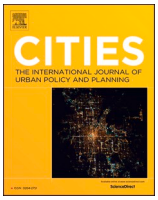
Cities 169 (2026) 106558

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities



Urban heatwaves and mortality: A socioeconomic and environmental study using a novel building heat vulnerability index

Jiwei Zou [a,b,c], Guowei Zhong [d,e], Liangzhu Leon Wang [a,*], Ali Katal [a], Abhishek Gaur [b], Shujie Yan [a], Maher Albettar [a], Ahmed Marey [a,b]

[a] Centre for Zero Energy Building Studies, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Canada
[b] Construction Research Centre, National Research Council Canada, 1200 Montreal Road, K1A 0R6, Ottawa, Ontario, Canada
[c] Department of Building and Real Estate, Hong Kong Polytechnic University, Hong Kong, China
[d] Department of Family Medicine, Faculty of Medicine and Health Sciences, McGill University, Quebec, Canada
[e] EvoClinical Inc., Montreal, Quebec, Canada

ARTICLE INFO

Keywords:
Heat vulnerable index
Heatwave
Overheating
Heat mortality
Logistic regression

ABSTRACT

Climate change has significantly increased the frequency, intensity, and magnitude of heatwaves, leading to numerous deaths in recent decades. While environmental parameters are known contributors to heat-related mortality, the specific impacts of socioeconomic factors remain less clear. This study introduces a new Building Heat Vulnerable Index (BHVI) to assess and map urban overheating mortality risk during heatwaves at the building level. Using data from the 2018 Montreal heatwave, we employed penalized logistic regression (PLR) to analyze the correlation between heat-related mortality and both environmental and socioeconomic parameters across Montreal. The City Building Energy Model (CityBEM) was used to simulate indoor overheating conditions, providing detailed exposure data. Socioeconomic variables were collected from Censusmapper and Geoportail Quebec. Our analysis revealed that "dwelling density" and "average income" are the most significant factors affecting heat-related mortality. Utilizing the BHVI, we generated a detailed heat vulnerability map identifying risk regions across Montreal, highlighting vulnerable areas with high dwelling density and low average income. Additionally, we thoroughly evaluated the impacts of increasing air conditioning (AC) capacity on mitigating heat vulnerability through bootstrap simulations. The results demonstrated that enhancing AC capacity significantly reduces heat-related mortality risk, particularly in high and critical risk areas. The findings underscore the importance of integrating socioeconomic factors and building-level data into heatwave mortality risk assessments. They suggest that targeted interventions, such as improving AC accessibility in vulnerable neighborhoods, can effectively mitigate heat-related health risks. This study provides valuable insights for policymakers to implement effective heat mitigation strategies in urban environments facing escalating climate challenges.

1. Introduction

The frequency of extreme weather events, such as heatwaves, floods, and droughts, has significantly increased in recent years due to global warming (Zou et al., n.d.-a; Change, 2001; Hamdy et al., 2017; Liu et al., 2021; Yau & Hasbi, 2013; Zou et al., 2022; Zou et al., 2023). These events exert substantial societal pressure on the economy, health, energy utilization, and infrastructure (Edenhofer, 2015; Yan, Wang, et al., 2022; Yan, Xiong, et al., 2022; Yang et al., 2023; Yau & Hasbi, 2013; Zou, Liu, et al., 2021; Zou, Yu, et al., 2021). A major impact of this trend is urban overheating, evidenced by more frequent and severe heatwaves (Zou et al., n.d.-b; Khan et al., 2021; Machard et al., 2024; Yang et al., 2024), which adversely affect citizens' lives and the environmental quality of cities (Mortezazadeh et al., 2022; Santamouris, 2020; Xie et al., 2021). Regions in cold climates are particularly vulnerable to heatwaves, often being unprepared for such events. For example, the 2003 European heatwave caused 30,000 deaths (Brücker, 2005; Luterbacher et al., 2004), and the 2021 heatwave in Canada resulted in 500 deaths (MPNRC, 2021). Projections of future climate indicate that by 2100, approximately half of the global population will be affected by

* Corresponding author at: S-EV 15105, 1455 de Maisonneuve Blvd. WestMontreal, Quebec, H3G1M8, Canada.
E-mail address: leon.wang@concordia.ca (L.L. Wang).

https://doi.org/10.1016/j.cities.2025.106558
Received 24 November 2024; Received in revised form 17 July 2025; Accepted 1 October 2025
Available online 10 October 2025

deadly heatwaves (Mora et al., 2017). Therefore, identifying heat-vulnerable communities and populations is crucial for implementing effective mitigation strategies (Zou et al., 2023). This involves evaluating factors contributing to heat-related deaths, considering both environmental and socioeconomic components, particularly among the poor, elderly, and minority groups in urban neighborhoods.

In recent years, multi-dimensional heat-related indicators, known as heat vulnerable index (HVI), have been proposed to account for climatic hazards and socioeconomic vulnerabilities (Niu et al., 2021) by a composite of socioeconomic status, physical environment, and climate-related factors (Cheng et al., 2021). These indicators can be divided into two conceptual frames: population vulnerability to environmental hazards (Cutter et al., 2003; Field & Barros, 2014) and risk triangle (Crichton, 1999; Field & Barros, 2014). For the former, vulnerability is calculated by the sum of sensitivity, exposure, and adaptive capacity. Here, sensitivity is defined as sociodemographic factors (Field & Barros, 2014), exposure refers to climatic and environmental factors (Zuhra et al., 2019), and adaptive capacity represents the accessibility of mitigation strategies (Mallen et al., 2019; Zuhra et al., 2019). The risk triangle states that risk is a function of hazard, exposure, and vulnerability, and these three elements must be spatially coincident for a risk to exist (Tomlinson et al., 2011). Hazard refers to severe extreme events (Hu et al., 2017), exposure represents the elements at risk (Hua et al., 2021), and vulnerability is the lack of material to mitigate impacts (Ho et al., 2015).

According to the literature on heat vulnerability (Bao et al., 2015; Inostroza et al., 2016; Johnson et al., 2012; Méndez-Lázaro et al., 2018; Nayak et al., 2018; Tomlinson et al., 2011), quantifying the severity of heat-related climate impacts is the key element of the exposure layer in population vulnerability or the hazard layer in the risk triangle. The outdoor environmental parameters are typically utilized, such as land surface temperature (Buscail et al., 2012; Zheng et al., 2020) from satellite images with a resolution of 60 m $\sim$ 1 km, daily and nighttime maximum and minimum air temperature (Hu et al., 2017; Kim et al., 2017) from weather stations, and daily or nighttime exceedance temperature (Hu et al., 2017; Kim et al., 2017). Besides, heat indices such as wet bulb global temperature (Zheng et al., 2020) and humidex (Aminipouri et al., 2016; Krstic et al., 2017) were also used to assess extreme heat impacts.

However, most heat-related exposure is likely to occur indoors as Canadians spend about 90 % of their time indoors (E. a. C. C. C. (ECCC), 2022). Based on our former study (Katal et al., n.d.), no strong spatial correlation was established between outdoor heat stress and mortality, and mortalities were often found due to the incapability of buildings to shelter occupants against harsh outdoor conditions (Hamdy et al., 2017; Hamdy & Hensen, 2015; Rosenthal et al., 2014; Uejio et al., 2011). Therefore, a better understanding of indoor heat vulnerabilities at the building level is necessary, whereas the conventional approach based on outdoor analysis falls short. This has been enabled by the recent advancement of Urban Building Energy Models (UBEMs).

Moreover, social and economic vulnerabilities have been reported to contribute significantly to heat-related morbidity and mortality during extreme heat events. The study by Johnson et al. (Johnson et al., 2005) specifically quantified the impact of the 2003 heat wave in England, reporting a 17 % increase in overall mortality with a staggering 59 % increase in the over 75 age group in London due to elevated temperatures combined with high ozone and particulate matter concentrations. Similarly, Stafoggia et al. (Stafoggia et al., 2006) analyzed multiple Italian cities using a case-crossover design, finding increased odds of mortality at higher temperatures, especially among women, the elderly, and those with psychiatric or cardiovascular conditions. They highlighted the compounded vulnerability of these groups under extreme heat conditions, suggesting that individual predispositions interact with socioeconomic factors to enhance risk.

Research by Santamouris and Kolokotsa (Santamouris & Kolokotsa, 2015) provided evidence of urban overheating impacting low-income populations by increasing energy consumption and degrading indoor environmental conditions, which disproportionately affects their health outcomes. Smoyer (Smoyer, 1998) reinforced these findings, indicating that heat-related mortality rates were generally higher in disadvantaged urban areas than in more affluent parts of cities, with spatial analysis pinpointing specific neighborhoods in St. Louis where these disparities were most pronounced. This place-based approach identified the combination of physical and social characteristics of locations that contributed to higher mortality rates during heat waves.

Despite the recognized importance of addressing urban overheating and its impacts on heat-related mortality, there remains a significant gap in evaluating heat vulnerability accounting for building-specific factors across an entire city. While the development of HVIs and UBEMs represents a significant advancement, previous studies have predominantly focused on broader district or city levels (Bell et al., 2008; Johnson et al., 2012; Reid et al., 2009; Rosenthal et al., 2014; Stafoggia et al., 2006; Uejio et al., 2011), with the availability of social-economic and climatic data at the same scale. This oversight limits our understanding of how building-specific factors, e.g., indoor overheating conditions, air conditioning accessibility, building age, value, and materials, contribute to indoor heat exposure and vulnerability. The motivation for emphasizing the building-scale factors is that people spend most of their lives indoors, and urban heat-related mortality is associated primarily with overheating conditions indoors, particularly in mild and cold climates. The ability to assess heat vulnerability comprehensively with building-scale factors encompassing the entire urban fabric is essential for pinpointing vulnerability hotspots and implementing precise, equitable interventions across diverse urban communities. The etiology of heat-related mortality is inherently complex and is influenced by an array of factors, including climatic conditions, socioeconomic vulnerabilities, and urban infrastructure. Traditional regression methods often fall short in handling the intricacies of these relationships due to limitations such as sparsity, non-linearity, and multicollinearity among predictors. Machine Learning (ML) techniques, such as random forests, present a robust alternative capable of addressing these challenges by uncovering non-linear relationships and interactions that traditional methods might miss. These techniques have been noted for their enhanced predictive accuracy, as evidenced in recent studies (Hirano et al., 2021). However, the interpretability of such models often remains limited, posing challenges for understanding the underlying etiological relationships and for the development of empirical indices. In contrast, methods like the Least Absolute Shrinkage and Selection Operator (LASSO) not only bolster predictive performance but also enhance model interpretability and prevent overfitting. The interpretability afforded by LASSO facilitates the creation of plasmode simulations, which use real data to define synthetic data-generating equations. This approach is crucial for evaluating assumptions under realistic conditions and for assessing the effectiveness of interventions, such as the promotion of air conditioning to mitigate heatwave mortality, thereby informing policy decisions (Bhatnagar et al., 2020; Greenwood et al., 2020). Despite these advancements, the application of ML models to explore the multifaceted impacts on heat-related mortality is still underrepresented in the literature. Furthermore, the considerable imbalance in data sets, such as the disparity between the few heat-related deaths and the numerous non-death cases during heat waves, significantly complicates the development of effective ML models for such evaluations (Krawczyk, 2016). Addressing these challenges through advanced ML methodologies could lead to more accurate and dynamic risk assessment tools, thus enabling policymakers and urban planners to better prioritize interventions in response to escalating climate challenges.

To bridge existing knowledge gaps and support more targeted urban heat mitigation planning, this study aims to address three interrelated scientific questions. First, it investigates the most influential environmental, socioeconomic, and building-specific factors contributing to heat-related mortality during the 2018 heatwave in Montreal. Second, it seeks to spatially identify and quantify the extent of heat vulnerability

across the city at the building level, thereby highlighting the urban areas most at risk. Third, the study evaluates the feasibility and effectiveness of potential mitigation strategies in reducing mortality risk during extreme heat events.

This study contributes a novel high-resolution vulnerability mapping method—the Building Heat Vulnerability Index (BHVI)—which integrates building-specific indoor overheating data with socioeconomic and environmental indicators using a penalized regression framework. Unlike traditional HVIs, this method leverages building-scale simulation and statistical modeling to enable fine-grained risk analysis. Although this study uses the 2018 Montreal heatwave as a case, the BHVI methodology is transferable to other urban settings, providing a replicable tool for policymakers and planners to evaluate and mitigate heat vulnerability at the building level.

In this paper, Section 2 describes our methodology, starting with the study area during the 2018 Montreal heatwave (Section 2.1), introducing the City Building Energy Model (CityBEM) simulations for indoor overheating data (Section 2.2), and detailing the socioeconomic variables (Section 2.3). This section also discusses the Building Heat Vulnerable Index (BHVI) that integrates diverse variables (Section 2.4) and explains the application of penalized logistic regression (PLR) models to analyze and predict heat-related mortality (Section 2.5). Section 3.1 evaluates the performance of our machine learning models, and Section 3.2 highlights the significant variables influencing heat-related mortality. Section 3.4 examines the spatial distribution of heat vulnerability in Montreal, identifying high-risk areas and discussing their socioeconomic and environmental characteristics. Section 3.4 analyzes the impacts of various air conditioning enhancement scenarios on heat risk mitigation, providing quantitative assessments of their potential to reduce vulnerabilities in identified high-risk areas.

2. Methodology

2.1. Study area and time period

Montreal, located on the Island of Montreal at the confluence of the Saint Lawrence and Ottawa Rivers, is distinguished by Mount Royal, a central, significant green space that serves as a natural landmark (Encyclopedia, 2022). The city encompasses multiple boroughs and neighborhoods, each with distinct characteristics. Notably, the Plateau Mont-Royal is recognized for its artistic flair and historic architecture, while Downtown serves as the economic hub with high-rise buildings, and Old Montreal is celebrated for its historic buildings and cobblestone streets. Home to over 1.7 million residents, making it the second-most

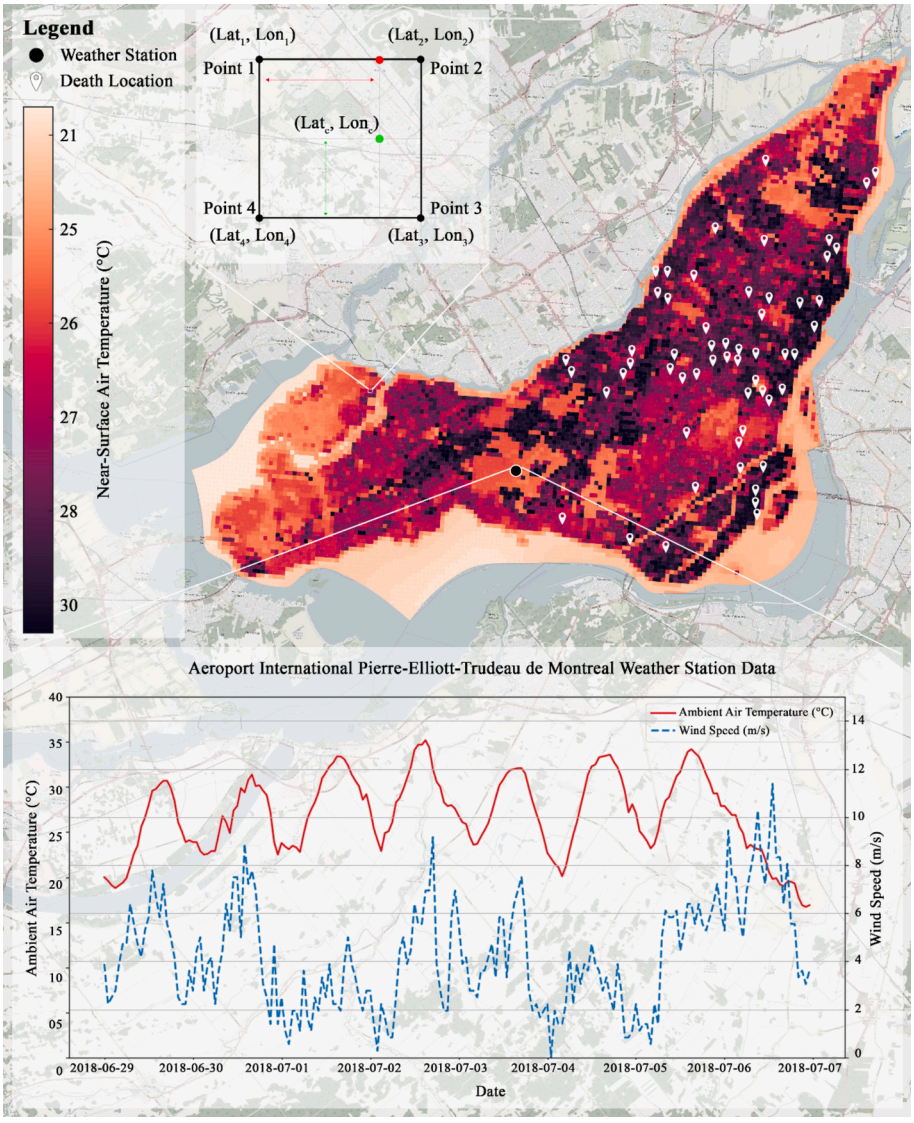


Fig. 1. Map of the Island of Montreal and heat-related death location (top), Ambient air temperature and wind speed weather station data (bottom), and a schematic of Surface Prediction System (SPS) data interpolation on the building center point (top left).

populous city in Canada, Montreal exhibits a wide range of income levels across neighborhoods. Affluent areas like Westmount and Outremont contrast sharply with the historically working-class neighborhood of Hochelaga-Maisonneuve. Housing within the city varies extensively, from high-rise apartments to single-family homes, reflecting its socioeconomic diversity.

The 2018 heatwave in Montreal, Canada with over 70 deaths in Quebec, was selected with the study period, from 00:00 EST June 30, 2018 to 00:00 EST July 5, 2018 (Oved, 2019; Press, 2018). In Fig. 1 (top), the white location symbols indicate the locations of heat-related mortality. Fig. 1 (bottom) also provides the weather conditions during the heatwave. The average temperature was around 28 °C with a peak of 35 °C and a low of 20 °C, and the heatwave was declared then for a daily maximum exceeding 31 °C on five consecutive days (Ding & Qian, 2011; Hemraj et al., 2020). Specifically, during the period from June 30 to July 8, 2018, extreme heat conditions led to 66 confirmed heat-related deaths on the Island of Montreal. The surveillance system's data collection by the Institut national de santé publique du Québec (INSPQ) focuses on mortality and morbidity associated with extreme heat events, providing valuable insights into the health impacts of such occurrences.

2.2. CityBEM simulations

2.2.1. CityBEM simulation coupling with surface prediction system (SPS)

Transient weather data serves as a crucial input dataset for CityBEM (Katal et al., 2022), employed to compute the thermal-energy performance of buildings. To bolster the precision of these simulations, it is imperative to utilize local climate data specific to each building. In this study, CityBEM is seamlessly integrated with SPS in a one-way configuration, with no feedback loop from CityBEM to SPS (Milbrandt et al., 2016). This process entails the extraction and subsequent interpolation of SPS data onto each building for CityBEM simulation. Key specifications include the city boundaries (latitude and longitude), along with all SPS grid points situated within the delineated domain.

In Fig. 1 (top left) Red circles represent the grid points of High-Resolution Deterministic Prediction System (HRDPS) weather data and the gray circle represents the SPS grid point (Kehler et al., 2016). Then, atmospheric elements required for building energy simulation (outdoor air temperature, solar radiation, wind speed, wind direction, and dew point temperature) are extracted for all SPS grid points inside the domain. Finally, four SPS grid points around each building are selected, and weather data are interpolated on the center of the building based on the distance between the building center point and grid points, assuming a linear interpolation [1].

The City Building Energy Model (CityBEM) is used to simulate indoor overheating degree hours during the 2018 heatwave for each building, capturing the physical exposure component of heat vulnerability. This exposure indicator, along with socioeconomic and building-related variables, forms the input dataset for the Penalized Logistic Regression (PLR) model. PLR is then used to quantify the relationship between these predictors and the probability of heat-related mortality. The trained PLR coefficients are subsequently applied to calculate the Building Heat Vulnerability Index (BHVI), ensuring that the index reflects statistically validated risk weights. Together, CityBEM and PLR operate within a unified framework—where CityBEM generates high-resolution exposure data, and PLR identifies and quantifies key mortality predictors—to support interpretable, spatially explicit vulnerability assessments at the building level.

2.2.2. Overheating hours

The severity of the indoor overheating risk in Montreal was evaluated by overheating degree hour (Lei et al., 2022; Porritt et al., 2012; Xie et al., 2022; Yaqubi et al., 2022) to better evaluate the severity of the overheating than using the number of overheating hours over threshold temperatures since it quantifies the extent of exceedance from the thresholds.

$$DH_{inOH} = \sum (T_h - T_s)H \quad (1)$$

where DH_{inOH} = overheating degree hour (°C·*hr*), T_h = the hourly indoor temperature (°C), T_s = the threshold temperature (°C), H = the time step of interest (*hr*), i.e., one hour in this study. Here, 26 °C (CIBSE, 2005; CIBSE, 2013; Goncalves et al., 2021; Zou et al., 2022) is chosen as the value of threshold temperature for the worst scenario in each building (Zou et al., n.d.-a; Zou et al., 2022).

2.3. Social and economic data

Social and economic vulnerabilities are significant subcomponents of BHVI (Bell et al., 2008; Cheng et al., 2021; Harlan et al., 2006; Rana et al., 2017; Reid et al., 2009). In this study, eight social economic subcomponents were considered. Normalized social-economic subcomponents extracted from Censusmapper, Geoportail Quebec, and CityBEM are shown in Fig. 2 (https://censusmapper.ca; https://www.inspq.qc.ca/geomatique/geoportail).

The population or **population density** is one of the most frequently mentioned subcomponents in the former heat vulnerability studies (Bao et al., 2015; Bell et al., 2008; Brücker, 2005; Harlan et al., 2006; Inostroza et al., 2016; Rosenthal et al., 2014; Stafoggia et al., 2006). It is normally defined as the density of inhabitants per living block or per square meter/km. The **density of the heat-vulnerable population** is quantified by assessing the prevalence of chronic diseases among individuals within a geographic area, typically measured in administrative divisions. Chronic diseases, such as cardiovascular conditions, diabetes, respiratory issues, and mental disorders, heighten individuals' susceptibility to heat (Kenny et al., 2010; Kjellstrom et al., 2010; Meade et al., 2020). High-density regions particularly face exacerbated risks under extreme heat events. Understanding the distribution of these vulnerable groups is crucial for implementing effective interventions and mitigating the impacts of heat waves. The second most considered component is the **average income** of the impoverished population (Cheng et al., 2021). It has been demonstrated that this subcomponent is associated with increased heat stress levels and heat-related mortality (Chen et al., 2022; Cheng et al., 2021; Oved, 2019), due to multiple reasons such as less affordability of mitigation methods and social services.

For the building aspect, according to Taylor et al. (Taylor et al., 2015), the **dwelling density** is one of the main factors of the population-attributable burden of heat death. Besides, it could be found that most heat-related death during the historical heatwave happened in the dwellings (Liu et al., 2017; Mitchell & Natarajan, 2019; Rinner et al., 2010). Normally, a house with high market value tends to equip with better heat stress resistance strategies (Harrington, 2014; Soriano, 2008). On the other hand, according to Hatvani-Kovacs et al. (Hatvani-Kovacs et al., 2018), the heat stress resistance of the building will significantly affect the **dwelling value** in the market. Thus, the high-value buildings will be less heat vulnerable than the low-value buildings during the heatwave. Moreover, the **building age** could be an important criterion to predict its ability to heat resilience as well as its accessibility to heat mitigation methods (Cheng et al., 2021). Older buildings tend to lack thermal insulation and AC systems, while the external wall of modern buildings could reflect more sunshine and absorb less heat (Laouadi et al., 2020). The **percentage of air conditioning (AC) distribution** across different areas is vital in assessing the adaptive capacity of communities to cope with heat waves. Areas with low AC penetration are generally more vulnerable to heat-related illnesses and mortality, particularly during prolonged periods of high temperatures (Peplinski et al., 2023; Sailor et al., 2019; Sarofim et al., 2016). Policymakers can use this information to prioritize infrastructure upgrades and subsidies for AC units in underserved communities, thereby reducing heat vulnerability.

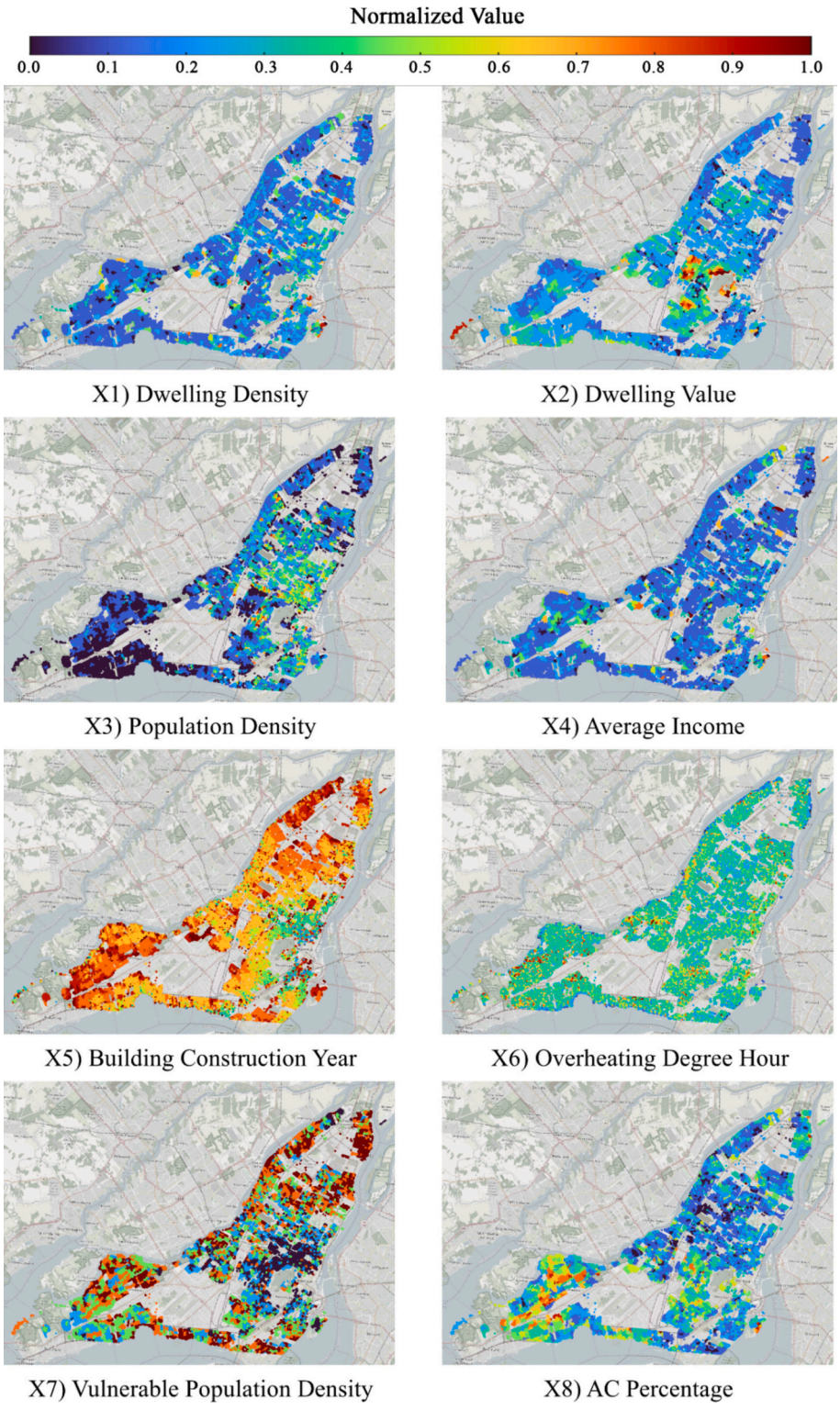


Fig. 2. Normalized subcomponents for building BHVI.

2.4. *Building heat vulnerable index (BHVI)*

The main risk assessment method of BHVI follows the concept of the risk triangle (Crichton, 1999; Field & Barros, 2014; Gwilliam et al., 2006; Pidgeon & Butler, 2009). Instead of assuming equal contributions from each risk component as in the traditional risk triangle (Crichton, 1999), the actual contributions are determined using machine learning methods. These contributions are then combined into the BVHI using the trained coefficients, as illustrated in Fig. 3.

Eq. 2 defines building heat vulnerable index (BHVI) of specific building n in the figure:

$$BHVI_n = \sum_{i=1}^{8} b_i X_{i,n} \tag{2}$$

where $X_{i,n}$ for the Min-Max normalized subcomponents X_i (Normalized

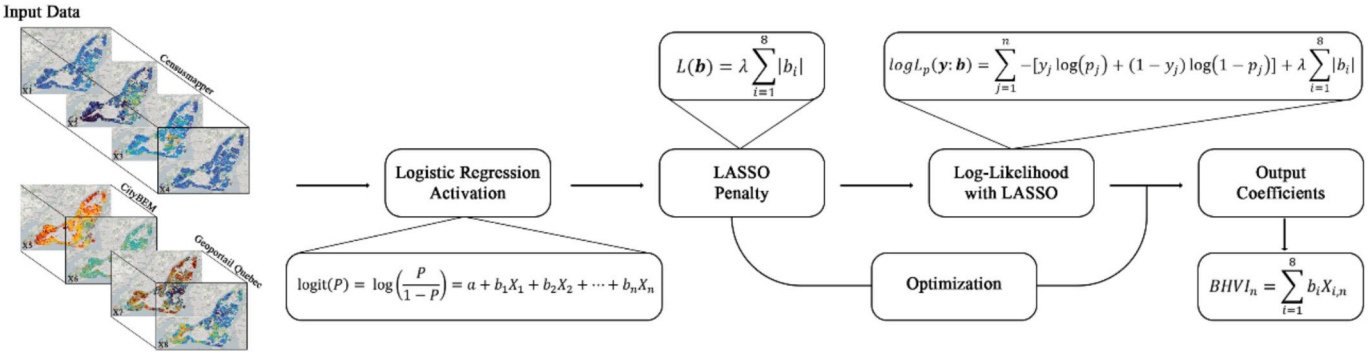


Fig. 3. Flowchart of calculating Building heat vulnerable index.

method is described in Section 2.5, information of X_i are shown in Fig. 2) for building n, b_i is the trained coefficient by PLR model for X_i. The overheating degree hour and building age were generated by CityBEM (Section 2.2), and the socioeconomic components were collected from Censusmapper and Geoportail Quebec (Section 2.3).

2.5. ML method for processing data

For the analysis and predictive modeling in this study, we employed the R Statistical Software (R. R Core Team, 2013), which facilitated the application of several advanced statistical techniques. For analyses that required PLR, the *glmnet* package was utilized, which implements regularization paths for generalized linear models as developed by Friedman et al. (Friedman et al., 2010).

2.5.1. Data pre-processing

The dataset was initially subjected to several pre-processing steps: Each variable was standardized using min-max standardization to allow for direct comparison of the magnitudes of the coefficients in the PLR model, and complete cases (i.e., all variables were non-missing) were retained for further analysis. Outliers, defined as observations lying more than plus or minus five standard deviations from the mean, were removed to enhance the generalizability and stability of the models. Although three standard deviations are commonly used for normally distributed data (per the empirical 68–95-99.7 rule), in high-dimensional and non-Gaussian datasets, especially where predictors span different scales and distributions, more stringent thresholds (e.g., ±5σ) have been adopted in prior literature to avoid loss of informative extreme values (Aggarwal, 2016; Iglewicz & Hoaglin, 1993; Singh & Upadhyaya, 2012).

The standardization formula applied was:

$$X_{i,n} = \frac{X_i - X_{i,min}}{X_{i,max} - X_{i,min}} \tag{3}$$

where $X_{i,n}$ is normalized subcomponents X_i, $X_{i,max}$ and $X_{i,min}$ are the maximum and minimum values of subcomponents X_i.

Additionally, the Variance Inflation Factor (VIF) was calculated to assess multicollinearity. The results of VIF are shown in Table 1.

Table 1
Results of VIF.

Variable	VIF
X1: Dwelling density	5.24
X2: Dwelling value	1.15
X3: Population density	1.33
X4: Average income	5.38
X5: Building construction year	1.19
X6: Overheating degree hour	1.03
X7: Density of heat-vulnerable population	1.28
X8: Percentage of AC distribution	1.08

Although high VIF values were noted for X1 and X4, variables were not removed due to the primary focus on prediction and the resilience of the chosen analytical method (PLR) to multicollinearity (Dormann et al., 2013; Neter et al., 1996).

2.5.2. Penalized logistic regression (PLR)

PLR was chosen over more complex machine-learning algorithms for three reasons. First, PLR simultaneously performs feature selection and yields easily interpretable coefficients, facilitating construction of the Building Heat Vulnerability Index (BHVI) and enabling direct policy insights. Second, regularization mitigates multicollinearity, which is common among socioeconomic and environmental covariates. Third, logistic models are less prone to overfitting on highly imbalanced data and require fewer resampling heuristics than ensemble or kernel methods.

$$\text{logit}(P) = \log\left(\frac{P}{1-P}\right) = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n \tag{4}$$

where, P is the probability of heat-related death occurrence for the building, a is the intercept (constant) term, b_i is the coefficient associated with the predictor variable X_i.

LASSO penalty is shown as

$$L(\boldsymbol{b}) = \lambda \sum_{i=1}^{8} |b_i| \tag{5}$$

where, $L(\boldsymbol{b})$ is the LASSO penalty, λ is the regularization parameter that controls the strength of the penalty (λ = in this study).

The log-likelihood of PLR use LASSO penalty (Eq. 5) is shown as follow:

$$logL_p(\boldsymbol{y}:\boldsymbol{b}) = \sum_{j=1}^{n} -\left[y_j \log\left(p_j\right) + \left(1-y_j\right)\log\left(1-p_j\right)\right] + \lambda \sum_{i=1}^{8} |b_i| \tag{6}$$

where $logL_p(\boldsymbol{y}:\boldsymbol{b})$is the log-likelihood of the penalized logistic regression model with the LASSO penalty, y_j is the actual binary outcome (1 if there is mortality, 0 if there is no mortality) for the *j-th* observation, p_j is the predicted probability of mortality occurring for the *j-th* observation, n is the number of observations.

The model performance is assessed by conducting a 10-fold cross-validation to ascertain the optimal penalty parameter λ, ensuring a balance between the model's complexity and its predictive accuracy. This cross-validation process was crucial for determining the appropriate level of regularization to avoid overfitting while maintaining a robust model performance. The threshold for decision-making was optimized based on the ROC curve, and both the sensitivity and specificity of the model were evaluated. The cutoff point determined by ROC curve is 0.4 for PLR, as shown in Fig. 4. This cutoff point was chosen to achieve a balance between sensitivity and specificity, as it yielded the highest Youden's Index during model evaluation (Fan et al., 2006). The

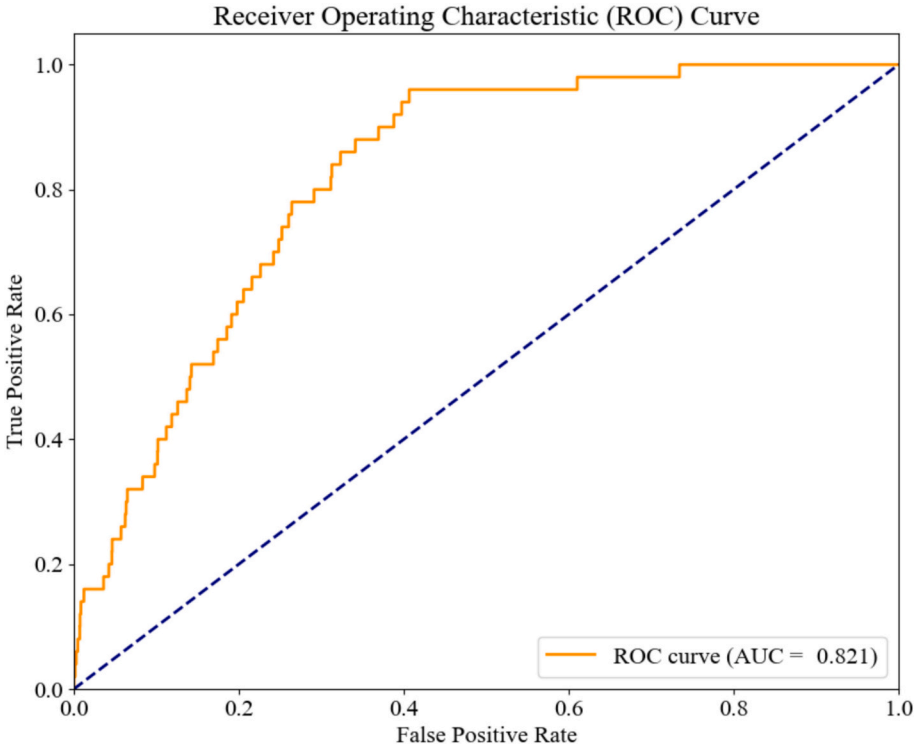


Fig. 4. Receiver operating characteristic (ROC) curve of PLR.

model's performance will be illustrated in Section 3.1.

3. Results and discussion

3.1. Performance of penalized logistic regression

The PLR was used to model the same dataset with k-fold (k = 10) cross-validation. The optimal threshold for classification determined based on ROC curve was found to be 0.39. The PLR model's performance metrics is assessed as shown in Table 3.

In the evaluation of the PLR, the model demonstrated a high sensitivity of 96 %, which effectively identified the most positive cases. However, specificity was lower at approximately 59 %, indicating a considerable rate of false positives. The Positive Predictive Value (PPV) was notably low at 0.202 %, suggesting that a large proportion of positive predictions were incorrect. Given the goal of predicting heatwave mortality, a less-than-optimal specificity and PPV do not pose a major concern. Conversely, the Negative Predictive Value (NPV) was extremely high at nearly 100 %, confirming the reliability of negative predictions. Despite these strengths, the model's overall accuracy, reflected by the Balanced Accuracy, was around 78. This performance indicates some limitations in the PLR approach for this dataset, primarily due to the high number of false positives.

Although PLR has relatively lower accuracy, it still holds significant advantages. The interpretability of its coefficients, which allows a detailed understanding of each predictor's impact on the outcome, is crucial in areas where elucidating the influence of variables is as important as predictive accuracy. This model inherently performs feature selection by applying an L1 penalty, which not only simplifies the model by eliminating non-contributive features but also enhances its generalizability. Additionally, the parsimony of logistic regression makes it easier to explain and validate, which is particularly important in settings where stakeholders require clear and understandable model mechanics. Finally, unlike more complex models, PLR facilitates statistical inference, enabling hypothesis testing and providing confidence intervals for estimates. These attributes justify its use in scenarios where the depth of insight into variable significance outweighs the need for the highest possible accuracy.

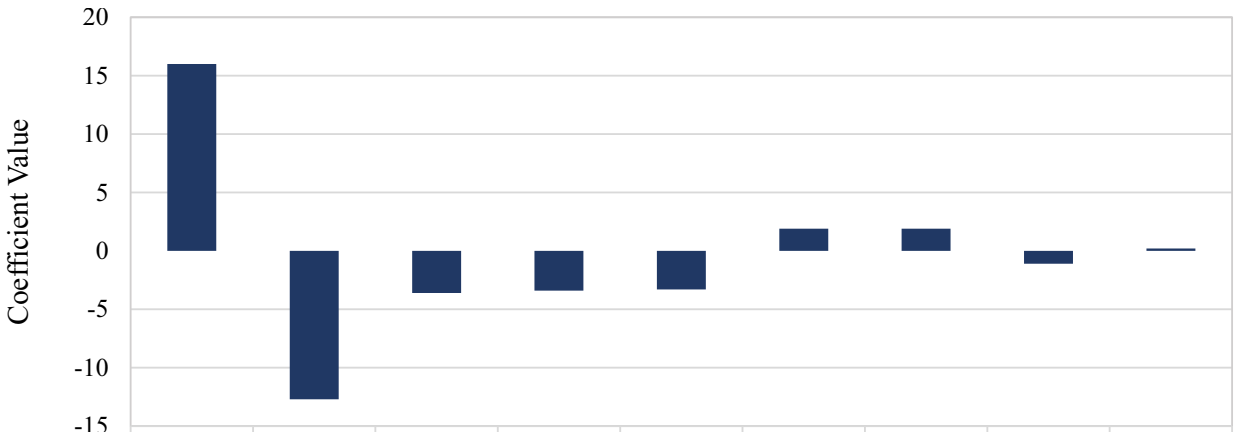
Table 3
Performance of PLR.

Index	Value
Sensitivity	0.96
Specificity	0.593401
Pos Pred Value	0.002024
Neg Pred Value	0.999942
Prevalence	0.000858
Detection Rate	0.000824
Detection Prevalence	0.407074
Balanced Accuracy	0.776701

3.2. Variable importance predicted by Penalized Logistic Regression

The results from the PLR offer a quantitative assessment of the impact of various predictors on the likelihood of outcomes in the dataset as shown in Fig. 5. Notably, the variable 'dwelling density' has a substantial positive coefficient of approximately 16.0 indicating a significant increase in the outcome's likelihood with each unit increase in 'dwelling density'. This suggests that factors related to housing density or quantity significantly elevate the probability of the outcome being studied. Conversely, 'Average income' presents substantial negative coefficient of −12.7 indicating that higher average incomes is associated with a reduced likelihood of the outcome. This reflects the potential mitigating effect of higher socioeconomic status on the risks associated with the outcomes.

Further analysis reveals that 'AC percentage' and 'Dwelling value' also have notable negative impacts on the outcome, with coefficients of −3.6 and − 3.3, respectively. These findings imply that improvements in living conditions, such as better-quality housing and increased access to air conditioning, significantly decrease the likelihood of the outcome. The variable 'Population' shows a positive coefficient of 1.9, which suggests that larger populations correlate with an increased probability



	Dwelling density	Average income	Percentage of AC distribution	Overheating degree hour	Dwelling value	Population density	Intercept	Building construction year	Density of heat vulnerable population
Parameter	16	-12.7	-3.6	-3.4	-3.3	1.9	1.9	-1.1	0.2

Fig. 5. Coefficient of social-economic parameters predicted by PLR.

of the outcome, possibly due to the amplification of risk factors in more densely populated areas. The model also indicates a slight negative trend over time with the 'year' variable showing a coefficient of −1.1, suggesting gradual improvements or adaptations over the years that reduce the likelihood of the outcome.

It should be noted that 'overheating degree hour' presents a negative coefficient of −3.4, indicating that a higher overheating degree hour leads to a lower probability of mortality occurrence, which is against the expected response. However, in the current study, all buildings were assumed to be naturally ventilated because the AC schedules and setting information were unavailable. So, the distribution of air conditioning was not implemented when calculating overheating hours, subject to strong uncertainties. Considering that the contribution of AC percentage is greater than the overheating degree hour, the analysis of building performance factors is therefore considered acceptable.

3.3. BHVI development and map of heat risk

BHVI is calculated through Eq. (2), with the values shown in Table 4. Then, the BHVI is normalized with Max-Min method, getting the following correlation:

$$\text{logit}(P) = \log\left(\frac{P}{1-P}\right) = -6.6 + 15.1nBHVI \tag{7}$$

The above equation could be transformed into:

$$P = \frac{1}{1+e^{-(-6.6+15.1nBHVI)}} \tag{8}$$

A correlation between nBHVI and probability of heat-related mortality occurrence, also known as relative heat risk (P) could then be shown in Fig. 6. Based on ROC curve of PLR, the optimal cut off point is determined as 0.4, shown as the red dash line inside the figure, which suggests that the building with predicted mortality probability larger than 0.4 is estimated to have mortality-risk by PLR, while the building

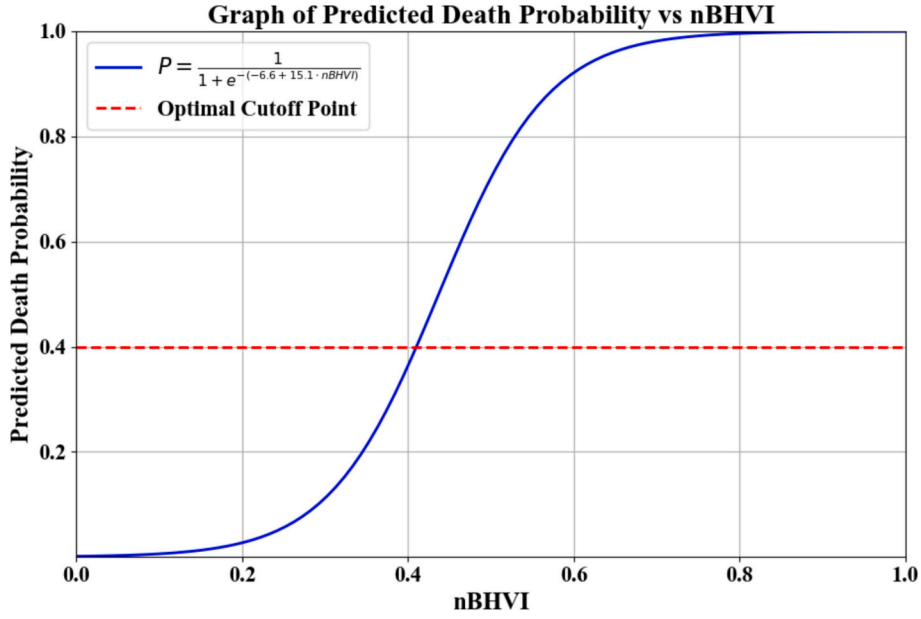


Fig. 6. Correlation between nBHVI and heat-related mortality probability predicted by PLR.

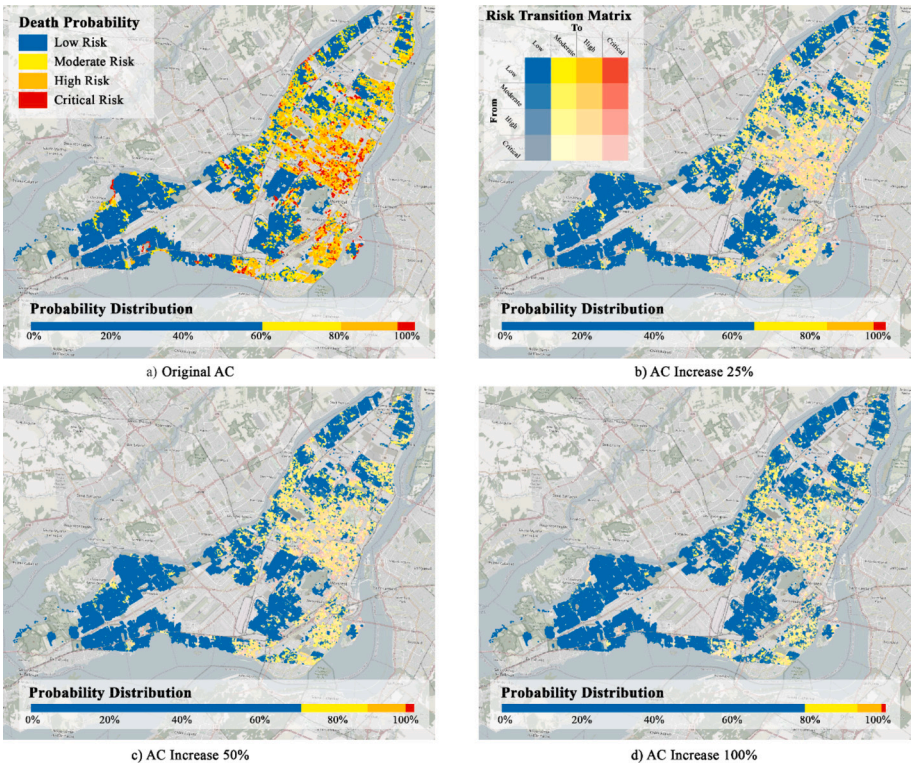


Fig. 7. Distribution (Probability distribution bar) of heat-related mortality risk with increased AC percentage and its corresponding change to risk level (Color map of Montreal in b, c, d). (a) Map of original death probability distribution without AC change. (b) Map of risk level change under 25 % AC increase; (c) Map of risk level change under 50 % AC increase; (d) Map of risk level change in 100 % AC increase.

with probability lower than 0.4 does not have risk of heat-related mortality.

Utilizing a cutoff point of 0.4 in the PLR model, the attached map (Fig. 7 a) delineates areas of varying heat risk. The legend categorizes the predicted mortality probabilities into four distinct heat risk levels: "Low risk" for probabilities ranging from 0 to 0.4, indicating minimal likelihood of heat-related mortality; "Moderate risk" for probabilities from 0.4 to 0.6, suggesting a low danger; "High risk" for probabilities from 0.6 to 0.8, reflecting a threat of heat-related mortality; and "Critical risk" for probabilities from 0.8 to1.0, signifying a high likelihood of adverse outcomes due to heat.

The central and eastern parts of Montreal, including densely populated boroughs such as Ville-Marie, Sud-Ouest, and Plateau-Mont-Royal, are identified as critical-risk areas (Red, probability 0.8–1.0). These zones are characterized by high urban density, which includes tall buildings, narrow streets, and minimal tree cover. This urban configuration exacerbates the urban heat island effect, where heat is trapped between buildings and hard surfaces like asphalt and concrete, significantly increasing temperatures. Additionally, industrial areas in boroughs like Saint-Léonard and Anjou also fall within this category due to their extensive impervious surfaces and heat emissions from industrial activities, which contribute further to local temperature increases.

High-risk areas (Orange, probability 0.6–0.8) are typically transitional zones that bridge densely populated urban centers with suburban outskirts. Neighborhoods such as Rosemont–La Petite-Patrie and Côte-des-Neiges are examples of high-risk areas. These regions possess a mix of residential and commercial developments with varying levels of green spaces and tree canopies. While these areas have better ventilation and less dense infrastructure than high-risk zones, they still retain enough urban characteristics to sustain elevated temperatures, albeit at a lesser intensity than the central boroughs.

The Moderate-risk areas (Yellow, probability 0.4 to 0.6) are predominantly found in suburban regions of Montreal like Pierrefonds-Roxboro and Pointe-Claire. These areas benefit from modern urban planning that often includes better building insulation standards, more prevalent air conditioning usage, and a substantial amount of green spaces and tree coverage. The urban sprawl in these areas is less intense, allowing for more open spaces and effective heat dissipation, which contributes to a cooler local climate compared to more urbanized regions.

Low-risk areas (Blue, probability 0 to 0.4) are primarily located on the outskirts of Montreal and near large bodies of water, such as L'Île-Bizard, Sainte-Anne-de-Bellevue, and along the shores of Verdun. These neighborhoods are characterized by low population density, extensive natural landscapes, and proximity to water bodies that provide significant cooling effects. The open spaces and abundant vegetation help in reducing the overall temperature, making these areas ideal for mitigating the urban heat island effect.

3.4. *Impacts of AC percentage increase on mitigating heat risk*

Enhancing air conditioning (AC) capacity within urban areas significantly affects the distribution of heat risk levels, thereby improving thermal comfort across different zones. To evaluate the impact of AC improvements on heat-related mortality risk, we employed a Predictive Logistic Regression (PLR) model. This model predicts the probability of heat-related mortality under various AC scenarios. To estimate the uncertainty associated with these predictions, we utilized a bootstrap simulation method. This method is advantageous as it does not make any assumptions about the underlying data distribution. Specifically, the bootstrap method involved conducting 1000 simulations, each drawing a 5 % subsample from the complete dataset using stratified sampling with replacement. This sampling approach was critical to maintain the original proportion of mortality and non-mortality cases in each subsample.

While the bootstrap method is robust for estimating empirical distributions without assuming a normal distribution, it has limitations. Firstly, the accuracy of bootstrap estimates can be compromised if the

original sample is not representative of the population, or if outliers heavily influence the results. Additionally, bootstrap methods can be computationally expensive, especially with a large number of simulations and complex models like PLR. These factors need to be considered when interpreting the results, as they might affect the generalizability and precision of the findings.

During each simulation iteration, three synthetic datasets were generated to represent potential real-world scenarios with incremental AC capacity increases of 25 %, 50 %, and 100 %. For instance, increasing a building's AC capacity by 50 % from an initial setting of 0.5 results in an adjusted AC level of 0.75. A maximum value of 100 % will be applied if the adjusted value is above 100 %. The impact of these adjustments on heat risk was then analyzed by computing 95 % confidence intervals for each risk category, both before and after the AC enhancements. This provided a quantifiable measure of statistical uncertainty.

The AC increases illustrated in Fig. 7 are relative to the original AC levels of each building as shown in Fig. 4 (X8). The data consistently reveal that increments in AC capacity leads to a significant rise in the areas classified as "Low risk" while concurrently reducing the extent of areas under higher risk categories. With the original AC settings (Fig. 7 a), a majority (60.35 %) of the area falls into the "Low risk" category, demonstrating that current AC levels are already mitigating heat-related risks to a substantial extent. However, as AC capacity is increased by 25 % (Fig. 7 b), the areas under "Low risk" rise to 65.71 %, indicating improved cooling efficacy. This enhancement in AC efficiency further reduces the "Moderate risk" and "High risk" categories by small margins and slightly lowers the "Critical risk" areas to 3.24 %. A more pronounced effect is seen with a 50 % increase in AC capacity (Fig. 7 c), where "Low risk" areas escalate to 70.54 %. Here, "High risk" and "Critical risk" categories experience more significant reductions, suggesting that moderate increases in AC are particularly effective at mitigating higher levels of thermal discomfort. The most dramatic shift is observed when AC capacity is doubled (100 % increase, Fig. 7 d), with "Low risk" areas covering 78.85 % of the surveyed zone. This scenario highlights a drastic reduction in all forms of heat risk, most notably with "Critical risk" areas plummeting to just 1.23 %.

The series of maps (Fig. 7) depicting various scenarios of AC capacity increases across Montreal offers a comprehensive view of how enhancements in cooling infrastructure impact urban heat risk distribution. Initially, under the original AC settings, critical-risk areas are predominantly located in densely populated and industrially active districts like Ville-Marie, Sud-Ouest, Plateau-Mont-Royal, Rosemont, and Mercier, characterized by high building density. These regions show a stark contrast in heat risk distribution compared to the peripheral and less densely populated areas, which exhibit lower risk levels.

As AC capacity is increased by 25 %, there is a noticeable expansion of "Low risk" zones with a general reduction of predicted death probability around 5 %, particularly in less urbanized areas such as Verdun and parts of Saint-Laurent, transitioning from moderate-risk statuses. This indicates that even modest increases in AC capacity can significantly enhance thermal comfort, especially in residential zones with lower building densities. Concurrently, central areas such as Ville-Marie and Sud-Ouest show a reduction in high-risk zones, suggesting improved conditions, though these areas continue to experience significant heat stress due to their urban makeup.

The impact of AC capacity increases becomes more pronounced with a 50 % enhancement. Critical and high-risk areas see a substantial decline with reduction of predicted death probability of 20 %, especially in transitional zones like Lachine and Lasalle, moving towards lower risk categories. This trend is more marked with a 100 % increase in AC capacity, where central zones around Montreal Downtown such as Ville-Marie, Mercier, and Rosemont witness a dramatic reduction in high-risk areas (with reduction up to 30–40 %), nearly eliminating them. Peripheral regions like Pierrefonds-Roxboro and Pointe-Claire largely transition to low-risk statuses, showcasing the potential of extensive AC deployment in mitigating urban heat risks effectively.

This progressive decrease in higher heat risk categories with increasing AC capacity underscores the critical role of adequate cooling systems in urban heat management, particularly under the escalating challenges posed by climate change. These findings advocate for policies that support the expansion and enhancement of AC systems, especially in densely populated or highly urbanized areas prone to severe heat stress. However, it is essential to balance this approach with considerations for energy consumption and environmental sustainability, potentially integrating renewable energy sources and improving building insulation to offset the increased energy demands of higher AC usage.

3.5. Discussion

Our study's building-level analysis of heat-related mortality risk in Montreal shares common ground with several recent works on urban heat vulnerability but extends them through finer spatial resolution and integration of indoor overheating data. For instance, Iqbal et al. (Iqbal et al., 2023) assessed human heat vulnerability across different Local Climate Zones (LCZs) in Lahore, highlighting lower vulnerability in compact urban zones due to better socioeconomic conditions and access to services. Similarly, our study identified "dwelling density" and "average income" as the two most influential factors affecting heat-related mortality. While compactness in Iqbal et al. (Iqbal et al., 2023) corresponded to reduced vulnerability due to better infrastructure, our findings show that high dwelling density in socioeconomically disadvantaged neighborhoods can amplify risk—underscoring that compactness alone does not determine vulnerability unless coupled with equitable resource access.

Other studies have focused on environmental mitigation strategies such as urban greening. Analyzed the cooling effect of urban parks in Xi'an and found that factors like park area, vegetation, and surrounding building density significantly influenced temperature reduction. While our study does not directly examine urban greenery, the spatial patterns identified in our Building Heat Vulnerability Index (BHVI) can inform city planners where green interventions, such as parks or vegetated buffers, may be most needed to alleviate heat risks. Similarly, the work of Hui et al. (2024), which showed that building-integrated vegetation can reduce aerodynamic wind loads and surface pressures, supports the broader implication from our findings that building-level characteristics should be a focus for future design-based heat mitigation strategies.

From a methodological standpoint, our approach also aligns with recent efforts to combine socioeconomic, environmental, and spatial data for localized vulnerability assessment. Rana et al. (2022) developed a heatwave vulnerability index for formal and informal settlements in Lahore and found informal communities to be more at risk due to lower coping capacity. This mirrors our finding that lower-income areas in Montreal show elevated mortality risk. Moreover, Nassar et al. (2024) employed GIS-integrated modeling to assess regenerative strategies for slums in Egypt, while Rahman (2024) applied machine learning to predict climatic trends like evapotranspiration. These studies reinforce the growing importance of using integrated, data-driven approaches for climate risk assessment. Our use of penalized logistic regression (PLR) adds to this toolkit by providing both interpretability and predictive capacity, particularly valuable in data-scarce settings. Altogether, our study contributes a scalable, high-resolution framework that combines statistical modeling with spatial mapping to inform localized and socially equitable heat mitigation strategies.

4. Conclusions

This study critically addresses urban heat vulnerability by focusing on three core objectives: identifying key climatic and socioeconomic variables contributing to heat-related mortality, pinpointing potential heat-vulnerable regions in Montreal during heatwaves, and assessing the impact of AC enhancements on mitigating heat risks.

Firstly, a machine learning model, PLR, is utilized in this study to analyze and quantify the influence of various climatic and socioeconomic factors on heat-related mortality. According to PLR, the variable 'dwelling density' has a substantial positive coefficient of approximately 15.995, indicating a significant increase in the outcome's likelihood with each unit increase in 'dwelling density'. Conversely, 'Average income' presents substantial negative coefficient of −12.735, indicating that higher average incomes is associated with a reduced likelihood of the outcome. This reflects the potential mitigating effect of higher socioeconomic status on the risks associated with the outcomes.

Secondly, the study identified and mapped heat-vulnerable regions across Montreal by employing predictions from our machine-learning models. These predictions enabled us to delineate areas with varying levels of heat risk, from critical to low risk, based on a composite index developed from the analyzed variables. This spatial analysis revealed critical hotspots, particularly in densely populated or socioeconomically disadvantaged areas, where interventions might be most needed to mitigate the effects of heatwaves.

Thirdly, the impact of AC modifications on heat vulnerability was thoroughly evaluated. Our findings demonstrated that increasing AC capacity significantly lowers the risk of heat-related mortality, with more pronounced benefits observed in high-risk areas. This was evidenced by a marked decrease in areas categorized as critical risk when AC capacity was increased, suggesting that strategic enhancements in cooling infrastructure could be a vital component of urban planning and public health initiatives aimed at reducing heat stress during extreme temperature events.

In conclusion, this study provides a comprehensive framework for understanding and addressing heat vulnerability in urban settings. By integrating climatic data, socioeconomic factors, and building-specific characteristics within a machine-learning context, we offer actionable insights for urban planners and policymakers. These insights facilitate targeted interventions, such as the optimization of AC systems and the development of cooling centers, particularly in identified vulnerable regions. While the study establishes a strong foundation, future research should aim to incorporate real-time environmental monitoring and consider the heterogeneity of AC distribution to refine the predictions further. This ongoing research will be crucial in enhancing urban resilience against the growing threat of heatwaves in the face of climate change.

CRediT authorship contribution statement

Jiwei Zou: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Guowei Zhong:** Writing – review & editing, Validation, Software, Methodology, Formal analysis, Conceptualization. **Liangzhu Leon Wang:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Ali Katal:** Methodology, Data curation, Conceptualization. **Abhishek Gaur:** Writing – review & editing, Supervision, Methodology, Investigation, Funding acquisition. **Shujie Yan:** Writing – review & editing, Software, Methodology, Formal analysis, Conceptualization. **Maher Albettar:** Software, Resources, Data curation. **Ahmed Marey:** Writing – review & editing, Visualization, Software, Formal analysis.

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.

Acknowledgment

This work was supported by "Climate Resilient Built Environment" Initiative to National Research Council of Canada (NRC) from Infrastructure Canada, Natural Sciences and Engineering Research Council (NSERC) of Canada through the Discovery Grants Program [#RGPIN-2018-06734], and the Advancing Climate Change Science in Canada Program [#ACCPJ 535986-18]; and Canada First Research Excellence Fund (CFREF) [IMPACT Project - Transforming Built and Urban Microclimates: Advancing Resilience Science for Vulnerable Populations in a Decarbonized and Electrified Canada].

Data availability

Data will be made available on request.

References

Aggarwal, C. C. (2016). An introduction to outlier analysis. In *Outlier analysis* (pp. 1–34). Springer.
Aminipouri, M., Knudby, A., & Ho, H. C. (2016). Using multiple disparate data sources to map heat vulnerability: Vancouver case study. *Canadian Geographies / Géographies canadiennes, 60*(3), 356–368.
Bao, J., Li, X., & Yu, C. (2015). The construction and validation of the heat vulnerability index, a review. *International Journal of Environmental Research and Public Health, 12*(7), 7220–7234.
Bell, M. L., O'neill, M. S., Ranjit, N., Borja-Aburto, V. H., Cifuentes, L. A., & Gouveia, N. C. (2008). Vulnerability to heat-related mortality in Latin America: A case-crossover study in Sao Paulo, Brazil, Santiago, Chile and Mexico City, Mexico. *International Journal of Epidemiology, 37*(4), 796–804.
Bhatnagar, S. R., et al. (2020). Simultaneous SNP selection and adjustment for population structure in high dimensional prediction models. *PLoS Genetics, 16*(5), Article e1008766.
Brücker, G. (2005). Vulnerable populations: lessons learnt from the summer 2003 heat waves in Europe. *Eurosurveillance, 10*(7). pp. 1-2%P 551.
Buscail, C., Upegui, E., & Viel, J.-F. (2012). Mapping heatwave health risk at the community level for public health action. *International Journal of Health Geographics, 11*(1), 1–9.
Change, I. P. O. C. (2001). *Climate change 2007: Impacts, adaptation and vulnerability*. Suíça: Genebra.
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 Climate, 41*, Article 101054.
Cheng, W., Li, D., Liu, Z., & Brown, R. D. (2021). Approaches for identifying heat-vulnerable populations and locations: A systematic review. *Science of the Total Environment, 799*, Article 149417.
CIBSE. (2005). *TM36: Climate change and the internal environment*.
CIBSE. (2013). The limits of thermal comfort: avoiding overheating in European buildings. In *TM52*. Chartered Institute of British Service Engineers.
Crichton, D. (1999). The risk triangle. *Natural disaster management, 102*(3).
Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly, 84*(2), 242–261.
Ding, T., & Qian, W. (2011). Geographical patterns and temporal variations of regional dry and wet heatwave events in China during 1960–2008. *Advances in Atmospheric Sciences, 28*(2), 322–337.
Dormann, C. F., et al. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27–46.
E. a. C. C. C. (ECCC). (2022). Indoor air quality. Available: https://www.canada.ca/en/environment-climate-change/campaigns/canadian-environment-week/clean-air-day/indoor-quality.html.
Edenhofer, O. (2015). *Climate change 2014: Mitigation of climate change*. Cambridge University Press.
Encyclopedia, N. W. (2022). Montreal. Available: https://www.newworldencyclopedia.org/entry/Montreal.
Fan, J., Upadhye, S., & Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine, 8*(1), 19–20.
Field, C. B., & Barros, V. R. (2014). *Climate change 2014–Impacts, adaptation and vulnerability: Regional aspects*. Cambridge University Press.
Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1.
Goncalves, V., Ogunjimi, Y., & Heo, Y. (2021). Scrutinizing modeling and analysis methods for evaluating overheating risks in passive houses. *Energy and Buildings, 234*, Article 110701.
Greenwood, C. J., et al. (2020). A comparison of penalised regression methods for informing the selection of predictive markers. *PLoS One, 15*(11), Article e0242730.
Gwilliam, J., Fedeski, M., Lindley, S., Theuray, N., & Handley, J. (2006). Methods for assessing risk from climate hazards in urban areas. In *, vol. 159, no. 4. Proceedings of the Institution of Civil Engineers-Municipal Engineer* (pp. 245–255). Thomas Telford Ltd.
Hamdy, M., Carlucci, S., Hoes, P.-J., & Hensen, J. L. (2017). The impact of climate change on the overheating risk in dwellings—A Dutch case study. *Building and Environment, 122*, 307–323.
Hamdy, M., & Hensen, J. (2015). Assessment of overheating risk in dwellings. In *Proceedings of the conference on healthy building Europe 2015, May 18-20th 2015* (pp. 1–9). The Netherlands: Eindhoven.

Harlan, S. L., Brazel, A. J., Prashad, L., Stefanov, W. L., & Larsen, L. (2006). Neighborhood microclimates and vulnerability to heat stress. *Social Science & Medicine, 63*(11), 2847–2863.
Harrington, P. (2014). *National energy efficient building project*. South Australia, Adelaide: Department of State Development.
Hatvani-Kovacs, G., Bush, J., Sharifi, E., & Boland, J. (2018). Policy recommendations to increase urban heat stress resilience. *Urban Climate, 25*, 51–63.
Hemraj, D. A., Posnett, N. C., Minuti, J. J., Firth, L. B., & Russell, B. D. (2020). Survived but not safe: marine heatwave hinders metabolism in two gastropod survivors. *Marine Environmental Research, 162*, Article 105117.
Hirano, Y., Kondo, Y., Sueyoshi, K., Okamoto, K., & Tanaka, H. (2021). Early outcome prediction for out-of-hospital cardiac arrest with initial shockable rhythm using machine learning models. *Resuscitation, 158*, 49–56.
Ho, H. C., Knudby, A., & Huang, W. (2015). A spatial framework to map heat health risks at multiple scales. *International Journal of Environmental Research and Public Health, 12*(12), 16110–16123.
Hu, K., Yang, X., Zhong, J., Fei, F., & Qi, J. (2017). Spatially explicit mapping of heat health risk utilizing environmental and socioeconomic data. *Environmental Science & Technology, 51*(3), 1498–1507.
Hua, J., Zhang, X., Ren, C., Shi, Y., & Lee, T.-C. (2021). Spatiotemporal assessment of extreme heat risk for high-density cities: A case study of Hong Kong from 2006 to 2016. *Sustainable Cities and Society, 64*, Article 102507.
Hui, Y., Tang, Y., Yang, Q., & Mochida, A. (2024). Numerical study on influence of surface vegetation on aerodynamics of high-rise buildings. *Sustainable Cities and Society, 107*, Article 105407.
Iglewicz, B., & Hoaglin, D. C. (1993). *Volume 16: How to detect and handle outliers*. Quality Press.
Inostroza, L., Palme, M., & de la Barrera, F. (2016). A heat vulnerability index: Spatial patterns of exposure, sensitivity and adaptive capacity for Santiago de Chile. *PLoS One, 11*(9), Article e0162464.
Iqbal, N., Aslam, A., Jamshed, A., Ravan, M., Birkmann, J., & Rana, I. A. (2023). Assessment of human heat vulnerability of different local climate zones in Lahore coupling remote sensing and socioeconomic data. In *2023 Joint Urban Remote Sensing Event (JURSE)* (pp. 1–4).
Johnson, D. P., Stanforth, A., Lulla, V., & Luber, G. (2012). Developing an applied extreme heat vulnerability index utilizing socioeconomic and environmental data. *Applied Geography, 35*(1–2), 23–31.
Johnson, H., et al. (2005). The impact of the 2003 heat wave on mortality and hospital admissions in England. *Health Statistics Quarterly, 25*, 6–11.
Katal, A., Mortezazadeh, M., Wang, L. L., & Yu, H. (2022). Urban building energy and microclimate modeling–From 3D city generation to dynamic simulations. *Energy, 251*, Article 123817.
A. Katal *et al.*, "Outdoor heat stress assessment using a multi-scale numerical weather prediction system: A case study of a heatwave in Montreal," Available at SSRN 4255607.
Kehler, S., Hanesiak, J., Curry, M., Sills, D., & Taylor, N. (2016). High resolution deterministic prediction system (HRDPS) simulations of Manitoba Lake breezes. *Atmosphere-Ocean, 54*(2), 93–107.
Kenny, G. P., Yardley, J., Brown, C., Sigal, R. J., & Jay, O. (2010). Heat stress in older individuals and patients with common chronic diseases. *Cmaj, 182*(10), 1053–1060.
Khan, H. S., Santamouris, M., Paolini, R., Caccetta, P., & Kassomenos, P. (2021). Analyzing the local and climatic conditions affecting the urban overheating magnitude during the heatwaves (HWs) in a coastal city: A case study of the greater Sydney region. *Science of the Total Environment, 755*, Article 142515.
Kim, D.-W., Deo, R. C., Lee, J.-S., & Yeom, J.-M. (2017). Mapping heatwave vulnerability in Korea. *Natural Hazards, 89*(1), 35–55.
Kjellstrom, T., Butler, A. J., Lucas, R. M., & Bonita, R. (2010). Public health impact of global heating due to climate change: Potential effects on chronic non-communicable diseases. *International Journal of Public Health, 55*, 97–103.
Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence, 5*(4), 221–232.
Krstic, N., Yuchi, W., Ho, H. C., Walker, B. B., Knudby, A. J., & Henderson, S. B. (2017). The heat exposure integrated deprivation index (HEIDI): A data-driven approach to quantifying neighborhood risk during extreme hot weather. *Environment International, 109*, 42–52.
Laouadi, A., Bartko, M., & Lacasse, M. (2020). A new methodology of evaluation of overheating in buildings. *Energy and Buildings, 226*, Article 110360.
Lei, M., van Hooff, T., Blocken, B., & Pereira Roders, A. (2022). The predicted effect of climate change on indoor overheating of heritage apartments in two different Chinese climate zones. *Indoor and Built Environment, 31*(7), 1986–2006, 1420326X221085861.
Liu, C., Kershaw, T., Fosas, D., Gonzalez, A. R., Natarajan, S., & Coley, D. A. (2017). High resolution mapping of overheating and mortality risk. *Building and Environment, 122*, 1–14.
Liu, S., Kwok, Y.-T., Lau, K., & Ng, E. (2021). Applicability of different extreme weather datasets for assessing indoor overheating risks of residential buildings in a subtropical high-density city. *Building and Environment, 194*, Article 107711.
Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., & Wanner, H. (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science (New York, N.Y.), 303*, 1499–1503, 04/01.
Machard, A., et al. (2024). Typical and extreme weather datasets for studying the resilience of buildings to climate change and heatwaves. *Scientific Data, 11*(1), 531.
Mallen, E., Stone, B., & Lanza, K. (2019). A methodological assessment of extreme heat mortality modeling and heat vulnerability mapping in Dallas, Texas. *Urban Climate, 30*, Article 100528.
Meade, R. D., et al. (2020). Physiological factors characterizing heat-vulnerable older adults: A narrative review. *Environment International, 144*, Article 105909.
Méndez-Lázaro, P., Muller-Karger, F. E., Otis, D., McCarthy, M. J., & Rodríguez, E. (2018). A heat vulnerability index to improve urban public health management in San Juan, Puerto Rico. *International Journal of Biometeorology, 62*(5), 709–722.
Milbrandt, J. A., Bélair, S., Faucher, M., Vallée, M., Carrera, M. L., & Glazer, A. (2016). The pan-Canadian high resolution (2.5 km) deterministic prediction system. *Weather and Forecasting, 31*(6), 1791–1816.
Mitchell, R., & Natarajan, S. (2019). Overheating risk in Passivhaus dwellings. *Building Services Engineering Research and Technology, 40*(4), 446–469.
Mora, C., et al. (2017). Global risk of deadly heat. *Nature Climate Change, 7*(7), 501–506.
Mortezazadeh, M., Zou, J., Hosseini, M., Yang, S., & Wang, L. (2022). Estimating urban wind speeds and wind power potentials based on machine learning with city fast fluid dynamics training data. *Atmosphere, 13*(2), 214.
MPNRC. (2021). Canada heat wave 2021 temperature, deaths, reason, map. Available: https://www.mpnrc.org/canada-heat-wave-2021-temperature-deaths-reason-map/.
Nassar, E. T., Elgazouly, H. G., Elnaggar, A. M., & Ayyad, S. M. (2024). Exploring climate change tackling in slums integrating ArcGIS data management and slums regenerative development model: case study of El Kharga city–Egypt. *Geology, Ecology, and Landscapes*, 1–11.
Nayak, S. G., et al. (2018). Development of a heat vulnerability index for New York State. *Public Health, 161*, 127–137.
Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models*.
Niu, Y., et al. (2021). A systematic review of the development and validation of the heat vulnerability index: Major factors, methods, and spatial units. *Current Climate Change Reports, 7*(3), 87–97.
Oved, M. C. (2019). Life and death under the dome. Available: https://projects.thestar.com/climate-change-canada/quebec/.
Peplinski, M., Kalmus, P., & Sanders, K. T. (2023). Investigating whether the inclusion of humid heat metrics improves estimates of AC penetration rates: A case study of Southern California. *Environmental Research Letters, 18*(10), Article 104054.
Pidgeon, N., & Butler, C. (2009). Risk analysis and climate change. *Environmental Politics, 18*(5), 670–688.
Porritt, S. M., Cropper, P. C., Shao, L., & Goodier, C. I. (2012). Ranking of interventions to reduce dwelling overheating during heat waves. *Energy and Buildings, 55*, 16–27.
Press, T. C. (2018). Quebec says up to 70 people may have died in connection with heat wave. Available: https://www.cbc.ca/news/canada/montreal/heat-wave-death-toll-1.4740031.
R. R Core Team. (2013). *R: A language and environment for statistical computing*.
Rahman, M. N. (2024). Seasonal and annual trends in reference evapotranspiration and prediction using machine learning models across seven climatic zones of Bangladesh. *Geology, Ecology, and Landscapes*, 1–16.
Rana, I. A., Bhatti, S. S., & Saqib, S. E. (2017). The spatial and temporal dynamics of infrastructure development disparity–From assessment to analyses. *Cities, 63*, 20–32.
Rana, I. A., Sikander, L., Khalid, Z., Nawaz, A., Najam, F. A., Khan, S. U., & Aslam, A. (2022). A localized index-based approach to assess heatwave vulnerability and climate change adaptation strategies: A case study of formal and informal settlements of Lahore, Pakistan. *Environmental Impact Assessment Review, 96*, Article 106820.
Reid, C. E., et al. (2009). Mapping community determinants of heat vulnerability. *Environmental Health Perspectives, 117*(11), 1730–1736.
Rinner, C., Patychuk, D., Bassil, K., Nasr, S., Gower, S., & Campbell, M. (2010). The role of maps in neighborhood-level heat vulnerability assessment for the city of Toronto. *Cartography and Geographic Information Science, 37*(1), 31–44.
Rosenthal, J. K., Kinney, P. L., & Metzger, K. B. (2014). Intra-urban vulnerability to heat-related mortality in New York City, 1997–2006. *Health & Place, 30*, 45–60.
Sailor, D. J., Baniassadi, A., O'Lenick, C. R., & Wilhelmi, O. V. (2019). The growing threat of heat disasters. *Environmental Research Letters, 14*(5), Article 054006.
Santamouris, M. (2020). Recent progress on urban overheating and heat island research. Integrated assessment of the energy, environmental, vulnerability and health impact. Synergies with the global climate change. *Energy and Buildings, 207*, Article 109482.
Santamouris, M., & Kolokotsa, D. (2015). On the impact of urban overheating and extreme climatic conditions on housing, energy, comfort and environmental quality of vulnerable population in Europe. *Energy and Buildings, 98*, 125–133.
Sarofim, M. C., et al. (2016). Temperature-related death and illness. In *US Global Change Research Program*.
Singh, K., & Upadhyaya, S. (2012). Outlier detection: applications and techniques. *International Journal of Computer Science Issues (IJCSI), 9*(1), 307.
Smoyer, K. E. (1998). Putting risk in its place: methodological considerations for investigating extreme event health risk. *Social Science & Medicine, 47*(11), 1809–1824.
Soriano, F. (2008). *Energy efficiency rating and house price in the ACT. Modelling the relationship of energy efficiency attributes to house price: the case of detached houses sold in the Australian Capital Territory in 2005 and 2006*. Canberra: Department of the Environment, Water, Heritage and the Arts.
Stafoggia, M., et al. (2006). Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. *Epidemiology*, 315–323.
Taylor, J., et al. (2015). Mapping the effects of urban heat island, housing, and age on excess heat-related mortality in London. *Urban Climate, 14*, 517–528.
Tomlinson, C. J., Chapman, L., Thornes, J. E., & Baker, C. J. (2011). Including the urban heat island in spatial heat health risk assessment strategies: a case study for Birmingham, UK. *International Journal of Health Geographics, 10*(1), 1–14.
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*(2), 498–507.
Xie, X., Luo, Z., Grimmond, S., Sun, T., & Morrison, W. (2022). Impact of inter-building longwave radiative exchanges on building energy performance and indoor overheating. *Building and Environment, 209*, Article 108628.
Xie, Z., et al. (2021). A field study on summertime overheating of six schools in Montreal Canada. In *, vol. 2069, no. 1. Journal of Physics: Conference Series* (p. 012168). IOP Publishing.
Yan, S., Wang, L. L., Birnkrant, M. J., Zhai, J., & Miller, S. L. (2022). Evaluating SARS-CoV-2 airborne quanta transmission and exposure risk in a mechanically ventilated multizone office building. *Building and Environment, 219*, Article 109184.
Yan, S., Xiong, J., Kim, J., & de Dear, R. (2022). Adapting the two-node model to evaluate sleeping thermal environments. *Building and Environment, 222*, Article 109417.
Yang, S., Wang, L. L., Stathopoulos, T., & Marey, A. M. (2023). Urban microclimate and its impact on built environment–a review. *Building and Environment, 238*, Article 110334.
Yang, S., Zhan, D., Stathopoulos, T., Zou, J., Shu, C., & Wang, L. L. (2024). Urban microclimate prediction based on weather station data and artificial neural network. *Energy and Buildings, 314*, Article 114283.
Yaqubi, O., Rodler, A., Guernouti, S., & Musy, M. (2022). Creation and application of future typical weather files in the evaluation of indoor overheating in free-floating buildings. *Building and Environment, 216*, Article 109059.
Yau, Y., & Hasbi, S. (2013). A review of climate change impacts on commercial buildings and their technical services in the tropics. *Renewable and Sustainable Energy Reviews, 18*, 430–441.
Zheng, M., Zhang, J., Shi, L., Zhang, D., Pangali Sharma, T. P., & Prodhan, F. A. (2020). Mapping heat-related risks in northern Jiangxi province of China based on two spatial assessment frameworks approaches. *International Journal of Environmental Research and Public Health, 17*(18), 6584.
J. Zou, A. Gaur, and L. L. Wang, A method of selecting future reference years for indoor overheating assessment.
Zou, J., Gaur, A., Wang, L. L., Laouadi, A., & Lacasse, M. (2022). Assessment of future overheating conditions in Canadian cities using a reference year selection method. *Building and Environment, 218*, Article 109102.
Zou, J., Liu, J., Niu, J., Yu, Y., & Lei, C. (2021). Convective heat loss from computational thermal manikin subject to outdoor wind environments. *Building and Environment, 188*, Article 107469.
Zou, J., Lu, H., Shu, C., Ji, L., Gaur, A., & Wang, L. L. (2023). Multiscale numerical assessment of urban overheating under climate projections: A review. *Urban Climate, 49*, Article 101551.
Zou, J., Yu, Y., Liu, J., Niu, J., Chauhan, K., & Lei, C. (2021). Field measurement of the urban pedestrian level wind turbulence. *Building and Environment, 194*, Article 107713.
J. Zou, Y. Yu, M. Mortezazadeh, H. Lu, A. Gaur, and L. L. Wang, "Evaluating climate change impacts on building level steady-state and dynamic outdoor thermal comfort," Available at SSRN 4782099.
Zuhra, S. S., Tabinda, A. B., & Yasar, A. (2019). Appraisal of the heat vulnerability index in Punjab: a case study of spatial pattern for exposure, sensitivity, and adaptive capacity in megacity Lahore, Pakistan. *International Journal of Biometeorology, 63*(12), 1669–1682.