Scilight

Regional Ecology and Management
https://www.sciltp.com/journals/rem



Check for updates

Article

Microclimatic Heterogeneity and Nonlinear Environmental Associations Across Contrasting Spaces in a Historical District: An Exploratory THI–XGBoost–SHAP Framework

Xuetong Song [1,2], Zhonghai Wang [1,2,*], Shaopeng Song [1,2] and Zhenjiang Tian [3,*]

[1] School of Architecture and Urban Planning, Chongqing University, Chongqing 400045, China
[2] Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University, Chongqing 400044, China
[3] Law School of Shihezi University, Shihezi University, Shihezi 832003, China
* Correspondence: 202615150852@stu.cqu.edu.cn (Z.W.); tzj_pol@shzu.edu.cn (Z.T.)

How To Cite: Song, X.; Wang, Z.; Song, S.; et al. Microclimatic Heterogeneity and Nonlinear Environmental Associations Across Contrasting Spaces in a Historical District: An Exploratory THI–XGBoost–SHAP Framework. *Regional Ecology and Management* **2026**, *1*(1), 12. https://doi.org/10.53941/rem.2026.100012

Received: 25 August 2026
Revised: 29 September 2026
Accepted: 12 October 2026
Published: 22 October 2026

Abstract: Historical districts contain spatial units with markedly different local microclimates, yet the temporal dynamics of their thermal-humidity environments and their associations with surrounding environmental conditions remain insufficiently resolved. This study examined three analytical settings in the Ciqikou Historical District, Chongqing: an Alley, the Main Street, and a Square. Using daytime multi-variable data collected on 4 August 2026, we integrated the temperature-humidity index (THI), an exact bivariate Shapley decomposition, and an XGBoost-SHAP framework to characterize thermal-humidity states and nonlinear environmental associations. Main Street and Square data were obtained from valid field measurements. Because the original Alley observations were lost owing to equipment failure, a stochastic reconstruction was generated within contemporaneous observational ranges and alley-specific environmental constraints; Alley results are therefore used only for exploratory comparison. Mean air temperature and THI were highest on Main Street (32.81 °C and 28.46 °C), followed by the Square (32.01 °C and 28.04 °C); the reconstructed Alley series yielded 31.16 °C and 27.73 °C, respectively. Exact Shapley decomposition showed that temperature accounted for 71.70–74.20% of the absolute contribution across the three analytical series, while the directions of the temperature and humidity contributions changed over time. These contribution shares depend on a common reference derived from both measured and reconstructed series and are not estimates based exclusively on field observations. Under grouped blocked cross-validation, out-of-fold R^2 values were 0.80 for Main Street and 0.73 for Square, compared with 0.57 for the reconstructed Alley series. SHAP analysis indicated setting-specific predictive association structures and nonlinear responses. The high importance assigned to atmospheric pressure more plausibly reflects background meteorological and temporal covariation within the monitoring window than a direct causal effect. Overall, the results reveal pronounced context dependence in historical-district microclimates and underscore the importance of data provenance, temporal dependence, and explicit limits on model interpretation in short-duration field studies.

Keywords: historical district; microclimate; XGBoost-SHAP; spatial heterogeneity; thermal environment

Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Against a background of global warming and increasingly frequent extreme-heat events, the urban thermal environment has evolved from a conventional climatological topic into a multidisciplinary issue involving urban design, human heat exposure, public-space quality, and climate adaptation [1]. Compared with suburban or natural surfaces, highly artificial urban fabrics can substantially reshape local thermal conditions by altering surface radiation budgets, heat storage, turbulent exchange, and near-surface ventilation [2]. These effects are far from spatially uniform. Relatively small changes in building height, street width, enclosure, orientation, shading, surface materials, vegetation, and ventilation can produce contrasting microclimates over distances of only tens of metres or less [3]. Consequently, city- or district-scale mean meteorological conditions often fail to represent the thermal environment actually experienced at pedestrian level [4]. In dense built environments in particular, research is therefore moving beyond broad urban-heat-island identification toward fine-scale diagnosis of streets, alleys, and nodal public spaces [5].

Historical districts are among the most representative yet complex settings in which this problem occurs [6]. Unlike regular modern urban blocks, they typically emerge through long-term incremental development and are characterized by high building density, narrow alleys, irregular street networks, continuous façades, and close interweaving of open and enclosed spaces [7]. Such morphology may reduce short-wave exposure through self-shading, but it can also restrict sky view and airflow, thereby limiting sensible-heat removal and nocturnal long-wave heat loss [8]. More open squares may allow stronger ventilation while simultaneously receiving greater direct solar radiation [9]. Main streets, as primary activity corridors, also differ from both narrow alleys and open squares in spatial interface, traffic activity, and patterns of use [10]. Thus, even under similar background weather, alleys, main streets, and squares may experience substantially different heat-exchange processes and microclimatic responses [11]. Treating a historical district as a single thermal unit can therefore obscure localized environments that may require targeted intervention [12].

This spatial heterogeneity is particularly important for climate-adaptive renewal in heritage environments. Thermal adaptation in historical districts is constrained by building conservation, street dimensions, landscape integrity, and established spatial functions, which limits the direct transfer of large-scale interventions commonly used in modern cities [13]. Additional shading, for example, may reduce radiant exposure but can also alter existing airflow pathways in narrow alleys; greater openness may improve air exchange while increasing daytime solar gain [14]. The key challenge is therefore not simply to maximize or minimize one environmental variable, but to identify which factors are most informative for thermal conditions in different spatial settings and over what environmental ranges their relationships change [15]. A simultaneous understanding of spatial context and response mechanisms is essential for moving from uniform cooling prescriptions toward fine-grained adaptation compatible with heritage conservation.

Air temperature is widely used to characterize urban thermal environments, but temperature alone cannot fully represent joint variations in heat and moisture [16]. Relative humidity modifies the thermal-moisture state of air and, together with temperature, contributes to the environmental basis of human thermal perception. Their diurnal variations are often inversely coupled: as daytime surface heating intensifies, air temperature generally increases while relative humidity decreases. High temperature therefore does not necessarily coincide with high humidity, and analysis of a single variable can miss important daytime state transitions. The temperature-humidity index (THI) combines temperature and relative humidity in an explicit mathematical form and provides a concise, interpretable metric for comparing thermal-humidity conditions among spaces [17]. THI does not, however, replace comprehensive human thermal-comfort models that include mean radiant temperature, wind speed, clothing, and metabolic rate [18]. Here, THI is used to characterize the joint air-temperature and humidity environment of historical spaces rather than to represent physiological heat risk or health outcomes directly [19].

Conventional THI analyses commonly stop at index calculation, mean comparisons, or time-series description and seldom quantify how strongly temperature and humidity separately contribute to a given THI state. Because THI is a nonlinear function of temperature and relative humidity, simple correlations with THI cannot rigorously partition their contributions, while one-at-a-time perturbations are sensitive to the selected reference state. An exact bivariate Shapley decomposition referenced to a common baseline provides a consistent solution [20]. By considering both possible orders in which temperature and humidity can enter the THI function, the deviation of each observation from the common reference THI can be decomposed exactly into temperature and humidity components. This permits comparison of their overall contribution magnitudes and tracking of directional changes through time, providing richer information than a single mean THI for rapidly varying historical-district microclimates.

Temperature and humidity are the direct mathematical components of THI, but their temporal and spatial variability is jointly conditioned by the surrounding microclimate [21]. Solar radiation is a major surface-energy

input; wind speed and direction regulate local air exchange and heat transport; and particulate matter, noise, and other monitored variables may encode human activity, pollution processes, or local environmental states that co-vary with time [22]. Complex covariance is also expected among predictors. $PM_{2.5}$ and PM_{10}, for example, often share temporal patterns, while radiation, wind conditions, and background weather may evolve concurrently during a monitoring period [23]. Quality control, distributional assessment, and correlation diagnostics are therefore essential before predictive driver analysis [24]. Variables with no variation, implausible sensor behaviour, or strong redundancy can otherwise add little information while impairing interpretability.

Linear regression and univariate correlation offer transparent parameter interpretation but are often poorly suited to the nonlinear responses, interactions, and potential thresholds characteristic of urban microclimates. Environmental associations need not be monotonic: the relationship with thermal conditions may change across ranges of wind speed, while increasing solar radiation may correspond to transitions from negative to positive model contributions followed by saturation [25]. Tree-based extreme gradient boosting (XGBoost) can learn such nonlinear structures without requiring a prespecified functional form [26]. Predictive performance alone, however, does not make the resulting model intrinsically interpretable; accuracy metrics and conventional feature importance do not indicate the ranges over which a variable raises or lowers the predicted thermal state.

SHapley Additive exPlanations (SHAP) provides a framework for interpreting such models [27]. SHAP decomposes individual predictions into feature-specific contributions with both magnitude and direction. At the global scale, mean absolute SHAP values rank predictive importance, while SHAP dependence relationships can reveal nonlinear response shapes, sign changes, and candidate turning regions. These properties are particularly relevant to historical districts, where the practical questions extend beyond which variable is most important to whether important variables differ among spaces, whether the same variable exhibits a consistent response, and over what ranges its association with predicted THI changes. Coupling the nonlinear learning capacity of XGBoost with local SHAP interpretation therefore provides a pathway from prediction toward interpretable diagnosis of complex historical-space microclimates [28].

An important caveat is that high SHAP importance is not equivalent to a causal effect [29,30]. A variable may carry substantial predictive information because it covaries with monitoring time, background weather, or unobserved processes. Small short-term changes in atmospheric pressure, for example, may closely track the temporal evolution of background meteorology, while noise may primarily reflect human activity and its temporal structure rather than direct heat exchange. Machine-learning analysis must therefore be grounded in rigorous quality control, defensible feature screening, and independent model validation, with SHAP values interpreted as contributions to model predictions rather than as physical causal strengths [24]. Continuous microclimate records also require safeguards against temporal information leakage; blocked out-of-fold validation offers a more conservative assessment than random train-test splitting [31].

Despite growing attention to thermal conditions in historical districts, three related gaps remain. First, city- or district-scale mean states cannot resolve microscale differences among alleys, primary streets, and open nodes. Second, temperature, humidity, and composite thermal indices are often described separately, with limited quantification of their time-varying contributions to the joint thermal-humidity state. Third, environmental analyses often emphasize aggregate correlations or single-variable importance while giving less attention to nonlinear responses, spatial context dependence, and the validation constraints inherent in continuous time-series observations. An integrated framework connecting thermal-humidity states, component contributions, and predictive associations with the external environment is therefore needed.

Accordingly, this study developed a multilevel framework comprising microclimatic-state identification, THI characterization, decomposition of temperature and humidity contributions, environmental-variable screening, and interpretation of nonlinear predictive associations. Alley, Main Street, and Square were treated as three analytical settings. We first characterized high-temporal-resolution air temperature and relative humidity and calculated THI; we then applied exact bivariate Shapley decomposition around a common reference state. Wind speed, wind direction, noise, atmospheric pressure, $PM_{2.5}$, PM_{10}, and total solar radiation were subsequently subjected to quality control and distributional and correlation diagnostics. Separate XGBoost models were finally evaluated using grouped out-of-fold validation, and TreeSHAP was used to examine predictive importance, contribution direction, and nonlinear dependence. Because the original Alley monitoring data were lost through equipment failure, its reconstructed series serves only as an exploratory scenario and is not treated as observational evidence equivalent to Main Street and Square.

The study addresses three questions: (1) what thermal-humidity states occurred in contrasting historical spaces during this summer daytime monitoring window; (2) how did the temperature and relative-humidity contributions to deviations in THI vary through time; and (3) did predictive associations between external environmental variables and THI exhibit setting-specific nonlinear structures? The work is explicitly positioned

as a single-day, high-temporal-resolution exploratory case study intended to establish a framework for subsequent repeated multi-day and multi-site testing, rather than to infer universal spatial causal relationships.

2. Study Area and Field Measurements

2.1. Study Area and Monitoring-Site Configuration

The Ciqikou Historical District in Chongqing, China, was selected as the case-study area (Figure 1). Functioning as a highly popular heritage tourism destination, the district experiences dense, transient daytime pedestrian traffic rather than maintaining a static residential demographic. Its urban morphology is characterized by high-density, low-to-mid-rise traditional commercial structures, leaving a predominantly impervious land cover dominated by historical stone paving with minimal natural vegetation. Its urban fabric includes narrow alleys, a principal pedestrian street, and relatively open public nodes. Based on observed enclosure and openness, the analytical settings were classified as Alley, Main Street, and Square. This classification organizes the present microclimatic comparison and should not be interpreted as a universal morphological typology quantitatively defined by parameters such as sky-view factor, street-canyon aspect ratio, or building height.

The Alley represents a comparatively enclosed internal lane, the Main Street represents the district's principal linear public-activity space, and the Square represents a relatively open public node. To capture the actual thermal environment experienced by pedestrians, all measurements were conducted using a tripod-mounted setup at a standard height of approximately 1.5 m above the historical hardscape pavement (e.g., stone slabs). Crucially, these settings presented distinct localized radiative exposures during the field campaign. The Main Street functions as a historical street canyon flanked by continuous traditional commercial building facades; consequently, the sensor was frequently shielded from direct sunlight through dynamic self-shading, resulting in significantly lower solar radiation inputs. In contrast, the Square is a relatively unshaded, open public node where the sensor was predominantly exposed to direct midday sunlight without overhead obstruction. Their thermal-humidity conditions and environmental variables were compared within a common analytical framework, followed by evaluation of predictive associations between external environmental indicators and THI.

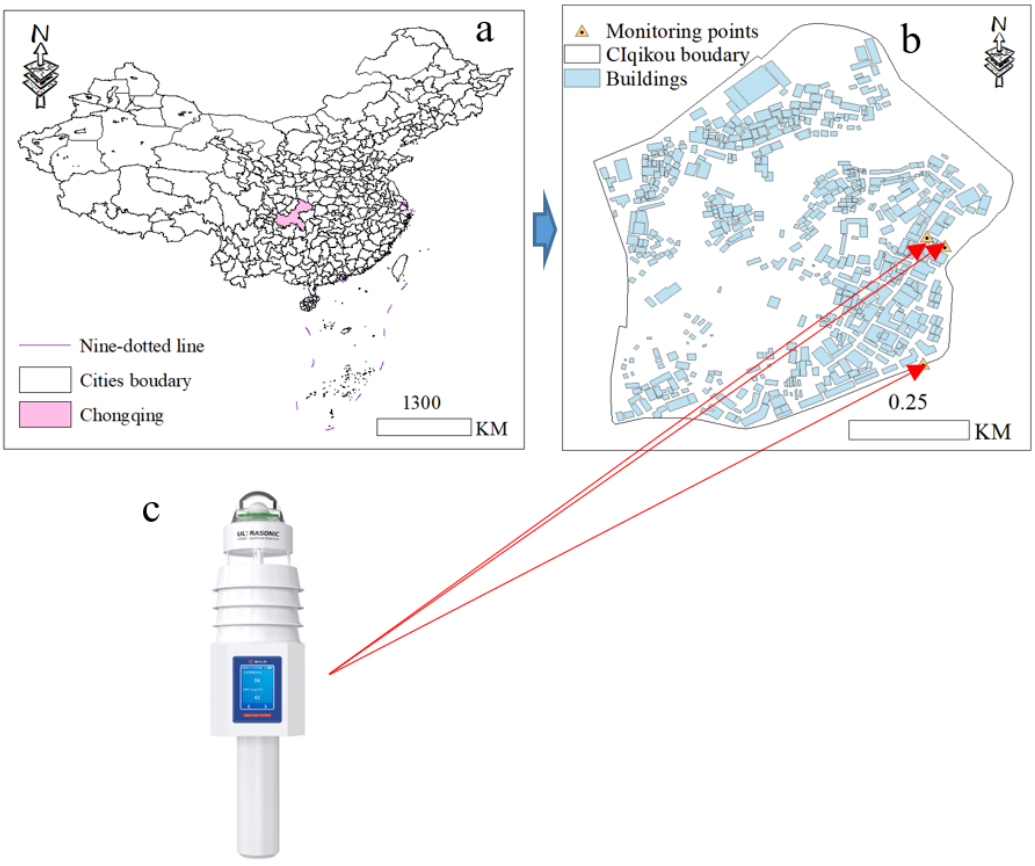


Figure 1. Study area, monitoring locations, and field measurement equipment: (**a**) location of Chongqing and the study area; (**b**) Ciqikou historical district and monitoring locations; (**c**) RS-HSQXZ-USB-1 field measurement equipment (Shandong Renke Control Technology Co., Ltd., Jinan, China).

2.2. Monitoring Equipment and Indicator Framework

Field measurements were obtained with an RS-HSQXZ-USB-1 handheld micro-weather station (Shandong Renke Control Technology Co., Ltd., Jinan, China; Figure 1). The portable configuration was selected to measure pedestrian-level conditions in the constrained spaces of the historical district. It was not intended to provide observations equivalent to those from a reference meteorological station. The measurement ranges and nominal accuracies reported below are manufacturer specifications and do not, by themselves, establish accuracy under the field conditions of this campaign. Variables used in the thermal-humidity assessment and subsequent predictive analysis included air temperature, relative humidity, wind speed, wind direction, noise level, atmospheric pressure, $PM_{2.5}$, PM_{10}, and total solar radiation. Air temperature and relative humidity were used to calculate THI and for exact bivariate Shapley decomposition; wind conditions, noise, pressure, particulate matter, and solar radiation formed the candidate external variables for the XGBoost-SHAP analysis.

Wind direction is a circular variable and was therefore treated using circular statistics for descriptive analysis and transformed into sine and cosine components for XGBoost modelling [32]. Temperature and relative humidity were excluded from the XGBoost predictor set because THI is calculated directly from them; reintroducing them as predictors would create deterministic information leakage and obscure the predictive information carried by external environmental variables.

According to the instrument specifications, air temperature has a measurement range of −40 to 80 °C and a nominal accuracy of ±0.5 °C at 25 °C; relative humidity has a range of 0–99% RH and a nominal accuracy of ±3% RH at 60% RH and 25 °C. The wind-speed and wind-direction ranges are 0–40 m s^{-1} and 0–359°, respectively. Atmospheric pressure is measured over 0–120 kPa with a nominal accuracy of ±0.15 kPa at 25 °C and 101 kPa. The noise range is 30–120 dB; $PM_{2.5}$ and PM_{10} each span 0–1000 μg m^{-3}; and total solar radiation spans 0–1800 W m^{-2} with a nominal accuracy of no more than ±3% at 150 W m^{-2}.

2.3. Field Data Acquisition, Data Provenance, and Quality Control

Field data were collected during daytime on 4 August 2026, covering approximately 08:48–15:01 local time at intervals of about 5 min.

The campaign was designed as a short-duration, exploratory field survey. Measurements at Main Street and Square covered approximately the same daytime window, which limited between-day weather differences in the site comparison. However, this observational design did not control solar exposure, local activity, or other time-varying influences and should not be interpreted as a controlled experiment. The findings describe conditions during the sampled window; the representativeness of this day within the local summer climate was not established through an independent meteorological comparison.

Main Street and Square each yielded 75 valid field records. The original Alley measurements were unusable because of equipment failure. To retain a three-setting exploratory comparison, 75 stochastic records were reconstructed within the ranges observed contemporaneously on Main Street and in the Square and under scenario constraints representing relatively greater shading, lower wind speed, and higher humidity in an alley. The reconstructed Alley series is used only as a methodological scenario. Its assumed shading, ventilation, and humidity characteristics cannot be treated as independently observed spatial effects. Comparisons involving this series describe the behaviour of the analytical workflow under the reconstruction assumptions and do not establish differences between real alleys and the measured sites. Direct observational statements are restricted to Main Street and Square. This reconstructed series is not treated as an independent field observation; all Alley statistics, THI decompositions, and XGBoost-SHAP results are used solely for methodological demonstration and sensitivity exploration.

All records were first checked for time format, numerical range, and missing values. Temperature and relative humidity were used to calculate THI. Candidate environmental variables were further screened for completeness, variability, physical plausibility, and redundancy [24]. Seven classes of environmental indicators were retained for the main analysis: wind speed, wind direction, noise, atmospheric pressure, $PM_{2.5}$, PM_{10}, and total solar radiation.

Rainfall was zero in every record and was excluded because it contained no variation. The carbon-dioxide field exactly duplicated $PM_{2.5}$ record by record and exhibited values inconsistent with conventional CO_2 concentrations; it was therefore treated as a field-mapping anomaly and excluded. Elevation showed implausible drift during the fixed monitoring period and was also omitted. Illuminance was highly redundant with total solar radiation; to reduce duplicate information, only total solar radiation, with clearer physical interpretation and complete range information, was retained. All exclusions followed predefined quality and redundancy criteria rather than the strength of association with THI.

Quality-controlled data were analysed at three levels: temporal variation in temperature, humidity, and THI; exact bivariate Shapley decomposition of THI relative to a common reference state; and distributional, wind-

direction, and correlation diagnostics for the seven retained external environmental indicators followed by separate XGBoost-SHAP models. Because Main Street and Square are measured series whereas Alley is stochastically reconstructed, all comparisons involving Alley are interpreted as exploratory.

3. Materials and Methods

3.1. Temporal Analysis of Air Temperature and Relative Humidity

Air-temperature and relative-humidity time series were plotted for the three analytical settings. A common y-axis range was used for each variable to allow direct visual comparison among settings. The time axis used hourly major ticks and 30-min minor ticks.

To enhance the information conveyed by the time-series plots, key turning points were identified from local extrema. For observation *i*, a change in the sign of consecutive first differences (from increasing to decreasing or vice versa) defined a candidate local extremum. Candidate importance was quantified by the absolute deviation of the observation from the mean of its two immediate neighbours:

$$A_i = |y_i - \frac{y_{i-1}+y_{i+1}}{2}| \quad (1)$$

where A_i is the local amplitude of candidate turning point i and y_i is the current observation. The global maximum and minimum were retained for every curve, and the remaining local extrema were ranked by decreasing local amplitude. To prevent excessive annotation, adjacent labelled points were separated by at least 4 monitoring records, with no more than 8 key points labelled in each panel.

3.2. Temperature-Humidity Index Calculation and Descriptive Analysis

The temperature-humidity index (THI) was used to represent the combined thermal-humidity state of air temperature and relative humidity [33]. THI was calculated as follows:

$$\mathrm{THI} = \mathrm{T} - (0.55 - 0.0055\mathrm{RH})(\mathrm{T} - 14.5) \quad (2)$$

where T is air temperature in degrees Celsius and RH is relative humidity entered as a percentage from 0 to 100.

Before calculation, relative humidity values were checked against the valid 0–100% range; records outside this range were treated as invalid. THI was calculated to two decimal places for every valid record. For each setting, we summarized the valid sample size, mean temperature, mean relative humidity, and the mean, standard deviation, minimum, median, and maximum of THI. THI time series were compared using a common y-axis range, and major turning points were identified and annotated using the same procedure applied to the temperature and humidity curves.

3.3. Decomposition of Temperature and Humidity Contributions to THI

To quantify the contributions of temperature and relative humidity to deviations in THI, an exact bivariate Shapley decomposition was applied [20]. The reported decomposition uses a common reference calculated from the pooled Main Street, Square, and reconstructed Alley series: T_0 = 31.995 °C and RH_0 = 60.183%. Because the reconstructed series contributes to this reference, the resulting contributions for all three settings are reference-dependent methodological quantities and are not estimates based exclusively on measured observations. The corresponding reference THI was:

$$\mathrm{THI}_0 = f(\mathrm{T}_0, \mathrm{RH}_0) \quad (3)$$

where *f*(T, RH) denotes the THI function. For any record (T, RH), the temperature contribution $\emptyset_{\mathrm{T}}$ was calculated as:

$$\emptyset_{\mathrm{T}} = \frac{1}{2}[f(\mathrm{T}, \mathrm{RH}_0) - f(\mathrm{T}_0, \mathrm{RH}_0) + f(\mathrm{T}, \mathrm{RH}) - f(\mathrm{T}_0, \mathrm{RH})] \quad (4)$$

The relative-humidity contribution φ_RH was calculated as:

$$\emptyset_{\mathrm{RH}} = \frac{1}{2}[f(\mathrm{T}_0, \mathrm{RH}) - f(\mathrm{T}_0, \mathrm{RH}_0) + f(\mathrm{T}, \mathrm{RH}) - f(\mathrm{T}, \mathrm{RH}_0)] \quad (5)$$

This decomposition satisfies the exact additivity relationship:

$$\mathrm{THI} - \mathrm{THI}_0 = \emptyset_{\mathrm{T}} + \emptyset_{\mathrm{RH}} \quad (6)$$

A positive contribution indicates that the corresponding variable shifts THI above the common reference state, whereas a negative contribution shifts THI below the reference. To compare the overall contribution

magnitudes of the two components, absolute temperature and humidity contributions were calculated and converted to relative contribution shares:

$$P_{\mathrm{T}} = \frac{|\emptyset_{\mathrm{T}}|}{|\emptyset_{\mathrm{T}}|+|\emptyset_{\mathrm{RH}}|} * 100\% \tag{7}$$

$$P_{\mathrm{RH}} = \frac{|\emptyset_{\mathrm{RH}}|}{|\emptyset_{\mathrm{T}}|+|\emptyset_{\mathrm{RH}}|} * 100\% \tag{8}$$

At the setting level, the overall contribution shares of temperature and humidity were calculated from their mean absolute contributions. Numerical closure was evaluated from the residual between each THI anomaly and the sum of its two Shapley components. This procedure is an exact bivariate algebraic decomposition of the THI equation and is distinct from the machine-learning SHAP analysis subsequently used to interpret XGBoost models.

3.4. Data Quality Control and Screening of Environmental Predictors

All candidate environmental variables were screened for completeness, dispersion, physical plausibility, and redundancy. Seven environmental indicators were retained: wind speed, wind direction, noise, atmospheric pressure, $PM_{2.5}$, PM_{10}, and total solar radiation.

Rainfall was zero in all records and was excluded because it lacked variability. The carbon-dioxide field duplicated $PM_{2.5}$ exactly record by record and its numerical range was inconsistent with conventional CO_2 concentrations, indicating a likely instrument field-mapping anomaly; it was therefore excluded. Elevation varied markedly over the approximately 6-h period despite fixed monitoring locations and was treated as sensor drift. Illuminance was almost perfectly monotonically associated with total solar radiation; to avoid strong redundancy, only the physically more explicit total solar radiation variable was retained.

Temperature and relative humidity were used only to calculate THI and were not included as XGBoost predictors. Because THI is directly determined by these two variables, including them as predictors would create deterministic information leakage and obscure associations between THI and other environmental variables.

Wind direction is circular and cannot be treated as an ordinary continuous variable over 0–360° [32]. For records with wind speed greater than zero, direction angle θ was transformed into sine and cosine components:

$$W_{sin} = sin(\theta) \tag{9}$$

$$W_{cos} = cos(\theta) \tag{10}$$

For calm records with wind speed equal to zero, wind direction has no well-defined physical meaning and both directional components were set to zero. Calm records were excluded from wind-direction distributions and circular statistics.

3.5. Descriptive and Exploratory Statistical Analysis of Environmental Variables

For wind speed, noise, atmospheric pressure, $PM_{2.5}$, PM_{10}, and total solar radiation, we calculated sample size, mean, standard deviation, coefficient of variation, minimum, 25th percentile, median, 75th percentile, maximum, and interquartile range. The coefficient of variation was calculated as:

$$\mathrm{CV} = \frac{s}{|\bar{x}|} * 100\% \tag{11}$$

where s is the sample standard deviation and $\bar{x}$ is the sample mean. The coefficient of variation was not calculated when the mean approached 0.

To describe differences among the three analytical series at matched sequential positions, Friedman tests were additionally calculated, followed by pairwise Wilcoxon signed-rank tests with Holm correction for multiple comparisons and Kendall's W as a sequence-level effect-size measure [34]. Each spatial setting, however, was represented by only one site and one daytime series, serial observations were temporally autocorrelated, and Alley was reconstructed. These tests are therefore treated solely as exploratory sequence-level sensitivity analyses and not as inferential tests of population-level differences among spatial types.

$$\mathrm{W} = \frac{X_F^2}{n(k-1)} \tag{12}$$

where X_F^2 is the Friedman test statistic, n is the number of paired records, and k is the number of analytical series. For comparisons reaching the nominal significance level, two-sided Wilcoxon signed-rank tests with Holm-adjusted p values were calculated. These p values describe consistency within the present monitoring sequences only and are not treated as confirmatory evidence of spatial-type effects.

Wind direction was described using circular statistics, including mean direction, mean resultant length, circular variance, and circular standard deviation; calm records were excluded [32]. Valid directions were divided into 16 sectors of 22.5° to calculate frequencies. A chi-square contingency test was used only as an exploratory description of directional-composition differences among the present series and not to infer general differences among historical-space types.

Spearman rank correlations were also calculated separately within each analytical series among wind speed, noise, atmospheric pressure, $PM_{2.5}$, PM_{10}, total solar radiation, and the sine and cosine components of wind direction. These coefficients were used to diagnose covariance structure and potential redundancy. Given serial autocorrelation, they describe covariation rather than causal relationships or significance under an independent-sample assumption.

3.6. XGBoost Predictive Modelling and Grouped Blocked Validation

Separate XGBoost regression models were fitted for Alley, Main Street, and Square to characterize nonlinear predictive associations between external environmental indicators and THI [26]. THI was the response variable. Inputs comprised wind speed, sine and cosine components of wind direction, noise, atmospheric pressure, $PM_{2.5}$, PM_{10}, and total solar radiation: eight numerical features representing seven conceptual environmental indicators. The three series were modelled independently to avoid averaging setting-specific covariance structures within a single model. XGBoost was used as an exploratory tool to describe potentially nonlinear predictive associations. Maximum tree depth was restricted to 1–3 during hyperparameter tuning to limit model complexity, although this does not eliminate overfitting risk in a small, temporally correlated dataset.

Models used the squared-error regression objective and the hist tree-building method. To account partly for temporal dependence among adjacent approximately 5-min observations, every five consecutive records were assigned to one group (approximately 25 min). Group identifiers constrained GroupKFold splits so that records from the same short block never appeared simultaneously in training and test sets. This strategy reduces direct within-block leakage, but GroupKFold is not strict forward temporal extrapolation and adjacent blocks may remain correlated [31,35]. The reported metrics were calculated from outer-fold held-out predictions and describe grouped out-of-fold performance within the sampled monitoring window. However, training sets may contain observations later than those in a test block, and neighbouring blocks may remain correlated. These metrics therefore do not establish forecasting performance for future periods or independent days and may be optimistic when interpreted as temporal generalization. Performance therefore represents grouped out-of-fold predictability within the monitoring window.

Predictive performance was evaluated using nested grouped cross-validation [36]. The outer loop used five-fold GroupKFold to generate out-of-fold predictions. Within each outer training set, a four-fold GroupKFold randomized hyperparameter search sampled 24 parameter combinations and selected the combination with the lowest validation RMSE. For final interpretation, a model was refitted to all valid records after parameter selection from at least 32 random combinations using five-fold grouped cross-validation. All stochastic procedures used the fixed random seed 20260805 to improve reproducibility.

The search space comprised 80, 120, 180, 250, or 350 trees; learning rates of 0.015, 0.03, 0.05, or 0.08; maximum tree depths of 1, 2, or 3; minimum child weights of 3, 5, 8, or 12; subsampling ratios of 0.65, 0.80, or 1.00; feature-subsampling ratios of 0.65, 0.80, or 1.00; L1 regularization coefficients of 0, 0.05, 0.20, or 0.50; L2 regularization coefficients of 1, 3, 8, or 15; and minimum loss-reduction thresholds of 0, 0.02, or 0.08. The final model for each setting was then refitted to all records using the parameter combination selected from at least 32 random combinations under five-fold grouped cross-validation.

Model performance was evaluated from out-of-fold predictions generated by the outer cross-validation loop using the coefficient of determination (R^2), root-mean-square error (RMSE), and mean absolute error (MAE):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{13}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{14}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{15}$$

where y_i and $\hat{y}_i$ are, respectively, the observed and out-of-fold predicted THI for record i, $\bar{y}$ is the mean observed THI, and n is the number of valid records. Model diagnostics additionally examined agreement between observed and out-of-fold predicted values and the distribution of residuals across predicted values.

3.7. Model Interpretation Using TreeSHAP

TreeSHAP was used to interpret the final XGBoost model for each analytical series [37]. SHAP values were obtained from the final model refitted to all valid records after grouped cross-validation and parameter selection; they therefore explain predictive structure within the observed/reconstructed sample domain and are not out-of-fold causal effects. For record *i*, the model prediction can be written as the baseline output plus the sum of feature-specific SHAP values:

$$\hat{y}_i = \emptyset_0 + \sum_{j=1}^{p} \emptyset_{ij} \tag{16}$$

where $\hat{y}_i$ is the model-predicted THI, $\emptyset_0$ is the model baseline output, $\emptyset_{ij}$ is the contribution of feature *j* to the prediction for record *i*, and *p* is the number of model input features. A positive SHAP value indicates that the variable increases the model-predicted THI, whereas a negative SHAP value indicates a decrease in predicted THI.

Overall importance for each environmental indicator was quantified as the sample mean of the absolute SHAP values:

$$I_j = \frac{1}{n}\sum_{i=1}^{n} |\emptyset_{ij}| \tag{17}$$

Mean absolute SHAP values were further normalized to relative importance shares. The record-level SHAP values for the sine and cosine components of wind direction were summed and recombined as the conceptual indicator 'Wind direction'. The sign of a SHAP value indicates whether a variable shifts predicted THI above or below the model baseline; it does not represent a physical causal effect on observed THI.

$$S_j = \frac{I_j}{\sum_{k=1}^{m} I_k} * 100\% \tag{18}$$

where I_j is the mean absolute SHAP value for indicator *j*, S_j is its relative importance share, and *m* is the number of conceptual environmental indicators. Record-level SHAP values for the sine and cosine components of wind direction were summed and recombined as the conceptual indicator 'Wind direction'.

3.8. SHAP Dependence Relationships and Visualization of Nonlinear Responses

To visualize nonlinear relationships between environmental values and SHAP contributions, a SHAP dependence scatterplot was produced for each conceptual indicator and locally smoothed using LOWESS [38]. The smoothing fraction was 0.42 for ordinary continuous variables. Wind direction was periodically extended over 360° before circular smoothing with a fraction of 0.28. A 2.5–97.5% envelope was obtained from 120 bootstrap resamples, and extrema and regions of marked curvature in the smoothed curve were annotated [39]. Because the observations are temporally autocorrelated and the sample size is limited, this envelope is used only as a visual uncertainty band, and annotated points are descriptive features of the fitted response rather than physical thresholds confirmed by independent statistical testing [40].

4. Results

Unless otherwise stated, results for Main Street and Square are based on field measurements, whereas Alley results are based on the stochastic series reconstructed after equipment failure. Alley is therefore used only for exploratory scenario comparison and not as independent observational evidence equivalent to the two measured sites.

4.1. Temporal Patterns of Temperature and Relative Humidity

All three analytical series showed an overall warming and drying tendency during the monitoring window, although their levels and short-term fluctuations differed (Figure 2). Mean measured air temperature was 32.81 °C on Main Street and 32.01 °C in the Square, while mean relative humidity was 57.62% and 59.65%, respectively. The reconstructed Alley series had a mean temperature of 31.16 °C and mean relative humidity of 63.28%; these values are included only for exploratory comparison.

Air temperature in the reconstructed Alley series increased overall from 27.5 °C at 08:50, with short-term fluctuations, and reached a series maximum of 35.5 °C at 14:38 (Figure 2a), giving a range of 27.5–35.5 °C. Relative humidity declined overall from 74.3% at the start of the reconstructed series to a minimum of 49.6% at 14:48, before recovering to 53.9% at the end (Figure 2d). These changes describe the temporal structure of the reconstructed scenario rather than actual Alley observations.

Measured Main Street air temperature increased overall from 28.9 °C at 08:48 to 38.0 °C at 14:41 and subsequently decreased to 36.1 °C (Figure 2b). Relative humidity declined overall from 69.7% to 42.4% at 14:41

and then recovered to 48.0% (Figure 2e). The measured ranges were 28.9–38.0 °C for temperature and 42.4–69.7% for relative humidity.

The measured Square series also showed overall warming and drying but more frequent short-term fluctuations (Figure 2c,f). Air temperature began at 28.0 °C at 08:52, fell to 27.6 °C at 09:52, then increased overall to 36.0 °C at 14:40 and ended at 34.1 °C. Relative humidity reached 73.5% at 09:52, subsequently declined overall to 47.6% at 14:50, and ended at 54.3%. Mean Square temperature and relative humidity were 32.01 °C and 59.65%, respectively.

Within this monitoring window, both measured series showed overall warming and drying, with more frequent short-term fluctuations in the Square. The reconstructed Alley series was configured as a relatively cooler and more humid alley scenario; consequently, differences between it and the measured series cannot be used to validate a real Alley spatial effect (Figure 2).

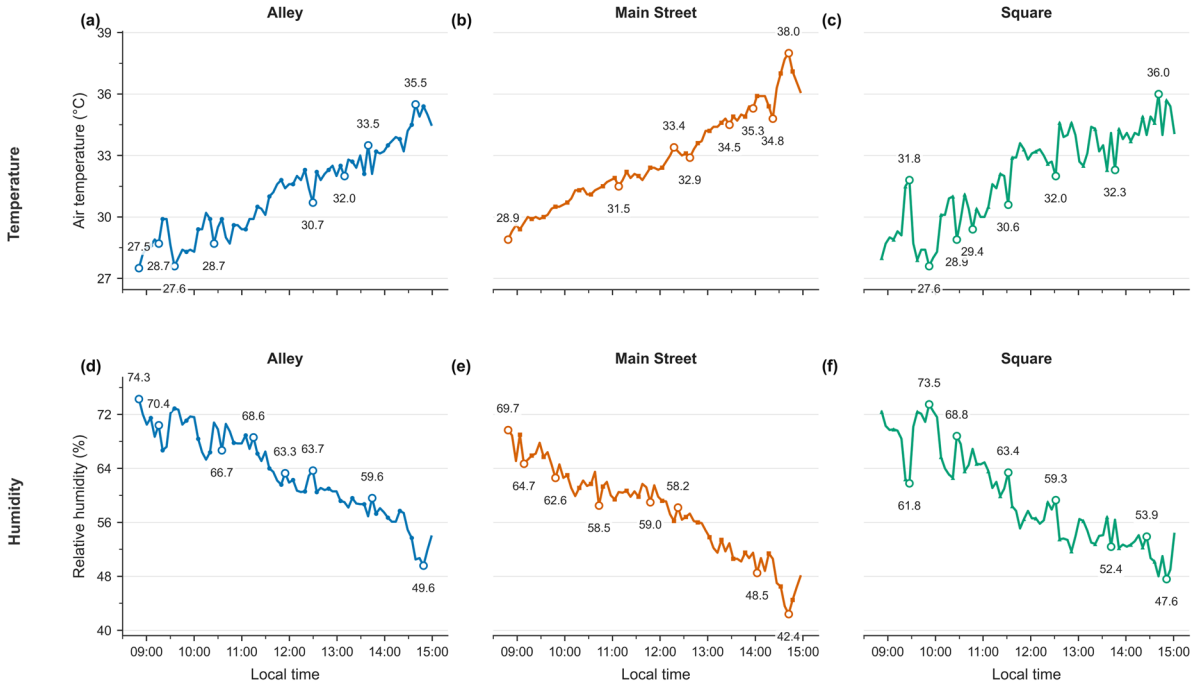


Figure 2. Temporal variations in air temperature and relative humidity across three analytical scenarios: (**a**–**c**) air temperature in the reconstructed Alley scenario, observed Main Street, and observed Square, respectively; (**d**–**f**) relative humidity in the corresponding scenarios.

4.2. Temporal Variation in THI

THI increased overall in all three analytical series during the monitoring window (Figure 3). Mean measured THI was 28.46 °C on Main Street and 28.04 °C in the Square; the reconstructed Alley series averaged 27.73 °C. THI standard deviations were approximately 1.07–1.11 °C across the three series, indicating substantial temporal variation within the monitoring window, although Alley values are not used for confirmatory inference of spatial differences.

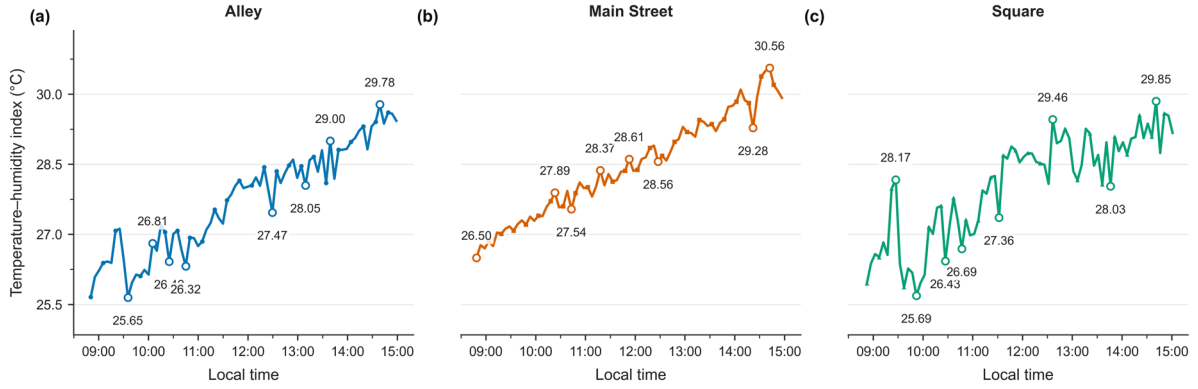


Figure 3. Temporal variations in the temperature–humidity index (THI) across three analytical scenarios: (**a**) reconstructed Alley scenario, (**b**) observed Main Street, and (**c**) observed Square.

THI in the reconstructed Alley series increased overall through time (Figure 3a), from 25.66 °C initially to a series minimum of 25.65 °C at 09:35 and a maximum of 29.78 °C at 14:38, ending at 29.43 °C; the median was 27.99 °C. These values describe only the reconstructed scenario.

Measured Main Street THI increased overall from 26.50 °C at 08:48 to 30.56 °C at 14:41 and subsequently fell to 29.92 °C at the end of the series (Figure 3b). The mean and median were 28.46 °C and 28.37 °C, respectively.

Measured Square THI increased overall with more frequent short-term fluctuations (Figure 3c). It began at 25.95 °C, decreased to 25.69 °C at 09:52, reached 29.85 °C at 14:40, and ended at 29.17 °C; the mean and median were 28.04 °C and 28.35 °C. Main Street exhibited the higher mean and peak THI of the two measured settings.

4.3. Contributions of Temperature and Relative Humidity to THI

The common reference state constructed from all three analytical series was 31.995 °C for air temperature and 60.183% for relative humidity, corresponding to a reference THI of 28.163 °C. Exact bivariate Shapley decomposition showed that temperature had a larger contribution magnitude than relative humidity in all three series under this reference definition and data range (Figure 4a,b). Because the common reference includes the reconstructed Alley series, these values are relative contributions within the present analytical framework rather than universal weights estimated solely from independent field observations.

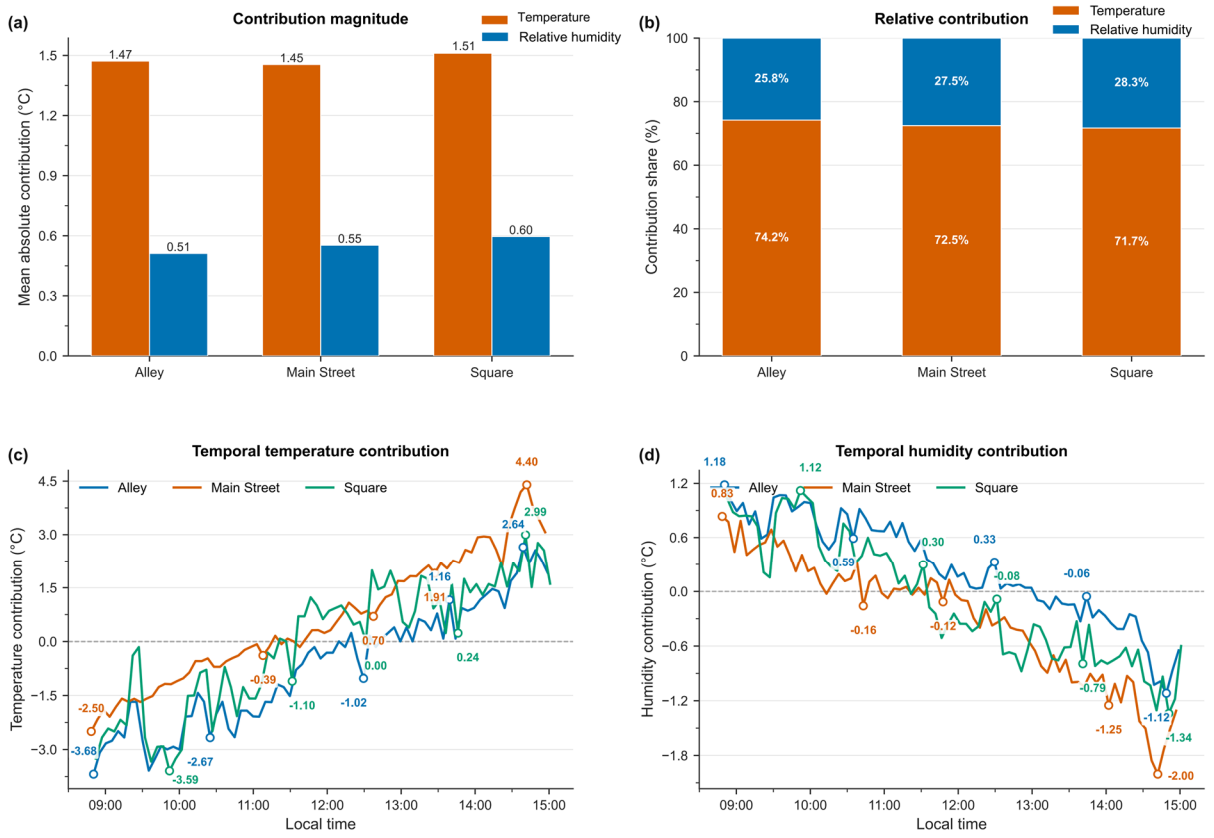


Figure 4. Shapley decomposition of temperature and relative humidity contributions to the temperature–humidity index (THI) across three analytical scenarios: (**a**) mean absolute contributions; (**b**) relative contribution shares; (**c**) temporal temperature contribution; and (**d**) temporal relative-humidity contribution. Alley represents a reconstructed exploratory scenario.

Mean absolute temperature contributions were 1.472, 1.455, and 1.511 °C for Alley, Main Street, and Square, respectively, compared with relative-humidity contributions of 0.512, 0.553, and 0.596 °C (Figure 4a). After normalization, temperature accounted for 74.20%, 72.46%, and 71.70%, and relative humidity for 25.80%, 27.54%, and 28.30%, respectively (Figure 4b). Thus, absolute THI deviations were dominated by the temperature component under the present THI formulation and common reference; Alley remains a reconstructed scenario.

The time-resolved Shapley decomposition further revealed systematic changes in contribution direction through the monitoring period (Figure 4c,d). Early in the period, temperature contributions were negative in all three series while relative-humidity contributions were predominantly positive. As daytime temperature increased and relative humidity declined, temperature contributions generally crossed from negative to positive and

strengthened, whereas humidity contributions decreased and became negative, producing opposite contribution patterns between the early and late monitoring periods.

In the reconstructed Alley series, the temperature contribution was lowest at −3.685 °C at 08:50, then increased overall to a maximum of 2.644 °C at 14:38 (Figure 4c). Conversely, the humidity contribution declined from a maximum of 1.184 °C at 08:50 to a minimum of −1.117 °C at 14:48 (Figure 4d). Mean signed contributions over the full series were −0.691 °C for temperature and 0.256 °C for humidity.

Main Street exhibited the strongest late-period positive temperature contribution. It rose from −2.498 °C at 08:48 to 4.397 °C at 14:41, the largest positive temperature contribution among the three analytical series. Over the same period, humidity contribution declined from an early maximum of 0.835 °C to −2.005 °C at 14:41, the largest negative humidity contribution observed (Figure 4c,d). Mean signed contributions over the monitoring period were 0.594 °C for temperature and −0.292 °C for humidity.

Square contributions showed greater short-term fluctuation. The temperature contribution reached a minimum of −3.593 °C at 09:52 and subsequently increased overall to a maximum of 2.994 °C at 14:40. The humidity contribution peaked at 1.120 °C at 09:52 and then declined overall to a minimum of −1.339 °C at 14:50 (Figure 4c,d). Mean signed contributions were −0.034 °C for temperature and −0.094 °C for humidity. Overall, all three analytical series showed temperature-dominated absolute contributions with a daytime reversal in contribution direction, while the Alley decomposition describes only the reconstructed scenario.

4.4. Environmental Characteristics of the Three Analytical Settings

Environmental indicators other than temperature and humidity also differed in distribution among the three analytical series (Figure 5). Median measured wind speed and noise were 1.00 m s^{-1} and 76.8 dB on Main Street and 0.70 m s^{-1} and 68.6 dB in the Square. Median measured $PM_{2.5}$ was 22 and 8 μg m^{-3}, and PM_{10} was 30 and 16 μg m^{-3}, respectively. Median total solar radiation was 396 W m^{-2} in the Square, compared with 129 W m^{-2} on Main Street. In the reconstructed Alley series, the corresponding medians for wind speed, noise, $PM_{2.5}$, PM_{10}, and solar radiation were 0.40 m s^{-1}, 73.3 dB, 17 μg m^{-3}, 25 μg m^{-3}, and 127 W m^{-2}. Given the single-site, single-day design and the provenance of the Alley series, between-setting tests are treated only as exploratory sequence-level sensitivity analyses; their *p* values are not used to infer population-level differences among historical-space types.

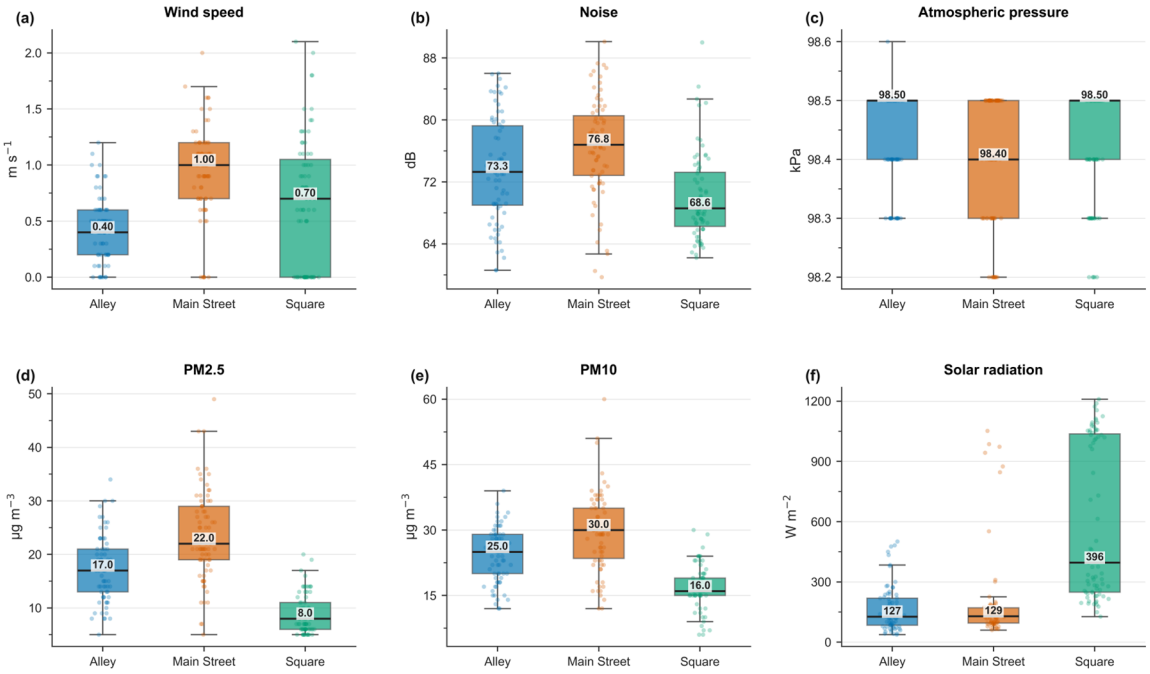


Figure 5. Distributions of environmental variables across three analytical scenarios: (**a**) wind speed, (**b**) noise, (**c**) atmospheric pressure, (**d**) $PM_{2.5}$, (**e**) PM_{10}, and (**f**) solar radiation. Alley represents a reconstructed exploratory scenario.

Wind-direction composition differed among the three analytical series (Figure 6). After calm records were excluded, 67, 69, and 49 valid direction records remained for Alley, Main Street, and Square, respectively. The reconstructed Alley series was concentrated in the 292.5–315.0° sector (34.33%), with a circular mean of 305.45°. Measured Main Street directions were dominated by 270.0–292.5° (60.87%), with a circular mean of 284.61°, whereas measured Square directions were concentrated at 337.5–360.0° (61.22%), with a circular mean of 359.60°.

Mean resultant lengths were 0.891, 0.954, and 0.552, respectively. These results describe directional structure in the present series and are not treated as a replicated test of general wind-direction differences among spatial types.

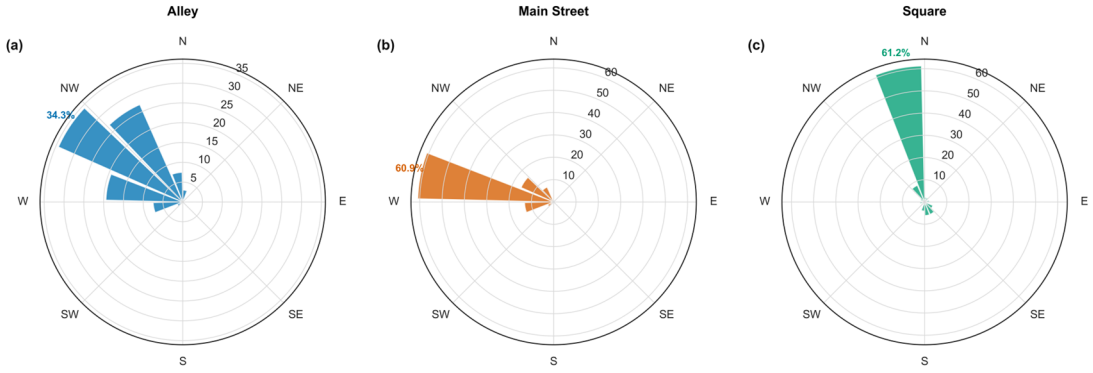


Figure 6. Wind-direction distributions across three analytical scenarios. Wind roses show the frequency of 16 wind-direction sectors for (**a**) the reconstructed Alley scenario, (**b**) observed Main Street, and (**c**) observed Square.

Spearman correlation structure among environmental variables also varied across analytical series (Figure 7). $PM_{2.5}$–PM_{10} correlations were 0.808, 0.849, and 0.615 for Alley, Main Street, and Square, respectively. On Main Street, wind speed correlated with noise at $\rho = 0.528$ and total solar radiation with atmospheric pressure at $\rho = -0.488$. In the Square, atmospheric pressure correlated with $PM_{2.5}$ at $\rho = 0.700$ and noise with $PM_{2.5}$ at $\rho = -0.475$. In the reconstructed Alley series, noise and atmospheric pressure correlated at $\rho = -0.455$. Given serial autocorrelation and the provenance of Alley, these coefficients are used only to diagnose covariance structure and not to infer causal relationships.

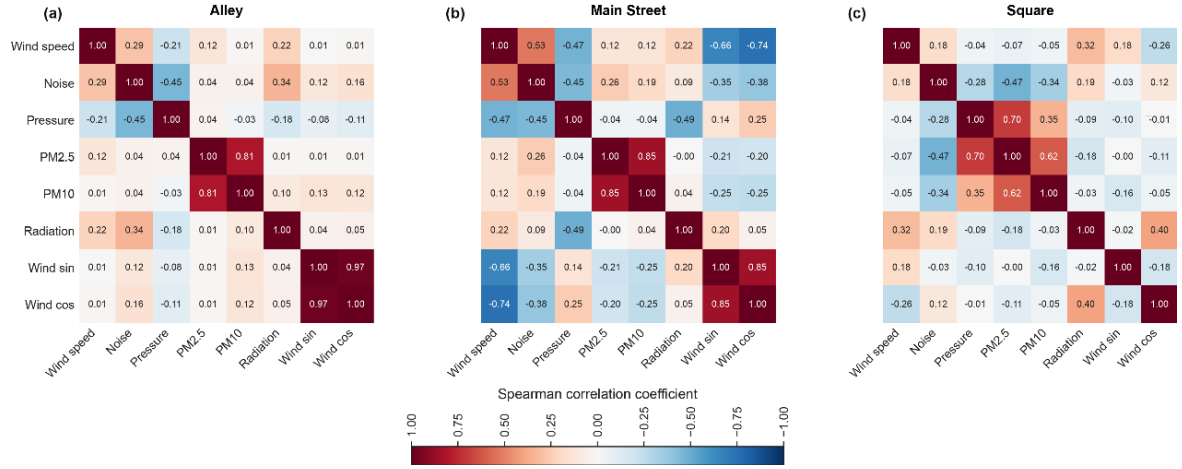


Figure 7. Spearman correlation structures among environmental predictors across three analytical scenarios. Correlation matrices are shown for (**a**) the reconstructed Alley scenario, (**b**) observed Main Street, and (**c**) observed Square; coefficients are used descriptively to diagnose covariance structure.

4.5. Predictive Associations with THI Across Analytical Settings

Grouped out-of-fold validation based on consecutive-record blocks yielded XGBoost R^2 values of 0.80 and 0.73 for the measured Main Street and Square series, respectively, with RMSE values of 0.47 and 0.58 and MAE values of 0.38 and 0.47 (Figure S1). The reconstructed Alley model yielded $R^2 = 0.57$, RMSE = 0.73, and MAE = 0.56; these metrics quantify fit and grouped out-of-fold prediction within the reconstructed scenario and do not validate generalization to a real Alley. Residuals from all three models occurred on both sides of zero. Because GroupKFold is not strict forward extrapolation, these metrics represent grouped out-of-fold predictability within the monitoring window [35].

SHAP analysis subsequently showed that the models relied on different predictive association structures across settings (Figures 8–13). Atmospheric pressure accounted for 38.2%, 55.3%, and 47.4% of relative importance in the Alley, Main Street, and Square models, respectively. The second-ranked predictor was noise in Alley (22.0%) and total solar radiation in Main Street and Square (14.6% and 27.5%). These percentages represent mean absolute SHAP contributions in the final XGBoost models and do not imply physical causal effects of equal magnitude [29]. Alley results remain exploratory features of the reconstructed scenario.

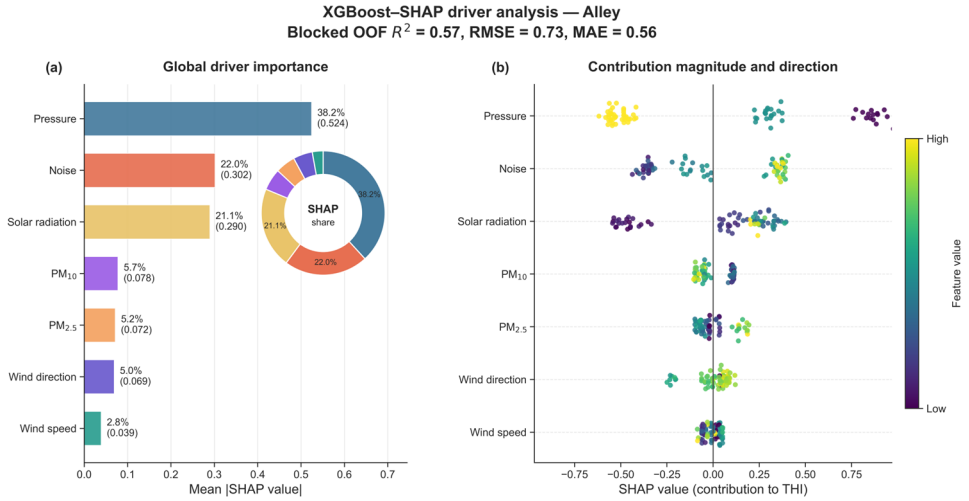


Figure 8. XGBoost–SHAP predictor importance for THI in the reconstructed Alley scenario: (**a**) global driver importance based on mean absolute SHAP values and relative SHAP contribution shares; (**b**) SHAP summary plot showing the magnitude and direction of feature contributions.

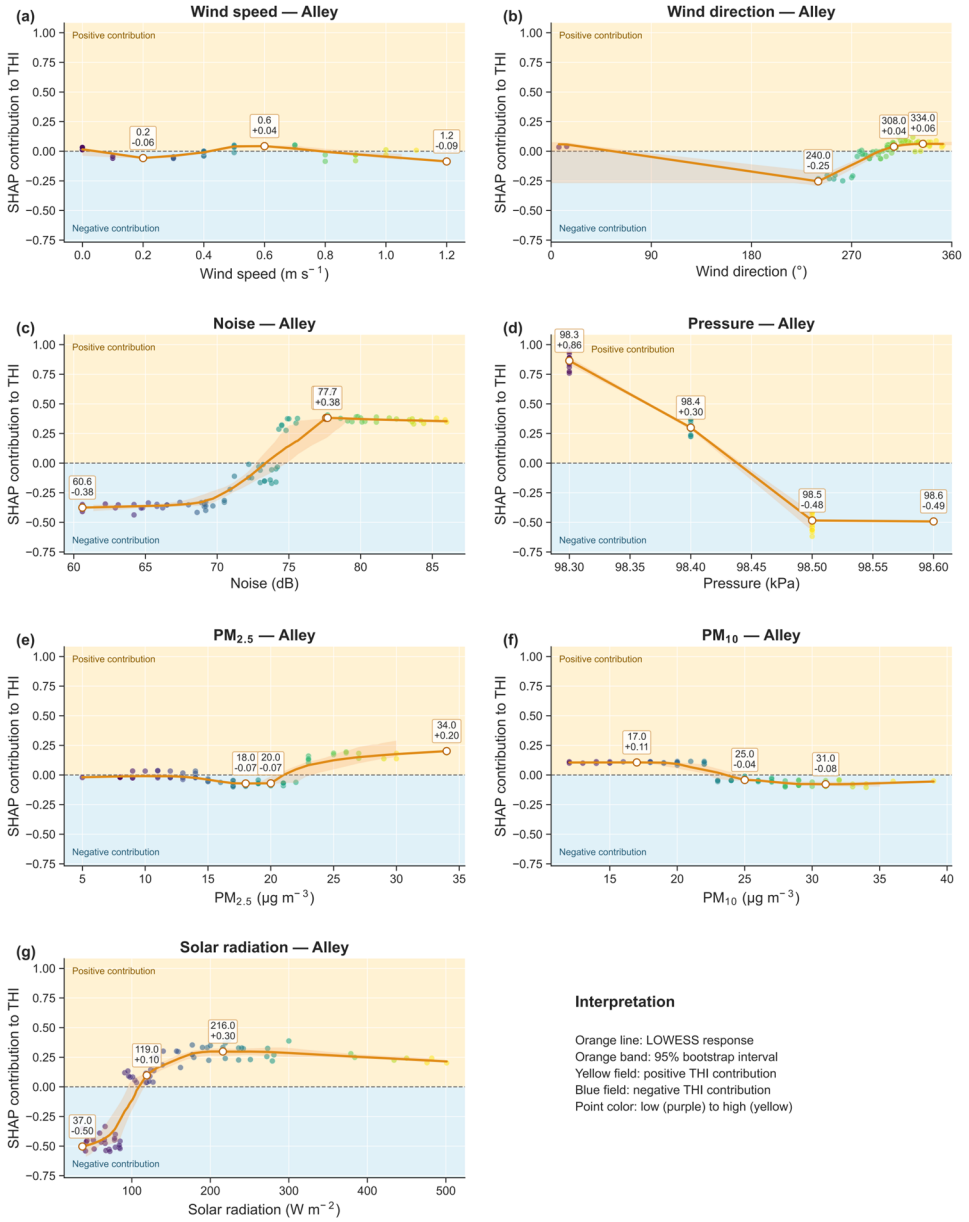


Figure 9. SHAP dependence patterns of environmental predictors in the reconstructed Alley scenario: (**a**) wind speed; (**b**) wind direction; (**c**) noise; (**d**) atmospheric pressure; (**e**) $PM_{2.5}$; (**f**) PM_{10}; and (**g**) total solar radiation.

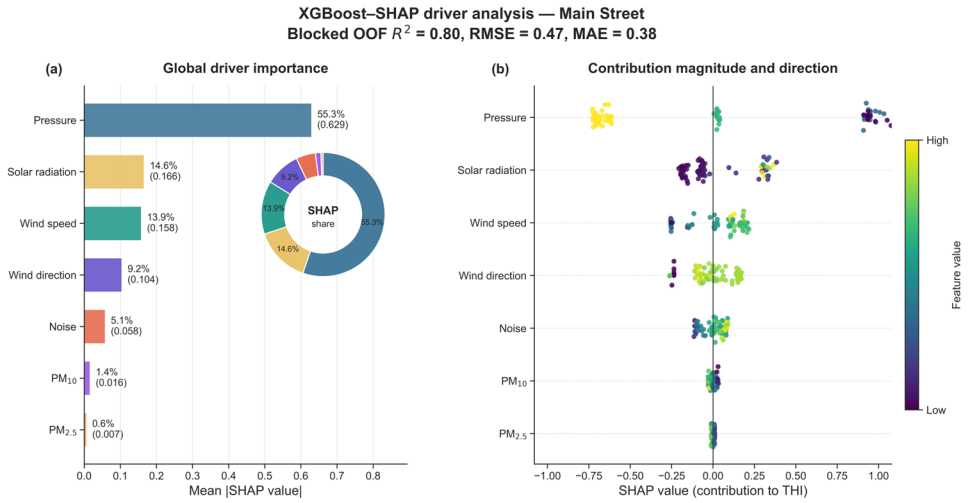


Figure 10. XGBoost–SHAP predictor importance for THI on Main Street: (**a**) global driver importance based on mean absolute SHAP values and relative SHAP contribution shares; (**b**) SHAP summary plot showing the magnitude and direction of feature contributions.

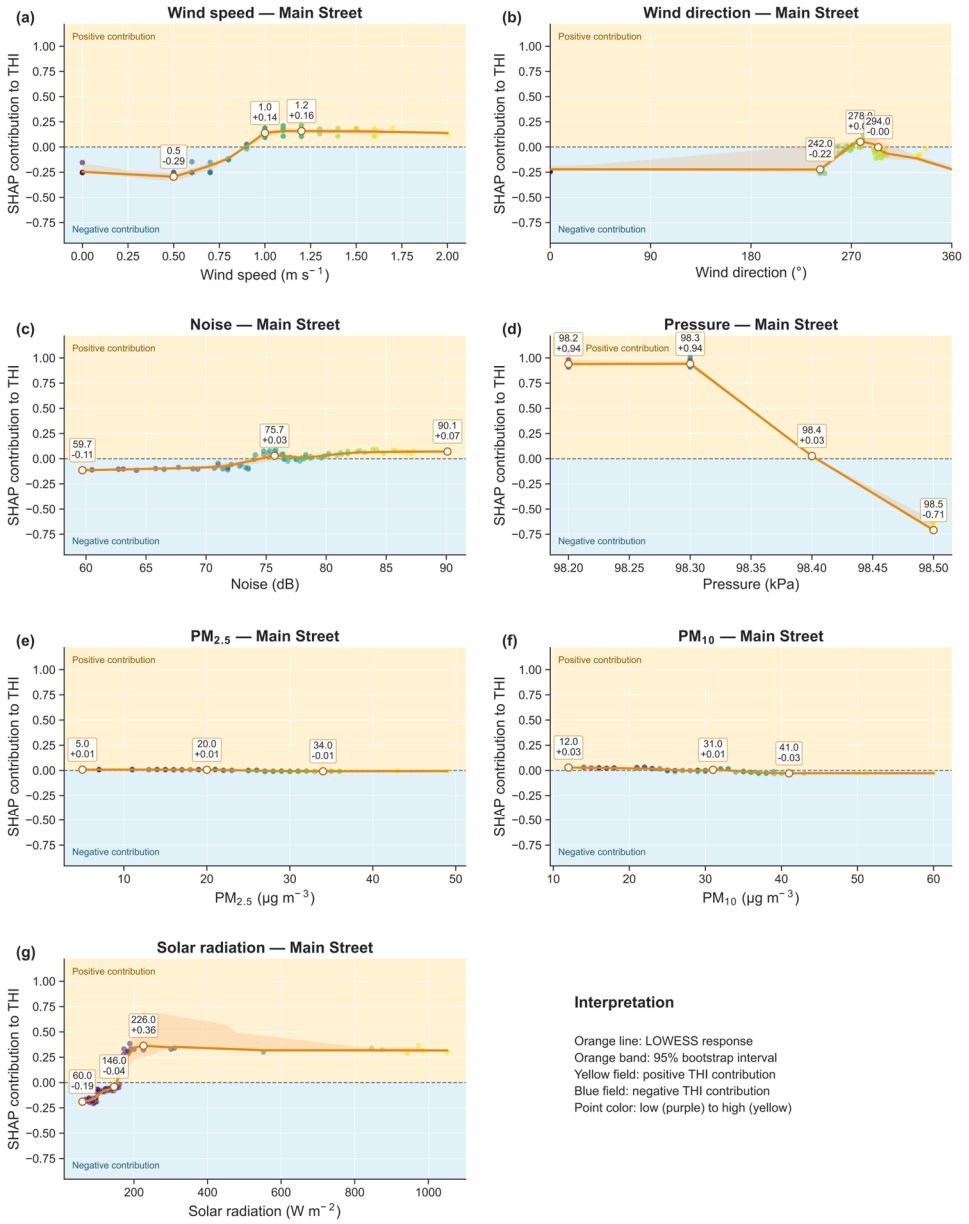


Figure 11. SHAP dependence patterns of environmental predictors in the Main Street: (**a**) wind speed; (**b**) wind direction; (**c**) noise; (**d**) atmospheric pressure; (**e**) $PM_{2.5}$; (**f**) PM_{10}; and (**g**) total solar radiation.

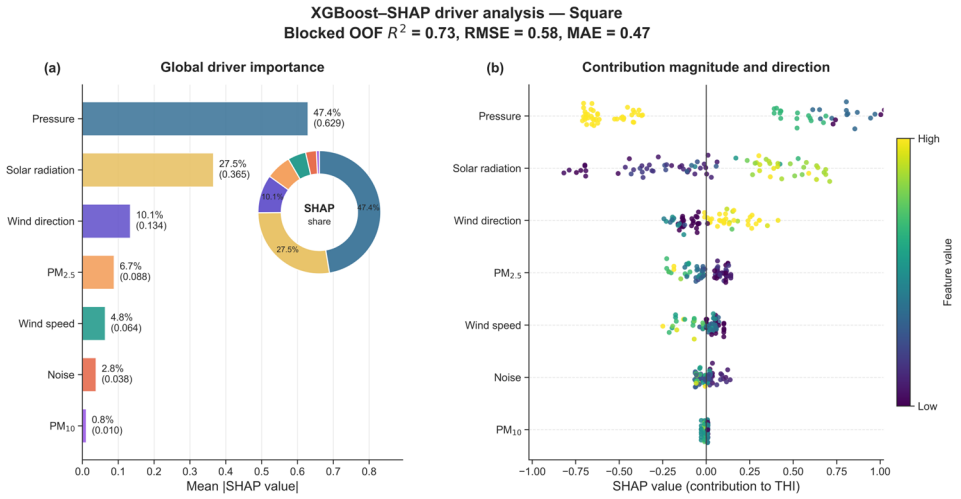


Figure 12. XGBoost–SHAP predictor importance for THI in the Square: (**a**) global driver importance based on mean absolute SHAP values and relative SHAP contribution shares; (**b**) SHAP summary plot showing the magnitude and direction of feature contributions.

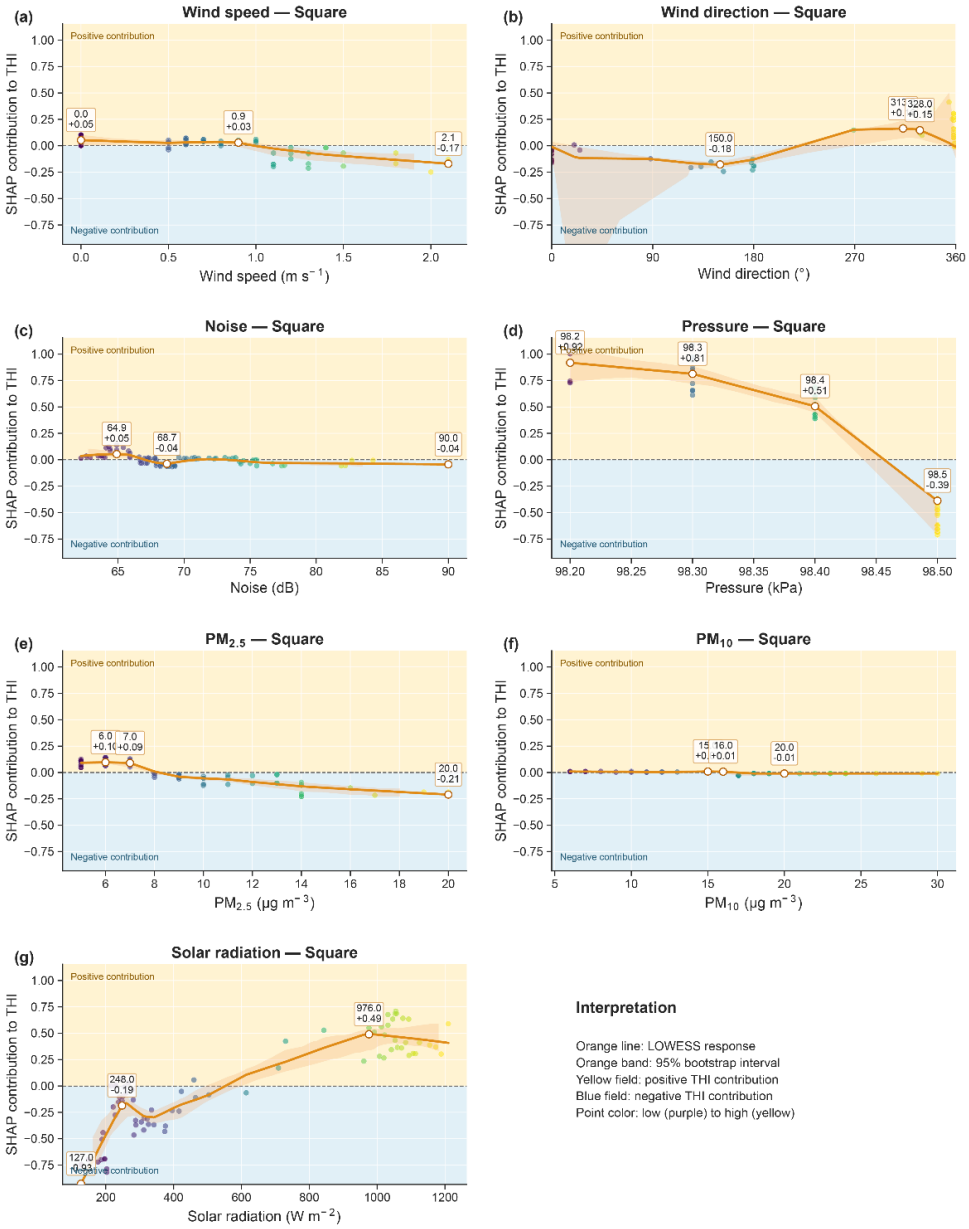


Figure 13. SHAP dependence patterns of environmental predictors in the Square: (**a**) wind speed; (**b**) wind direction; (**c**) noise; (**d**) atmospheric pressure; (**e**) $PM_{2.5}$; (**f**) PM_{10}; and (**g**) total solar radiation.

4.5.1. Predictive Associations with THI in the Reconstructed Alley Scenario

The grouped out-of-fold performance of the reconstructed Alley model was R^2 = 0.57, RMSE = 0.73, and MAE = 0.56 (Figure S1a,b). In the final model, atmospheric pressure had a mean absolute SHAP value of 0.524 and accounted for 38.2% of total importance; noise and total solar radiation contributed 0.302 (22.0%) and 0.290 (21.1%), respectively. PM_{10}, $PM_{2.5}$, wind direction, and wind speed accounted for 5.7%, 5.2%, 5.0%, and 2.8%, respectively (Figure 8a). These results describe only the model structure of reconstructed data and are not field evidence of causal drivers in a real Alley.

SHAP dependence relationships for the reconstructed Alley series showed nonlinear model responses (Figure 9). The smoothed pressure contribution declined from approximately +0.86 at 98.3 kPa to +0.30 at 98.4 kPa and became negative near 98.5–98.6 kPa (Figure 9d). The noise contribution increased from −0.38 at 60.6 dB to +0.38 at 77.7 dB (Figure 9c). For total solar radiation, the contribution rose from −0.50 at 37 W m^{-2} to +0.10 at 119 W m^{-2}, reached approximately +0.30 near 216 W m^{-2}, and then levelled off (Figure 9g). Annotated points describe LOWESS curve geometry and are not interpreted as independently confirmed thresholds [40].

Other variables exhibited smaller SHAP response amplitudes. For wind speeds of 0.2–1.2 m s^{-1}, annotated SHAP values ranged from −0.09 to +0.04 (Figure 9a). Wind-direction contribution increased from −0.25 at 240° to +0.04 and +0.06 at 308° and 334°, respectively (Figure 9b). $PM_{2.5}$ showed weakly negative contributions around 18–20 μg m^{-3} but +0.20 at the high-end value of 34 μg m^{-3} (Figure 9e); PM_{10} contributions were likewise small (Figure 9f). Overall, pressure, noise, and total solar radiation carried most of the predictive importance within the reconstructed Alley model.

4.5.2. Predictive Associations with THI on Main Street

The Main Street model achieved the highest grouped out-of-fold performance among the three models (R^2 = 0.80, RMSE = 0.47, MAE = 0.38; Figure S1c,d). Atmospheric pressure had a mean absolute SHAP value of 0.629 and accounted for 55.3% of total importance (Figure 10a). Total solar radiation and wind speed contributed 0.166 (14.6%) and 0.158 (13.9%), respectively; wind direction contributed 0.104 (9.2%), noise 0.058 (5.1%), PM_{10} 0.016 (1.4%), and $PM_{2.5}$ 0.007 (0.6%). This ranking describes predictive information within the model rather than physical causal strength.

The dependence relationship for atmospheric pressure showed a marked negative nonlinear response (Figure 11d). Values of 98.2 and 98.3 kPa corresponded to relatively high positive SHAP contributions; the contribution declined to +0.03 at 98.4 kPa and to −0.71 at 98.5 kPa. Total solar radiation shifted primarily from negative to positive contribution: SHAP values were −0.19 at 60 W m^{-2} and −0.04 at 146 W m^{-2}, but +0.36 at 226 W m^{-2} and remained generally positive thereafter (Figure 11g). Wind speed also showed a directional transition, from −0.29 at 0.5 m s^{-1} to +0.14 and +0.16 at 1.0 and 1.2 m s^{-1}, respectively (Figure 11a).

Responses of the remaining variables were weaker. Noise contribution approached zero from −0.11 at 59.7 dB and became slightly positive at higher levels (Figure 11c), whereas $PM_{2.5}$ and PM_{10} SHAP responses remained largely centred near zero (Figure 11e,f). The final Main Street model therefore relied primarily on predictive information carried by atmospheric pressure, total solar radiation, and the wind environment; the importance of pressure should be interpreted cautiously because of its potential role as a temporal proxy.

4.5.3. Predictive Associations with THI in the Square

The Square model yielded grouped out-of-fold R^2 = 0.73, RMSE = 0.58, and MAE = 0.47 (Figure S1e,f). Atmospheric pressure had a mean absolute SHAP value of 0.629 and accounted for 47.4% of importance; total solar radiation contributed 0.365 (27.5%), wind direction 0.134 (10.1%), $PM_{2.5}$ 0.088 (6.7%), wind speed 0.064 (4.8%), noise 2.8%, and PM_{10} 0.8% (Figure 12a). Pressure and total solar radiation together accounted for 74.9%, reflecting concentration of predictive information in the final model rather than their causal contributions.

SHAP dependence relationships showed a strong negative pressure response in the Square: contribution declined from approximately +0.92 at 98.2 kPa to +0.81 at 98.3 kPa and +0.51 at 98.4 kPa, becoming −0.39 at 98.5 kPa (Figure 13d). Total solar radiation showed a pronounced nonlinear positive transition. A value of 127 W m^{-2} corresponded to a strongly negative SHAP contribution and 248 W m^{-2} to −0.19, after which the contribution generally increased with radiation and reached +0.49 at 976 W m^{-2} (Figure 13g). This comparatively large SHAP range is consistent with the 27.5% overall importance of total solar radiation in the final model.

Wind and particulate-matter responses in the Square also differed. Wind-speed contribution declined from +0.05 at 0 m s^{-1} and +0.03 at 0.9 m s^{-1} to −0.17 at 2.1 m s^{-1} (Figure 13a). Wind direction contributed −0.18 at 150° and became positive around 313–328° (Figure 13b). $PM_{2.5}$ shifted generally from positive to negative with increasing concentration and reached −0.21 at 20 μg m^{-3} (Figure 13e), whereas PM_{10} remained close to zero

throughout (Figure 13f). These relationships describe final-model responses and do not establish direct thermal effects of wind or particulate matter.

5. Discussion

5.1. Contrasting Daytime Thermal-Humidity States Across Historical-Space Settings

The daytime observations revealed different thermal-humidity states on Main Street and in the Square despite broadly similar background-weather periods. Measured mean air temperature and relative humidity were 32.81 °C and 57.62% on Main Street, compared with 32.01 °C and 59.65% in the Square. The reconstructed Alley scenario was cooler and more humid. Because sky-view factor, street-canyon aspect ratio, building height, and surface thermal properties were not measured directly, these contrasts may be associated with local spatial conditions but cannot demonstrate causal effects of specific morphological parameters [41].

The contrast between Main Street and Square is particularly informative. Median total solar radiation was 396 W m^{-2} in the Square, substantially higher than the 129 W m^{-2} measured on Main Street, yet neither mean air temperature nor THI was higher in the Square. Thus, within this observation window, greater incident radiation did not correspond one-to-one with the highest near-surface air temperature [42]. Open space may combine stronger radiant exposure with different air-exchange conditions [43], while street environments are additionally influenced by heat storage, shading, activity intensity, and ventilation pathways [44]. Because these processes were not measured independently, they are treated here as testable mechanistic hypotheses rather than demonstrated explanations.

Measured differences in the wind environment further indicate that the local background was not identical between settings [45]. Median wind speed was 1.00 m s^{-1} on Main Street and its wind directions were relatively concentrated, whereas directions in the Square were more dispersed. The reconstructed Alley scenario was intentionally assigned lower-wind characteristics and therefore cannot independently validate the field proposition that narrow alleys are less ventilated [46]. This morphological hypothesis requires repeated monitoring at multiple real Alley sites.

Spatial differences in noise and particulate matter are more plausibly interpreted as differences in environmental context than as direct thermal forcing. Median measured noise, $PM_{2.5}$, and PM_{10} were all higher on Main Street than in the Square. Such variables may encode human activity, pollution sources, and temporal structure simultaneously; even when they carry predictive information in XGBoost-SHAP, they should not be described directly as physical causes of THI variation [29].

5.2. THI Deviations Were Dominated by Temperature, Whereas Humidity Contributions Were Time-Dependent

THI compresses joint variation in air temperature and relative humidity into a single thermal-humidity metric [16]. In the measured series, mean THI was 28.46 °C on Main Street and reached a monitoring-window maximum of 30.56 °C, whereas the Square averaged 28.04 °C; the reconstructed Alley scenario averaged 27.73 °C. Exact Shapley decomposition shows that interpretation of THI should move beyond absolute temperature and humidity values to consider the direction and magnitude of their contributions relative to a common reference state [20].

Within the present reference state and data range, mean absolute temperature contribution shares were 74.20%, 72.46%, and 71.70% for Alley, Main Street, and Square, compared with 25.80%, 27.54%, and 28.30% for relative humidity. Under the present THI equation and reference definition, temperature variation therefore generated the larger share of THI deviation. However, Alley is reconstructed and the reference state was calculated jointly from all three series, so these proportions cannot be interpreted as universal physiological weights derived from independent field observations.

The temporal decomposition showed that neither contribution was constant. Early in the monitoring period, temperature was generally below the common reference and therefore contributed negatively to THI, while higher relative humidity contributed positively. As daytime temperature rose and relative humidity declined, the temperature contribution became positive and the humidity contribution shifted toward negative values. Main Street showed the strongest late-period contrast at 14:41, when temperature and humidity contributions reached 4.397 °C and −2.005 °C, respectively.

Humidity is therefore not an invariant positive term in this THI formulation; its sign depends on the simultaneous temperature-humidity state and the common reference. Lower afternoon humidity can algebraically offset part of the THI increase associated with warming. This ‘offset’ refers strictly to the mathematical decomposition of the THI equation and does not imply that lower humidity has an equivalent effect across all human thermal-comfort or health contexts [47].

Accordingly, the approximately 72–74% temperature share is specific to the monitoring window, THI formulation, and chosen reference state and should not be generalized as a physiological weighting of temperature versus humidity. Studies aimed at human thermal comfort should additionally incorporate mean radiant temperature, wind speed, clothing, and metabolic rate and apply more comprehensive comfort indices [19,48].

5.3. XGBoost-SHAP Revealed Setting-Specific Nonlinear Predictive Associations

Grouped blocked out-of-fold validation yielded R^2 values of 0.80 and 0.73 and RMSE values of 0.47 and 0.58 for the measured Main Street and Square models, respectively; the reconstructed Alley scenario yielded R^2 = 0.57 and RMSE = 0.73. Residuals were distributed on both sides of zero, but these metrics should not be interpreted as fully independent temporal generalization. GroupKFold prevents records within the same approximately 25-min block from appearing simultaneously in training and test sets, yet it is not strict forward temporal validation and adjacent blocks may remain correlated [35]. In particular, Alley performance reflects only predictability internal to the reconstructed data.

Atmospheric pressure received high SHAP importance in all three final models, accounting for 38.2%, 55.3%, and 47.4% in Alley, Main Street, and Square, respectively, and dependence curves generally decreased with increasing pressure. This result requires particular caution. Pressure varied over an extremely narrow and discrete range during the approximately 6-h monitoring period and may therefore encode both background meteorology and temporal progression. Its high SHAP importance is better interpreted as that of an informative proxy than as evidence that small pressure changes directly dominated THI [29].

This distinction illustrates the fundamental difference between machine-learning importance and causal effect. SHAP quantifies how variables modify model predictions under the covariance structure of the present data and fitted model; it does not establish counterfactual causal relationships [30]. Moreover, the SHAP values analysed here were obtained from final models refitted to the full samples rather than independently from each outer test fold. They therefore explain the structure of the final models and cannot substitute for out-of-fold effect estimation.

Beyond pressure, the two measured settings showed different secondary predictive structures [49]. On Main Street, total solar radiation and wind speed accounted for 14.6% and 13.9%, respectively, whereas in the Square total solar radiation accounted for 27.5% and wind direction for 10.1%. The higher radiation distribution in the Square is internally consistent with its larger radiation SHAP importance, but this remains a predictive association rather than an independent test of a causal radiation effect. In the reconstructed Alley scenario, noise and total solar radiation accounted for 22.0% and 21.1%, respectively, and can only be regarded as exploratory model responses to the reconstruction.

SHAP dependence relationships further suggested nonlinearities and directional transitions [50]. For example, the smoothed Main Street wind-speed contribution changed from negative at 0.5 m s^{-1} to positive at 1.0–1.2 m s^{-1}, while Square solar radiation shifted from negative contribution at low values toward positive contribution at high values. Because environmental variables co-varied with time and both the LOWESS curves and bootstrap envelopes were estimated from limited samples, these turning regions should be treated as candidate response ranges for future repeated observations rather than established physical thresholds.

Overall, the evidence more strongly supports the limited conclusion that different spatial settings possess different predictive association structures than any universal control of historical-district THI by a single variable. The measured Main Street and Square models already differ in their secondary predictor rankings; Alley requires genuine repeated field measurements before it can enter a spatial comparison at the same evidential level.

Atmospheric pressure varied over a narrow range during the monitoring period, and its high SHAP importance may partly reflect temporal covariation or measurement-related variability. A model refitted without atmospheric pressure was not evaluated in this study. Consequently, the present results do not establish how model performance or predictor rankings would change after pressure exclusion. The importance of radiation and wind reported here refers only to the fitted models that include pressure and should not be interpreted as evidence from a pressure-exclusion sensitivity analysis. Such an analysis is needed to assess the robustness of these predictive associations.

5.4. Implications for Microclimate-Sensitive Adaptation in Historical Districts

Although the evidence derives from a single-day case study, it provides testable spatial hypotheses for future microclimate adaptation in historical districts. Individual spatial units should not be diagnosed solely from a district-wide mean thermal environment; their principal radiation exposure, ventilation conditions, and thermal-humidity states should first be identified [51].

For open squares, the combination of higher solar radiation and greater model importance of radiation suggests that shading interventions merit priority field testing. Potential measures include tree canopy, reversible

lightweight shading, and continuous shaded pedestrian routes, provided that heritage character and existing ventilation pathways are protected. This planning implication follows from consistency between variable distributions and predictive associations, not from a demonstrated causal intervention effect.

On Main Street, relatively high mean temperature and THI co-occurred with higher wind speed, indicating that ventilation alone cannot be equated with a lower thermal-humidity state. Future intervention assessments should jointly consider shading, surface heat storage, continuity of airflow, and human activity. For narrow alleys, valid field measurements are currently unavailable; maintaining shade while avoiding obstruction of existing airflow can therefore be proposed only as a hypothesis for testing, not as a deterministic design recommendation derived from the reconstructed series.

Methodologically, treating a historical district as a mosaic of microclimatic units rather than a homogeneous space can support more targeted field sampling and adaptation assessment. Before these findings are used for planning decisions, independent sites within each spatial category and repeated observations across multiple hot days are required to test the stability of the reported spatial contrasts and nonlinear responses.

5.5. Limitations and Boundaries of Interpretation

First, the field observations cover only approximately six daytime hours on 4 August 2026, with 75 valid records at each of the two measured sites. They do not represent a complete diurnal cycle or establish conditions across other days, seasons, or weather regimes. Each setting is represented by one measured site, and repeated records within a site are not independent spatial replicates. Concurrent sampling limits between-day differences but does not eliminate temporal confounding or isolate causal effects of urban morphology. The reported contrasts and predictive associations are therefore specific to the sampled sites and monitoring window. Repeated monitoring across independent sites and multiple dates is needed to assess their generalizability.

Second, the original Alley field data were unusable because of equipment failure, and a stochastic reconstruction was used to preserve the completeness of the analytical framework. Reconstructed data cannot substitute for observations and cannot validate actual Alley microclimate, statistical significance, or model generalization. Alley field measurements must be repeated before any confirmatory three-setting comparison. If remeasurement is impossible, Alley results should remain explicitly separated from Main Street and Square observations as exploratory or sensitivity material.

Third, temporal dependence may still affect both statistical and machine-learning evaluation. Grouping consecutive observations into approximately 25-min blocks before GroupKFold reduces direct leakage within blocks, but adjacent blocks may remain correlated and training folds may include blocks occurring after a test block. Out-of-fold R^2, RMSE, and MAE therefore quantify grouped predictability within the present window rather than strict forecasting into future time. Friedman, Wilcoxon, Spearman, and wind-direction contingency analyses are similarly constrained by serial autocorrelation and the lack of spatial replication and should be interpreted exploratorily.

Fourth, SHAP dependence relationships describe the predictive structure of final full-sample models rather than causal effects, and the ordinary bootstrap envelopes around LOWESS curves do not explicitly preserve temporal blocks. Atmospheric pressure, noise, particulate matter, and apparent turning points in dependence curves may therefore partly encode temporal progression or unobserved covariates. Future studies should use strict forward temporal validation, aggregate SHAP values from outer test folds, and apply block bootstrap methods to assess the stability of variable rankings and nonlinear responses.

Fifth, several candidate indicators were excluded for quality reasons: rainfall had no variation, the carbon-dioxide field exactly duplicated $PM_{2.5}$, elevation showed implausible short-term drift, and illuminance was highly redundant with total solar radiation. In addition, sky-view factor, street-canyon aspect ratio, building height, surface materials, shading fraction, mean radiant temperature, and pedestrian/traffic intensity were not quantified concurrently, limiting direct attribution of differences to urban morphology and heat-exchange mechanisms. Future work should integrate these morphological and radiative parameters with repeated multi-day, multi-site monitoring to distinguish physical mechanisms from temporal proxy variables.

Future work should evaluate the robustness of the reported associations using stricter temporal validation, such as forward-chaining or a contiguous temporal holdout. This is needed to assess sensitivity to temporal dependence even when the intended use is exploratory interpretation rather than operational forecasting.

6. Conclusions

This study developed an exploratory framework linking temperature-humidity dynamics, THI, exact Shapley decomposition, and XGBoost-SHAP interpretation using measured Main Street and Square series and a reconstructed Alley scenario in the Ciqikou Historical District, Chongqing. During daytime monitoring on 4

August 2026, measured mean air temperature and THI were 32.81 °C and 28.46 °C on Main Street, exceeding the corresponding Square values of 32.01 °C and 28.04 °C; the reconstructed Alley scenario yielded 31.16 °C and 27.73 °C.

Under the shared reference derived from the measured and reconstructed series, temperature accounted for 71.70–74.20% of the absolute Shapley contributions, compared with 25.80–28.30% for relative humidity. Temperature contributions generally changed from negative early in the monitoring period to positive in the afternoon, while humidity contributions changed in the opposite direction. Because the reconstructed Alley series contributes to the reference, these proportions are reference-dependent methodological results and are not estimates based exclusively on field observations or universal weights for human thermal comfort.

Under grouped out-of-fold validation, R^2 was 0.80 for the measured Main Street model and 0.73 for the measured Square model, compared with 0.57 for the reconstructed Alley scenario. SHAP results showed that predictor rankings and nonlinear responses differed across settings. Although atmospheric pressure received high importance in all three final models, it more plausibly carried information about background meteorology and temporal progression during the monitoring window and should not be interpreted as direct causal control. Total solar radiation and the wind environment showed more setting-specific predictive contributions across the two measured spaces.

Overall, this study supports an analytical strategy that combines spatial-unit differentiation, time-resolved assessment, and interpretable modelling for thermal-environment research in historical districts, while the present evidence remains limited to a single-day case and exploratory predictive associations. Generalizable conclusions about spatial types will require new field measurements in a real Alley, independent replication across multiple sites, dates, and seasons, and concurrent measurement of key physical variables including sky-view factor, street-canyon geometry, shading, surface materials, and mean radiant temperature.

Supplementary Materials

The additional data and information can be downloaded at: https://media.sciltp.com/articles/others/2610200945416993/REM-26080197-SI.pdf.

Author Contributions

X.S.: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing—original draft preparation; Z.W.: Conceptualization, Methodology, Formal analysis, Visualization, Writing—original draft preparation, Writing—review & editing; S.S.: Investigation, Data curation, Formal analysis, Visualization, Writing—review & editing; Z.T.: Supervision, Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Data derived from the current study can be provided to readers upon request.

Acknowledgments

The authors gratefully acknowledge Zihan Dai, Qiulin Zhan, Xibing Chen, and Linsen Huang for their valuable assistance with the field investigation and data collection.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) solely for language polishing and improving readability. All AI-assisted content was reviewed and revised by the authors, who take full responsibility for the final manuscript.

References

1. Tuholske, C.; Caylor, K.; Funk, C.; et al. Global urban population exposure to extreme heat. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2024792118. https://doi.org/10.1073/pnas.2024792118.
2. Li, H.; Li, Y.; Wang, T.; et al. Quantifying 3D building form effects on urban land surface temperature and modeling seasonal correlation patterns. *Build. Environ.* **2021**, *204*, 108132. https://doi.org/10.1016/j.buildenv.2021.108132.
3. Yu, J.; Hu, Q.; Li, J. Combining visual intelligence and social-physical urban features facilitates fine-scale seasonality characterization of urban thermal environments. *Build. Environ.* **2024**, *266*, 112088. https://doi.org/10.1016/j.buildenv.2024.112088.
4. Ma, X.; Leung, T.M.; Chau, C.K.; et al. Analyzing the influence of urban morphological features on pedestrian thermal comfort. *Urban Clim.* **2022**, *44*, 101192. https://doi.org/10.1016/j.uclim.2022.101192.
5. Li, Y.; Huang, N.; He, J. Analytical evaluation of thermal comfort in the pedestrian environment using pedestrian shade space distribution. *Urban Clim.* **2023**, *51*, 101665. https://doi.org/10.1016/j.uclim.2023.101665.
6. Qi, L.; Li, T.; Chang, B.; et al. Research on Thermal Comfort Evaluation and Optimization of Green Space in Beijing Dashilar Historic District. *Buildings* **2024**, *14*, 3121. https://doi.org/10.3390/buildings14103121.
7. Zheng, J.; Ou, Z.; Xiang, Y.; et al. How can street interface morphology effect pedestrian thermal comfort: A case study of the old town of Changsha, China. *Urban Clim.* **2025**, *60*, 102341. https://doi.org/10.1016/j.uclim.2025.102341.
8. Cheng, R.; Wang, S.; Gao, W.; et al. Evaluation of courtyard outdoor thermal comfort: A case study of Qingdao Riyuan historic district in hot summer. *Case Stud. Therm. Eng.* **2025**, *74*, 106804. https://doi.org/10.1016/j.csite.2025.106804.
9. Behzad, Z.; Guilandoust, A. Enhancing outdoor thermal comfort in a historic site in a hot dry climate (Case study: Naghsh-e-Jahan Square, Isfahan). *Sustain. Cities Soc.* **2024**, *102*, 105209. https://doi.org/10.1016/j.scs.2024.105209.
10. Lai, B.; Fu, J.M.; Guo, C.K.; et al. Street Geometry Factors Influencing Outdoor Pedestrian Thermal Comfort in a Historic District. *Buildings* **2025**, *15*, 613. https://doi.org/10.3390/buildings15040613.
11. Cherchi, G.; Santus, A.; Fiorino, D.R.; et al. Microclimate Assessment and Outdoor Human Comfort Enhancement of a Historic Village in Sardinia, Italy. *Urban Plan.* **2025**, *10*, 10708. https://doi.org/10.17645/up.10708.
12. Wang, Y.; Liu, X.; Xie, R.; et al. Thermal comfort and urban microclimate response: A new thermal environment assessment model for waterfront spaces in historic ancient towns. *Energy Build.* **2025**, *331*, 115393. https://doi.org/10.1016/j.enbuild.2025.115393.
13. Rezaie, P.; Lopez-Cabeza, V.P.; Sola-Caraballo, J.; et al. Cooling Heritage Scenarios: Transforming Historic Squares for Thermal Comfort. *Buildings* **2025**, *15*, 564. https://doi.org/10.3390/buildings15040564.
14. Yin, S.; Wang, F.; Xiao, Y.; et al. Comparing cooling efficiency of shading strategies for pedestrian thermal comfort in street canyons of traditional shophouse neighbourhoods in Guangzhou, China. *Urban Clim.* **2022**, *43*, 101165. https://doi.org/10.1016/j.uclim.2022.101165.
15. Haeri, T.; Hassan, N.; Ghaffarianhoseini, A. Evaluation of microclimate mitigation strategies in a heterogenous street canyon in Kuala Lumpur from outdoor thermal comfort perspective using Envi-met. *Urban Clim.* **2023**, *52*, 101719. https://doi.org/10.1016/j.uclim.2023.101719.
16. Yang, Y.; Lee, X. Contributions of temperature and humidity to intra-city variations in humid heat. *Urban Clim.* **2024**, *58*, 102201. https://doi.org/10.1016/j.uclim.2024.102201.
17. Ioannou, L.G.; Mantzios, K.; Tsoutsoubi, L.; et al. Indicators to assess physiological heat strain—Part 1: Systematic review. *Temperature* **2022**, *9*, 227–262. https://doi.org/10.1080/23328940.2022.2037376.
18. Ji, Y.; Song, J.; Shen, P. A review of studies and modelling of solar radiation on human thermal comfort in outdoor environment. *Build. Environ.* **2022**, *214*, 108891. https://doi.org/10.1016/j.buildenv.2022.108891.
19. Mahdavinejad, M.; Shaeri, J.; Nezami, A.; et al. Comparing universal thermal climate index (UTCI) with selected thermal indices to evaluate outdoor thermal comfort in traditional courtyards with BWh climate. *Urban Clim.* **2024**, *54*, 101839. https://doi.org/10.1016/j.uclim.2024.101839.
20. Shapley, L.S. A Value for n-Person Games. In *Contributions to the Theory of Games (AM-28), Volume II*; Princeton University Press: Princeton, NJ, USA, 1953; pp. 307–318. https://doi.org/10.1515/9781400881970-018.
21. Roy, T.B.; Middey, A.; Krupadam, R.J. Unveiling the microclimate: A comprehensive review of tools, techniques, and future directions for sustainable cities. *Build. Environ.* **2025**, *274*, 112726. https://doi.org/10.1016/j.buildenv.2025.112726.
22. Kousis, I.; Pigliautile, I.; Pisello, A.L. Intra-urban microclimate investigation in urban heat island through a novel mobile monitoring system. *Sci. Rep.* **2021**, *11*, 9732. https://doi.org/10.1038/s41598-021-88344-y.

23. Ahmad, M.; Panyametheekul, S.; Thaveevong, P.; et al. Long-term monitoring of PM2.5 and PM10: Implications for air quality and public health in urban Bangkok, Thailand. *Environ. Chall.* **2025**, *21*, 101312. https://doi.org/10.1016/j.envc.2025.101312.
24. Samad, N.; Farnaghi, M.; Ostermann, F.O.; et al. A systematic literature review of machine learning and deep learning for urban heat island modelling. *Discov. Cities* **2026**, *3*, 43. https://doi.org/10.1007/s44327-026-00223-1.
25. He, J.; Shi, Y.; Xu, L.; et al. Exploring the scale effect of urban thermal environment through XGBoost model. *Sustain. Cities Soc.* **2024**, *114*, 105763. https://doi.org/10.1016/j.scs.2024.105763.
26. Chen, T.; Guestrin, C. XGBoost. In Proceedings of the KDD'16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. https://doi.org/10.1145/2939672.2939785.
27. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30, pp. 4765–4774.
28. Guo, R.; Yang, B.; Guo, Y.; et al. Machine learning-based prediction of outdoor thermal comfort: Combining Bayesian optimization and the SHAP model. *Build. Environ.* **2024**, *254*, 111301. https://doi.org/10.1016/j.buildenv.2024.111301.
29. Molnar, C.; König, G.; Herbinger, J.; et al. General Pitfalls of Model-Agnostic Interpretation Methods for Machine Learning Models. In *xxAI—Beyond Explainable AI*; Springer International Publishing: Cham, Switzerland, 2022; Volume 13200, pp. 39–68. https://doi.org/10.1007/978-3-031-04083-2_4.
30. Aas, K.; Jullum, M.; Løland, A. Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. *Artif. Intell.* **2021**, *298*, 103502. https://doi.org/10.1016/j.artint.2021.103502.
31. Roberts, D.R.; Bahn, V.; Ciuti, S.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. https://doi.org/10.1111/ecog.02881.
32. Fisher, N.I. *Statistical Analysis of Circular Data*; Cambridge University Press: Cambridge, UK, 1993. https://doi.org/10.1017/cbo9780511564345.
33. Thom, E.C. The Discomfort Index. *Weatherwise* **1959**, *12*, 57–61. https://doi.org/10.1080/00431672.1959.9926960.
34. Hollander, M.; Wolfe, D.A.; Chicken, E. *Nonparametric Statistical Methods*, 3rd ed.; Wiley: Hoboken, NJ, USA, 2015. https://doi.org/10.1002/9781119196037.
35. Cerqueira, V.; Torgo, L.; Mozetič, I. Evaluating time series forecasting models: An empirical study on performance estimation methods. *Mach. Learn.* **2020**, *109*, 1997–2028. https://doi.org/10.1007/s10994-020-05910-7.
36. Sweet, L.B.; Müller, C.; Anand, M.; et al. Cross-Validation Strategy Impacts the Performance and Interpretation of Machine Learning Models. *Artif. Intell. Earth Syst.* **2023**, *2*, e230026. https://doi.org/10.1175/aies-d-23-0026.1.
37. Lundberg, S.M.; Erion, G.; Chen, H.; et al. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. https://doi.org/10.1038/s42256-019-0138-9.
38. Cleveland, W.S. Robust Locally Weighted Regression and Smoothing Scatterplots. *J. Am. Stat. Assoc.* **1979**, *74*, 829–836. https://doi.org/10.1080/01621459.1979.10481038.
39. Efron, B. Bootstrap Methods: Another Look at the Jackknife. *Ann. Stat.* **1979**, *7*, 1–26. https://doi.org/10.1214/aos/1176344552.
40. Lahiri, S.N. *Resampling Methods for Dependent Data*; Springer: New York, NY, USA, 2003. https://doi.org/10.1007/978-1-4757-3803-2.
41. Solanki, P.; Amirtham, L.R.; Deb, C. Optimizing sky view factor and vegetation to mitigate urban heat in hot-humid climates. *Discov. Cities* **2025**, *2*, 79. https://doi.org/10.1007/s44327-025-00108-9.
42. Zhang, H.; Ao, M.; Ardabili, N.G.; et al. Impact of urban sunken square design on summer outdoor thermal comfort using machine learning. *Urban Clim.* **2024**, *58*, 102214. https://doi.org/10.1016/j.uclim.2024.102214.
43. Mandić, L.; Đjukić, A.; Marić, J.; et al. A Systematic Review of Outdoor Thermal Comfort Studies for the Urban (Re)Design of City Squares. *Sustainability* **2024**, *16*, 4920. https://doi.org/10.3390/su16124920.
44. Zhong, H.; Feng, J.; Lam, C.K.C.; et al. The impact of semi-open street roofs on urban pollutant exposure and pedestrian-level thermal comfort in 2-D street canyons. *Build. Environ.* **2023**, *239*, 110387. https://doi.org/10.1016/j.buildenv.2023.110387.
45. Zou, J.; Yu, Y.; Liu, J.; et al. Field measurement of the urban pedestrian level wind turbulence. *Build. Environ.* **2021**, *194*, 107713. https://doi.org/10.1016/j.buildenv.2021.107713.
46. Nasrollahi, N.; Namazi, Y.; Taleghani, M. The effect of urban shading and canyon geometry on outdoor thermal comfort in hot climates: A case study of Ahvaz, Iran. *Sustain. Cities Soc.* **2021**, *65*, 102638. https://doi.org/10.1016/j.scs.2020.102638.
47. Zhou, Z.; Nguyen-Xuan, T.; Liao, H.; et al. Characterization of temperature and humidity effects on extreme heat stress under global warming and urban growth in the Pearl and Yangtze River Deltas of China. *Weather Clim. Extrem.* **2024**, *44*, 100659. https://doi.org/10.1016/j.wace.2024.100659.
48. Li, Y.; Fei, W.; Yang, M.; et al. How to realize large-scale outdoor thermal comfort studies? A systematic review based on OTC characterization, methods and research trends. *Front. Sustain. Cities* **2025**, *7*, 1552994. https://doi.org/10.3389/frsc.2025.1552994.

49. Jiang, J.; Li, Z.; Bedra, K.B.; et al. Predicting outdoor thermal comfort in traditional villages: An explainable machine learning framework integrating model optimization, seasonal variability, and tourist-resident insights. *Build. Environ.* **2025**, *282*, 113315. https://doi.org/10.1016/j.buildenv.2025.113315.
50. Zhang, M.; Zhou, D.; Xu, D.; et al. How do built environment factors influence urban heat islands across local climate zones? Evidence from an interpretable XGBoost-SHAP model. *Sustain. Cities Soc.* **2026**, *136*, 107078. https://doi.org/10.1016/j.scs.2025.107078.
51. Villaverde, A.; Álvarez, I.; Rojí, E.; et al. Categorisation of urban open spaces for heat adaptation: A cluster based approach. *Build. Environ.* **2024**, *263*, 111861. https://doi.org/10.1016/j.buildenv.2024.111861.