with urban vitality in Shenzhen. Transp. Res. A Policy Pract. 144, 74–88. https://doi.org/10.1016/j.tra.2020.11.014.

Yao, X., Zhu, Z., Zeng, X., Huang, S., Liu, Q., Yu, K., Zhou, X., Chen, Z., Liu, J., 2022. Linking maximum-impact and cumulative-impact indices to quantify the cooling effect of waterbodies in a subtropical city: a seasonal perspective. Sustain. Cities Soc. 82. https://doi.org/10.1016/j.scs.2022.103902.

Yu, S., Chen, Z., Yu, B., Wang, L., Wu, B., Wu, J., Zhao, F., 2020. Exploring the relationship between 2D/3D landscape pattern and land surface temperature based on explainable eXtreme gradient boosting tree: a case study of Shanghai, China. Sci. Total Environ. 725. https://doi.org/10.1016/j.scitotenv.2020.138229.

Yuan, Y., Li, C., Geng, X., Yu, Z., Fan, Z., Wang, X., 2022. Natural-anthropogenic environment interactively causes the surface urban heat island intensity variations in global climate zones. Environ. Int. 170. https://doi.org/10.1016/j.envint.2022.107574.

Zemtsov, S., Shartova, N., Varentsov, M., Konstantinov, P., Kidyaeva, V., Shchur, A., Timonin, S., Grischchenko, M., 2020. Intraurban social risk and mortality patterns during extreme heat events: a case study of Moscow, 2010-2017. Health Place 66. https://doi.org/10.1016/j.healthplace.2020.102429.

Zeng, P., Sun, F., Shi, D., Liu, Y., Zhang, R., Tian, T., Che, Y., 2022. Integrating anthropogenic heat emissions and cooling accessibility to explore environmental justice in heat-related health risks in Shanghai, China. Landsc. Urban Plan. 226. https://doi.org/10.1016/j.landurbplan.2022.104490.

Zhang, W., Zheng, C., Chen, F., 2019. Mapping heat-related health risks of elderly citizens in mountainous area: a case study of Chongqing, China. Sci. Total Environ. 663, 852–866. https://doi.org/10.1016/j.scitotenv.2019.01.240.

Zhao, Q., Li, Z., Shah, D., Fischer, H., Solís, P., Wentz, E., 2023. Understanding the interaction between human activities and physical health under extreme heat environment in Phoenix, Arizona. Health Place 79. https://doi.org/10.1016/j.healthplace.2021.102691.

Zhou, X., Chen, H., 2018. Impact of urbanization-related land use land cover changes and urban morphology changes on the urban heat island phenomenon. Sci. Total Environ. 635, 1467–1476. https://doi.org/10.1016/j.scitotenv.2018.04.091.

Zhu, W., Yuan, C., 2023. Urban heat health risk assessment in Singapore to support resilient urban design — by integrating urban heat and the distribution of the elderly population. Cities 132. https://doi.org/10.1016/j.cities.2022.104103.