ISSN 2616-5775 Vol. 9, Issue 7: 1-8, DOI: 10.25236/AJCIS.2026.090701

Identifying the Optimal Mitigation Zone around the Main Urban Area via Feature Importance and Response Pattern Analysis

Guozhu Wang*

School of Information Science and Technology, Yunnan Normal University, Kunming, China
**Corresponding author:2324410012@ynnu.edu.cn*

Abstract: *The rapid urbanization process, marked by extensive land-use changes and built-up area expansion, has considerably aggravated the Urban Heat Island (UHI) effect, presenting a grave challenge to the ecological integrity of cities and the well-being of their inhabitants. Constrained by the scarcity of urban land resources, cooling strategies relying on internal greening face practical bottlenecks, whereas the peri-urban areas possess significant yet long-overlooked cooling potential. To address the core questions—"What are the key driving factors?" and "How should the optimization range be determined?"—this study conducts a systematic investigation. Taking Kunming, China, as the study area, this study integrates multi-source remote sensing data to construct an analytical framework consisting of spectral indices, landscape metrics, and land surface temperature (LST). Eight machine learning models are employed to model the driving factors of LST in the main urban area, and the SHapley Additive exPlanations (SHAP) method is used to identify key variables. Furthermore, a progressively accumulated concentric buffer zone strategy is adopted, and SHAP together with Accumulated Local Effects (ALE) is applied to evaluate the response patterns of key features to LST across different surrounding buffer ranges, thereby identifying the optimal mitigation zone with the greatest cooling potential. The results show that: (1) spectral indices exhibit a stronger influence on LST than landscape metrics, making them the dominant driving factors; (2) the peri-urban area extending outward from the boundary of the main urban area to the equivalent radius constitutes the optimal mitigation zone that most significantly affects the thermal environment of the main urban area.*

Keywords: *UHI, Key Driving Factors, SHAP, ALE, Peripheral Zones*

1. Introduction

The Urban Heat Island is a phenomenon arising from the context of rapid and intense urbanization, driven primarily by changes in land surface properties. Rapid urban development leads to a continuous increase in urban land use[1], accelerating drastic changes in land cover[2][3]. Related studies indicate that the extents of urban built-up areas, vegetation, and water bodies exhibit strong correlations with UHI[4][5], exerting significant impacts on the thermal environment[6]. The widespread replacement of green spaces and water bodies with impervious surfaces substantially reduces surface evapotranspiration, thereby exacerbating UHI and making it a critical issue threatening urban ecological security and resident health[7]. Therefore, scientific land planning and management play an important role in effectively regulating the urban thermal environment, enhancing urban adaptability and resilience, and promoting sustainable development[8][9].

In UHI research, the land surface temperature (LST) difference between urban and rural areas is the most commonly used indicator for studying UHI intensity[10]. Numerous manuscripts on urban thermal environments employ LST as a synonym for UHI[11][12], analyzing the correlation and spatial dependence between LST and land cover types[13]. Consequently, we use the relationship between land use and land surface temperature as an entry point to identify key influencing factors and further determine the optimal areas for improvement.

Existing studies primarily achieve cooling effects by increasing cooling elements[14][15]. However, urban land resources are highly scarce[4], and solely relying on new cooling sources to reduce UHI intensity faces significant constraints in realistic planning. In contrast, peri-urban areas typically possess a higher proportion of natural land cover[5][6][7], including forests, rivers, grasslands, and farmlands[16][17], along with substantial room for land type adjustment, offering great regulatory

ISSN 2616-5775 Vol. 9, Issue 7: 1-8, DOI: 10.25236/AJCIS.2026.090701

potential in spatial restructuring and ecological function optimization. Current literature has begun to recognize the advantages and significance of peri-urban areas in mitigating UHI[18][19]. Yao et al.[20] pointed out that the greening level of rural areas is a key driver of the diurnal variation in urban surface thermal island intensity, accounting for up to 22.5% of the variation. Yuan et al.[21] noted that the optimal layout of rural settlements can guide their sustainable development, and that different factors exhibit large differences in their influence on the suitability assessment and layout of rural settlements. This underscores the necessity of feature selection for drivers of the urban thermal environment based on their importance. Peng et al.[22] explored the patterns of continuous interannual variation in surface thermal island intensity, elucidated the potential physical mechanisms by which rural landscapes influence UHI changes, and indicated that mitigating UHI requires region-specific assessment thresholds tailored to diverse rural landscapes. However, these studies have not discussed the influence of peri-urban land cover on urban land surface temperature nor the response patterns of landscape parameters to UHI. Therefore, based on identifying the main driving factors, further determining the peri-urban areas with the highest cooling potential is of great significance for addressing urban high-temperature hazards from a rural perspective and promoting coordinated urban-rural development.

2. Research Methods

This study aims to identify the most promising optimization zones within the peri-urban areas of the central city. The research comprises the following components: (1) assess the influence magnitude of different characteristics on the UHI of the central urban area, screen out the primary characteristics, and perform a spatial heterogeneity analysis on them; (2) analyze the response patterns between the primary characteristics and land surface temperature, thereby identifying the areas with the greatest cooling potential as the optimal optimization zones.

2.1 Study Area

Kunming City (24°23′–26°33′ N, 102°10′–103°40′ E) is located in the central part of the Yunnan-Guizhou Plateau and serves as the core of the Central Yunnan city cluster. Its central urban area is situated within the fault-bound basin of Dianchi Lake. In 2022, the built-up area of Kunming was 460 km², compared with 22 km² in 1978, representing an approximately 20-fold increase over more than 40 years. Notably, following the official establishment of the Chenggong New District, the built-up area expanded by 115.3 km². In recent years, urban land use has become more intensive; however, temperatures within the urban area and its surroundings have continued to rise over the past decade. Therefore, the central urban area of Kunming is selected as the study area.

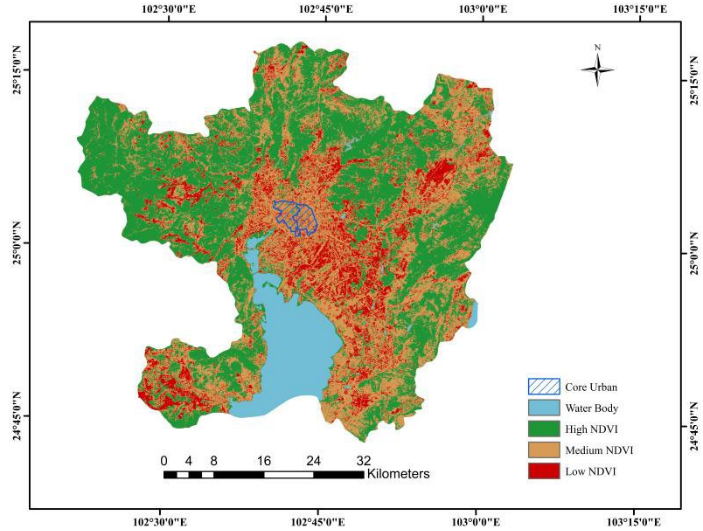

Figure 1: Study area

2.2 Data Sources and Feature Extraction

The land cover imagery uses the 2024 Landsat 8 image, clipped to the area shown in Figure 1, where the red boundary represents the central urban area. The image consists of 2175×2373 cells, each

ISSN 2616-5775 Vol. 9, Issue 7: 1-8, DOI: 10.25236/AJCIS.2026.090701

measuring 30m×30m. The initial feature set comprises 15 variables: NDVI, NDBI, MNDWI, NP, LPI, LSI, CONTAG, PLADJ, COHESION, DIVISION, SPLIT, PR, SHDI, SHEI, and AI.

2.3 Cumulative Expanding Zonal Partitioning of the Peri-Urban Area

To encapsulate the thermal environment gradient around the central urban area into quantifiable and comparable statistical samples, this study proposes a "cumulative expanding buffer ring" partitioning strategy. Using the continuous mask of the central urban area as the seed region, the calculation starts from its boundary. The maximum influence range is set as the equivalent circle diameter of the city (approximately 5.4 km), providing a physical boundary for the subsequent regression-based interpretation framework that uses "rings as samples." This zonal range is further divided equidistantly according to the 30 m pixel resolution into 184 concentric rings (Ring 0 to Ring 183).

2.4 Model Fitting and Explanatory Power Assessment for Different Cumulative Zones

To evaluate the influence intensity of various characteristics on LST variation in the central urban area across different cumulative zones, a regression model is constructed within each cumulative zone, and its explanatory power is measured using the R^2 metric. R^2 is a parameter reflecting the goodness of fit of a regression model, representing the proportion of LST variance jointly explained by the relevant feature variables within the subzone. Related studies have used R^2 as an important basis for feature importance screening.

In this study, eight regression models (Ridge, BayesianRidge, Huber, RF, ExtraTrees, KNN, GBDT, and HistGB) are used to train the dataset. These models differ significantly in their training approaches and are recognized for their ability to handle high-dimensional data. During sample construction, each ring is treated as an independent observational unit. For each ring, the mean values of the 15 features (taken as independent variables) and the mean LST of all pixels within that ring (taken as the dependent variable) are calculated. To prevent model overfitting, the ring samples are randomly split into training and validation sets at a ratio of 7:3. Ultimately, each individual ring yields eight different R^2 values derived from the eight independent regression models. By comparing these R^2 values and selecting the highest among them, this maximum R^2 is considered the maximum explainable potential of the peri-urban area on LST under different modeling conditions, termed the "explanatory power." Subsequent feature contribution analyses are based on the model corresponding to the maximum R^2.

2.5 SHAP-Based Feature Contribution Calculation and Key Feature Screening Using the Optimal Model

After identifying the optimal regression model, this study employs the SHAP method to systematically evaluate the marginal contribution of each feature to the predicted mean LST of each ring, thereby revealing the relative importance of each driving factor in the formation of the urban thermal environment. First, a local interpretation of feature contributions is constructed for each ring sample, i.e., the contribution of each feature in that sample to the difference between the model's predicted value and the baseline value is calculated. Subsequently, the weighted average across all samples yields the average SHAP value for each feature over the entire dataset.

2.6 Analysis of the Response Patterns between Features and Land Surface Temperature

To deeply reveal the nonlinear response mechanisms of key parameters to LST variations in the central urban area, this study introduces ALE to deconstruct the effects and identify patterns of the established optimal model. This approach is used to examine the response patterns between key features and the mean LST within cumulative zones (including all pixels inside each ring's outer boundary). The ALE curves allow for the intuitive identification of the following critical information: (1) Direction of influence: Whether the curve lies entirely above or below the zero axis directly indicates a warming or cooling effect of the feature on LST; (2) Response morphology: The monotonicity or nonlinear transitions of the curve reflect the variation pattern of the influence intensity; (3) Effect intensity: The slope changes of the curve over different value intervals quantify the magnitude of temperature change induced by a unit change in the feature.

ISSN 2616-5775 Vol. 9, Issue 7: 1-8, DOI: 10.25236/AJCIS.2026.090701

3. Results and Analysis

3.1 Optimal Model

For the complete set of 184 cumulative rings, the training and validation sets were randomly partitioned at a ratio of 7:3. Among all models, ExtraTrees exhibited the best performance, achieving an R^2 value of 0.99 on the validation set, which is consistently superior to the other models and demonstrates robust explanatory power. As shown in Figure 2, the ExtraTrees model adequately explains the variance in the cumulative LST rings, providing a reliable model foundation for subsequent SHAP and ALE interpretations.

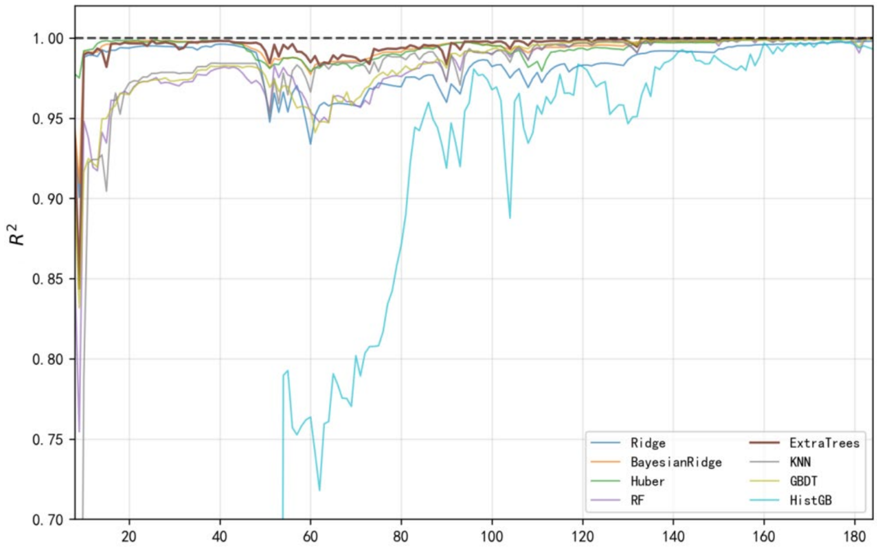

Figure 2: Performance Comparison of Eight Machine Learning Models

3.2 Screening and Distribution Characteristics of Key Driving Variables

Based on the ExtraTrees model, the average SHAP value of each feature across all rings was calculated, as shown in Figure 3. Ranked by the magnitude of these values, the three most influential features are NDBI, NDVI, and MNDWI, which constitute the primary driving factors. The univariate contribution values of the remaining 12 landscape structure features are substantially lower than those of the spectral features, placing them in a clearly secondary position. Among these secondary features, NP, CONTAG, and SHEI exhibit the highest influence levels.

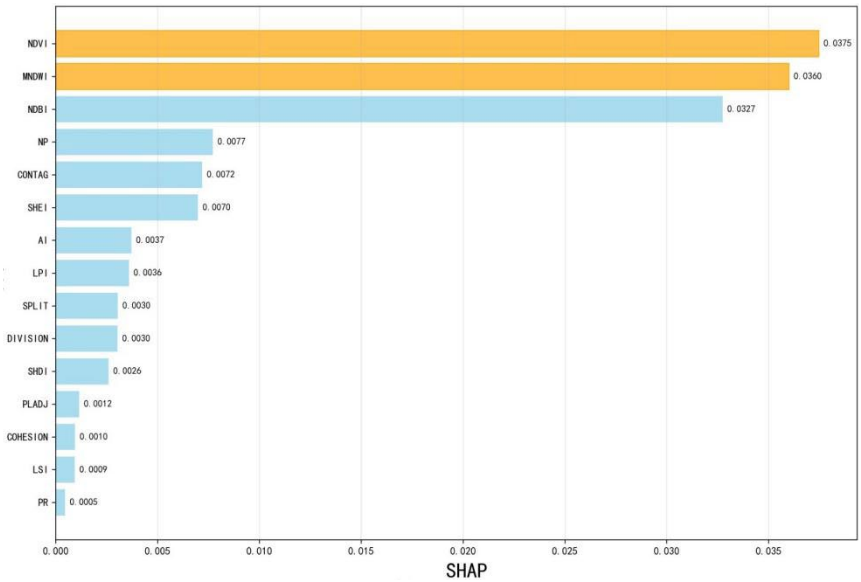

Figure 3: SHAP Mean Value Ranking of Feature Importance

ISSN 2616-5775 Vol. 9, Issue 7: 1-8, DOI: 10.25236/AJCIS.2026.090701

3.3 Trend Analysis of Feature Contributions across Cumulative Rings

The 184 valid rings were arranged spatially, and SHAP interpretation from the optimal ExtraTrees model was deconstructed to obtain SHAP-Ring curves for each major feature (Figure 4).

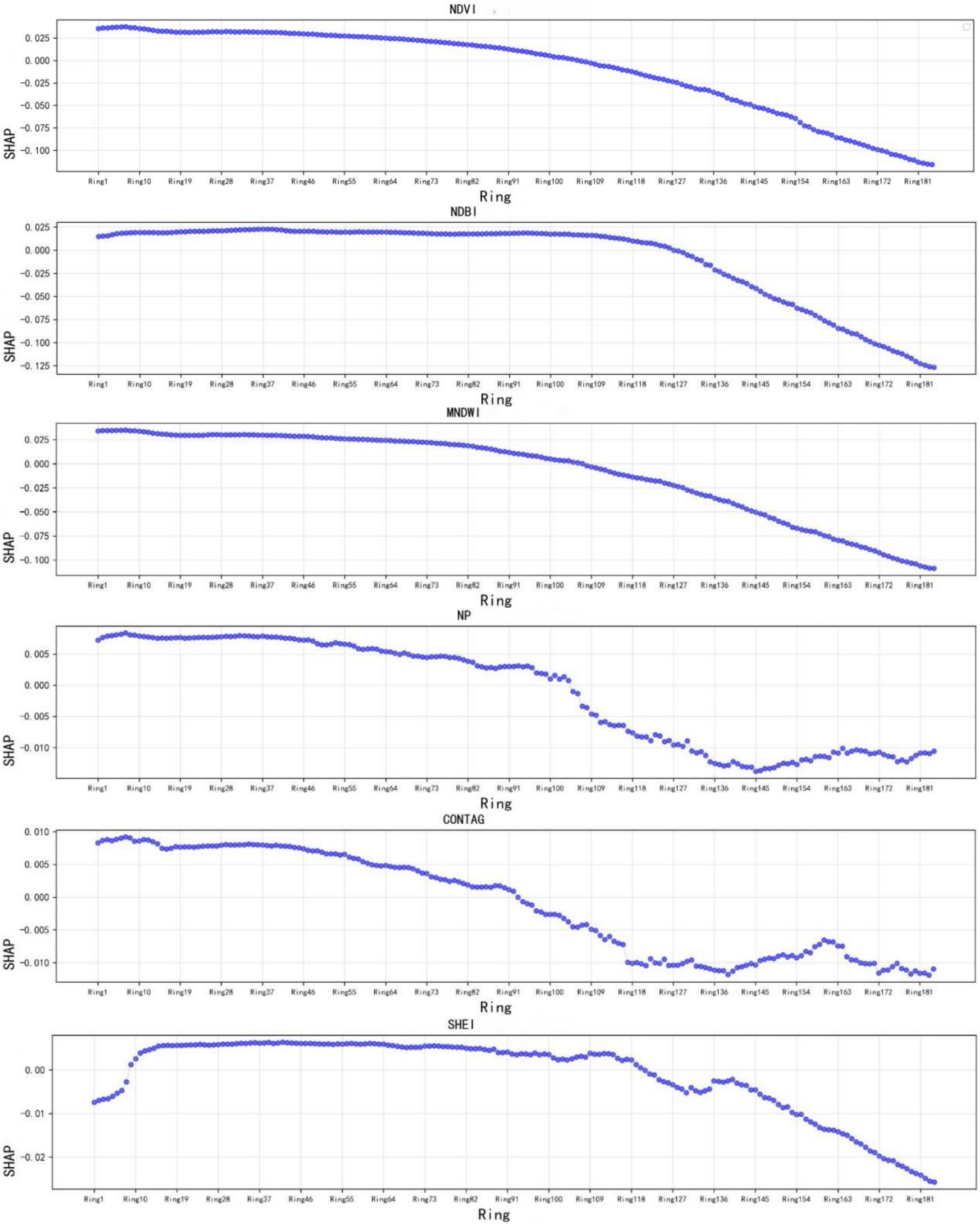

Figure 4: Dynamic Curves of SHAP Values for Major Features with Expanding Buffer Zones

The NDVI contribution curve shows a significant downward trend. The mean SHAP remains positive from Ring 0 to Ring 100, transitions to negative between Ring 100 and Ring 110, and declines notably after Ring 110, with the decrease magnitude expanding outward to the study area boundary. This indicates that, under the original peri-urban configuration, rings near the central urban area fall within a "positive contribution interval": vegetation cover has not reached the effective cooling threshold, or a stable evaporative cooling network has not been established due to high imperviousness and heat diffusion. Thus, within Ring 0–Ring 100, NDVI exhibits the largest adjustable space rather

ISSN 2616-5775 Vol. 9, Issue 7: 1-8, DOI: 10.25236/AJCIS.2026.090701

than the strongest cooling effect. Beyond Ring 100, SHAP values gradually turn negative, indicating that peripheral green spaces begin to exert actual cooling.

The contribution trends of MNDWI and NDVI are generally consistent. NDBI begins to decline around Ring 110, turns negative after Ring 120, and remains negative to the boundary, suggesting that the positive contribution ranges of NDVI and MNDWI near the central urban area are limited to the proximal zone, with actual cooling emerging only beyond a certain distance.

Within the positive contribution interval of the main effects (Ring 0–Ring 90), the three landscape indices (NP, CONTAG, SHEI) all exhibit positive contributions. After Ring 90, these effects gradually turn negative. Their SHAP values remain below 0.01, while those of the three spectral indices peak at approximately 0.025. Given the physical meanings of landscape indices and the general consistency of their contribution curves with those of spectral indices, a close relationship exists between landscape and spectral features in driving LST variations, indicating a synergistic relationship between land cover types and their spatial patterns.

Overall, the contribution structure exhibits a distinct primary–secondary two-layer hierarchy. Spectral indices show significantly stronger explanatory power for LST than landscape indices, indicating that thermal regulation depends first on the functional and quantitative attributes of land surface cover—i.e., specific cover types and their abundances. However, landscape indices are not marginal; thermal regulation is also synergistically driven by spatial landscape structure and morphological integration. Although their SHAP magnitude is less than half that of primary features, they provide stable contributions throughout the spatial accumulation process. Thus, landscape structure variables do not directly alter heat source intensity but secondarily modulate primary functional effects by regulating heat transfer pathways and efficiency. Consequently, focusing solely on individual cover types or area-based optimization while neglecting spatial integration of cooling types may fail to achieve expected systemic cooling outcomes.

3.4 ALE Response Patterns of Individual Features

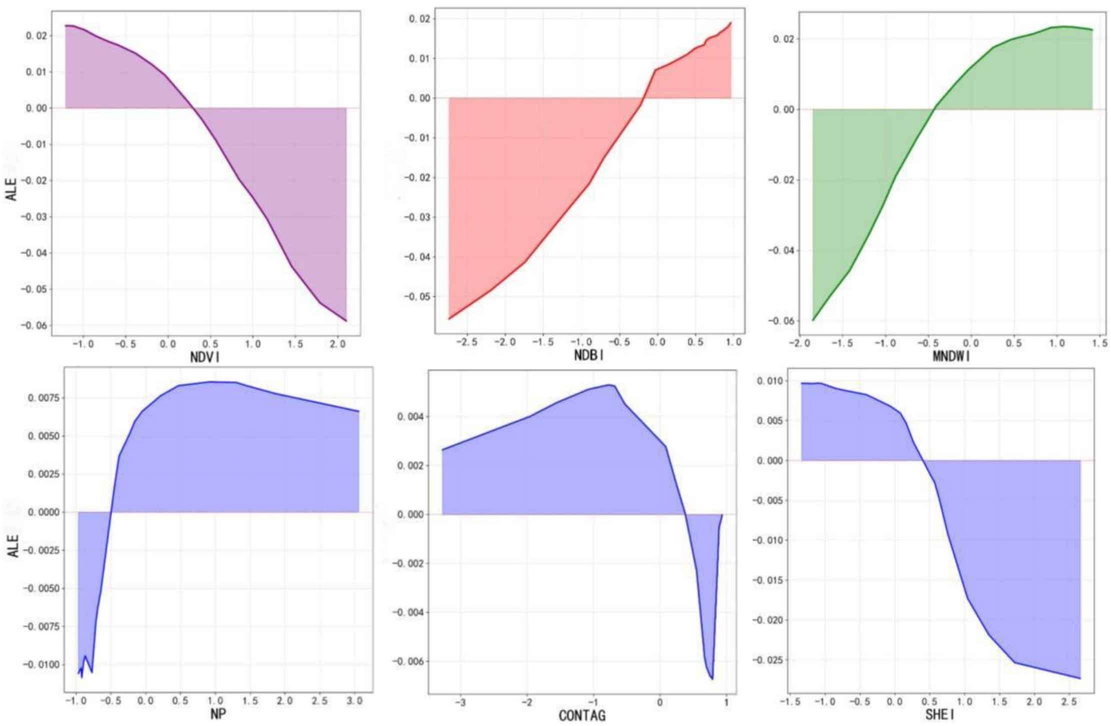

Figure 5: ALE Curves of Major Features

As shown in Figure 5, the NDVI ALE curve is flat at both ends and steepest in the middle, showing an overall decreasing LST trend with increasing NDVI, indicating cooling effects with diminishing marginal returns. The NDBI curve shows the opposite trend, with LST increasing as NDBI rises, indicating a warming effect. The MNDWI curve follows the same trend as NDBI, primarily because the buffer zone lacks true water bodies, with mean MNDWI values falling in the negative range, and pixels with highly negative MNDWI values mostly corresponding to non-water land cover. Thus, this warming pattern reflects a statistical relationship rather than actual warming by water bodies, and this effect rapidly exhibits diminishing returns as MNDWI increases.

Compared with these three key drivers, the ALE curves of NP, CONTAG, and SHEI show narrower variation ranges and stronger marginal effects, presenting a unified pattern. This indicates that, under a 30 m pixel resolution and 184 equidistant rings, LST is still determined by the physical components—vegetation, built-up areas, and water—while landscape configuration plays an auxiliary role. In other words, cooling interventions should prioritize precise pixel-type optimization while also attending to macro-scale pattern adjustments.

4. Conclusions

Based on the deconstruction of the SHAP–Ring relationships for each feature across 184 cumulative rings using the ExtraTrees model, a functional-structural coupling transition zone with a significant threshold can be identified around the central urban area. The three primary driving factors—NDVI, NDBI, and MNDWI—exhibit highly consistent spatial evolutionary trends. NDVI and MNDWI show positive mean SHAP values in the Ring0–Ring100 interval, transition from positive to negative after Ring100, and display a gradually increasing negative contribution with distance after Ring110. NDBI begins to decline around Ring110, turns negative after Ring120, and remains so until the outer boundary, demonstrating a larger positive contribution range and more stable positive contribution influence. This indicates that, under the original configuration, the Ring0–Ring100 interval has not yet crossed the effective cooling threshold; impervious-surface-dominated heat diffusion remains predominant, and green spaces and water bodies have failed to form a continuous cold source network.The landscape indices NP, CONTAG, and SHEI exhibit a positive-negative transition around Ring90, which is spatially close to the transition nodes of the spectral features, and their contribution curves follow similar trends, confirming that the Ring0–Ring100 interval belongs to a highly plastic cooling potential zone. Beyond this range, the peripheral cooling structure gradually stabilizes, and the optimization potential and returns decline.

Therefore, the outer boundary of the optimal optimization zone is delineated near the equivalent radius of the central urban area's outer boundary (Ring92), targeting the area where both primary and secondary factors contribute strongly to LST warming. Within this optimization zone, cooling strategies should focus on the synergistic advancement of functional restructuring and structural integration: enhancing vegetation cover and restraining impervious surface expansion to push functional variables across the cooling threshold, while simultaneously regulating heat transfer pathways and efficiency. Specifically, when the cover level has not yet reached the threshold, priority should be given to reducing the fragmentation of green spaces and prioritizing allocation in areas closer to the central urban area to facilitate the diffusion of cold sources toward the urban core. After the cover level has significantly improved and reached the threshold, further optimization of patch morphology and connectivity can increase the embedding rate and interspersion of green spaces within impervious patches, thereby enhancing ecological permeability and spatial synergy. This forms a systemic regulation pathway characterized by the synergistic interaction of cover optimization and structural reshaping, making Ring92 a critical spatial interface for achieving efficient and stable cooling of the central urban thermal environment..

References

[1] Giannaros C, Agathangelidis I, Papavasileiou G, et al. The extreme heat wave of July–August 2021 in the Athens urban area (Greece): Atmospheric and human-biometeorological analysis exploiting ultra-high resolution numerical modeling and the local climate zone framework[J]. Science of The Total Environment, 2023, 857: 159300.

[2] Mayou L A, Alamdari N, Ahmadisharaf E, et al. Impacts of future climate and land use/land cover change on urban runoff using fine-scale hydrologic modeling[J]. Journal of Environmental Management, 2024, 362: 121284.

[3] Guan R, Huang W, Li Y, et al. Can land use/land cover optimization improve the environmental flow guarantee in regulated basins?[J]. Journal of Hydrology, 2023, 626: 130309.

[4] Zhang A, Xia C. Scale effect on the relationship between urban landscape patterns and land surface temperature[J]. Sustainable Cities and Society, 2024, 117: 105942.

ISSN 2616-5775 Vol. 9, Issue 7: 1-8, DOI: 10.25236/AJCIS.2026.090701

[5] Zhu D, Zhou X, Cheng W. Water effects on urban heat islands in summer using WRF-UCM with gridded urban canopy parameters—A case study of Wuhan[J]. Building and Environment, 2022, 225: 109528.

[6] Deng X, Cao Q, Wang L, et al. Characterizing urban densification and quantifying its effects on urban thermal environments and human thermal comfort[J]. Landscape and Urban Planning, 2023, 237: 104803.

[7] Lin L, Zhao Y, Zhao J, et al. Comprehensively assessing seasonal variations in the impact of urban greenspace morphology on urban heat island effects: A multidimensional analysis[J]. Sustainable Cities and Society, 2025, 118: 106014.

[8] Qiu G Y, Yan C, Liu Y. Urban evapotranspiration and its effects on water budget and energy balance: Review and perspectives[J]. Earth-Science Reviews, 2023, 246: 104577.

[9] Gupta A, De B, Das S, et al. Thermal hazards in urban spaces: A review of climate-resilient planning and design to reduce the heat stress[J]. Urban Climate, 2025, 59: 102296.

[10] Patel S, Indraganti M, Jawarneh R N. A comprehensive systematic review: Impact of Land Use/Land Cover (LULC) on Land Surface Temperatures (LST) and outdoor thermal comfort[J]. Building and Environment, 2024, 249: 111130.

[11] Heinl M, Hammerle A, Tappeiner U, et al. Determinants of urban–rural land surface temperature differences–A landscape scale perspective[J]. Landscape and Urban Planning, 2015, 134: 33-42.

[12] Taloor A K, Parsad G, Jabeen S F, et al. Analytical study of land surface temperature for evaluation of UHI and UHS in the city of Chandigarh India[J]. Remote Sensing Applications: Society and Environment, 2024, 35: 101206.

[13] Wang C, Zhan W, Li L, et al. Urban heat islands characterized by six thermal indicators[J]. Building and Environment, 2023, 244: 110820.

[14] Cao S J, Zhang C K, Wang J Q, et al. Low-carbon design towards sustainable city development: Integrating glass space with natural greenery[J]. Science China Technological Sciences, 2024, 67(9): 2659-2674.

[15] Adilkhanova I, Santamouris M, Yun G Y. Green roofs save energy in cities and fight regional climate change[J]. Nature Cities, 2024, 1(3): 238-249.

[16] Motlagh Z K, Lotfi A, Pourmanafi S, et al. Spatial modeling of land-use change in a rapidly urbanizing landscape in central Iran: Integration of remote sensing, CA-Markov, and landscape metrics[J]. Environmental Monitoring and Assessment, 2020, 192(11): 695.

[17] Yang J, Huang X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019[J]. Earth System Science Data, 2021, 13(8): 3907-3925.

[18] Gong P, Liu H, Zhang M, et al. Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017[J]. Sci. Bull, 2019, 64(6): 370-373.

[19] Li Z, He W, Cheng M, et al. SinoLC-1: The first 1 m resolution national-scale land-cover map of China created with a deep learning framework and open-access data[J]. Earth system science data, 2023, 15(11): 4749-4780.

[20] Angel S, Parent J, Civco D L, et al. The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050[J]. Progress in planning, 2011, 75(2): 53-107.

[21] Yao R, Wang L, Huang X, et al. Greening in rural areas increases the surface urban heat island intensity[J]. Geophysical Research Letters, 2019, 46(4): 2204-2212.

[22] Guan S, Chen Y, Wang T, et al. Mitigating urban heat island through urban-rural transition zone landscape configuration: Evaluation based on an interpretable ensemble machine learning framework[J]. Sustainable Cities and Society, 2025, 123: 106272.