SJST
sjst.scst.edu.ly

Surman **J**ournal for **S**cience and **T**echnology
ISSN: Online (2790-5721) - Print (2790-5713)

مجلة صرمان للعلوم والتقنية
Vol **8**, No.2. Jun – Dec. 20xx
Pages: 399 ~ 408

Spatially Validated Explainable GeoAI for Urban Heat Attribution in Benghazi, Libya: Quantifying Validation Leakage, Built-Up Effects, and Coastal Moderation

Esam Miftah Abdulnabi Aboudoumat[1]*, Aiad F. Mohamed Fareg[2]

[1]Computer Department, College of Science and Technology Qumins, Libya
[2]Department of Architectural Engineering, College of Science and Technology Qumins, Libya

**Corresponding author: esam.mouftah@gmail.com*

Abstract

Urban-heat studies in data-scarce cities frequently use random train-test splits that can overstate machine-learning performance because neighboring cells share spatial structure. We develop a reproducible, explainable GeoAI framework for Benghazi, Libya, using Landsat Collection 2 Level-2 land surface temperature with NDVI, NDBI, MNDWI, local spectral heterogeneity, and distance from the Mediterranean coast. The primary 240 m analysis contains 12,381 valid non-water cells and is repeated at 150 and 510 m. Ridge, Random Forest, and histogram gradient boosting are evaluated under both random five-fold validation and geographically blocked five-fold validation. At 240 m, the best R^2 falls from about 0.958 under random validation to 0.829 under spatial validation, while RMSE increases from about 1.56 °C to 3.14 °C. The same optimism gap persists at all three scales. SHAP attribution identifies NDBI as the strongest model driver, followed by MNDWI and spatial-context variables; high-NDBI cells are 5.70 °C warmer in median LST than low-NDBI cells. The results show that spatial validation is a substantive requirement for credible urban-heat inference and that Benghazi's thermal pattern reflects coupled built-up, moisture, and coastal controls rather than a single greenness indicator.

Keywords: *Benghazi; GeoAI; land surface temperature; spatial cross-validation; SHAP; urban heat*

الملخص

تواجه دراسات الحرارة الحضرية في المدن الشحيحة بالبيانات مشكلة منهجية متكررة تتمثل في الاعتماد على تقسيمات عشوائية قد تبالغ في تقدير قدرة نماذج التعلم الآلي بسبب الترابط المكاني بين الخلايا المتجاورة. تطبق هذه الدراسة إطار GeoAI تفسيريًا وقابلاً لإعادة الإنتاج على بنغازي باستخدام درجة حرارة سطح الأرض من Landsat Collection 2 Level-2 ومؤشرات NDVI وNDBI وMNDWI والمسافة من الساحل. تم تحليل 12,381 خلية صالحة عند مقياس 240 مترًا، مع اختبارات حساسية عند 150 و510 أمتار، ومقارنة Ridge وRandom Forest وHistGradientBoosting تحت تحقق عشوائي وتحقق مكاني بخمس كتل جغرافية. عند 240 مترًا انخفض أفضل R^2 من نحو 0.958 في التحقق العشوائي إلى 0.829 في التحقق المكاني، وارتفع RMSE من نحو 1.56 إلى 3.14 درجة مئوية، واستمر النمط نفسه عبر المقاييس الثلاثة. أظهر SHAP أن NDBI هو المتغير الأكثر تأثيرًا، بينما كان MNDWI والمسافة من الساحل من أهم محددات السياق الحراري. وكانت الخلايا ذات NDBI المرتفع أدفأ بوسيط 5.70 درجة مئوية من الخلايا ذات NDBI المنخفض. توضح النتائج أن التحقق المكاني ليس تفصيلاً إجرائيًا بل شرطًا لتقدير واقعي لقابلية تعميم نماذج الحرارة الحضرية، وأن التفسير التخطيطي لبنغازي يجب أن يميز بين أثر البناء والسياق الساحلي والغطاء السطحي بدلاً من الاعتماد على مؤشر خضرة منفرد.

Submitted: *11/08/2026* ***Accepted****: 10/09/2026*

1. Introduction

Urban heat is increasingly studied with satellite data and machine learning because remote sensing can provide consistent observations where ground sensor networks are incomplete. Yet prediction accuracy alone is not a sufficient basis for planning. Spatial data violate the independence assumptions behind ordinary random cross-validation: adjacent cells share land cover, coastal exposure, development patterns, and unobserved neighborhood context. When nearby observations are split between training and test folds, a model can appear to generalize while partly exploiting spatial similarity. This problem is especially consequential in data-scarce cities, where a seemingly excellent model may be used to justify neighborhood-scale design recommendations without a realistic out-of-area test.

Benghazi offers a demanding case. It is a Mediterranean coastal city with strong land-water gradients, semi-arid vegetation, heterogeneous development, and recent outward urban expansion. Our preceding evidence-fusion study documented a 3.89% increase in Copernicus artificial surfaces between 2017 and 2020 and a 34.36% decline in the green-to-urban ratio in an independent Sentinel-2 assessment through 2021 (Aboudoumat et al., 2026). That study intentionally stopped short of claiming a city-specific thermal mechanism. The present paper is a distinct empirical extension: it directly estimates the spatial relationship between Landsat surface temperature and spectral/coastal predictors and tests whether that relationship survives spatially separated validation.

Three research questions structure the analysis. RQ1 asks how much random cross-validation inflates urban-heat prediction performance relative to geographically blocked validation. RQ2 asks whether the validation gap persists when the analytical grid changes from 150 to 240 and 510 m. RQ3 asks which observable surface and contextual variables dominate the fitted nonlinear model after water masking and spatial validation. The contribution is therefore methodological and empirical: we combine a leakage audit, multi-scale sensitivity analysis, and explainable machine learning in a North African coastal city, then translate the findings into planning statements that are constrained by the evidence rather than by model accuracy alone.

2. Related Work and Research Gap

Contemporary urban-heat research increasingly combines remotely sensed LST with nonlinear machine learning. Wang et al. (2025) used multiple regression, XGBoost, and SHAP to show that morphology-LST relationships differ across plain and plateau cities. Wan et al. (2026) extended this logic across seven cities and demonstrated that spatial-configuration variables materially improve thermal explanation. These studies support explainable nonlinear modeling, but their reported accuracy cannot be transferred directly to Benghazi because climate, coastal exposure, vegetation scarcity, and urban form differ by context.

A second literature establishes that validation design matters for spatial prediction. Roberts et al. (2017) showed that cross-validation should reflect the dependence structure of the data rather than defaulting to random folds. In remote-sensing and ecological modeling, spatial blocking is now widely used to estimate performance at locations that are genuinely separated from the training sample. The implication for urban heat is straightforward: if the intended use is neighborhood transfer, a model should be tested on held-out geographic regions rather than intermingled cells.

A third line of work concerns mechanisms. Global evidence shows substantial but unequal cooling from urban green infrastructure, particularly between Global South and Global North cities (Li et al., 2024). At the local scale, however, NDVI is not a universal causal variable: it can covary with soil moisture, irrigated land, water proximity, land-use type, and seasonal surface conditions. Built-up indices such as NDBI and moisture/water indices such as MNDWI can therefore provide complementary information. Benghazi's coastline further makes distance to the Mediterranean an important confounder and potentially a genuine thermal control. The gap addressed here is the joint treatment of these factors under spatially honest validation rather than a single-index correlation or a random-split benchmark.

3. Data and Methodology

3.1 Study area and Landsat scene

The study window covers Benghazi and its immediate urban/peri-urban surroundings (19.88°-20.25° E; 31.94°-32.23° N). We queried the Microsoft Planetary Computer STAC catalog for Landsat Collection 2 Level-2 scenes acquired during June-September 2020 with less than 20% scene-level cloud cover. The selected scene was Landsat 8 LC08_L2SP_184038_20200830_02_T1, acquired on 30 August 2020, with reported scene cloud cover of 0%. The analysis uses the Level-2 surface-temperature product rather than reconstructing temperature from top-of-atmosphere brightness temperature.

3.2 Preprocessing and predictor construction

QA_PIXEL bits for dilated cloud, cirrus, cloud, and cloud shadow were masked. Surface temperature digital numbers were converted using the Collection 2 scale and offset (0.00341802 and 149.0 K) and then expressed in degrees Celsius. Surface reflectance bands were scaled using the Level-2 reflectance transformation. NDVI was computed from near-infrared and red reflectance, NDBI from SWIR1 and near-infrared reflectance, and MNDWI from green and SWIR1 reflectance. The largest connected water component detected by MNDWI was treated as sea, and water-dominated cells were excluded. A Euclidean distance-to-sea surface was then derived at the 30 m raster resolution.

The primary grid aggregates valid pixels to nominal 240 m cells (8 × 8 Landsat pixels). For each cell we retain median LST, median NDVI, median NDBI, median MNDWI, within-cell standard deviations of NDVI and NDBI, and median distance to the detected coast. Cells require at least 35 valid non-water pixels. Identical logic is repeated at 150 m (5 × 5 pixels) and 510 m (17 × 17 pixels) for scale sensitivity. The final 240 m dataset contains 12,381 valid cells.

3.3 Models, leakage audit, and spatial validation

We compare Ridge regression, Random Forest, and histogram gradient boosting. The purpose is not to conduct unrestricted hyperparameter search but to test whether nonlinear gains survive a change in validation regime. Two five-fold estimators are used. Random CV shuffles cells before splitting and represents the common but potentially optimistic benchmark. Spatial Block CV first partitions cell centroids into five geographically compact K-means regions and then holds out each region once. Thus, every reported spatial prediction is generated for a geographic block that is absent from model fitting. Performance is summarized by mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2) computed from all out-of-fold predictions. Because the K-means regions are contiguous but unbuffered, training cells can remain geographically close to test cells along shared block boundaries. This residual border spillover makes the blocked protocol more conservative than random CV but does not guarantee complete spatial separation. Future validation should therefore add exclusion buffers around held-out block boundaries and compare several buffer distances before model transfer is claimed.

3.4 Explainability and robustness

For interpretability, a Random Forest is refit on the complete 240 m dataset and explained using SHAP (Lundberg & Lee, 2017). Mean absolute SHAP values rank predictors by their average contribution magnitude. SHAP is used for attribution rather than causal identification. To examine nonlinear response regions rather than importance rankings alone, SHAP dependence plots are additionally generated for NDBI and MNDWI, the two highest-ranked predictors. For efficient visualization, SHAP values are computed on a reproducible random sample of 2,500 cells drawn from the full fitted dataset (random seed 42), and local median trends are shown only where the predictor distribution is sufficiently populated. We also report Spearman correlations and quartile contrasts as descriptive diagnostics. The multi-scale rerun is treated as a robustness test: conclusions are considered stable only when the random-versus-spatial validation gap remains visible at 150, 240, and 510 m.

4. Results

4.1 Thermal distribution and descriptive contrasts

Across the 12,381 valid 240 m cells, mean LST was 46.63 °C and median LST was 48.48 °C. The 10th and 90th percentiles were 28.09 and 53.72 °C, confirming a broad within-city thermal range in the selected summer scene. Cells in the highest NDBI quartile were 5.70 °C warmer in median LST than cells in the lowest NDBI quartile. By contrast, the simple low-versus-high NDVI quartile contrast was -1.54 °C (low-NDVI cells were cooler), an intentionally reported counterexample to simplistic interpretation of greenness. This bivariate sign is consistent with confounding by coastal and moisture gradients and is not interpreted as evidence that vegetation increases temperature.

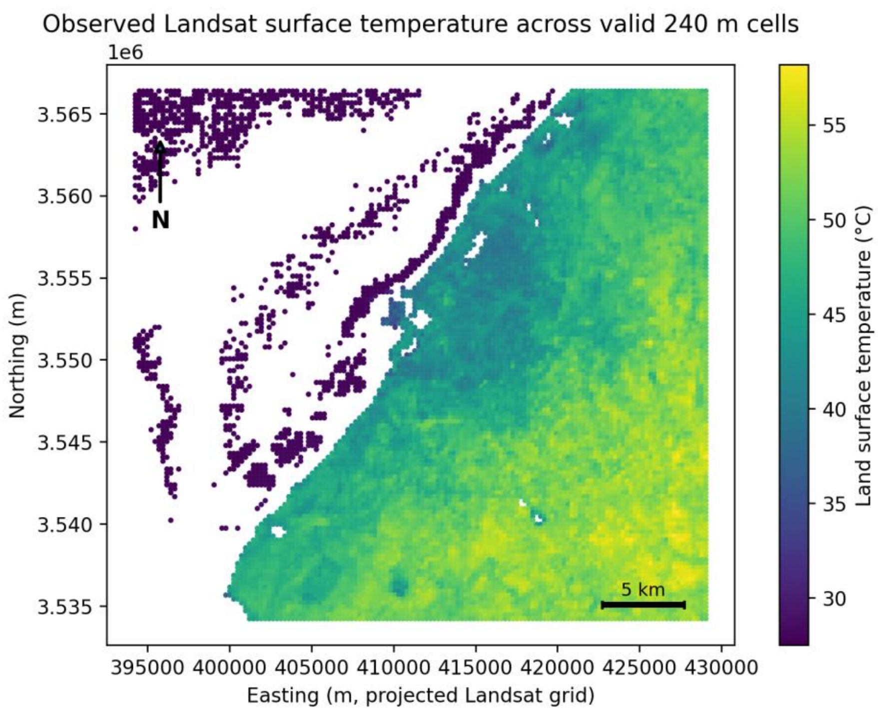

Figure 1. Spatial distribution of Landsat-derived land surface temperature for the valid 240 m analysis cells, with north arrow and a 5 km scale bar.

sjst.scst.edu.ly
Surman Journal for Science and Technology
ISSN: Online (2790-5721) - Print (2790-5713)
مجلة صرمان للعلوم والتقنية
Vol **8**, No.2. Jun – Dec. 20xx
Pages: 399 ~ 408

Table 1. Primary 240 m dataset and descriptive thermal diagnostics

Metric	Value
Valid non-water cells	12,381
Mean LST	46.63 °C
Median LST	48.48 °C
10th-90th percentile LST	28.09-53.72 °C
High-NDBI minus low-NDBI median LST	+5.70 °C
Low-NDVI minus high-NDVI median LST	-1.54 °C

4.2 Random validation materially overstates model performance

At 240 m, both nonlinear models achieved R^2 values near 0.958 under random CV. Under spatial block CV, Random Forest fell to R^2=0.829 with RMSE=3.14 °C and HistGradientBoosting to R^2=0.829 with RMSE=3.14 °C. The corresponding random-CV RMSE values were approximately 1.55-1.56 °C. In other words, geographically honest testing increased RMSE by roughly 100% while reducing R^2 by about 0.13. Ridge regression showed the same direction of change, falling from R^2=0.902 to 0.799.

Table 2. Model performance at the primary 240 m scale

Model	Validation	MAE (°C)	RMSE (°C)	R²
Ridge	Random CV	1.94	2.38	0.902
Ridge	Spatial block CV	2.93	3.40	0.799
Random Forest	Random CV	1.11	1.55	0.958
Random Forest	Spatial block CV	2.50	3.14	0.829
HistGradientBoosting	Random CV	1.14	1.56	0.958
HistGradientBoosting	Spatial block CV	2.53	3.14	0.829

4.3 The optimism gap is stable across analysis scales

The leakage result is not an artifact of the 240 m grid. At 150 m, the best nonlinear random-CV R^2 was 0.959 whereas the best spatial R^2 was 0.838. At 510 m, the best random-CV R^2 was 0.959 and the best spatial R^2 was 0.827. Thus, spatial validation reduced R^2 by approximately 0.12-0.13 at every tested scale. Spatial RMSE ranged from 3.01 to 3.23 °C for the best nonlinear model, compared with 1.46-1.62 °C under random CV.

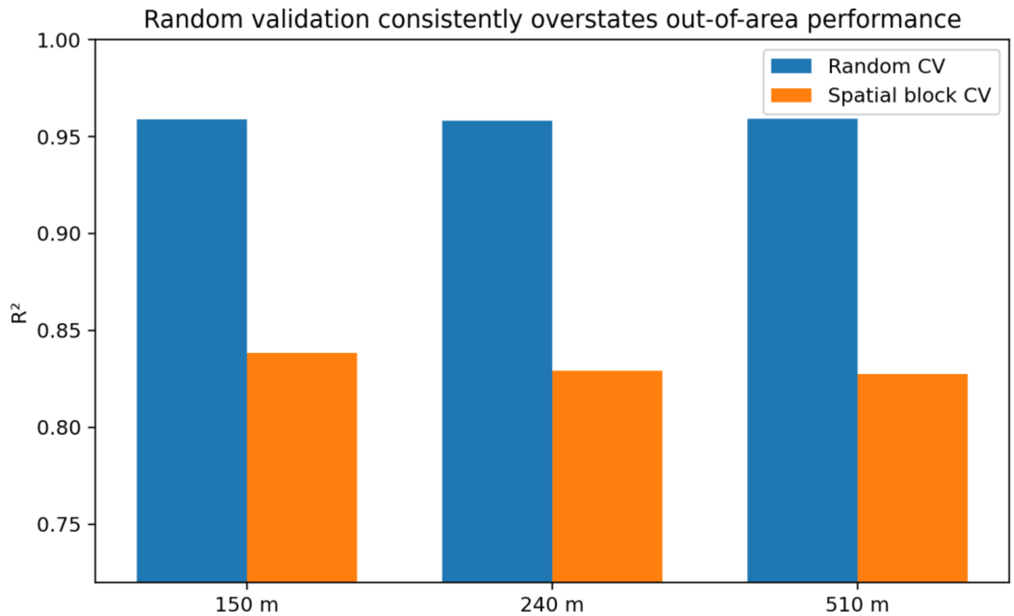

Figure 2. Best nonlinear model R^2 under random versus geographically blocked five-fold validation at three grid scales.

Table 3. Multi-scale sensitivity of the best nonlinear validation result

Scale	Best random R^2	Best spatial R^2	Best random RMSE	Best spatial RMSE
150 m	0.959	0.838	1.62 °C	3.23 °C
240 m	0.958	0.829	1.55 °C	3.14 °C
510 m	0.959	0.827	1.46 °C	3.01 °C

4.4 Explainable attribution emphasizes built-up and moisture context

Random-Forest SHAP attribution at 240 m ranked NDBI first (mean absolute SHAP 3.55), followed by MNDWI (1.52), NDVI within-cell heterogeneity (0.57), distance from the coast (0.46), NDVI (0.26), and NDBI heterogeneity (0.21). The dominance of NDBI is consistent with the +5.70 °C high-versus-low built-up quartile contrast. The importance of MNDWI and coastal distance confirms that Benghazi's temperature field is strongly structured by blue/coastal context and that a vegetation-only model would be incomplete.

The dependence plots reveal nonlinear transition regions that are hidden by mean absolute importance alone. In the densely populated portion of the NDBI distribution, the local median SHAP contribution changes from negative to positive at approximately NDBI ≈ 0.07 and then rises steeply through the main urban spectral range. For MNDWI, the local median SHAP response changes sign at approximately MNDWI ≈ -0.35, with progressively less negative MNDWI values associated with increasingly negative temperature contributions over the principal data range. These values are empirical, scene- and model-specific transition points rather than universal physical thresholds; sparse extreme-index clusters are therefore not interpreted as architectural breakpoints.

Descriptive rank correlations reinforce this interpretation: MNDWI correlated negatively with LST (Spearman ρ=-0.816), distance from the coast correlated positively (ρ=0.564), and NDBI correlated

SJST
sjst.scst.edu.ly
Surman **J**ournal for **S**cience and **T**echnology
ISSN: Online (2790-5721) - Print (2790-5713)
مجلة صرمان للعلوم والتقنية
Vol **8**, No.2. Jun – Dec. 20xx
Pages: 399 ~ 408

positively (ρ=0.290). NDVI alone had a near-zero marginal rank correlation (ρ=0.051). These values do not establish causal effects, but they explain why the simple NDVI quartile comparison can be misleading when coast and moisture are not controlled.

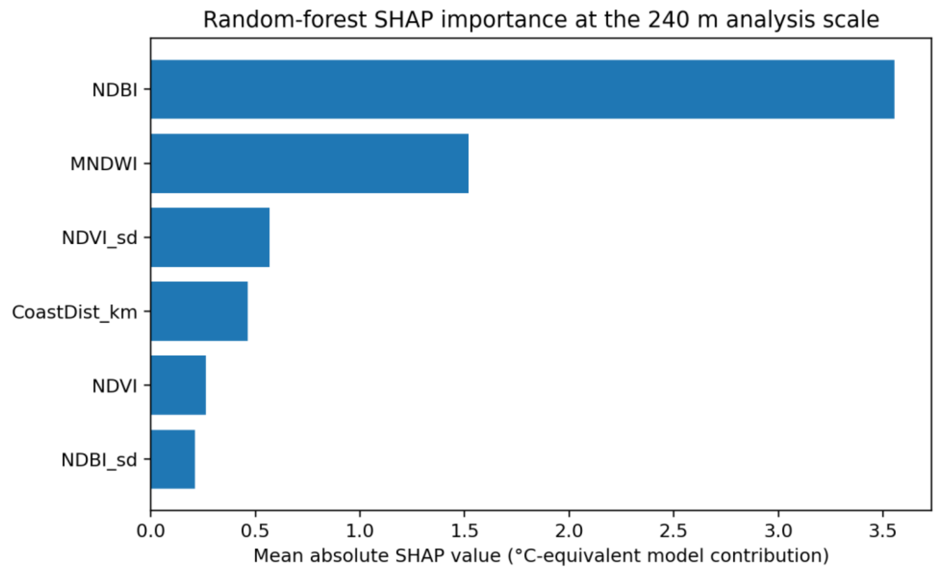

Figure 3. Mean absolute SHAP feature importance for the 240 m Random Forest model.

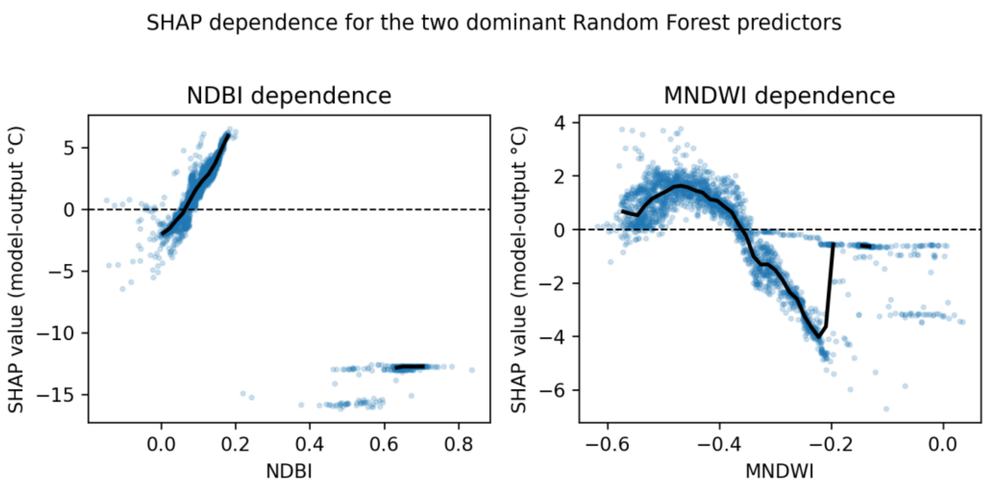

Figure 4. SHAP dependence plots for NDBI and MNDWI in the 240 m Random Forest model. Black lines show local median SHAP trends across sufficiently populated predictor ranges.

5. Discussion

5.1 Spatial validation changes the scientific claim

The main result is not that machine learning can fit Benghazi LST; random CV already makes that conclusion easy. The stronger result is that model performance remains useful but materially lower when entire geographic regions are withheld. A random-split R^2 near 0.96 could encourage overconfident neighborhood-level prediction, whereas the spatially validated R^2 near 0.83 and RMSE

Surman Journal for Science and Technology
ISSN: Online (2790-5721) - Print (2790-5713)
مجلة صرمان للعلوم والتقنية
Vol 8, No.2. Jun – Dec. 20xx
Pages: 399 ~ 408

near 3 °C give a more realistic estimate of transfer to unseen parts of the city. For planning applications, this difference changes the uncertainty attached to any proposed intervention map.

The blocked estimates should nevertheless be interpreted as spatially stricter, not spatially perfect. K-means creates compact held-out regions, but cells on opposite sides of a shared boundary can still be close enough to share local surface and neighborhood structure. Such border spillover can leave a residual optimistic component even in blocked CV. A stronger deployment-oriented design would exclude a spatial buffer around each test block from model training, repeat the analysis across multiple buffer widths, and compare the resulting loss of performance. If the model remains stable as the buffer widens, the evidence for true out-of-area transfer would be substantially stronger.

The scale sensitivity test strengthens this interpretation. If the gap had disappeared at 510 m, it could have been dismissed as a fine-resolution sampling artifact. Instead, the same pattern survives at all three scales. This supports a general recommendation for urban-heat GeoAI: the validation geometry should approximate the intended deployment geometry, and random CV should be reported, at most, as an optimistic benchmark rather than as the primary evidence of spatial generalization.

5.2 Built-up intensity, coastal moderation, and the NDVI caution

NDBI is the strongest explanatory variable in the fitted model and the highest-NDBI quartile is substantially warmer than the lowest. This is consistent with physical expectations for dry, impervious, and strongly absorbing surfaces in a hot summer scene. Yet the results also show why single-variable narratives are risky. NDVI has little marginal rank association with LST and the raw quartile contrast is counterintuitive. In Benghazi, sparse vegetation, irrigated patches, exposed soil, moisture, and proximity to the sea can occupy overlapping spectral and geographic regimes. The correct interpretation is therefore conditional: vegetation may contribute to cooling, as the broader literature establishes, but the local Landsat snapshot does not support treating NDVI alone as a sufficient design rule.

Coastal context is particularly important. MNDWI is the second-ranked SHAP predictor and distance from the coast has a clear positive rank association with LST. These variables partly encode marine moderation, surface moisture, and land-water adjacency. For Benghazi planning, this implies that inland neighborhoods and highly built-up surfaces should not be compared directly with coastal cells without accounting for baseline thermal context.

The predictor ranking is also season-dependent. The selected 30 August scene represents a clear-sky late-summer thermal extreme, when dry surfaces, low soil moisture, irrigation contrasts, and Mediterranean proximity can dominate the spectral-temperature relationship. During wetter winter or transitional seasons, NDVI may increase over temporarily vegetated or rain-responsive surfaces, while MNDWI may respond more strongly to soil moisture and seasonal water conditions; the relative SHAP importance and even the location of dependence-plot transition regions may therefore shift. Multi-season and multi-year Landsat sampling is required before the present late-summer ranking is generalized as a year-round thermal mechanism for Benghazi.

5.3 Planning implications

The results support three evidence-bounded planning priorities. First, high built-up intensity should be treated as a thermal-risk screening signal, particularly in inland areas where coastal moderation is weaker. Second, blue-green interventions should be evaluated as spatial systems rather than as isolated NDVI increases: moisture availability, continuity, and coastal/lake adjacency matter. Third, any GeoAI heat-priority map used for architecture or public-space investment should carry uncertainty derived from spatial validation. A model that performs well under random folds but deteriorates under blocked folds should not be presented as uniformly reliable across the city.

These conclusions also complement, rather than duplicate, our 2026 urban-growth study. The earlier paper established that urban expansion was not accompanied by proportional green-space compensation. The present study tests a separate question: how reliably can a summer thermal field be explained and transferred spatially, and which surface/context variables dominate that explanation? No figures or tables from the prior article are reused as new results.

6. Limitations

The analysis is deliberately conservative but has several limitations. First, it uses a single clear-sky summer Landsat scene; the results characterize daytime surface temperature on that date rather than seasonal mean air temperature or pedestrian heat stress. Second, Landsat LST represents radiometric surface temperature and should not be interpreted as near-surface air temperature. Third, the current executed predictor stack is spectral and coastal; verified building-height, street-canyon, material, road-intensity, and parcel-level variables were not inserted merely to create a richer model. Fourth, NDBI in a semi-arid landscape can also respond to dry exposed soil and should not be interpreted as a pure building-density measure. Fifth, the water mask is algorithmic and can misclassify some wet or bright surfaces. Sixth, SHAP explains fitted model behavior but does not identify causal effects. Finally, geographically compact K-means blocks remain vulnerable to boundary proximity because no exclusion buffer was imposed; future work should compare buffered blocks, directional holdouts, repeated spatial partitions, and temporally separated scenes across wet and dry seasons.

7. Conclusion

This study shows that the credibility of urban-heat GeoAI depends as much on validation design as on model choice. Using 12,381 valid 240 m cells in Benghazi, nonlinear models achieved $R^2 \approx 0.96$ under random cross-validation but only ≈0.83 when entire geographic blocks were withheld; RMSE approximately doubled from about 1.55 °C to 3.14 °C. The same optimism gap persisted at 150 and 510 m, demonstrating that it is not a single-scale artifact. Explainable attribution ranked NDBI as the dominant predictor, while MNDWI and coastal distance showed that Benghazi's thermal field is strongly conditioned by moisture and Mediterranean proximity. High-NDBI cells were 5.70 °C warmer in median LST than low-NDBI cells, whereas NDVI alone provided an unstable marginal signal. The practical conclusion is twofold: urban-heat models should be validated spatially before they support neighborhood decisions, and climate-responsive planning in Benghazi should integrate built-up intensity with blue/coastal context rather than rely on greenness as a standalone proxy.

8. Data Availability and Reproducibility

The analysis uses publicly accessible Landsat Collection 2 Level-2 products queried through a STAC endpoint. The selected scene identifier, study bounding box, preprocessing rules, grid sizes, model definitions, and validation logic are stated in this manuscript. Derived cell-level tables and model-result tables were generated programmatically from the executed workflow. No unexecuted performance value is reported.

9. Funding

This research received no external funding.

10. Conflict of Interest

The authors declare no conflict of interest.

References

Aboudoumat, E. M. A., Fareg, A. F. M., & Bader, A. I. R. (2026). Urban growth without green compensation: A multi-source geospatial evidence-fusion framework for climate-responsive architectural planning in Benghazi, Libya. Al-Farooq Journal of Sciences, 2(Suppl. 3), 2077-2086. https://doi.org/10.65405/41md9s79

Beele, E., Aerts, R., Reyniers, M., & Somers, B. (2024). Spatial configuration of green space matters: Associations between urban land cover and air temperature. Landscape and Urban Planning, 249, 105121. https://doi.org/10.1016/j.landurbplan.2024.105121

Bufarwa, M. T. (2022). Assessing the spatio-temporal changes in urban green spaces of Benghazi using remote sensing technique (NDVI). Libyan Journal of Basic Sciences, 19(5), 68-79.

Copernicus Emergency Management Service. (2020). EMSN080: Land use mapping and population dynamics analysis in Benghazi, Libya. Risk and Recovery Mapping.

Eneche, P. S. U., Atun, F., Zeng, Y., & Pfeffer, K. (2024). Robust drivers of urban land surface temperature dynamics across diverse landscape characters: An augmented systematic literature review. Ecological Indicators, 163, 112056. https://doi.org/10.1016/j.ecolind.2024.112056

Li, H., Zhao, Y., Wang, C., Ürge-Vorsatz, D., Carmeliet, J., & Bardhan, R. (2024). Cooling efficacy of trees across cities is determined by background climate, urban morphology, and tree trait. Communications Earth & Environment. https://doi.org/10.1038/s43247-024-01908-4

Li, Y., Svenning, J.-C., Zhou, W., et al. (2024). Green spaces provide substantial but unequal urban cooling globally. Nature Communications, 15, 7108. https://doi.org/10.1038/s41467-024-51355-0

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765-4774.

Microsoft. (2026). Planetary Computer: STAC catalog and geospatial data access platform. Microsoft Research.

Roberts, D. R., Bahn, V., Ciuti, S., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913-929. https://doi.org/10.1111/ecog.02881

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. Bulletin of the American Meteorological Society, 93(12), 1879-1900. https://doi.org/10.1175/BAMS-D-11-00019.1

U.S. Geological Survey. (2024). Landsat Collection 2 Level-2 science products: Surface temperature and surface reflectance product documentation. U.S. Department of the Interior.

Wan, B., Shen, N., Zhang, H., & Sheng, Q. (2026). Analyzing the impact of urban morphology on urban land surface temperature from the perspective of spatial configuration and explainable machine learning: A case study of seven cities. Sustainable Cities and Society, 138, 107165. https://doi.org/10.1016/j.scs.2026.107165

Wang, Z., Zhou, R., Rui, J., & Yu, Y. (2025). Revealing the impact of urban spatial morphology on land surface temperature in plain and plateau cities using explainable machine learning. Sustainable Cities and Society, 118, 106046. https://doi.org/10.1016/j.scs.2024.106046

Zeng, J., Cai, Y., Yang, C., Liu, Y., Liu, W., & Xing, H. (2025). Revealing the impact patterns of urban morphology on land surface temperature in the Guangdong-Hong Kong-Macao Greater Bay Area, China. Cities, 165, 106147. https://doi.org/10.1016/j.cities.2025.106147