

A Field-Calibrated Composite Indicator of Urban Vegetation Vulnerability to Compound Heat–Drought Stress: Spatially Blocked Validation and an Exploratory Ecological Application to Cardiovascular Hospitalizations in Isfahan, Iran

Pouriya Javari

Faculty of Medicine Isfahan University of Medical Sciences

Majid Javari

majid_javari@pnu.ac.ir

PayameNoor University

Hossein Ladvar

PayameNoor University

Research Article

Keywords: DUVRI-Eco; multi-dimensional heat–drought stress; urban vegetation; landscape structure; cardiovascular hospitalization; spatial vulnerability; semi-arid cities

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11154147/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A Field-Calibrated Composite Indicator of Urban Vegetation Vulnerability to Compound Heat–Drought Stress: Spatially Blocked Validation and an Exploratory Ecological Application to Cardiovascular Hospitalizations in Isfahan, Iran

Pouriya Javari

Medical Student, Faculty of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
Email: pouriya.javari77@gmail.com

Majid Javari

Associate Professor, College of Social Sciences, PayameNoor University, P.O. Box 19395-3697, Tehran, Iran
Email: majid_javari@pnu.ac.ir
Corresponding Author

Hossenin Ladvar

Ph.D. Student, PayameNoor University, Tehran, Iran
Email: hossenladvar@stu.pnu.ac.ir

Abstract: Extreme heat, prolonged drought, vegetation decline, air pollution, and landscape structure are all implicated in cardiovascular disease, yet their combined influence remains poorly understood in arid and semi-arid cities. Urban vegetation in these settings faces pressure from compounding heat and drought stress, and the remote-sensing indicators currently in use tend to capture only one dimension of this problem at a time, rarely bringing hydroclimatic exposure, vegetation condition, landscape structure, and field-observed mortality together within a single calibrated framework. We developed DUVRI-Eco, a construct-first composite indicator that integrates extreme heat (TXx), drought persistence (CDD), vegetation condition (NDVI), and landscape patch dominance, quantified using the Largest Patch Index (LPI) and labeled Dominant Patch Cover (DPC) as a study-specific term. The indicator was calibrated against field-measured tree mortality at 34 sites nested within Isfahan's 15 municipal districts, drawing on 72 years of homogenized climate records (1951–2022) and multi-decadal satellite imagery (1991–2022). Because each component was defined at the district level, the effective number of independent exposure units was bounded by these 15 districts, so the primary calibration model was specified at the district level, with mortality summarized as a pooled proportion. The constrained maximum-likelihood solution placed the largest weight on CDD (0.31), followed by TXx (0.28), NDVI (0.26), and DPC (0.15). Profile-likelihood diagnostics revealed weak identifiability of the individual weights, particularly for CDD and TXx, so we treat the weight vector as one admissible calibration solution rather than a precise ranking. Site-level validation yielded an AUC of 0.81 (95% CI: 0.73–0.88), and district-block validation yielded 0.79 (95% CI: 0.73–0.85). As a secondary, exploratory application, an ecological district-level panel analysis (2019–2022) found that a 0.10unit increase in DUVRI-Eco was associated with a 5.2% higher recorded cardiovascular hospitalization rate (IRR = 1.052; 95% CI: 1.018–1.087) within the participating university-affiliated hospital network, a hypothesis-generating association that cannot support individual-level or causal inference. Because the calibrated weights function as domain-specific regularization parameters rather than estimable population-level effects, external validation with local recalibration is a precondition, not merely a recommendation, for applying the indicator

elsewhere.

Keywords: DUVRI-Eco; multi-dimensional heat–drought stress; urban vegetation; landscape structure; cardiovascular hospitalization; spatial vulnerability; semi-arid cities.

1. Introduction

Urban vegetation helps cities withstand rising temperatures: it cools the air, filters pollutants, sustains local ecosystems, and indirectly supports cardiovascular health. These benefits depend on the vegetation staying alive, a condition that grows harder to sustain in dry and semi-arid cities, where climate stress and human pressure reinforce each other (Bowler et al., 2010; Nowak et al., 2014; Grimm et al., 2008). The physiological chain linking climate exposure to vegetation decline is well established. Heat increases atmospheric water demand and draws down soil moisture; drought, in turn, reduces hydraulic conductivity and photosynthesis (Allen et al., 2010; McDowell et al., 2008; Choat et al., 2012). When extreme heat coincides with prolonged drought, trees face compounded stress that can trigger hydraulic failure and carbon starvation. Urbanization adds further pressure, replacing pervious surfaces with impervious materials that trap heat and disrupt the surface energy balance (Kalnay & Cai, 2003; Santamouris, 2020). Fine particulate pollution introduces yet another stressor—one that affects both plant physiology and human cardiovascular outcomes (Grantz et al., 2003; Pope et al., 2009; Rajagopalan et al., 2018).

Individual pieces of this problem have received considerable attention: heat–drought interactions, the cooling benefits of green space, urbanization’s effect on thermal exposure, and the role irrigation and landscape pattern play in sustaining vegetation under stress (Aram et al., 2019; Li et al., 2021; Chen et al., 2018; Kumar et al., 2020). What has been slower to emerge is a framework that pulls these dimensions together. The Vegetation Condition Index (VCI; Kogan, 1995) and the Vegetation Health Index (VHI) capture vegetation condition and thermal stress but stop short of linking these dimensions to field-anchored mortality calibration or landscape structural information. Drought indices built on the Standardized Precipitation Evapotranspiration Index (SPEI) capture meteorological drought comprehensively, but they too fall short of incorporating landscape structural continuity or the urban thermal modifications that govern vegetation survival in irrigated semi-arid cities.

A second gap concerns how vegetation responds to climate stress. These responses are likely nonlinear: the shift from stable vegetation to high vulnerability probably depends on antecedent moisture, species identity, root environment, irrigation, and local urban exposure. Yet most remote-sensing studies treat vegetation response as continuous, relying on standard spectral or thermal indicators rather than estimating empirical transition points from field data.

A third gap concerns calibration and validation. NDVI and similar spectral measures offer broad spatial coverage, but coverage is not the same as validation. An indicator meant to capture vegetation vulnerability should be anchored against field-observed outcomes, and that validation should account for spatial structure: random train–test splits tend to place geographically close observations on both sides of the split, which can make performance look better than it really is (Roberts et al., 2017).

Against this backdrop, we developed DUVRI-Eco: a field-calibrated, study-domain-specific indicator of multi-dimensional hydroclimatic stress and urban vegetation vulnerability. It comprises four construct-defined components: extreme heat (TXx), drought persistence (CDD), vegetation condition (NDVI), and landscape patch dominance, quantified by the Largest Patch Index (LPI) and labeled Dominant Patch Cover (DPC) as a study-specific term.

DUVRI-Eco is construct-first by design: it is built to represent four environmental dimensions that the theory identifies as relevant to vegetation vulnerability, and its weights are calibrated against

field-observed mortality to anchor the construct rather than to optimize prediction. Two consequences follow. First, DPC is retained as a construct-defined component representing patch dominance even though its incremental predictive contribution was not statistically distinguishable in this calibration domain. Second, the calibrated weights function as predictive parameters within this construct; they are not causal effect sizes and should not be read as independent measures of environmental importance. This construct-first framing differs from a prediction-first approach, in which components would be chosen purely for their incremental predictive performance; such an approach would likely exclude DPC altogether. We chose the construct-first route because the indicator is meant to serve as a screening framework for urban environmental vulnerability, not as a purely predictive model.

DUVRI-Eco is explicitly an operational, outcome-anchored indicator of urban vegetation vulnerability under multi-dimensional hydroclimatic stress. It does not attempt to represent the full vulnerability construct of exposure, sensitivity, and adaptive capacity of an urban ecosystem. Instead, it operationalizes vulnerability as a compound environmental state empirically associated with field-observed vegetation mortality. Here, “vulnerability” is used in an operational ecological sense: the susceptibility of urban vegetation to hydroclimatic stress, distinct from the broader social-vulnerability framework of exposure, sensitivity, and adaptive capacity. The distinction matters: DUVRI-Eco is a calibrated stress-state indicator, not a comprehensive vulnerability-assessment framework.

The novelty of DUVRI-Eco lies in its architecture rather than in any single component. It brings together four elements typically treated in isolation: (1) a long-term hydroclimatic exposure record spanning seven decades; (2) multi-decadal satellite-derived vegetation dynamics; (3) outcome-calibrated component weights estimated against field-observed tree mortality; and (4) spatially blocked internal validation that explicitly tests geographic consistency, paired with explicit weight-identifiability diagnostics and index-score robustness analysis.

DPC is a structural derivative of NDVI. DPC is operationally defined as the Largest Patch Index (LPI) calculated by FRAGSTATS v4.2; we use the term DPC to emphasize its interpretation as a measure of dominant-patch structure, though mathematically it is identical to LPI. Because DPC is derived from the same NDVI-based vegetation mask used to compute NDVI itself, it is not an independent environmental dimension but a structurally transformed derivative of NDVI. The strong correlation with NDVI ($\rho = 0.76$) reflects this shared derivation. DPC was retained as a construct-defined component representing patch dominance, but it was not an independent predictor of vegetation mortality; its incremental contribution was modest ($\Delta\text{AUC} = -0.02$; 95% CI including zero).

The health dimension of this study is secondary and exploratory. Green space may influence cardiovascular health through several pathways, including heat exposure, air quality, physical activity, and psychosocial mechanisms (Nieuwenhuijsen et al., 2020). PM_{2.5} is a well-established cardiovascular risk factor and may serve as a shared exposure linking urban form, vegetation condition, and hospitalizations spatially (Pope et al., 2009; Rajagopalan et al., 2018). District-level ecological associations between multi-dimensional environmental vulnerability and cardiovascular hospitalizations have received limited attention in semi-arid cities specifically. Because only a limited number of distinct exposure units are available here, the health analysis aims to characterize spatial associations and generate hypotheses—not to estimate individual-level risk (Morgenstern, 1995). Earlier work in Isfahan has already reported associations between heat stress and cardiovascular hospitalizations (Javari & Javari, 2026).

This study pursues three objectives:

1. Characterize long-term hydroclimatic change in Isfahan and quantify the corresponding changes in vegetation condition and landscape structure.
2. Characterize the empirical contribution of each indicator component and calibrate the component-weighting structure against observed vegetation-mortality outcomes.
3. Internally validate DUVRI-Eco using spatially structured cross-validation and examine its exploratory association with district-level cardiovascular hospitalizations.

2. Materials and Methods

2.1. Study Design and Analytical Framework

We link long-term climate observations with satellite-derived vegetation measures and field observations to assess vegetation vulnerability to multi-dimensional heat and drought. The work proceeded in four stages. Stage 1 characterized hydroclimatic conditions over the full 1951–2022 record using TXx and CDD; CMAI was computed for descriptive purposes but excluded from the final indicator because it correlated strongly with CDD ($\rho = 0.69$) and did not improve predictive discrimination. Stage 2 characterized the urban vegetation condition and landscape structure over 1991–2022 using NDVI and DPC.

Stage 3 calibrated DUVRI-Eco against field-observed tree mortality at 34 sites across Isfahan's 15 municipal districts. A primary calibration model was specified at the district level, with mortality summarized as a pooled proportion; a site-level hierarchical model with district random effects was fitted as a sensitivity analysis. We assessed predictive performance using two complementary cross-validation schemes, site-level leave-one-site-out and district-block leave-one-district-out validation, treating the district-block estimate as the primary measure of internal geographic discrimination. Statistical identifiability of the weight vector was assessed using profile likelihood and two-dimensional likelihood surfaces.

To characterize the stability of the composite indicator score, rather than the precision of the individual weights, we generated 10,000 weight vectors from the likelihood-supported region (the sampling algorithm appears in Section 2.4.5). For each weight vector, we calculated DUVRI-Eco for all 15 municipal districts and assessed five features: (i) the rank correlation between the resulting indicator and the primary indicator (Spearman's ρ); (ii) the maximum absolute rank change for any district; (iii) the stability of the candidate transition-point classification; (iv) the distribution of district-level DUVRI-Eco scores; and (v) the absolute score change relative to the maximum-likelihood solution.

Stage 4 applied the calibrated indicator, exploratorily, to district-level cardiovascular hospitalizations. The primary health analysis was restricted to 2019–2022. Direct Landsat observations were available for 2019, 2020, and 2022; a reconstructed value was used for 2021. The extended analysis (2019–2024) used reconstructed NDVI values for 2023 and 2024 and is reported as a sensitivity analysis, a structure that keeps the primary health estimate from depending on temporal extrapolation beyond the observed range (the last direct Landsat year was 2022). The primary health model is a district-level ecological panel analysis with a time-invariant exposure,

including month and year fixed effects and a district random intercept; the ICAR spatial model is reported as a sensitivity analysis. The health analysis treats DUVRI-Eco as the primary exposure and PM_{2.5} as a measured potential confounder.

Two features of the design deserve emphasis. First, although we surveyed 34 field sites, they are nested within Isfahan's 15 municipal districts. Because the DUVRI-Eco components are defined at the district level, the effective number of independent exposure units for both calibration and health analysis is bounded by these 15 districts, which constitute the primary spatial exposure units throughout, a constraint we return to in Sections 2.2.6 and 2.4.6, since it governs how both the calibration and the health model are identified.

2.2. Data Sources and Preprocessing

2.2.1. Meteorological Data and Quality Control

Daily meteorological observations for 1951–2022 came from Isfahan's principal synoptic station, supplemented by seven neighboring stations within roughly 50 km of the study area. Before quality control and homogenization, we compiled station metadata, geographic coordinates, elevation, observation period, variable availability, and the proportion of missing observations, and used it to distinguish observational artifacts from physically plausible climatic variability (Alexandersson, 1986; Peterson et al., 1998). Quality control was performed separately for temperature and precipitation, combining distributional diagnostics with station metadata to flag observations warranting closer examination (Easterling et al., 2000; Aguilar et al., 2003); statistical extremeness alone was never grounds for removing an observation. This mattered most for maximum temperature and precipitation, where unusually large values may reflect genuine climatic extremes rather than measurement error (Zhang et al., 2005; Alexander et al., 2006).

Temporal homogeneity was evaluated using the Standard Normal Homogeneity Test (SNHT), applied at $\alpha = 0.05$, alongside diagnostics based on neighboring reference series (Alexandersson, 1986; Alexandersson & Moberg, 1997). Breakpoints were identified in two steps: candidate breakpoints were first detected using SNHT, and only those independently supported by reference-series diagnostics were then considered for adjustment. We used a pre-specified operational criterion, a Pearson correlation of at least 0.70 between the daily series and neighboring stations, to identify sufficiently correlated reference series, following Peterson and Easterling (1994), and confirmed the results were stable under alternative thresholds (0.60 and 0.80). Confirmed discontinuities were adjusted using the reference-series procedure (Peterson & Easterling, 1994; Wijngaard et al., 2003).

At the principal station, we identified two breakpoints at 1972 and 1996. The 1972 breakpoint was adjusted because it was supported by reference-series diagnostic and metadata evidence, and was inconsistent with regional climate signals. The 1996 breakpoint was retained as a climatic breakpoint for three reasons: it was independently detected at multiple neighboring stations, it coincided with documented regional climate shifts in the Middle East, and metadata showed no instrumental changes at that time. Table 1 outlines the complete evidence for each candidate breakpoint, SNHT statistic, reference-series correlation, metadata review, regional climate evidence, pre/post means, and the final adjustment decision. We distinguish climatic breakpoints, interpreted as genuine shifts in the climate system, from measurement-inhomogeneity breakpoints, interpreted as instrumental or observational artifacts; only the latter were adjusted.

Missing daily observations were reconstructed using inverse-distance weighting (IDW) with a power parameter of $p = 2$ (Gotway et al., 1996; Daly et al., 2008), using a fixed 50 km search radius and between 3 and 6 neighboring stations. When a target station itself had missing data on a given day, we excluded it from the interpolation to avoid biasing the reconstructed value.

Reconstructed observations made up 8.3% of the temperature record and 11.7% of the precipitation record; missingness was episodic rather than trending, and extreme events (above the 95th percentile) were not preferentially missing ($\chi^2 = 1.24$, $df = 1$, $P = 0.27$). We evaluated reconstruction performance separately for temperature, precipitation, and consecutive dry-day metrics. Leave-one-out cross-validation showed lower errors for IDW than for nearest-neighbor interpolation, with a root mean square error (RMSE) of 0.41 versus 0.53°C for temperature and 2.1 versus 2.8 mm for precipitation. Validating the reconstruction specifically on days above the 95th percentile of temperature and precipitation yielded RMSE values of 0.62°C and 3.4 mm. A separate, temporally blocked validation that withheld entire years produced similar error metrics (RMSE range: 0.39–0.43°C for temperature; 2.0–2.3 mm for precipitation). Sensitivity analyses using ordinary kriging on a subset of years gave comparable RMSE values (temperature RMSE = 0.39°C; precipitation RMSE = 2.0 mm), supporting IDW for this application. We ultimately retained IDW because it delivered lower prediction error while preserving the observed distribution and temporal variability of extreme observations (Dirks et al., 1998; Lloyd, 2005).

2.2.2. Multi-Dimensional Heat–Drought Stress

To assess hydroclimatic stress, we relied on three complementary indices: the annual maximum temperature (TXx), the maximum number of consecutive dry days (CDD), and the Composite Moisture Anomaly Index (CMAI). TXx represents the highest daily maximum temperature recorded in a given year (Frich et al., 2002; Alexander et al., 2006; Klein Tank et al., 2009). CMAI, in turn, is defined as the standardized anomaly of the ratio between actual and potential evapotranspiration (AET/PET). Daily values of AET and PET were derived from station-level daily temperature and precipitation data using the Thornthwaite water-balance model (Thornthwaite, 1948). Since the Thornthwaite approach does not explicitly account for radiation, humidity, or wind, we conducted a sensitivity analysis with the Hargreaves–Samani method over a subset of years. The resulting CMAI series was strongly correlated with the primary one ($r = 0.94$), and the descriptive findings remained unchanged. Importantly, this limitation does not influence the DUVRI-Eco weights, as these were calibrated independently of the CMAI. For each station and day, we calculated the AET/PET ratio and then applied an exponentially weighted moving average (EWMA) to the daily CMAI series. These smoothing captures persistence rather than isolated daily anomalies (Hunter, 1986; Hyndman, 2008):

$$S_t = \lambda \cdot \text{CMAI}_t + (1 - \lambda) \cdot S_{t-1}$$

where λ was derived from a 10-day half-life ($\lambda = 1 - 2^{(-1/10)} \approx 0.067$) (Zhao et al., 2011). The 10-day half-life was selected a priori based on physiological evidence concerning vegetation responses to moisture anomalies (Ji & Peters, 2003; Brown et al., 2008), rather than being optimized against NDVI. Sensitivity analyses using alternative temporal windows of 7, 14, and 30 days yielded highly stable DUVRI-Eco weights, with variation limited to ± 0.02 , and healthy IRRs, with variation limited to ± 0.005 . The resulting smoothed daily CMAI estimates were subsequently aggregated to the monthly scale for district-level interpolation and analysis. Importantly, the temporal window was defined solely on the basis of hydroclimatic characterization and prior to both DUVRI-Eco calibration and specification of the health model, thereby separating parameter selection from outcome-driven model optimization.

The construct-first framework defined the candidate conceptual dimensions a priori, after which empirical redundancy assessment was used to determine whether alternative operationalizations contributed sufficiently distinct information to warrant separate inclusion. CMAI was excluded because it satisfied both pre-specified redundancy criteria: its inclusion did not improve predictive discrimination ($\Delta\text{AUC} \leq 0$), and its correlation with an existing component exceeded the

predefined threshold of $|\rho| > 0.65$. Specifically, CMAI showed a correlation of $\rho = 0.69$ with CDD and a ΔAUC of -0.01 (95% CI: -0.04 to 0.02), thereby meeting both exclusion criteria. This exclusion should be interpreted specifically as evidence of informational redundancy with CDD within the present calibration framework, not as an indication that antecedent moisture lacks climatic or physiological relevance.

2.2.3. Remote Sensing Data

We characterized vegetation conditions using Landsat imagery. For the long-term climate–vegetation panel analysis, direct Landsat observations were available for 1991, 2000, 2005, 2010, 2015, 2020, and 2022 (TM, ETM+, and OLI-TIRS sensors) (Chander et al., 2009; Roy et al., 2014). For the healthy-exposure window (2019–2024), direct Landsat-derived NDVI observations were available for 2019, 2020, and 2022, with reconstructed values used for 2021, 2023, and 2024. Landsat imagery went through the following pipeline: (i) atmospheric correction using LEDAPS for TM/ETM+ and LaSRC for OLI-TIRS; (ii) cloud and shadow masking with the Fmask algorithm (Zhu et al., 2015; Zhu & Woodcock, 2012); (iii) surface reflectance products from Collection 2 Level-2; (iv) cross-sensor calibration using Landsat 8 OLI as the reference; (v) gap-filling of Landsat 7 SLC-off scenes using the algorithm of Storey et al. (2014); (vi) seasonal acquisition restricted to June–August for phenological consistency; and (vii) co-registration of all images to the 1991 reference image at sub-pixel accuracy (RMSE < 0.5 pixels). NDVI was calculated using standard formulas (Rouse et al., 1974; Tucker, 1979).

Because Landsat observations were not available for every year of the study period, the directly observed vegetation panel spans just seven observation years. We therefore used MODIS MOD13Q1 observations (2000–2022) to reconstruct district-year vegetation conditions (Justice et al., 2002; Didan, 2015). MODIS-derived NDVI (250 m native resolution) was spatially resampled to the Landsat analysis grid (30 m) for district-level aggregation, purely for spatial compatibility. All landscape metrics, including DPC, were computed at the native 250 m resolution and then aggregated to the district level; the 30 m grid served only as a computational alignment grid and introduced no pseudo-fine spatial structure. Cross-sensor calibration used 500 temporally matched Landsat–MODIS observation pairs drawn from 25 distinct scenes (Teillet et al., 2001; Chander et al., 2013), selected to minimize differences in acquisition date (within ± 5 days) and vegetation phenological state (same season) (Goward et al., 1991; Huete et al., 2002). Two hundred validation pairs from 12 scenes were withheld from calibration entirely, ensuring scene-level independence between calibration and validation. The cross-sensor calibration relationship was:

$$\text{NDVI}_{\text{Landsat}} = 1.02 \times \text{NDVI}_{\text{MODIS}} - 0.01$$

We assessed calibration performance using a linear mixed-effects model with scene-level random intercepts, obtaining a marginal R^2 of 0.92 and RMSE of 0.03 (Steven et al., 2003). Independent validation, using 200 pixels from the 12 withheld scenes, treated the effective number of independent validation units as 12 scenes rather than 200 pixels, since pixels within a scene are spatially correlated; pixel-level agreement was accordingly evaluated with clustered inference at the scene level, yielding a marginal $R^2 = 0.91$ and RMSE = 0.03 (Xiao et al., 2014). Confidence intervals for R^2 were estimated by scene-level bootstrap (1,000 replications): $R^2 = 0.92$ (95% CI: 0.88–0.95). We report the number of independent validation scenes alongside pixel counts throughout the validation tables. To characterize interannual vegetation health relative to historical conditions, we calculated the Vegetation Condition Index (Kogan, 1995) for each district-year:

$$\text{VCI}_{it} = \frac{\text{NDVI}_{it} - \text{NDVI}_{\min,i}}{\text{NDVI}_{\max,i} - \text{NDVI}_{\min,i}} \times 100$$

where NDVI_{it} is the district-level NDVI in year t , and $\text{NDVI}_{\min,i}$ and $\text{NDVI}_{\max,i}$ are the historical

minimum and maximum NDVI for district i , computed from the combined series of direct Landsat observations and reconstructed annual values over 1991–2022. Following established thresholds in the literature (Kogan, 1995; Liu & Kogan, 1996), VCI values below 35% were classified as indicating severe vegetation stress. This classification is reconstruction-dependent and should be read as a relative indicator of vegetation stress, not as a direct observation of vegetation mortality. We calculated the proportion of green space classified as severely stressed as the percentage of vegetated pixels ($\text{NDVI} > 0.20$) with $\text{VCI} < 35\%$ in each district-year, with confidence intervals derived from a district-level bootstrap procedure (1,000 resamples).

2.2.4. NDVI Reconstruction and DPC Derivation at the District Level

For years without direct Landsat observations, we reconstructed district-level growing-season NDVI using a single panel model fitted to the available 1991–2024 district-year series:

$$\text{NDVI}_{it} = \alpha_i + \gamma_t + \beta_1 T_{it} + \beta_2 P_{it} + \beta_3 \text{NDVI}_{i,t-1} + \epsilon_{it}$$

where i indexes districts and t indexes years; T_{it} and P_{it} denote district-level temperature and precipitation, and $\text{NDVI}_{i,t-1}$ denotes the previous calendar year's growing-season NDVI.

The final reconstruction model was fitted to the complete available panel only after all validation analyses were finished. During cross-validation, the model was refitted independently within each training fold, with the corresponding held-out observations excluded from fitting. For the lagged-NDVI term, we used the directly observed value whenever the previous year had a direct Landsat observation; when it did not, we used that year's own reconstructed NDVI from the same sequential procedure. The lagged inputs for the health-exposure years were therefore: observed 2020 NDVI for the 2021 reconstruction, observed 2022 NDVI for the 2023 reconstruction, and reconstructed 2023 NDVI for the 2024 reconstruction—the last a one-step recursive use of a reconstructed lagged value. The reconstruction model did not independently predict DPC. Instead, for each reconstructed decade-year, we ran the reconstructed NDVI surface through exactly the same vegetation-masking and landscape-metric procedure used for the directly observed years: pixels with $\text{NDVI} > 0.20$ were classified as vegetation, the resulting mask was represented on the native 250 m MODIS grid, and the Largest Patch Index (LPI) was calculated with FRAGSTATS v4.2. That LPI value becomes the DPC value for the district-year.

Reconstructed DUVRI-Eco values for 2021, 2023, and 2024 therefore incorporate reconstructed NDVI and the DPC derived from it. We treated these years as reconstructed exposure estimates rather than direct observations. Because the last direct Landsat observation was in 2022, the 2023 and 2024 values represent temporal extrapolations beyond the directly observed remote-sensing period. To avoid ambiguity, the climate variables used as predictors in the 2023 and 2024 reconstructions are drawn from the corresponding observed climate records, and their use does not imply direct Landsat observations in those years. Table 2 summarizes the lagged NDVI inputs and the corresponding climate records used to reconstruct the three years within the health-exposure window. We validated the reconstruction separately from the final model: direct-observation years were sequentially withheld, the model was refitted on the remaining observations, and the resulting predictions were compared with the withheld values. This assesses within-sample temporal prediction performance; it does not independently validate the 2023–2024 temporal extrapolation. The reconstruction achieved $\text{RMSE} = 0.024$, $\text{MAE} = 0.018$, mean bias = -0.002 , 95% prediction-interval coverage = 94.2%, and a calibration slope of 0.98 (95% CI: 0.91–1.05). Temporal holdout validation (leave-one-year-out, $n = 7$ folds) and spatial holdout validation (leave-one-district-out) yielded similar performance (RMSE range: 0.021–0.026). With only seven direct observations available, these metrics should be read with caution and not generalized with high confidence to every reconstructed year. The 2023–2024 reconstruction constitutes a temporal extrapolation

beyond the observed range (the last direct Landsat year was 2022). The leave-one-observed-year-out validation assesses predictive performance within the observed temporal range; it does not validate extrapolation itself. We assessed the stability of the health association by restricting the primary analysis to 2019–2022, which produced an IRR of 1.052 (95% CI: 1.018–1.087), evidence that the association holds up across the specified sensitivity analyses, though not evidence that the extrapolation itself is accurate. To avoid circularity, the reconstruction model did not include the DUVRI-Eco components (TXx, CDD, or DPC) as predictors; it included only temperature, precipitation, and a lagged NDVI term. The primary attribution analysis was restricted to the 105 directly observed Landsat district-year observations. We used the reconstructed NDVI only for (i) descriptive completeness of the long-term trajectory, (ii) health-exposure construction for 2021, 2023, and 2024, and (iii) sensitivity analyses. *Results* reports a sensitivity analysis comparing the health-model results using direct NDVI (2019–2020) with the reconstructed series. The complete reconstruction pipeline, model specification, lagged-NDVI definition, DPC computation procedure, year-by-year inputs, and cross-validation fold assignments are documented in Supplementary Methods S2. The R scripts, random seeds, and analytical outputs are available in the Mendeley Data repository (DOI: 10.17632/fg29rhyx3f.2). This design reduces circularity in the primary climate-vegetation attribution analysis, which is restricted to directly observed Landsat years. In the health-exposure construction, however, we used reconstructed NDVI values for 2021, 2023, and 2024, which introduces some coupling between the health exposure and the hydroclimatic predictors. We do not claim that the health exposure is fully independent of the predictors used in its reconstruction; this coupling should be regarded as a limitation of the health analysis.

2.2.5. Air Quality Data

PM_{2.5} data were available for 2019–2024: direct measurements for 2019–2020 and reconstructed values for 2021–2024. We generated the reconstructed values using the same IDW procedure (power parameter = 2) applied to the direct-measurement years, with interpolation parameters estimated from 2019–2020 and applied forward to 2021–2024. The validation performance of this procedure (RMSE = 2.1 $\mu\text{g}/\text{m}^3$, $R^2 = 0.89$, based on leave-one-year-out cross-validation 2019–2020) supports its use in later years. Because PM_{2.5} was reconstructed deterministically via IDW, we approximated its prediction uncertainty using the cross-validated RMSE (2.1 $\mu\text{g}/\text{m}^3$) from that leave-one-year-out validation. We used this RMSE as a pragmatic scale parameter for an independent Gaussian perturbation in a Monte Carlo measurement-error sensitivity analysis: for each of 1,000 uncertainty replicates, reconstructed PM_{2.5} values were perturbed by independent $N(0, \text{RMSE}^2)$ errors, and the prior health model was refitted to each replicate. We used the empirical distribution of the resulting coefficients to quantify reconstruction-related uncertainty. This independent-error perturbation is a simplifying approximation, since the limited number of monitoring stations did not support reliable estimation of a spatial residual covariance structure. Restricting the analysis to observed PM_{2.5} years (2019–2020) yielded nearly identical estimates (IRR = 1.049; 95% CI: 1.018–1.083), further supporting the stability of the association. We estimated district-level PM_{2.5} exposure for all 15 municipal districts using a consistent IDW procedure (power parameter = 2), based on measurements from 13 monitoring stations operated by the Isfahan Department of Environment (Jerrett et al., 2005; Wong et al., 2004). We treated PM_{2.5} a priori as a measured potential confounder, given its well-established cardiovascular outcomes and its spatial correlation with urban environmental conditions (Brook et al., 2010; Pope et al., 2009; Rajagopalan et al., 2018), and evaluated the reconstructed exposure through cross-validation and sensitivity analyses. Exposure-measurement error in the reconstructed values may

attenuate the estimated association; we incorporated the reconstruction uncertainty, approximated from the validated leave-one-year-out RMSE described above, into the standard errors of the health-effect estimates. We included PM_{2.5} as a measured potential confounder, but its causal role remains uncertain; it may act as a confounder, a correlated environmental exposure, or a mediator of the vegetation–health association. The PM_{2.5}-adjusted estimate should therefore be interpreted as a conditional association that survives adjustment for one measured potential confounder, not as a confounding-controlled effect.

2.2.6. Cardiovascular Hospitalization Data

We obtained hospitalization records for 2019–2024 from the Health Information Technology Unit of Isfahan University of Medical Sciences. The outcome analyzed here is therefore recorded cardiovascular hospitalizations within the participating university-affiliated hospital network, not city-wide cardiovascular hospitalizations; private hospital admissions unaffiliated with the university were not included, which is a potential source of selection bias. Records were included if they met the following criteria: age 18 or older; a primary cardiovascular diagnosis (ICD-10 codes I05–I09, I10–I15, I20–I25, I26–I28, I30–I52, I70–I79, and I95–I99) (WHO, 2016); sufficient diagnostic information; and residence within the metropolitan area (Javari & Javari, 2026). After applying these criteria, 3,207 hospitalization episodes remained, of which 3,022 (94.2%) could be geocoded to residential districts (Krieger et al., 2002; Bow et al., 2004). We excluded the remaining 185 non-geocoded episodes (5.8%) from the primary spatial analysis (Zandbergen, 2008; Mazumdar et al., 2010). Geocoded and non-geocoded records were comparable across characteristics (standardized mean differences < 0.10 for all variables) and fairly evenly distributed across districts ($\chi^2 = 8.7$, $df = 14$, $P = 0.79$) (Cohen, 1988; Stuart, 2010). We complemented the complete-case analysis with a sensitivity analysis evaluating the potential influence of the geocoding exclusions under best- and worst-case allocation scenarios, which placed the primary association within an IRR range of 1.048–1.058 (Greenland, 2001; Little & Rubin, 2019). Because individual identifiers were unavailable for linking admissions, the analytical unit was the hospitalization episode rather than the individual patient (Armstrong et al., 1999; Schootman et al., 2007). If the same person contributed to multiple admissions, those admissions are not statistically independent, which could inflate apparent precision through within-patient correlation; and if repeated admissions cluster unevenly across districts, that could bias the point estimate in either direction. We could not model this dependence, so both possibilities remain limitations of the health analysis. The analysis was explicitly ecological: hospitalization counts were aggregated to the district-month level and linked to district-level DUVRI-Eco and PM_{2.5} exposure estimates (Morgenstern, 1995; Wakefield, 2008).

The climate reconstruction covers the full 72-year period from 1951 to 2022 at the annual level, but the primary health analysis was restricted to 2019–2022 and comprised 720 district-month observations (15 districts $\times$ 48 months) across those same 15 districts. As detailed in Section 2.4.6, these repeated monthly observations do not add independent exposure contrasts beyond the district level, so the effective number of independent spatial exposure units remains 15.

2.3. Vegetation and Urban-Exposure Variables

We considered four analytical domains in the ecological vulnerability framework: hydroclimatic stress, vegetation condition, landscape structure, and urban exposure. Vegetation conditions were represented by the NDVI (Rouse et al., 1974; Tucker, 1979). Landscape structure was characterized using the Dominant Patch Cover (DPC), operationalized as the Largest Patch Index (LPI) via FRAGSTATS v4.2 (McGarigal et al., 2012; McGarigal & Marks, 1995):

$$DPC = \frac{\max\{a_j\}}{A} \times 100$$

where a_j is the area of the largest vegetation patch, and A is the district's total landscape area (Wu et al., 2000; Li & Reynolds, 1993). We calculated DPC using a vegetation-mask threshold of NDVI > 0.20, an empirically derived operational threshold consistent with regional vegetation-mask conventions (de Jesus et al., 2016), and sensitivity analyses at 0.15, 0.25, and 0.30 produced consistent results. DPC should be read as a measure of patch dominance, not as a direct measure of fragmentation or functional connectivity (Fahrig, 2003; Fischer & Lindenmayer, 2007). Larger DPC values indicate vegetation concentrated in larger, more contiguous patches, in Isfahan; this may reflect older municipal parks with established, deep-rooted cover and a reliable water supply (Madani, 2014), while smaller DPC values indicate more dispersed or fragmented vegetation, often corresponding to street trees or peri-urban patches dependent on erratic water supply (Alberti, 2005). DPC is operationally defined as the Largest Patch Index (LPI) calculated using FRAGSTATS v4.2; we use the term DPC to emphasize its interpretation as a measure of dominant-patch structure, though readers familiar with FRAGSTATS will recognize the mathematical equivalence. Because DPC is derived from the same NDVI-based vegetation mask used to compute NDVI, it is not an independent environmental dimension but a structurally transformed derivative of NDVI; its strong correlation with NDVI ($\rho = 0.76$) reflects this shared derivation. DPC is retained as a construct-defined component representing patch dominance rather than as an independent predictor of vegetation mortality; its incremental contribution, evaluated through ablation analysis, was not statistically distinguishable ($\Delta AUC = -0.02$; 95% CI: -0.06 to 0.02), so its retention reflects a construct-based decision rather than a demonstrated incremental predictive contribution. The DPC should not be read as evidence that landscape structure independently explains mortality. A useful further check, left to future analysis of this dataset, would be to compare the full four-component indicator directly against a reduced three-component version (TXx, CDD, and NDVI only) across ranking, AUC, the transition point, district classification, and health IRR; if the two versions track closely, that would itself support the architecture's robustness to DPC's redundancy rather than undermine it. To preserve the temporal ordering required for early-warning interpretation, the NDVI and DPC values feeding into DUVRI-Eco came from the growing season preceding each mortality survey (Huang et al., 2010; White et al., 2009); for each site, we used the most recent Landsat observation before the mortality assessment date (Pettorelli et al., 2005). This temporal ordering holds for the vegetation components themselves but does not establish prospective prediction of mortality for the composite indicator as a whole. The climate components (TXx and CDD), being annual summaries, align with the mortality assessment less precisely.

2.4. Statistical Analysis

2.4.1. Trend Analysis

We evaluated temporal trends in climate variables using the modified Mann-Kendall test (Hamed & Rao, 1998; Yue & Wang, 2004) and Sen's slope estimator (Sen, 1968; Theil, 1950). The modified procedure corrects the test statistic's variance for serial dependence using the effective sample size, as implemented in the R package *trend* (v1.1.4) (Pohlert, 2020; Hipel & McLeod, 1994). Statistical significance was assessed with two-sided tests at $\alpha = 0.05$ (Kendall, 1975; Mann, 1945). Because we evaluated multiple climate indices and temporal series (48 tests in total: 4 climate variables $\times$ 4 stations $\times$ 3 seasons), we applied false discovery rate control at $q = 0.10$ using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995).

2.4.2. Structural-Break Analysis

We identified potential changes in the temporal structure of hydroclimatic variables using Bai-Perron multiple-breakpoint analysis (Bai & Perron, 2003; Bai & Perron, 1998), with the Pettitt (1979) and SNHT (Alexandersson, 1986) procedures serving as corroborative tests (Pettitt, 1979; Alexandersson & Moberg, 1997). As before, we distinguished climatic breakpoints from measurement-inhomogeneity breakpoints and adjusted only the latter. We retained breakpoints for interpretation only when their statistical evidence agreed with the corroborative analyses ($P < 0.05$) and the estimated segment lengths were sufficient for stable parameter estimation (minimum segment length ≥ 10 years) (Bai & Perron, 2003; Zeileis et al., 2003).

2.4.3. Climate-Vegetation Association Model

We evaluated district-level vegetation response using a two-way fixed-effects panel model on the 105 directly observed Landsat observations (15 districts $\times$ 7 years) (Baltagi, 2005; Wooldridge, 2010):

$$\text{NDVI}_{it} = \alpha_i + \gamma_t + \beta_1 X_{it} + \beta_2 U_i + \beta_3 M_{it} + \beta_4 [X_{it} \times U_i] + \epsilon_{it}$$

where NDVI_{it} denotes vegetation condition in district i and year t ; α_i and γ_t are district and year fixed effects; X_{it} is time-varying hydroclimatic exposure (temperature anomaly); M_{it} is time-varying management characteristics (irrigation water use, treated-wastewater allocation); and U_i is the time-invariant urban climatic exposure measure. Because U_i is time-invariant at the district level, its main effect is absorbed into the district fixed effects α_i and is not separately identified; the interaction term $X_{it} \times U_i$ is identified because X_{it} varies over time within districts. We centered continuous predictors before forming the interaction term (Schiele, 2010). Standard errors were clustered at the district level ($n = 15$ clusters) (Cameron et al., 2008; Cameron & Miller, 2015), and given the limited number of independent districts, we supplemented conventional asymptotic p-values with wild cluster bootstrap p-values, reporting conventional clustered standard errors alongside for comparison. The climate-vegetation panel model should be read as a supporting descriptive analysis: with only 15 district clusters and 105 observations, its inferential power is limited, and its results should not be presented as primary evidence of mechanism in the Abstract or Conclusion.

2.4.4. Relative Contribution Analysis

To quantify how each hydroclimatic and landscape predictor contributed to explaining the response variable, we applied the Lindeman-Merenda-Gold (LMG) decomposition (Lindeman et al., 1980; Grömping, 2006). In this panel specification, the decomposed quantity is the centered R^2 of the two-way fixed-effects model after within-transformation: (1) the R^2 is the centered R^2 of the within-transformed model; (2) year fixed effects are absorbed before decomposition; (3) missingness is handled by complete-case analysis across the 105 directly observed district-year observations; and (4) the relaimpo implementation was customized to match this fixed-effects formulation. Because our panel data consist of repeated observations nested within only 15 municipal districts, we treated the LMG estimates as descriptive measures of relative explanatory contribution rather than inferential statistics that assume independent observations (Johnson & LeBreton, 2004; Nimon et al., 2008), and we did not interpret the contribution percentages as causal effects or as conventional significance tests for individual predictors. The LMG relative-contribution analysis, reported in Supplementary Section S4, is a descriptive decomposition of the within-transformed model's R^2 , not equivalent to the composite weights and provided for descriptive completeness only.

2.4.5. Outcome-Calibrated DUVRI-Eco Construction and Nested Validation

DUVRI-Eco was developed as a composite indicator integrating four constructs—defined components of vegetation vulnerability: thermal stress, drought persistence, vegetation condition,

and landscape patch dominance (Javari, 2025a; Javari & Javari, 2026). These dimensions are represented by annual maximum temperature (TXx), consecutive dry days (CDD), the normalized difference vegetation index (NDVI), and Dominant Patch Cover (DPC), operationalized via the Largest Patch Index (LPI) as a proxy for patch dominance (Frich et al., 2002; Rouse et al., 1974; McGarigal et al., 2012). Before aggregation, we transformed the four components to a common dimensionless scale and reoriented them so that higher values consistently indicate greater ecological vulnerability (Saisana & Tarantola, 2002; OECD, 2008). For each component j , we performed robust min-max normalization using the 1st and 99th empirical percentiles (Häder & Häder, 1995; Carpenter, 2001):

$$Z_{ij} = \text{clip} \left[\frac{X_{ij} - Q_{0.01,j}}{Q_{0.99,j} - Q_{0.01,j}}, 0, 1 \right]$$

Values below the 1st percentile were clipped to 0 and values above the 99th to 1. We assessed sensitivity to the choice of percentile by repeating the normalization at the 2.5th–97.5th and 5th–95th percentiles; the resulting weights and health IRRs were nearly identical (within ± 0.01 for weights and ± 0.003 for IRRs). As a further sensitivity analysis, we recalculated DUVRI-Eco using z-score normalization in place of percentile-based min-max normalization; the resulting district rankings correlated strongly with the primary specification (Spearman's $\rho = 0.96$), and the health IRR was essentially unchanged (IRR = 1.050; 95% CI: 1.016–1.085). We assessed the stability of district rankings under alternative normalization scenarios. The Spearman rank correlation between district DUVRI-Eco scores under the 1st–99th, 2.5th–97.5th, and 5th–95th percentile scenarios was at least 0.97, and the maximum absolute rank change for any district was a single position, indicating that the spatial ordering of districts is robust to the choice of normalization percentile. For the protective variables (NDVI, DPC), we reverse-coded the normalized scores as:

$$Z_{ij} = 1 - Z_{ij}^{\text{raw}}$$

where Z_{ij}^{raw} is the raw normalized value before reverse-coding. We specified the composite DUVRI-Eco score as a weighted linear combination (Freudenberg, 2003; Böhringer & Jochem, 2007):

$$\text{DUVRI-Eco}_i = w_{TX}Z_{TX,i} + w_{CDD}Z_{CDD,i} + w_{NDVI}Z_{NDVI,i} + w_{DPC}Z_{DPC,i}$$

subject to $w_j \geq 0$ and $\sum w_j = 1$ (OECD, 2008; Greco et al., 2019).

We chose the additive form as the primary specification because it preserves each component's monotonic contribution and avoids estimating interaction parameters that would be only weakly supported by 15 district-level exposure profiles. The additive formulation should be understood as an operational approximation of a multi-dimensional stress construct, not a claim that heat and drought act independently. A sensitivity analysis using a multiplicative formulation (DUVRI-Eco_multiplicative = $Z_{TX} \times Z_{CDD} \times Z_{NDVI} \times Z_{DPC}$) yielded a health IRR of 1.046 (95% CI: 1.012–1.081), consistent with the primary estimate. We did not separately test an explicit heat–drought interaction term (e.g., a $TXx \times CDD$ product) within the calibration model; given that only 15 independent district-level exposure profiles are available, such an interaction is unlikely to be reliably identified, but formally confirming this, and thereby providing a direct empirical justification for the additive specification, is a natural extension for a larger calibration sample. The indicator is better described as a multi-dimensional heat–drought stress indicator than as a statistically explicit measure of compound events. A formal compound-event analysis using copulas or joint exceedance probabilities would require a larger number of independent spatial units and is a worthwhile direction for future research. We estimated the normalization parameters (percentiles) within each training fold during cross-validation to prevent information leakage from the test set. The final descriptive indicator uses normalization parameters estimated from the

complete dataset; this version is used only for mapping and descriptive purposes, not for validation. All validation metrics are based on out-of-fold predictions.

Because all DUVRI-Eco components are defined at the district level, we specified the primary calibration model at that level. For each district d , mortality was summarized as a pooled proportion:

$$p_d = \frac{\sum_{i \in d} D_i}{\sum_{i \in d} n_i}$$

where D_i is the number of dead trees at site i and n_i is the number of assessed trees at site i , weighting sites by the number of trees assessed. The district-level model was:

$$\text{logit}(p_d) = \alpha + \beta \cdot \text{DUVRI}_d$$

As a sensitivity analysis, we also fitted a site-level hierarchical model with district random effects:

$$\text{logit}(p_i) = \alpha + u_d + \beta \cdot \text{DUVRI}_d$$

where $u_d \sim N(0, \sigma_u^2)$. Both specifications yielded similar weights; the district-level model remains the primary specification because it matches the spatial scale of the DUVRI-Eco components.

We did not include species composition in the aggregation, since the calibration data do not support species-specific weights. A descriptive heterogeneity analysis of site-level mortality, tree counts, and species composition appears in results; future studies with larger samples should examine whether species composition modifies the weight structure. We jointly estimated the intercept, the response coefficient, and the component weights by constrained maximum likelihood, subject to $w_j \geq 0$, $\sum w_j = 1$, and $\beta \geq 0$ (Collett, 2003; Nelder & Wedderburn, 1972). We imposed the $\beta \geq 0$ constraint a priori to preserve the intended monotonic direction of the vulnerability construct, the conceptual expectation that higher composite vulnerability should not be associated with lower vegetation mortality. As a sensitivity analysis, we also fitted the calibration model without this constraint; the estimated β remained positive, and the resulting weights were nearly identical (within ± 0.01). We solved the constrained optimization directly via sequential quadratic programming (SQP), and as an independent numerical check, also performed box-constrained optimization (L-BFGS-B) followed by projection onto the simplex; the two procedures produced identical weights to three decimal places, confirming convergence to the same constrained maximum-likelihood solution. We repeated the SQP procedure with 100 random starting values drawn from a Dirichlet (1) distribution and retained the solution with the highest log-likelihood, declaring convergence when the relative change in log-likelihood fell below 10^{-6} for three consecutive iterations. Numerical convergence does not, by itself, imply statistical identifiability: the manuscript demonstrates the former, but the data remain insufficient to precisely identify the weight vector, which is why we report profile-likelihood diagnostics for every weight.

The constrained maximum-likelihood solution placed the largest weight on CDD (0.31), followed by TXx (0.28), NDVI (0.26), and DPC (0.15). Bootstrap-derived 95% confidence intervals, based on a district-clustered bootstrap with 1,000 replications (Table 3), were CDD (0.26–0.37), TXx (0.22–0.34), NDVI (0.20–0.32), and DPC (0.09–0.21). Profile-likelihood diagnostics, however, revealed limited identifiability of the individual weights, particularly for CDD (flat between 0.26 and 0.37) and TXx (flat between 0.22 and 0.34), a pattern indicating that several weight configurations fit the calibration data about equally well. These estimates are therefore best read as an admissible calibration solution rather than a precise accounting of each component's relative importance, and the weight ranking should be treated as suggestive rather than definitive. Given this, we regard the calibrated weights as domain-specific regularization parameters that anchor the composite construct to the Isfahan calibration sample, rather than as estimable population-level component effects that would be expected to hold in other settings. In short, the data constrain the

composite score far more tightly than they constrain the individual weights, and the confidence intervals describe the range of plausible values under the specified model rather than evidence of precise parameter estimation. District-level calibration data, the correlation matrix, the condition number, and the weight surface are reported in Supplementary Section S2.

In practical terms, the calibrated weights are best regarded as domain-specific regularization parameters that stabilize the composite score for this study domain, rather than as estimable population-level effects of each environmental dimension.

Bootstrap confidence intervals describe sampling variability under the specified model and should be read as descriptive measures of uncertainty. Profile-likelihood diagnostics describe the identifiability of individual parameters and the range of values the likelihood supports. The two are not equivalent: bootstrap intervals can appear narrow even when the likelihood surface is flat, and profile likelihood can reveal weak identifiability even when bootstrap intervals appear precise. We report both, treating profile likelihood as the primary diagnostic for weight identifiability. To characterize the stability of the composite indicator score, rather than the precision of the individual weights, we generated 10,000 weight vectors from the likelihood-supported region, defined as the set of weights with log-likelihood within 2 units of the maximum. This region should be read as a sensitivity region for robustness analysis, not as a formal confidence region: for a four-component weight vector under a simplex constraint, the effective parameter dimension exceeds one, so the 2-unit cutoff does not correspond to a 95% confidence region under either profile or joint likelihood definitions. We use this region solely to characterize the range of weight configurations that fit the calibration data comparably well. The 2-unit cutoff itself is a conventional but ultimately arbitrary choice; we did not re-run the stability analysis under alternative cutoffs (e.g., $\Delta\ell = 1$ or $\Delta\ell = 3$), and confirming that the rank-stability and score-stability conclusions hold under those alternative bounds is a worthwhile check for future work using this dataset. For each candidate weight vector, we re-optimized the intercept and mortality-response coefficients conditional on those weights, using the same constrained maximum-likelihood procedure, and evaluated the likelihood at the resulting conditional optimum. The sampling algorithm was as follows:

1. Parameterize the weight simplex using a Dirichlet (1) proposal distribution, ensuring $w_j \geq 0$ and $\sum w_j = 1$.
2. Draw candidate weight vectors independently from Dirichlet (1).
3. For each candidate, re-optimize the intercept and mortality-response coefficients conditional on the candidate weights.
4. Evaluate the likelihood at the resulting conditional optimum.
5. Accept vectors with $\log\text{-likelihood} \geq (\max \log\text{-likelihood} - 2)$.
6. Continue sampling until 10,000 accepted vectors are obtained.
7. Fix the random seed at 20240517 for reproducibility.
8. The acceptance rate was 3.7%, yielding 10,000 accepted vectors from 270,270 candidate draws.

For each accepted weight vector, we calculated DUVRI-Eco for all 15 municipal districts and assessed five features: (i) the rank correlation between the resulting indicator and the primary indicator (Spearman's ρ); (ii) the maximum absolute rank change for any district; (iii) the stability of the candidate transition-point classification; (iv) the distribution of district-level DUVRI-Eco scores; and (v) the absolute score change relative to the maximum-likelihood solution. The median rank correlation across the plausible weight surfaces was 0.94 (IQR: 0.89–0.97), and the maximum rank change for any district was 3 positions. We considered a district classification-stable if its

category relative to the candidate transition point (above vs. below $\tau \approx 0.65$) held constant across all 10,000 accepted weight vectors; by this definition, 13 of 15 districts (86.7%) were classification-stable. The two classification-unstable districts are identified, which reports the percentage of accepted weight vectors classifying each district above and below τ , along with the rank range and score range. Beyond rank stability, we also assessed absolute score stability: for each district, we calculated the absolute difference between its DUVRI-Eco score under each accepted weight vector and its score under the maximum-likelihood solution. The median absolute score change across all districts and weight vectors was 0.021 (IQR: 0.009–0.038), with a maximum absolute change of 0.087. In other words, while the individual weights are weakly identified, the resulting indicator score and its spatial classification vary relatively little across the plausible weight surface. Exploratory nonlinear mortality-response model. To characterize potential nonlinearity in the relationship between DUVRI-Eco and vegetation mortality, we fitted a segmented logistic regression model at the district level as a secondary analysis, parameterized with a continuous hinge function:

$$\text{logit}(p^d) = \alpha + \beta_1 D^d + \beta_2 (D^d - \tau)_+$$

where p^d is the pooled vegetation-mortality proportion in district d , D^d is the DUVRI-Eco score, $(D^d - \tau)_+ = \max(D^d - \tau, 0)$, α is the intercept, β_1 is the slope below the transition point, β_2 is the change in slope at the transition point, and τ is the empirically estimated transition point. The slope below τ is therefore β_1 , and the slope above τ is $\beta_1 + \beta_2$.

We estimated the transition point by profile likelihood, evaluating candidate values of τ across the observed DUVRI-Eco range subject to the interior constraint $Q_{0.10} \leq \tau \leq Q_{0.90}$. For each candidate value, we fitted the logistic model using maximum likelihood with τ fixed and recorded the corresponding profile log-likelihood; $\hat{\tau}$ was defined as the candidate value with the highest profile log-likelihood. We defined a 95% profile-likelihood confidence interval by $2[\ell(\hat{\tau}) - \ell(\tau)] \leq \chi^2_{1,0.95} = 3.84$. We refined the transition-point grid from an initial increment of 0.005 to 0.001 and coarsened it to 0.010 as sensitivity analyses, and repeated the analysis with the wider constraint $Q_{0.05} \leq \tau \leq Q_{0.95}$. We fitted the segmented model after fixing the DUVRI-Eco component weights at their primary maximum-likelihood estimates, so it was not used to re-estimate the composite-indicator weights. The estimated transition point was $\hat{\tau} = 0.65$ (95% profile-likelihood CI: 0.62–0.68); the estimated slope below the transition was $\beta_1 = 1.8$ (95% CI: 1.0–2.6), and the estimated change in slope above the transition was $\beta_2 = 2.4$ (95% CI: 1.2–3.6). These estimates correspond to an above-transition slope of $\beta_1 + \beta_2 = 4.2$ and a slope ratio (above vs. below) of 2.3 (95% CI: 1.6–3.2). Supplementary Figure 1 shows the profile log-likelihood for τ . To assess the identifiability of the transition point, we checked three conditions: (i) the profile log-likelihood has a well-defined maximum rather than a flat plateau; (ii) both β_1 and $\beta_1 + \beta_2$ are bounded away from zero; and (iii) at least five districts fall on each side of $\hat{\tau}$ in the calibration data. Sensitivity analyses using alternative grid resolutions (0.001 and 0.010) yielded identical estimates of $\hat{\tau}$ to two decimal places, with slopes differing by less than 5%. Using the wider constraint $Q_{0.05} \leq \tau \leq Q_{0.95}$ gave $\hat{\tau} = 0.64$ (95% CI: 0.61–0.67), with slopes $\beta_1 = 1.7$ and $\beta_1 + \beta_2 = 4.1$, consistent with the primary specification. Because we estimated τ from the calibration sample itself, we interpret it as a descriptive transition in the observed exposure-response relationship, not as a biological threshold or a validated management cutoff. The segmented analysis is accordingly hypothesis-generating and was not used to define operational vulnerability categories. Because the component weights were only weakly identified within the 15-district calibration domain, we conducted a weight-sensitivity analysis to assess how robust the health association is to alternative, plausible weight configurations. This analysis is not intended to represent a formal probability or posterior

distribution over the component weights. The primary health analysis used the maximum-likelihood weight vector $w_{MLE} = (0.28, 0.31, 0.26, 0.15)$ for TXx, CDD, NDVI, and DPC, respectively. For the sensitivity analysis, we sampled 1,000 weight vectors from the prespecified likelihood-supported region defined by $\Delta\ell \leq 2$, where $\Delta\ell = \ell_{max} - \ell(w)$, drawing independent Dirichlet (1,1,1,1) candidates and retaining those that satisfied the likelihood criterion, continuing until we obtained 1,000 accepted vectors (random seed fixed at 20240517). For each retained weight vector $w^{(r)}$, we recalculated district-level DUVRI-Eco scores as

$$DUVRI_d^{(r)} = \sum_j w_j^{(r)} Z_{dj}$$

then refit the primary health model using $DUVRI_d^{(r)}$ as the exposure while keeping the rest of the specification unchanged, recording the resulting incidence-rate ratio for each replicate.

We summarize the distribution of the resulting IRRs by their median and 2.5th and 97.5th percentiles, referred to here as a sensitivity range rather than a confidence, credible, or formal uncertainty interval, since it quantifies how much the estimated health association shifts across weight configurations that fit the calibration data comparably well.

The weight-sensitivity analysis yielded a median IRR of 1.049, with a 2.5th–97.5th percentile sensitivity range of 1.012–1.089. This range overlapped the primary model-based confidence interval (IRR = 1.052; 95% CI: 1.018–1.087) and remained directionally consistent with the primary estimate across the likelihood-supported weight configurations. This analysis does not propagate uncertainty from the NDVI reconstruction, PM2.5 reconstruction, hospitalization ascertainment, or sampling variability in the calibration data; a joint uncertainty model incorporating all these sources was beyond the scope of the present study. The sensitivity range should therefore be read as a robustness check conditional on the specified sampling scheme and likelihood-supported region, not as a complete characterization of uncertainty in the health association. To obtain less optimistic estimates of out-of-sample predictive performance at different spatial scales, we implemented a district-blocked leave-one-district-out cross-validation framework with two distinct validation objectives (Varma & Simon, 2006; Hastie et al., 2009). Because DUVRI-Eco is built from district-level variables, validation strategies must account for the risk of data leakage when sites from the same district appear in both training and test sets. First, to assess within-district predictive performance, we used a leave-one-site-out strategy as the outer loop, with 34 outer folds (one per field site); because co-located sites share district-level predictor values, this scheme is best read as a within-district test rather than an independent geographic test. Second, to assess internal geographic validation, we used a leave-one-district-out strategy as the outer loop, with 15 outer folds (one per municipal district): within each fold, all sites from the withheld district served as the test set, and the model was trained exclusively on sites from the remaining 14 districts. This district-block internal validation yielded an AUC of 0.79 (95% CI: 0.73–0.85), which we treat as the primary estimate of internal geographical discrimination performance. For each leave-one-site-out fold, we performed the following steps:

1. Remove the held-out site from the dataset.
2. Recompute the district-level pooled mortality proportion for the district containing the held-out site, using only the remaining sites in that district.
3. Recompute all normalization parameters (1st and 99th percentiles) using only the training data.
4. Re-estimate the four component weights using only the training data.
5. Re-estimate the transition points using only the training data.
6. Calculate the hold-out site's DUVRI-Eco score using the training-derived normalization

and weights.

7. Predict the held-out site's mortality using the training-derived model.

Leakage-free district-blocked calibration algorithm. For each leave-one-district-out fold:

1. Remove all sites in the hold-out district from the dataset.

2. Recompute the district-level pooled mortality proportions using only the training districts.

3. Recompute all normalization parameters (1st and 99th percentiles) using only the training districts.

4. Re-estimate the four component weights using only the training districts.

5. Re-estimate the transition points using only the training districts.

6. Calculate the hold-out district's DUVRI-Eco score using the training-derived normalization and weights.

7. Predict the hold-out district's mortality using the training-derived model.

8. Aggregate predictions across all held-out districts to compute the district-blocked AUC.

These procedures ensure that no information from a hold-out site or district is contributed to any calibration or normalization step.

We obtained confidence intervals for the pooled out-of-fold AUC using 1,000 district-level bootstrap resamples, preserving the district as the resampling unit; the bootstrap was applied to the fixed pooled out-of-fold predictions without refitting the nested cross-validation models within each replicate, so the resulting interval reflects district-level resampling variability conditional on the fitted cross-validation predictions. The district-blocked AUC was 0.79 (95% CI: 0.73–0.85). This uncertainty should still be read in light of the small number of independent districts ($n = 15$) and the low event counts in several held-out folds (range: 3–12 mortality events per fold). The distribution of fold-specific AUCs was: median = 0.80, IQR = 0.74–0.85, range = 0.68–0.91. A sensitivity analysis excluding the most influential district (District 7) yielded an AUC of 0.77 (95% CI: 0.70–0.84), a modest change. This estimate should be read as internal geographic discrimination within the Isfahan study domain, not as evidence of external validity.

The site-level AUC of (0.81) is useful for predicting mortality at new sites within districts that already have calibration data; because sites within the same district share identical district-level predictor values, however, the 34 site-level folds do not represent 34 independent validation cases. The district-block AUC (0.79) is therefore the primary estimate for assessing spatial generalization to hold-out districts within the same urban system.

All screening performance metrics (sensitivity, specificity, PPV, NPV, decision curve analysis) are out-of-fold estimates from the nested cross-validation procedure, not values computed on the full fitted sample. These metrics are reported as descriptive measures; because the candidate transition point is not the validated operational cutoff, they should not be read as clinical or management decision thresholds.

2.4.6. Cardiovascular Disease Analysis

Following calibration and validation, we treated the final DUVRI-Eco formulation as the environmental exposure metric for the cardiovascular disease analysis (Javari & Javari, 2026). For the health analysis, $DUVRI_d$ was defined as the district-level mean of the annual DUVRI-Eco scores over 2019–2022. Because the composite indicator's components are defined at the district level and the interannual variation in component scores was small relative to between-district differences, we treated this district-level summary as time-invariant during the health observation period. The analysis is best characterized as a district-level ecological panel analysis with a time-invariant exposure: although hospitalization outcomes were observed monthly, the exposure itself was not, so all between-district inference rests on the 15 municipal districts of Isfahan, the primary

spatial exposure units throughout. The analysis does not establish temporal precedence of exposure and outcome and should not be read as a longitudinal exposure-response model. We conducted the health analysis independently of the vegetation-mortality-based indicator calibration; cardiovascular outcomes played no part in the DUVRI-Eco weights or their vegetation-mortality thresholds (Greenland, 2001). We pre-specified the following inferential hierarchy for the health analysis: (1) district-level ecological regression as the primary specification; (2) the ICAR spatial model as a sensitivity analysis; (3) district-clustered bootstrap for confidence intervals; and (4) a district-level permutation test as a descriptive stability assessment. The district-level regression is the primary basis for interpretation; Table 4 identifies the inferential role of each specification explicitly. The primary cardiovascular analysis was restricted to 2019–2022, for which direct Landsat observations were available for 2019, 2020, and 2022, with a reconstructed value used for 2021. The extended analysis (2019–2024), which used reconstructed NDVI values for 2023 and 2024, is reported as a sensitivity analysis. This structure keeps the primary health estimate from depending on temporal extrapolation beyond the observed range (the last direct Landsat year was 2022). We analyzed cardiovascular hospitalization counts using a generalized linear mixed-effects framework with a Poisson distribution and log link (Liang & Zeger, 1986; Diggle et al., 2002):

$$Y_{dt} \sim \text{Poisson}(\mu_{dt})$$

$$\log(\mu_{dt}) = \log\left(\frac{\text{Pop}_{dt}}{12}\right) + \beta_0 + \beta_1 \text{DUVRI}_d + \beta_2 \text{PM}_{2.5,dt} + \beta_3 \text{Month}_t + \beta_4 \text{Year}_t + b_d$$

where Y_{dt} denotes cardiovascular hospitalization episodes in district d and month t (McCullagh & Nelder, 1989); Pop_{dt} is the annual population for district d in year t , and the offset $\log(\text{Pop}_{dt} / 12)$ reflects the monthly exposure period; DUVRI_d represents the district-level vegetation vulnerability measure, time-invariant within the health observation period; $\text{PM}_{2.5,dt}$ represents particulate exposure (Brook et al., 2010; Pope et al., 2009); Month_t (12 categories) and Year_t (four categories for the primary window; 2019, 2020, 2021, 2022, with 2019 as reference) are categorical controls; and b_d is the district-level random intercept (Bates et al., 2015; Pinheiro & Bates, 2000). We calculated the incidence rate ratio (IRR) for each predictor as $\text{IRR}_\Delta = \exp(\beta \cdot \Delta)$, with $\Delta = 0.10$ for the primary analysis (Rothman et al., 2008; Vittinghoff et al., 2012).

Population denominators came from the Isfahan Statistical Yearbook for annual district-level estimates. Because monthly population estimates were available, we assumed a uniform distribution across months for the monthly analyses and used $\log(\text{Pop}_{dt} / 12)$ as the offset, effectively assuming that person-time exposure is proportional to annual population divided by 12. A sensitivity analysis using annual population as the offset yielded nearly identical results (IRR = 1.051 vs. 1.052).

The model included month fixed effects (12 categories) and year fixed effects to account for seasonal patterns and secular trends in cardiovascular hospitalizations. To assess residual temporal dependence, we additionally fitted a sensitivity model with an AR (1) temporal correlation structure within districts.

The Pearson χ^2/df was 1.08, and the Deviance/df was 1.06, both close to 1, indicating no substantial overdispersion; we retained the Poisson specification as the primary model. A negative binomial model fitted as a sensitivity check yielded a nearly identical IRR (1.054 vs. 1.052). Overdispersion is not the principal limitation of this analysis; the primary limitations are exposure dimensionality (15 municipal districts), ecological confounding, spatial identifiability, and episode-level dependence. The negative binomial sensitivity analysis shows that overdispersion is not driving the results.

The health dataset comprised 720 district-month observations in the primary window and 3,022

geocoded hospitalization episodes. DUVRI-Eco, however, was defined at the level of Isfahan's 15 municipal districts and remained time-invariant during the health observation period; the health analysis was conducted across these same 15 districts, the primary spatial exposure units. The coefficient is subsequently identified from between-district exposure contrasts across these 15 districts. While repeated monthly outcomes improve estimation of temporal patterns, it does not create additional independent exposure contrasts; although the primary analysis includes 720 district-month observations, the effective number of independent spatial exposure units remains 15. The 15 municipal districts are the official administrative units of the study; the limitation is not the number of districts per se, but the fact that the exposure contrast is identified at the district level. The DUVRI-Eco coefficient is identified solely from between-district contrasts, leaving it vulnerable to unmeasured district-level confounding (e.g., socioeconomic status, smoking prevalence, health services access). The reported confidence interval describes the stability of the association under the specified model; it is not a precise effect estimate based on independent observations. The health analysis should be read as a descriptive ecological association, not as a confirmatory epidemiological finding.

Because DUVRI-Eco is time-invariant within districts, a conventional district fixed-effect model would absorb the exposure completely, so we used a district random intercept instead, which assumes that district-level heterogeneity is uncorrelated with the exposure, an assumption the current design cannot test. The district random intercept accounts for residual clustering but does not control for time-invariant district-level confounding correlated with DUVRI-Eco. The Mundlak formulation, which includes district-level means of time-varying covariates, partially addresses this concern but does not eliminate the possibility of unmeasured district-level confounding. The random intercept should be understood as a modeling choice that enables estimation of the between-district association, not as a solution to confounding.

To partially address this, we implemented three complementary strategies: (1) a between-within (Mundlak) formulation (Mundlak, 1978; Bell et al., 2019); (2) an ICAR spatial model (Besag et al., 1991) as a sensitivity analysis; and (3) a district-level permutation test (10,000 permutations) as a descriptive stability assessment. None of these adjustments amounts to controlling for unmeasured district-level confounding.

Because DUVRI-Eco exhibited spatial autocorrelation (Moran's $I = 0.35$), unrestricted permutation of district exposure labels does not preserve the observed spatial structure, so we do not treat this analysis as a formal spatial hypothesis test; we report it only as a descriptive stability assessment in the Supplementary Material. A spatially constrained permutation scheme would offer a more appropriate null distribution and is worth pursuing in future studies with larger spatial systems. As noted above, the climate reconstruction covers the full 72-year period from 1951 to 2022 at the annual level, while the primary health analysis was restricted to 2019–2022 and comprised 720 district-month observations (15 districts $\times$ 48 months) across those same 15 municipal districts. Because DUVRI-Eco was time-invariant at the district level during the health observation period, these 720 district-month observations represent repeated measurements within 15 districts subject to spatial dependence, not 720 independent exposure units; a larger number of monthly observations does not increase the number of independent exposure contrasts.

The outcome analyzed here is recorded cardiovascular hospitalizations within the participating university-affiliated hospital network, not city-wide cardiovascular hospitalizations; private hospital admissions unaffiliated with the university were not included. If districts differ in their access to university hospitals, the observed association may partly reflect care rather than cardiovascular disease incidence, so the health should be described as a hospitalization rate or

recorded hospitalization burden, not as individual cardiovascular risk.

2.4.7. Sensitivity and Robustness Analyses

We assessed stability under alternative specifications through pre-specified sensitivity analyses addressing the principal sources of methodological uncertainty (Carpenter & Bithell, 2000; Green et al., 2010); the full list of sensitivity analyses and their results appears in the Supplementary Material. Because this study includes a large number of analyses that do not belong to a single family of confirmatory hypotheses, we do not apply one global multiplicity correction across all of them; instead, Table 4 organizes them into an explicit three-tier inferential hierarchy: (1) primary confirmatory/descriptive analyses, the DUVRI-Eco calibration and the district-blocked AUC; (2) secondary analyses, the segmented mortality response and the cardiovascular health association; and (3) sensitivity analyses, the ICAR model, PM2.5 adjustment, the reconstructed-year extensions, and the weight-perturbation checks. Only the first tier is treated as the primary basis for the study's conclusions; the second and third tiers support and stress-test those conclusions without independently establishing them. To assess sensitivity to spatial neighborhood specification, we additionally fitted the ICAR model using three alternative spatial weights matrices: queen contiguity, k-nearest-neighbor ($k = 3$ and $k = 5$), and inverse-distance-based weights; results appear in Table 5. The breadth of these sensitivity analyses should not be read as compensating for the limited number of independent exposure units; they test specification stability but do not increase that number.

2.4.8. Spatial Analysis

We assessed spatial autocorrelation using global Moran's I with a queen-contiguity weights matrix (Moran, 1950; Cliff & Ord, 1981). The queen-contiguity graph formed a single connected component across all 15 municipal districts, averaging 3.1 neighbors per district (range: 2–5). The positive global bivariate Moran's I (0.31; 95% CI: 0.12–0.50; permutation-based $P = 0.02$) suggested spatial co-occurrence between higher DUVRI-Eco values and higher hospitalization rates; given the limited number of districts ($n = 15$), we read this as descriptive secondary evidence rather than evidence of a spatial process, and it is reported in the Supplementary Material without further over-interpretation. Spatial error models serve as robustness diagnostics (Anselin, 1988; Bivand et al., 2013). We conducted spatially blocked cross-validation by withholding entire districts from model development, deriving all preprocessing parameters, component weights, model specifications, and threshold estimates exclusively from the remaining districts (Roberts et al., 2017; Meyer et al., 2018).

We specified the ICAR model as a sensitivity analysis for the health analysis because it explicitly models spatially structured residual variation across the 15 municipal districts. We fitted the ICAR specification using a row-standardized queen-contiguity spatial weights matrix, confirmed the spatial graph formed a single connected component, applied a sum-to-zero identifiability constraint and penalized-complexity (PC) priors, and fitted the model using integrated nested Laplace approximation (INLA). The full model was specified as:

$$Y_{dt} \sim \text{Poisson}(\mu_{dt})$$

$$\log(\mu_{dt}) = \log\left(\frac{\text{Pop}_{dt}}{12}\right) + \beta_0 + \beta_1 \text{DUVRI}_d + \beta_2 \text{PM}_{2.5,dt} + \gamma^T \text{Month}_t + \delta^T \text{Year}_t + b_d + u_d$$

with:

$$b_d \sim N(0, \sigma_b^2), u \sim \text{ICAR}(\sigma_u^2), \sum_{d=1}^{15} u_d = 0$$

where Y_{dt} denotes cardiovascular hospitalization episodes in district d and month t ; Pop_{dt} is the annual population for district d in year t ; DUVRI_d represents the district-level vegetation

vulnerability measure; $PM2.5_{dt}$ represents particulate exposure; $Month_t$ (12 categories) and $Year_t$ (four categories for the primary window) are categorical controls; bd is an unstructured district-level random intercept; and ud is a spatially structured random effect following an intrinsic conditional autoregressive (ICAR) prior. We specified the PC prior for the spatial standard deviation σ_u as $P(\sigma_u > 1) = 0.01$, following the recommendation of Simpson et al. (2017), and achieved identifiability between bd and ud through the sum-to-zero constraint $\sum ud = 0$. Results are reported in Table 6. The ICAR model is a sensitivity analysis that accounts for spatially structured residual variation; it does not constitute control for unmeasured spatial confounding. Because DUVRI-Eco, the district random intercept, and the ICAR spatial effect are all defined over the same 15 municipal districts, the exposure coefficient is not fully separable from the latent spatial structure. The ICAR estimate should be treated as a conditional spatial association under the specified prior, not as a fully adjusted exposure effect or an independent confirmation of the primary model. Throughout this manuscript, we therefore treat the DUVRI-Eco coefficient as a conditional spatial association under the specified prior/model, not as a fully separable exposure effect.

To assess sensitivity to the prior specification for the spatial variance component, we fitted the ICAR model under three PC prior scales: (1) the default, $P(\sigma_u > 1) = 0.01$; (2) a wider prior, $P(\sigma_u > 1.5) = 0.01$; and (3) a stronger prior, $P(\sigma_u > 0.5) = 0.01$. Results are reported in Table 6. To assess whether the choice of spatial neighborhood definition materially affected the ICAR results, we fitted the model using three alternative spatial weights matrices, k-nearest-neighbor ($k = 3$ and $k = 5$) and inverse-distance-based weights, while retaining the queen-contiguity matrix as the primary specification. After fitting the ICAR model, we assessed residual spatial autocorrelation using Moran's I with the queen-contiguity structure. The residual Moran's I was 0.06 (95% CI: -0.11 to 0.23; $P = 0.29$), indicating no substantial spatially structured residual variation; the spatial random effect was accounted for. The global bivariate Moran's I analysis identified a positive spatial association between DUVRI-Eco and hospitalization rates ($I = 0.31$, 95% CI: 0.12–0.50; permutation-based $P = 0.02$). Local bivariate Moran's I results are reported in the Supplementary Material as exploratory and descriptive; because we conducted 15 local tests without multiplicity adjustment, the local p-values should not be read as confirmatory evidence of district-specific clustering. Globally, Moran's I remains the primary spatial association.

3. Results

3.1. Hydroclimatic Trends and Vegetation Changes

3.1.1. Long-Term Hydroclimatic Trends

Analysis of the 72-year homogenized climate record (1951–2022) revealed persistent warming and declining moisture availability across the Isfahan study area (Figure 2). Minimum temperatures rose by 0.29°C per decade (95% CI: 0.21–0.37; $P = 0.003$, FDR-adjusted), and maximum temperatures rose by 0.21°C per decade (95% CI: 0.15–0.27; $P = 0.004$, FDR-adjusted). Over the same period, growing-season moisture availability, measured by the Composite Moisture Anomaly Index (CMAI), fell by 26% on a period-mean basis, compared to 1951–1990 (mean = 0.42; 95% CI: 0.38–0.46) to 1991–2022 (mean = 0.31; 95% CI: 0.27–0.35; $P < 0.001$). The endpoint change from 1951 to 2022 was larger still (–35.6%), reflecting the cumulative effect over the full study period.

3.1.2. Vegetation Responses to Climatic Stress

Consistent with this drying trajectory, NDVI declined an estimated 15–20% in the most severely affected districts between 1991 and 2022. This estimate rests on direct Landsat observations supplemented by reconstructed values for descriptive completeness; because part of the historical

series was reconstructed from a statistical model, it should be read as an inferred temporal trend rather than a direct observation of vegetation loss. The primary attribution analysis, restricted to the 105 directly observed district-year observations, confirmed a statistically significant negative trend ($\beta = -0.0032$ per year; 95% CI: -0.0041 to -0.0023 ; $P < 0.001$). The proportion of green space classified as severely stressed ($VCI < 35\%$) rose from 24% to 58% over the same interval. This VCI-based classification is reconstruction-dependent and should be read as a relative indicator of vegetation stress, not as a direct observation of vegetation mortality.

3.2. DUVRI-Eco Calibration and Validation

3.2.1. Outcome-Calibrated Component Weights

The primary district-level calibration model yielded non-negative component weights summing to unity. The constrained maximum-likelihood solution placed the largest weight on CDD (0.31), followed by TXx (0.28), NDVI (0.26), and DPC (0.15). Bootstrap-derived 95% confidence intervals, based on a district-clustered bootstrap with 1,000 replications (Table 3), were CDD (0.26–0.37), TXx (0.22–0.34), NDVI (0.20–0.32), and DPC (0.09–0.21). Profile-likelihood diagnostics, however, indicated limited identifiability of the individual weights, particularly for CDD (flat between 0.26 and 0.37) and TXx (flat between 0.22 and 0.34), a pattern that several weight configurations fit the calibration data about equally well (Figure 4). These estimates are therefore best read as an admissible calibration solution rather than a precise accounting of each component's relative importance, and the weight ranking should be treated as suggestive rather than definitive. The data constrain the composite score far more tightly than the individual weights, and the confidence intervals describe the range of plausible values under the specified model rather than evidence of precise parameter estimates. As reported in Methods, removing the a priori monotonicity constraint ($\beta \geq 0$) left the estimated mortality-response coefficient positive and changed the component weights by no more than ± 0.01 , indicating that the outcome direction was not an artifact of the constraint.

The median rank correlation across the plausible weight surfaces was 0.94 (IQR: 0.89–0.97), and the maximum rank change for any district was 3 positions. We considered a district classification-stable if its category relative to the candidate transition point (above vs. below $\tau \approx 0.65$) held constant across all 10,000 accepted weight vectors; by that definition, 13 of 15 districts (86.7%) were classification-stable, with the two unstable districts identified. The median absolute score change across all districts and weight vectors was 0.021 (IQR: 0.009–0.038), with a maximum absolute change of 0.087. In short, while individual weights are weakly identified, the resulting indicator score and its spatial classification show relatively limited variation.

The relatively low weight assigned to DPC (0.15), together with its strong correlation with NDVI (Spearman's $\rho = 0.76$, 95% CI: 0.68–0.82), suggests partial redundancy between the two indicators. These weights were calibrated for vegetation mortality and should not be assumed to reflect the relative importance of these variables for cardiovascular outcomes.

The site-level hierarchical sensitivity model, which modeled mortality at the site level with district random effects, produced similar weights: CDD = 0.32 (95% CI: 0.27–0.38), TXx = 0.27 (95% CI: 0.21–0.33), NDVI = 0.26 (95% CI: 0.19–0.32), and DPC = 0.15 (95% CI: 0.08–0.22), with an AUC of 0.78 (95% CI: 0.72–0.84).

3.2.2. Exploratory Nonlinear Mortality Response

The relationship between DUVRI-Eco and vegetation mortality followed a segmented nonlinear pattern, with an empirically estimated candidate transition point at $\hat{\tau} = 0.65$ (95% profile-likelihood CI: 0.62–0.68). The estimated slope below the transition was $\beta_1 = 1.8$ (95% CI: 1.0–2.6), and the estimated change in slope above the transition was $\beta_2 = 2.4$ (95% CI: 1.2–3.6); together these give

an above-transition slope of $\beta_1 + \beta_2 = 4.2$ and a slope ratio (above vs. below) of 2.3 (95% CI: 1.6–3.2). This candidate transition point marks a substantial increase in the mortality response's slope increase substantially on the logistic scale; it is not a biological invariant or a validated health cutoff. The candidate transition point held across all 15 district-blocked validation folds, but it should be read as a descriptive feature of the calibration sample, not as a biological invariant or validated management threshold, and should not be used for operational decision-making without external validation.

3.2.3. Predictive Performance via Nested Cross-Validation

Pooled out-of-fold predictions from the 34 outer validation folds yielded a nested site-level out-of-sample AUC of 0.81 (95% CI: 0.73–0.88). Because sites within the same district share identical district-level predictor values, the 34 site-level folds do not represent 34 independent validation cases.

When entire districts were withheld from model development to assess internal geographic generalization, the district-block spatial AUC was 0.79 (95% CI: 0.73–0.85; $\Delta\text{AUC} = -0.02$; 95% CI: -0.06 to 0.02). This district-blocked AUC is the primary estimate of the indicator's predictive performance in held-out districts within the same urban system; with only 15 municipal districts available, it should be read as internal geographic discrimination within the Isfahan study domain rather than evidence of external validity. The distribution of fold-specific AUCs was: median = 0.80, IQR = 0.74–0.85, range = 0.68–0.91, with mortality events per fold ranging from 3 to 12. A sensitivity analysis excluding the most influential district (District 7) yielded an AUC of 0.77 (95% CI: 0.70–0.84)—a modest change.

Discrimination (AUC) and rank stability are not the whole of validation for a screening indicator, and we report what the calibration sample can support without overstating it. The index-score stability analysis above (median rank correlation = 0.94; median absolute score change = 0.021) speaks to ranking stability. Formal calibration diagnostics for the mortality-response model itself—a calibration slope and intercept, a Brier score, and sensitivity/specificity or predictive values at the candidate transition point—would ordinarily accompany AUC, but with only 15 independent district-level exposure profiles and 3–12 mortality events per hold-out fold, such threshold-specific estimates would carry very wide, potentially uninformative uncertainty. We therefore report screening-relevant metrics (sensitivity, specificity, PPV, NPV, decision-curve analysis) as descriptive, out-of-fold values rather than in the main text, and we do not compute a Brier score or calibration slope for the mortality model itself; both choices reflect a considered judgment about what this sample size can support, not an oversight. Readers should treat the AUC values in this section as evidence of discrimination, not as a substitute for the calibration and decision-analytic evidence that a larger calibration sample would allow.

3.2.4. Weight Stability and Ablation Analysis

Nested model comparison showed incremental AUC gains: NDVI alone (AUC = 0.71), NDVI + TXx (AUC = 0.75), NDVI + CDD (AUC = 0.77), NDVI + TXx + CDD (AUC = 0.79), and NDVI + TXx + CDD + DPC (AUC = 0.81). Ablation analysis showed that removing CDD produced the largest numerical reduction in AUC ($\Delta\text{AUC} = -0.05$), followed by TXx ($\Delta\text{AUC} = -0.04$), NDVI ($\Delta\text{AUC} = -0.03$), and DPC ($\Delta\text{AUC} = -0.02$; 95% CI: -0.06 to 0.02). The ablation AUC ranking broadly matches the calibrated weight ranking, as expected given that both are from the same calibration data; this consistency should not be taken as independent confirmation (Table 7). With only 15 independent exposure profiles, the numerical differences between component contributions should not be read as a hierarchy of environmental importance.

The results suggest limited incremental predictive information from DPC in this calibration

domain ($\Delta\text{AUC} = -0.02$; 95% CI: -0.06 to 0.02); because the confidence interval includes zero, we cannot establish a statistically distinguishable incremental contribution (Figure3). DPC should not be interpreted as an independent predictor of vegetation mortality; it is a structurally transformed derivative of NDVI that overlaps substantially with vegetation condition in this study domain. Its inclusion reflects the indicator's design as a multidimensional construct rather than an empirical finding of independent predictive value. The primary geographic validation estimate remains the full-indicator leave-one-district-out AUC.

3.3. Cardiovascular Hospitalization Analysis

3.3.1. Primary Ecological Association

Following calibration and validation, we examined whether district-level DUVRI-Eco vulnerability was associated with cardiovascular hospitalization rates across Isfahan's 15 municipal districts. We conducted this analysis entirely independent of the vegetation-mortality-based indicator calibration. Consistent with the pre-specified inferential hierarchy, the primary estimate comes from the district-level ecological regression. After accounting for month and year fixed effects and a district random intercept, a 0.10unit increase in district-level DUVRI-Eco was associated with an IRR of 1.052 (95% CI: 1.018–1.087), a conditional spatial association, not a fully separable exposure effect. Higher district-level DUVRI-Eco was associated with higher recorded cardiovascular hospitalization rates across the 15 districts. The 3,022 geocoded hospitalization episodes sharpened the precision of these rate estimates but, as discussed in Section 2.4.6, did not increase the effective number of independent exposure units, which remains bounded by the 15 districts. The health analysis should be read as a descriptive ecological association, not as a confirmatory epidemiological finding (Figure7).

Extending the analysis to 2019–2024 using reconstructed NDVI values for 2023–2024 produced a nearly identical estimate (IRR = 1.048; 95% CI: 1.016–1.082) under this extended-period sensitivity analysis, though this does not validate the accuracy of the temporal extrapolation (Table 8).

The ICAR spatial model yielded a posterior mean IRR of 1.048 (95% CrI: 1.016–1.082), conditional on the ICAR spatial prior and the calibrated DUVRI-Eco weights. Because DUVRI-Eco and the ICAR spatial effect are defined over the same 15 districts, the coefficient is not fully separable from the spatial latent structure, and the reported credible interval should be read with caution given that exposure varied across only 15 districts.

The weight-sensitivity analysis yielded a median IRR of 1.049, with a 2.5th–97.5th percentile sensitivity range of 1.012–1.089; the primary model-based confidence interval (IRR = 1.052; 95% CI: 1.018–1.087) remained directionally consistent with the primary estimate across the likelihood-supported weight configurations.

3.3.2. Dose-Response Relationship

Restricted cubic spline analysis revealed a monotonic positive association across the full observed DUVRI-Eco range (0.21–0.78), with no statistically detectable departure from the prespecified linear association (P for nonlinearity = 0.54). This should be read as insufficient evidence to reject linearity, not as evidence of linearity. The interquartile contrast ($\Delta = 0.27$) yielded an IRR of 1.15 (95% CI: 1.06–1.24), and the 10th–90th percentile contrast ($\Delta = 0.47$) yielded an IRR of 1.27 (95% CI: 1.12–1.44). The nonlinear transition identified in the vegetation-mortality calibration does not imply a nonlinear cardiovascular exposure-response relationship (Figure5). The AR (1) temporal correlation sensitivity model yielded an IRR of 1.046 (95% CI: 1.014–1.080), consistent with the primary estimate.

3.3.3. Sensitivity Analyses

Across the pre-specified sensitivity analyses, the DUVRI-Eco–cardiovascular hospitalization association remained directionally consistent. The detailed results for all sensitivity analyses. Key results: alternative formulations produced IRRs ranging from 1.044 to 1.052, and estimates adjusted for PM2.5, a measured potential confounder, yielded an IRR of 1.045 (95% CrI: 1.010–1.082). This PM2.5-adjusted estimate should be read as a conditional association that survives adjustment for one measured potential confounder, not as a confounding-controlled effect (Table 4).

3.4. Spatial Heterogeneity

3.4.1. Geographic Distribution of Vulnerability

The spatial distribution of DUVRI-Eco across the 15 municipal districts revealed a descriptive east–west pattern (Figure1). Districts in the central-eastern part of the city (Districts 7, 10, 11, 12, and 15) exhibited the highest vulnerability values (DUVRI-Eco $\geq$ 0.65), while western districts (Districts 1, 2, 3, and 4) showed lower vulnerability (DUVRI-Eco $<$ 0.40).

3.4.2. Spatial Clustering

Global Moran’s I for DUVRI-Eco alone was 0.35 (95% CI: 0.18–0.52; $P = 0.01$), and the global bivariate Moran’s I between DUVRI-Eco and cardiovascular hospitalization rates was 0.31 (95% CI: 0.12–0.50; permutation-based $P = 0.02$). Given the limited number of districts ($n = 15$), we read this as descriptive secondary evidence rather than evidence of a spatial process (Figure6). Local bivariate Moran’s I results are reported in the Supplementary Material as exploratory and descriptive; because we conducted 15 local tests without multiplicity adjustment, the local p-values should not be read as confirmatory evidence of district-specific clustering. Globally, Moran’s I remains the primary spatial association result.

4. Discussion

4.1. Principal Findings and Scope of Inference

In this study, we developed DUVRI-Eco, a construct-first composite indicator integrating thermal stress (TXx), drought persistence (CDD), vegetation condition (NDVI), and landscape patch dominance (DPC). Field calibration against observed vegetation mortality at 34 sites, using a district-blocked cross-validation framework, yielded an internal evaluation of discrimination within the Isfahan study domain. The calibration sample is modest: because the indicator components are defined at the district level, the 34 field-site outcomes are nested within only 15 unique district-level exposure profiles, so the effective number of independent exposure units for component-weight calibration was bounded by Isfahan’s 15 municipal districts.

The indicator achieved an AUC of 0.81 in site-level validation (95% CI: 0.73–0.88) and 0.79 in leave-one-district-out internal geographic validation (95% CI: 0.73–0.85). The spatially-blocking validation is a more stringent test of internal geographic generalization, since it evaluates the indicator on entirely withheld districts. The leave-one-district-out validation provides evidence of internal geographic consistency within the Isfahan study domain; it should not be equated with external validation throughout different cities or climatic regions. External validation with local recalibration is required before the indicator, or its thresholds, can be applied operationally elsewhere.

The cardiovascular analysis identified a positive association at the district level. In the primary window (2019–2022), a 0.10-unit increase in DUVRI-Eco was associated with a 5.2% higher recorded cardiovascular hospitalization rate after accounting for temporal structure and a district random intercept (IRR = 1.052; 95% CI: 1.018–1.087)—an estimate that held up under a battery of pre-specified sensitivity checks: adjustment for PM2.5 as a measured potential confounder (IRR = 1.045; 95% CrI: 1.010–1.082), extension of the analysis window to 2019–2024 (IRR = 1.048;

95% CI: 1.016–1.082), and the weight-sensitivity analysis (median IRR = 1.049; 2.5th–97.5th percentile range: 1.012–1.089). We also fitted an ICAR spatial model as a supplementary robustness check (posterior mean IRR = 1.048; 95% CrI: 1.016–1.082); because DUVRI-Eco and the ICAR spatial effect are defined over the same 15 districts, however, we do not treat this as an independent confirmation and report it only as a secondary diagnostic. As detailed in Section 2.4.6, the time-invariant DUVRI exposure is identified mainly from between-district contrasts, so the effective number of independent spatial exposure units remains 15 regardless of the 720 district-month observations in the primary analysis.

The district-block validation framework is the methodological centerpiece of this study. Recognizing that 34 field sites nested within 15 municipal districts do not constitute 34 independent exposure profiles gives a more realistic picture of internal geographic generalization, a design choice that conforms with recommendations that validation respect the spatial dependence and geographic structure of the data (Roberts et al., 2017; Wenger & Olden, 2012).

Four boundaries of inference merit emphasis.

First, DUVRI-Eco is a study-domain screening framework. The calibrated weights are only weakly identified within the present 15-district calibration domain and are best as one admissible calibration solution under the specified constraint, not as a precise estimate of relative environmental importance; they should not be treated as universal constants. The data constrain the composite score more tightly than they constrain the individual weights.

Second, the index-score stability analysis shows that individual weights are weakly identified; the resulting indicator score and its spatial classification vary relatively little across the plausible weight surface. This distinction matters for applications: the indicator can support spatial screening even when the precise weights remain uncertain.

Third, the health analysis is a district-level ecological panel analysis with a time-invariant exposure, based on Isfahan's 15 municipal districts as the primary spatial exposure units. It is hypothesis-generating only and cannot support individual-level or causal inference.

Fourth, external validation with local recalibration is required before broader application.

4.2. Positioning DUVRI-Eco Relative to Existing Indices

DUVRI-Eco differs from established indices in three main respects. First, it integrates multi-dimensional stressors (heat, drought, vegetation state, and landscape structure) while relying on a single dimension. Second, the weights are outcome-calibrated against field-observed tree mortality rather than assigned equally or derived from covariance structure. Third, it is validated using spatially blocked procedures that explicitly test internal geographic consistency within the study domain. We empirically benchmarked DUVRI-Eco against VCI and VHI in the Isfahan study domain (Table 9), using identical training/validation folds, normalization procedures, and spatial blocking for all three indicators, with Δ AUC calculated via district-level bootstrap resampling (1,000 replications). DUVRI-Eco showed numerically higher site-level AUC than VCI (Δ AUC = 0.06; 95% CI: -0.02 to 0.14) and VHI (Δ AUC = 0.04; 95% CI: -0.03 to 0.11). Because these confidence intervals include zero, the differences should be read as numerically higher discrimination in this dataset, not as evidence of superior performance.

4.3. Ecological Interpretation and Geographic Consistency

CDD and TXx received the largest point estimates within the composite indicator under the specified calibration model (CDD = 0.31, TXx = 0.28). Profile-likelihood diagnostics, however, indicated limited identifiability of the individual weights, so the weight ranking should be treated as suggestive rather than definitive, the data do not support a precise ranking of component importance. The low weight for DPC (0.15) does not imply that the landscape configuration is

ecologically unimportant. In Isfahan, DPC correlated strongly with NDVI (Spearman's $\rho = 0.76$, 95% CI: 0.68–0.82), and ablation analysis confirmed that removing DPC reduced site-level AUC by only 0.02 (95% CI: –0.06 to 0.02). We retained DPC in the final indicator as a construct-defining component, reflecting its role as a structural descriptor of patch dominance rather than a primary predictive driver. Cities with different spatial planning regimes may find that alternative landscape metrics better capture functional connectivity and management heterogeneity, which points to the need for local recalibration when applying DUVRI-Eco elsewhere. This context-dependency is itself a major limitation: independent external ecological validation is required before the indicator is transferred to other settings. The candidate transition point $\tau = 0.65$ held across the internal validation folds (0.62–0.68), reflecting the homogeneity of the Isfahan calibration domain. Strong caution is warranted before applying this candidate transition point or the exact weighting vector to cities with different species pools, soil substrates, or irrigation regimes. This value is a descriptive feature of the calibration sample, not a biological invariant or a validated management threshold, and should not be used for operational decision-making without external validation.

4.4. NDVI Reconstruction and Circularity Avoidance

We took explicit, methodologically rigorous steps to reduce the risk of circularity between the NDVI reconstruction and subsequent inference. The reconstruction model did not include the DUVRI-Eco components as predictors, and all primary attribution analyses were restricted to the 105 directly observed district-year observations. We used reconstructed NDVI only for descriptive completeness of the long-term trajectory, health-exposure construction for 2021, 2023, and 2024, and sensitivity analyses.

This design reduces circularity in the primary climate-vegetation attribution analysis, which is restricted to directly observed Landsat years. In the health-exposure construction, however, we used reconstructed NDVI values for 2021, 2023, and 2024, which introduces a degree of coupling between health exposure and the hydroclimatic predictors. We do not claim that the health exposure is fully independent of the hydroclimatic predictors used in its reconstruction; this coupling should be regarded as a limitation of the health analysis.

The 2023–2024 reconstruction constitutes a temporal extrapolation beyond the observed range. The leave-one-observed-year-out validation assesses predictive performance within the observed temporal range; it does not validate the extrapolation to 2023–2024. The stability of the health association when the primary analysis is restricted to 2019–2022 indicates that the association holds up across the specified sensitivity analyses, though it does not validate the accuracy of the extrapolation itself.

4.5. DPC as a Measure of Patch Dominance

In Isfahan, DPC is best interpreted as a proxy for patch dominance under a centralized irrigation regime, not as a direct measure of ecological resilience or functional connectivity. The ablation analysis confirmed that removing DPC did not materially change site-level predictive performance ($\Delta\text{AUC} = -0.02$; 95% CI: –0.06 to 0.02; a separate district-blocked ablation was not conducted). We retained DPC for construct completeness as a structural descriptor of patch dominance, while acknowledging that its contribution is context-specific and may differ in cities with decentralized green infrastructure. DPC is operationally defined as the Largest Patch Index (LPI) calculated using FRAGSTATS v4.2; we use the term DPC to emphasize its interpretation as a measure of dominant-patch structure, though mathematically it is identical to LPI. Because the DPC is derived from the same NDVI-based vegetation mask used to compute NDVI itself, it is not an independent environmental dimension but a structurally transformed derivative of NDVI, an equivalence that

readers familiar with FRAGSTATS will recognize.

4.6. Cardiovascular Health: Interpretation and Boundaries

The positive DUVRI-Eco–cardiovascular association is biologically plausible through several pathways: increased thermal stress, reduced vegetation-mediated cooling, altered outdoor physical activity, correlated air pollution, and socioeconomic disadvantage concentrated in high-exposure districts. The present design cannot attribute the association to any single pathway, nor can it establish causality. These four limitations, taken together, fundamentally constrain how the health analysis can be interpreted.

- As detailed in Section 2.4.6, DUVRI-Eco was defined at the level of Isfahan’s 15 municipal districts and remained time-invariant during the health observation period, so its coefficient is identified from between-district contrasts across these 15 districts; the 720 district-month observations and 3,022 geocoded episodes sharpen temporal and rate estimation but do not add independent exposure contrasts. The limitation is not the number of districts per se, but that the exposure contrast itself is identified at the district level. The reported confidence interval accordingly describes the stability of the association under the specified model rather than a precise effect estimate based on independent observations, which is why we treat the health findings as hypothesis-generating rather than confirmatory.
- Analytical unit is the hospitalization episode, not the patient, since patient identifiers were unavailable for linkage. If the same individual contributed to multiple admissions, those admissions are not statistically independent, and within-patient correlation could inflate apparent precision; if repeated admissions cluster differentially across districts, that could also bias the point estimate in either direction. We could not model this dependence, so both possibilities remain limitations of the health analysis.
- District-level aggregation precludes individual-level inference. We cannot conclude that individuals residing in high-DUVRI-Eco districts face greater personal cardiovascular risk, nor can we rule out that unmeasured district-level confounders, income, access to medical care, baseline disease prevalence, or health-seeking behavior, drive the observed pattern.
- PM2.5 should not be treated as the sole or definitive confounder in this ecological analysis; important district-level determinants of cardiovascular hospitalization were not available for adjustment, and PM2.5’s own causal role remains uncertain; it may be a correlated environmental exposure or a mediator rather than a confounder. The PM2.5-adjusted estimate should therefore be read as a conditional association that survives adjustment for one measured potential confounder. Because the exposure is time-invariant, the analysis cannot disentangle the DUVRI-Eco association from unmeasured district characteristics correlated with both vegetation vulnerability and hospitalization burden.
- The Pearson χ^2/df was 1.08, and the Deviance/df was 1.06, supporting the Poisson model; the negative binomial sensitivity analysis yielded nearly identical results (IRR = 1.054 vs. 1.052). The principal limitations of the health analysis are not overdispersion but exposure dimensionality (15 municipal districts), ecological confounding, and spatial identifiability.
- The weight-sensitivity analysis yielded a median IRR of 1.049 (2.5th–97.5th percentile sensitivity range: 1.012–1.089), compared with the primary estimate of 1.052 (95% CI: 1.018–1.087). This sensitivity range overlapped the primary model-based confidence interval and remained directionally consistent with the primary estimate across the likelihood-supported weight configurations. This analysis provides a robustness check to calibrate weight uncertainty; it does not eliminate the fundamental limitation that exposure is identified from only 15 district-level contrasts. A fully joint uncertainty framework

would offer a more complete characterization of uncertainty and represents a priority for future multi-city validation studies.

Because DUVRI-Eco exhibited spatial autocorrelation (Moran's $I = 0.35$), unrestricted permutation of district exposure labels does not preserve the observed spatial structure, so we do not treat this analysis as a formal spatial hypothesis test; it is reported only as a descriptive stability assessment. A spatially constrained permutation scheme would offer a more appropriate null distribution and is worth pursuing in future studies with larger spatial systems.

DUVRI-Eco and the ICAR spatial effects are defined over the same 15 municipal districts, so their coefficients are not fully separable. The DUVRI-Eco coefficient should be regarded as a conditional spatial association under the specified prior.

The outcome analyzed here is recorded cardiovascular hospitalizations within the participating university-affiliated hospital network, not city-wide cardiovascular hospitalizations; private hospital admissions unaffiliated with the university were not included. If districts differ in access to university hospitals, the observed association may partly reflect access to medical care rather than cardiovascular disease incidence.

Given these limitations, we frame our health findings as hypothesis-generating. The exploratory local analysis identified nominal high-high patterns in Districts 7, 10, 11, 12, and 15, a pattern that warrants testing in individual-level cohorts with residential histories, behavioral covariates, and high-resolution exposure models (Nieuwenhuijsen et al., 2020; Rajagopalan et al., 2018).

4.7. Potential Applications and Research-Oriented Implications

Despite its limitations, DUVRI-Eco offers information that may be useful for Isfahan's municipal planners and land managers. Municipal authorities could consider three steps: (i) using DUVRI-Eco as a screening tool to identify candidate locations for prospective ecological assessment; (ii) monitoring vegetation mortality and health outcomes in high-vulnerability districts over time; and (iii) collaborating with researchers to design prospective validation studies. The indicator is not a substitute for on-the-ground monitoring; it is a complementary screening tool that can support future evaluations of resource allocation and early-warning applications, and its use for health intervention planning requires independent validation. The value $\tau \approx 0.65$ is a candidate transition point identified in this calibration sample, not a management target, a regulatory threshold, or a validated operational cutoff; any use in management prioritization requires prospective validation.

4.8. Future Research Directions

Seven priorities would strengthen the framework. First, multi-site validation with local recalibration in contrasting semi-arid cities, drawing on a larger number of independent spatial exposure units, would test whether DUVRI-Eco generalizes beyond Isfahan. Second, individual-level longitudinal studies with residential histories and behavioral covariates would allow for stronger causal inference. Third, higher-resolution environmental data would enable analysis at finer spatial scales. Fourth, formal causal mediation analysis using longitudinal exposure structures could test whether vegetation decline operates through heat exposure, air pollution, or psychosocial pathways. Fifth, a fully joint uncertainty framework would provide a more complete characterization of uncertainty. Sixth, given DPC's strong correlation with NDVI, a three-component variant (TXx, CDD, and NDVI alone, without DPC) is worth calibrating and validating explicitly against the full four-component indicator, comparing district rankings, AUC, the candidate transition point, and the downstream health IRR, to establish directly whether DPC's retention changes any operational conclusion. Seventh, because the additive form was chosen partly for parsimony given only 15 independent exposure profiles, a constrained interaction model (e.g., a single TXx $\times$ CDD term) fitted as an explicit sensitivity analysis, once a larger multi-city

sample is available, would let the data speak more directly to whether heat and drought act additively or interactively.

5. Conclusion

This study developed DUVRI-Eco, a field-calibrated multidimensional indicator for assessing urban vegetation vulnerability under multidimensional hydroclimatic stress in a semi-arid city. The indicator is construct-first by design: predictive calibration anchors the weighting structure without attempting to optimize prediction. It achieved an AUC of 0.81 in site-level validation and 0.79 in leave-one-district-out internal geographic validation, supporting internal geographic discrimination within Isfahan without establishing external validity across cities or climatic regions.

The constrained maximum-likelihood solution placed the largest weight on CDD (0.31), followed by TXx (0.28), NDVI (0.26), and DPC (0.15). Profile-likelihood diagnostics, however, revealed weak identifiability of the individual weights, and we accordingly treat them as domain-specific regularization parameters that anchor the construct to the Isfahan calibration sample rather than as estimable population-level component effects. The data constrain the composite score more tightly than the individual weights: even with weakly identified weights, the resulting indicator score and its spatial classification showed relatively limited variation (median rank correlation = 0.94; median absolute score change = 0.021), and the candidate transition-point classification held for 13 of 15 districts.

The secondary cardiovascular analysis identified a positive association between DUVRI-Eco and recorded hospitalization rates, identified mainly between-district contrasts across Isfahan's 15 municipal districts—an effective sample size of 15 that the 720 underlying district-month observations do not (Section 2.4.6). The weight-sensitivity analysis remained directionally consistent with the primary estimate across the likelihood-supported weight configurations. These findings are hypothesis-generating and cannot be interpreted as individual-level cardiovascular risk or causal effects.

The empirically identified vegetation-mortality candidate transition point near 0.65 is a candidate point for hypothesis generation, not for management prioritization. Given how weakly the component weights are identified within a single 15-district calibration domain, external validation with local recalibration is not an optional refinement but a precondition for applying DUVRI-Eco, or its thresholds, outside Isfahan. Within these limits, DUVRI-Eco offers a practical framework for integrating hydroclimatic stress, vegetation condition, and landscape structure into locally calibrated urban environmental vulnerability assessment.

Ethics Approval

The study protocol was approved by the Ethics Committee of PayameNoor University and the Ethics Committee of Isfahan University of Medical Sciences. The ethics committee waived the requirement for individual patient consent because the study used secondary administrative data. All hospitalization data were de-identified before analysis.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors acknowledge administrative and logistical support from Isfahan University of Medical Sciences, PayameNoor University, the Isfahan Meteorological Office, and the Municipality of Isfahan. The supporters had no role in study design, data collection, analysis, interpretation, manuscript preparation, or the decision to submit this work for publication.

Data Availability

The complete R scripts, derived district-level datasets, variable dictionaries, model specification

files, random seeds, and cross-validation fold assignments generated during this study are stored in the Mendeley Data repository (DOI: 10.17632/fg29rhyx3f.2). Raw hospitalization records are subject to privacy and confidentiality restrictions under Iranian health data protection regulations and are therefore not publicly available; access to raw hospitalization data requires institutional approval from Isfahan University of Medical Sciences. Aggregated data and analytical code are freely accessible through the Mendeley Data repository. All preprocessing decisions, exclusion criteria, calibration parameters, and sensitivity-analysis definitions are documented in the supplementary materials. Station-level meteorological metadata, field-sampling information, and quality-controlled hydroclimatic data are also available through the Mendeley Data repository.

Competing Interests

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

Author Contributions

Pouriya Javari: Investigation, data curation, writing – original draft, review & editing. Majid Javari: Conceptualization, methodology, formal analysis, supervision, project administration, writing – review & editing. Hossein Ladvar: Methodology, formal analysis, review & editing. All authors contributed critically to the drafts and gave final approval for publication.

Acknowledgments

The authors thank the Isfahan Department of Environment for providing air-quality monitoring data, the Meteorological Office for providing historical climate records, the Health Information Technology Unit of Isfahan University of Medical Sciences for facilitating access to anonymized hospitalization records, and the municipal authorities for providing green-space management data.

References

- Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V., & van der Sluis, S. (2014). A solution to dependency: Using multilevel analysis to accommodate nested data. *Nature Neuroscience*, 17(4), 491–496. <https://doi.org/10.1038/nn.3648>
- Aguilar, E., Auer, I., Brunet, M., Peterson, T. C., & Wieringa, J. (2003). *Guidelines on climate metadata and homogenization* (WMO-TD No. 1186). World Meteorological Organization.
- Alberti, M. (2005). The effects of urban patterns on ecosystem function. *International Regional Science Review*, 28(2), 168–192. <https://doi.org/10.1177/0160017605275160>
- Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., ... Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, 111(D5), D05109. <https://doi.org/10.1029/2005JD006290>
- Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology*, 6(6), 661–675. <https://doi.org/10.1002/joc.3370060607>
- Alexandersson, H., & Moberg, A. (1997). Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends. *International Journal of Climatology*, 17(1), 25–34. [https://doi.org/10.1002/\(SICI\)1097-0088\(199701\)17:1<25::AID-JOC103>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1097-0088(199701)17:1<25::AID-JOC103>3.0.CO;2-J)
- Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J. H., Allard, G., Running, S. W., Semerci,

- A., & Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, 259(4), 660–684. <https://doi.org/10.1016/j.foreco.2009.09.001>
- Allison, P. D. (2009). *Fixed effects regression models*. SAGE Publications. <https://doi.org/10.4135/9781412993869>
- Anselin, L. (1988). *Spatial econometrics: Methods and models*. Springer. <https://doi.org/10.1007/978-94-015-7799-1>
- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Aram, F., Higuera García, E., Solgi, E., & Mansouri Daneshvar, M. R. (2019). Urban green space cooling effect in cities. *Heliyon*, 5(4), e01339. <https://doi.org/10.1016/j.heliyon.2019.e01339>
- Armstrong, M. P., Rushton, G., & Zimmerman, D. L. (1999). Geographically masking health data to preserve confidentiality. *Statistics in Medicine*, 18(5), 497–525. [https://doi.org/10.1002/\(SICI\)1097-0258\(19990315\)18:5<497::AID-SIM45>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0258(19990315)18:5<497::AID-SIM45>3.0.CO;2-6)
- Arnold, C. L., & Gibbons, C. J. (1996). Impervious surface coverage: The emergence of a key environmental indicator. *Journal of the American Planning Association*, 62(2), 243–258. <https://doi.org/10.1080/01944369608975688>
- Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1), 47–78. <https://doi.org/10.2307/2998540>
- Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1–22. <https://doi.org/10.1002/jae.659>
- Baltagi, B. H. (2005). *Econometric analysis of panel data* (3rd ed.). John Wiley & Sons.
- Bastin, J. F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C. M., & Crowther, T. W. (2019). The global tree restoration potential. *Science*, 365(6448), 76–79. <https://doi.org/10.1126/science.aax0848>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Becker, W., Saisana, M., Paruolo, P., & Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological Indicators*, 80, 12–22. <https://doi.org/10.1016/j.ecolind.2017.03.056>
- Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: Parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10), 3001–3023. <https://doi.org/10.1002/joc.3887>
- Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity*, 53(2), 1051–1074. <https://doi.org/10.1007/s11135-018-0802-x>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1), 1–20. <https://doi.org/10.1007/BF00116466>

- Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-4614-7618-4>
- Blangiardo, M., & Cameletti, M. (2015). *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley & Sons. <https://doi.org/10.1002/9781118950203>
- Böhringer, C., & Jochem, P. E. P. (2007). Measuring the immeasurable: A survey of sustainability indices. *Ecological Economics*, 63(1), 1–8. <https://doi.org/10.1016/j.ecolecon.2007.03.008>
- Bow, C. J., Waters, N. M., Faris, P. D., Seidel, J. E., Galbraith, P. D., Knudtson, M. L., & Ghali, W. A. (2004). Accuracy of city postal code coordinates as a proxy for location of residence. *International Journal of Health Geographics*, 3(1), 5. <https://doi.org/10.1186/1476-072X-3-5>
- Bowler, D. E., Buyung-Ali, L., Knight, T. M., & Pullin, A. S. (2010). Urban greening to cool towns and cities: A systematic review of the empirical evidence. *Landscape and Urban Planning*, 97(3), 147–155. <https://doi.org/10.1016/j.landurbplan.2010.05.006>
- Briber, B. M., Hutya, L. R., Reinmann, A. B., Raciti, S. M., Dearborn, V. K., Holden, C. E., & Dunn, A. L. (2013). Tree productivity enhanced with conversion from forest to urban land covers. *PLoS ONE*, 8(7), e69416. <https://doi.org/10.1371/journal.pone.0069416>
- Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1), 1–3. [https://doi.org/10.1175/1520-0493\(1950\)078<0001:voefit>2.0.co;2](https://doi.org/10.1175/1520-0493(1950)078<0001:voefit>2.0.co;2)
- Brook, R. D., Rajagopalan, S., Pope, C. A., III, Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Jr., Whitsel, L., & Kaufman, J. D. (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*, 121(21), 2331–2378. <https://doi.org/10.1161/CIR.0b013e3181dbcee1>
- Brown, M. E., Lewis, H. G., & Wiseman, J. (2008). Using satellite remote sensing to monitor global vegetation. *Remote Sensing of Environment*, 112(6), 2855–2870. <https://doi.org/10.1016/j.rse.2008.02.009>
- Bryan, M. L., & Jenkins, S. P. (2016). Multilevel modelling of country effects: A cautionary tale. *European Sociological Review*, 32(1), 3–22. <https://doi.org/10.1093/esr/jcv059>
- Buckland, S. T., Burnham, K. P., & Augustin, N. H. (1997). Model selection: An integral part of inference. *Biometrics*, 53(2), 603–618. <https://doi.org/10.2307/2533961>
- Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190–1208. <https://doi.org/10.1137/0916069>
- Caesar, J., Alexander, L., & Vose, R. (2006). Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *Journal of Geophysical Research: Atmospheres*, 111(D5), D05101. <https://doi.org/10.1029/2005JD006280>
- Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3), 414–427. <https://doi.org/10.1162/rest.90.3.414>
- Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372. <https://doi.org/10.3368/jhr.50.2.317>
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (2nd ed.).

- Cambridge University Press. <https://doi.org/10.1017/CBO9781139013567>
- Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: When, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, 19(9), 1141–1164. [https://doi.org/10.1002/\(SICI\)1097-0258\(20000515\)19:9<1141::AID-SIM479>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F)
- Carpenter, S. R. (2001). *Ecological forecasting and the problem of scaling*. Princeton University Press.
- Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5), 893–903. <https://doi.org/10.1016/j.rse.2009.01.007>
- Chander, G., Mishra, N., Helder, D. L., Aaron, D. B., Angal, A., Choi, T., Xiong, X., & Doelling, D. R. (2013). Applications of spectral band adjustment factors (SBAF) for cross-calibration. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3), 1267–1281. <https://doi.org/10.1109/TGRS.2012.2226581>
- Chen, J., Jin, S., & Du, P. (2018). Urban green space cooling effects in Beijing. *Landscape and Urban Planning*, 179, 45–56. <https://doi.org/10.1016/j.landurbplan.2018.07.006>
- Choat, B., Jansen, S., Brodribb, T. J., Cochard, H., Delzon, S., Bhaskar, R., Bucci, S. J., Feild, T. S., Gleason, S. M., Hacke, U. G., Jacobsen, A. L., Lens, F., Maherali, H., Martínez-Vilalta, J., Mayr, S., Mencuccini, M., Mitchell, P. J., Nardini, A., Pittermann, J., ... Zanne, A. E. (2012). Global convergence in the vulnerability of forests to drought. *Nature*, 491(7426), 752–755. <https://doi.org/10.1038/nature11688>
- Clark, T. S., & Linzer, D. A. (2015). Should I use fixed or random effects? *Political Science Research and Methods*, 3(2), 399–408. <https://doi.org/10.1017/psrm.2014.32>
- Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models and applications*. Pion.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Collett, D. (2003). *Modelling binary data* (2nd ed.). Chapman & Hall/CRC. <https://doi.org/10.1201/9780203163199>
- Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data* (2nd ed.). Chapman & Hall. <https://doi.org/10.1007/978-1-4899-3356-6>
- Croissant, Y., & Millo, G. (2008). Panel data econometrics in R: The plm package. *Journal of Statistical Software*, 27(2), 1–43. <https://doi.org/10.18637/jss.v027.i02>
- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, 28(15), 2031–2064. <https://doi.org/10.1002/joc.1688>
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511802843>
- de Jesus, J., Mello, K., & de Oliveira, R. (2016). Vegetation mask thresholds for urban landscape analysis. *Remote Sensing Applications: Society and Environment*, 4, 1–10. <https://doi.org/10.1016/j.rsase.2016.05.001>
- Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K., & Arnold, J. G. (2008). Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. *Journal of Applied Meteorology and Climatology*, 47(2), 475–497. <https://doi.org/10.1175/2007JAMC1356.1>
- Didan, K. (2015). *MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN*

- Grid V006* [Data set]. NASA EOSDIS Land Processes DAAC. <https://doi.org/10.5067/MODIS/MOD13Q1.006>
- Diggle, P. J., Heagerty, P., Liang, K. Y., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780198524847.001.0001>
- Dirks, K. N., Hay, J. E., Stow, C. D., & Harris, D. (1998). High-resolution rainfall estimation using interpolation methods. *Journal of Hydrology*, 210(1–4), 94–109. [https://doi.org/10.1016/S0022-1694\(98\)00175-9](https://doi.org/10.1016/S0022-1694(98)00175-9)
- Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., & Samet, J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*, 295(10), 1127–1134. <https://doi.org/10.1001/jama.295.10.1127>
- Easterling, D. R., Meehl, G. A., Parmesan, C., Changnon, S. A., Karl, T. R., & Mearns, L. O. (2000). Climate extremes: Observations, modeling, and impacts. *Science*, 289(5487), 2068–2074. <https://doi.org/10.1126/science.289.5487.2068>
- Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on cross-validation. *Journal of the American Statistical Association*, 78(382), 316–331. <https://doi.org/10.1080/01621459.1983.10477973>
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Chapman & Hall/CRC. <https://doi.org/10.1007/978-1-4899-4541-5>
- Elvidge, C. D., Tuttle, B. T., Sutton, P. C., Baugh, K. E., Howard, A. T., Milesi, C., Bhaduri, B. L., & Nemani, R. (2007). Global distribution and density of constructed impervious surfaces. *Sensors*, 7(9), 1962–1979. <https://doi.org/10.3390/s7091962>
- Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 34, 487–515. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132419>
- Fischer, J., & Lindenmayer, D. B. (2007). Landscape modification and habitat fragmentation: A synthesis. *Global Ecology and Biogeography*, 16(3), 265–280. <https://doi.org/10.1111/j.1466-8238.2007.00287.x>
- Freudenberg, M. (2003). *Composite indicators of country performance: A critical assessment* (OECD Science, Technology and Industry Working Papers, 2003/16). OECD Publishing. <https://doi.org/10.1787/405566708255>
- Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Klein Tank, A. M. G., & Peterson, T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, 19(3), 193–212. <https://doi.org/10.3354/cr019193>
- Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1), 168–182. <https://doi.org/10.1016/j.rse.2009.08.016>
- Genizi, A. (1993). Decomposition of R^2 in multiple regression with correlated regressors. *Statistica Sinica*, 3(2), 407–420.
- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378. <https://doi.org/10.1198/016214506000001437>
- Gotway, C. A., Ferguson, R. B., Hergert, G. W., & Peterson, T. A. (1996). Comparison of

- kriging and inverse-distance methods for mapping soil parameters. *Soil Science Society of America Journal*, 60(4), 1237–1247. <https://doi.org/10.2136/sssaj1996.03615995006000040040x>
- Goward, S. N., Markham, B., Dye, D. G., Dulaney, W., & Yang, J. (1991). Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer. *Remote Sensing of Environment*, 35(2–3), 257–277. [https://doi.org/10.1016/0034-4257\(91\)90017-Z](https://doi.org/10.1016/0034-4257(91)90017-Z)
- Grantz, D. A., Garner, J. H. B., & Johnson, D. W. (2003). Ecological effects of particulate matter. *Environment International*, 29(2–3), 213–239. [https://doi.org/10.1016/S0160-4120\(02\)00181-2](https://doi.org/10.1016/S0160-4120(02)00181-2)
- Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, 141(1), 61–94. <https://doi.org/10.1007/s11205-017-1832-9>
- Green, D. P., Ha, S. E., & Bullock, J. G. (2010). Enough already about "black box" experiments: Studying mediation is more difficult than most scholars suppose. *The ANNALS of the American Academy of Political and Social Science*, 628(1), 200–208. <https://doi.org/10.1177/0002716209351526>
- Greene, W. H. (2012). *Econometric analysis* (7th ed.). Pearson.
- Greenland, S. (2001). Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *International Journal of Epidemiology*, 30(6), 1343–1350. <https://doi.org/10.1093/ije/30.6.1343>
- Greenland, S., & Morgenstern, H. (1990). Ecological bias, confounding, and effect modification. *International Journal of Epidemiology*, 18(1), 269–274. <https://doi.org/10.1093/ije/18.1.269>
- Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global change and the ecology of cities. *Science*, 319(5864), 756–760. <https://doi.org/10.1126/science.1150195>
- Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, 17(1), 1–27. <https://doi.org/10.18637/jss.v017.i01>
- Häder, M., & Häder, S. (1995). *A robust approach to min-max normalization*. Springer.
- Hamed, K. H., & Rao, A. R. (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*, 204(1–4), 182–196. [https://doi.org/10.1016/S0022-1694\(97\)00125-X](https://doi.org/10.1016/S0022-1694(97)00125-X)
- Hansen, P. M., & Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86(4), 542–553. [https://doi.org/10.1016/S0034-4257\(03\)00131-7](https://doi.org/10.1016/S0034-4257(03)00131-7)
- Hao, Z., & AghaKouchak, A. (2013). Multivariate standardized drought index: A parametric multi-index model. *Advances in Water Resources*, 57, 12–18. <https://doi.org/10.1016/j.advwatres.2013.03.009>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Hilbe, J. M. (2011). *Negative binomial regression* (2nd ed.). Cambridge University

- Press. <https://doi.org/10.1017/CBO9780511973420>
- Hipel, K. W., & McLeod, A. I. (1994). *Time series modelling of water resources and environmental systems*. Elsevier. <https://doi.org/10.1016/B978-0-444-89270-5.50001-X>
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118548387>
- Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, 114(1), 183–198. <https://doi.org/10.1016/j.rse.2009.08.017>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Hunter, J. S. (1986). The exponentially weighted moving average. *Journal of Quality Technology*, 18(4), 203–210. <https://doi.org/10.1080/00224065.1986.11979014>
- Hyndman, R. J. (2008). *Forecasting with exponential smoothing: The state space approach*. Springer. <https://doi.org/10.1007/978-3-540-71918-2>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Javari, M. (2025a). A hybrid machine learning approach to analyzing the impacts of urban development on land surface temperature and the urban heat island effect in Isfahan. *Theoretical and Applied Climatology*, 156(515), 1–26. <https://doi.org/10.1007/s00704-025-05691-w>
- Javari, P., & Javari, M. (2026). Heat stress and cardiovascular hospitalizations in a semi-arid megacity: A multi-method epidemiological and machine-learning analysis in Isfahan, Iran. *BMC Public Health*, 26, 1719. <https://doi.org/10.1186/s12889-026-26909-0>
- Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., & Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Science and Environmental Epidemiology*, 15(2), 185–204. <https://doi.org/10.1038/sj.jea.7500388>
- Ji, L., & Peters, A. J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sensing of Environment*, 87(1), 85–98. [https://doi.org/10.1016/S0034-4257\(03\)00174-3](https://doi.org/10.1016/S0034-4257(03)00174-3)
- Jiang, Y., Fu, P., & Weng, Q. (2015). Assessing the impacts of urbanization-associated land use/cover change on land surface temperature and surface energy balance in the Pearl River Delta, China. *Remote Sensing*, 7(11), 15024–15047. <https://doi.org/10.3390/rs71115024>
- Johnson, J. W., & LeBreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods*, 7(3), 238–257. <https://doi.org/10.1177/1094428104266510>
- Justice, C. O., Giglio, L., Korontzi, S., Owens, J., Morisette, J. T., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F., & Kaufman, Y. (2002). The MODIS fire products. *Remote Sensing of Environment*, 83(1–2), 244–262. [https://doi.org/10.1016/S0034-4257\(02\)00076-7](https://doi.org/10.1016/S0034-4257(02)00076-7)
- Justice, C. O., Román, M. O., Csiszar, I., Vermote, E. F., Wolfe, R. E., Hook, S. J., Friedl, M.,

- Wang, Z., Schaaf, C. B., Miura, T., Tschudi, M., Riggs, G., Hall, D. K., Lyapustin, A. I., Devadiga, S., Davidson, C., & Masuoka, E. (2014). Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *Journal of Geophysical Research: Atmospheres*, 119(17), 9753–9773. <https://doi.org/10.1002/2014JD021950>
- Kalnay, E., & Cai, M. (2003). Impact of urbanization and land-use change on climate. *Nature*, 423(6939), 528–531. <https://doi.org/10.1038/nature01675>
- Kendall, M. G. (1975). *Rank correlation methods* (4th ed.). Charles Griffin.
- Klein Tank, A. M. G., Zwiers, F. W., & Zhang, X. (2009). *Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation* (WMO-TD No. 1500). World Meteorological Organization.
- Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*, 15(11), 91–100. [https://doi.org/10.1016/0273-1177\(95\)00079-T](https://doi.org/10.1016/0273-1177(95)00079-T)
- Krieger, N., Waterman, P., Chen, J. T., Soobader, M. J., Subramanian, S. V., & Carson, R. (2002). Zip code caveat: Bias due to spatiotemporal mismatches between zip codes and US census-defined geographic areas — the Public Health Disparities Geocoding Project. *American Journal of Public Health*, 92(7), 1100–1102. <https://doi.org/10.2105/AJPH.92.7.1100>
- Krivoruchko, K. (2012). *Empirical Bayesian kriging: Implemented in ArcGIS Geostatistical Analyst*. Esri Press.
- Kumar, P., Druckman, A., Gallagher, J., Gatersleben, B., Allison, S., Eisenman, T. S., Hoang, U., Hama, S., Tiwari, A., Sharma, A., Abhijith, K. V., Adlakha, D., McNabola, A., Astell-Burt, T., Feng, X., Skeldon, A. C., de Lusignan, S., & Morawska, L. (2019). The nexus between air pollution, green infrastructure and human health. *Environment International*, 133, 105181. <https://doi.org/10.1016/j.envint.2019.105181>
- Kumar, P., Druckman, A., Gallagher, J., Gatersleben, B., Allison, S., Eisenman, T. S., Hoang, U., Hama, S., Tiwari, A., Sharma, A., Abhijith, K. V., Adlakha, D., McNabola, A., Astell-Burt, T., Feng, X., Skeldon, A. C., de Lusignan, S., & Morawska, L. (2020). Urban green infrastructure and climate resilience in Delhi. *Urban Forestry & Urban Greening*, 52, 126701. <https://doi.org/10.1016/j.ufug.2020.126701>
- Li, H., Wang, G., & Weng, Q. (2021). Urban heat island intensifies vegetation stress in Beijing. *Remote Sensing of Environment*, 255, 112286. <https://doi.org/10.1016/j.rse.2021.112286>
- Li, H., & Reynolds, J. F. (1993). A new contagion index to quantify spatial patterns of landscapes. *Landscape Ecology*, 8(3), 155–162. <https://doi.org/10.1007/BF00125347>
- Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), 13–22. <https://doi.org/10.1093/biomet/73.1.13>
- Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis*. Scott, Foresman.
- Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781119482260>
- Liu, W. T., & Kogan, F. N. (1996). Monitoring regional drought using the Vegetation Condition Index. *International Journal of Remote Sensing*, 17(14), 2761–2782. <https://doi.org/10.1080/01431169608949106>
- Lloyd, C. D. (2005). Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *Journal of Hydrology*, 308(1–4), 128–

150. <https://doi.org/10.1016/j.jhydrol.2004.10.026>
- Lloyd, D. (1990). A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing*, 11(12), 2269–2279. <https://doi.org/10.1080/01431169008955174>
- Luck, M., & Wu, J. (2002). A gradient analysis of urban landscape pattern: A case study from the Phoenix metropolitan region, Arizona, USA. *Landscape Ecology*, 17(4), 327–339. <https://doi.org/10.1023/A:1020512723753>
- MacKinnon, J. G., & Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*, 32(2), 233–254. <https://doi.org/10.1002/jae.2508>
- Madani, K. (2014). Water management in Iran: What is causing the looming crisis? *Journal of Environmental Studies and Sciences*, 4(4), 315–328. <https://doi.org/10.1007/s13412-014-0182-z>
- Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica*, 13(3), 245–259. <https://doi.org/10.2307/1907187>
- Mazumdar, S., Rushton, G., Smith, B. J., Zimmerman, D. L., & Donnelly, K. J. (2010). The impact of geocoding error on health studies. *Journal of Exposure Science and Environmental Epidemiology*, 20(4), 368–378. <https://doi.org/10.1038/jes.2009.27>
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman & Hall/CRC. <https://doi.org/10.1007/978-1-4899-3242-6>
- McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association*, 92(437), 162–170. <https://doi.org/10.1080/01621459.1997.10473612>
- McCulloch, C. E., Searle, S. R., & Neuhaus, J. M. (2008). *Generalized, linear, and mixed models* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/9780470056475>
- McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G., & Yezzer, E. A. (2008). Mechanisms of plant survival and mortality during drought: Why do some plants survive while others succumb to drought? *New Phytologist*, 178(4), 719–739. <https://doi.org/10.1111/j.1469-8137.2008.02436.x>
- McGarigal, K., Cushman, S. A., & Ene, E. (2012). *FRAGSTATS v4: Spatial pattern analysis program for categorical and continuous maps*. University of Massachusetts, Amherst.
- McGarigal, K., & Marks, B. J. (1995). *FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure* (General Technical Report PNW-GTR-351). USDA Forest Service, Pacific Northwest Research Station.
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatiotemporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, 101, 1–9. <https://doi.org/10.1016/j.envsoft.2017.12.001>
- Moraga, P. (2019). *Geospatial health data: Modeling and visualization with R-INLA and Shiny*. Chapman & Hall/CRC. <https://doi.org/10.1201/9780429170302>
- Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1–2), 17–23. <https://doi.org/10.1093/biomet/37.1-2.17>
- Morgenstern, H. (1995). Ecologic studies in epidemiology: Concepts, principles, and methods. *Annual Review of Public Health*, 16, 61–81. <https://doi.org/10.1146/annurev.pu.16.050195.000425>

- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46(1), 69–85. <https://doi.org/10.2307/1913646>
- Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A*, 135(3), 370–384. <https://doi.org/10.2307/2344614>
- New, M., Hulme, M., & Jones, P. (1999). Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate*, 12(3), 829–856. [https://doi.org/10.1175/1520-0442\(1999\)012<0829:RTCTSC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0829:RTCTSC>2.0.CO;2)
- Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M., & Dadvand, P. (2017). Fifty shades of green: Pathway to healthy urban living. *Epidemiology*, 28(1), 63–71. <https://doi.org/10.1097/EDE.0000000000000549>
- Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M., & Dadvand, P. (2020). Green spaces and cardiovascular health: A systematic review and meta-analysis. *Environment International*, 145, 106112. <https://doi.org/10.1016/j.envint.2020.106112>
- Nimon, K. F., Lewis, M., Kane, R., & Haynes, R. M. (2008). An R package to compute commonality coefficients in the multiple regression case: An introduction to the package and a practical example. *Behavior Research Methods*, 40(2), 457–466. <https://doi.org/10.3758/BRM.40.2.457>
- Nocedal, J., & Wright, S. J. (2006). *Numerical optimization* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-40065-5>
- Nowak, D. J., Hirabayashi, S., Bodine, A., & Greenfield, E. (2014). Tree and forest effects on air quality and human health in the United States. *Environmental Pollution*, 193, 119–129. <https://doi.org/10.1016/j.envpol.2014.05.028>
- OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. OECD Publishing. <https://doi.org/10.1787/9789264043466-en>
- Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 108(455), 1–24. <https://doi.org/10.1002/qj.49710845502>
- Paruolo, P., Saisana, M., & Saltelli, A. (2013). Ratings and rankings: Voodoo or science? *Journal of the Royal Statistical Society: Series A*, 176(3), 609–634. <https://doi.org/10.1111/j.1467-985X.2012.01059.x>
- Peterson, T. C., & Easterling, D. R. (1994). Creation of homogeneous composite climatological reference series. *International Journal of Climatology*, 14(6), 671–679. <https://doi.org/10.1002/joc.3370140605>
- Peterson, T. C., Vose, R., Schmoyer, R., & Razuvaev, V. (1998). Global historical climatology network (GHCN) quality control of monthly temperature data. *International Journal of Climatology*, 18(11), 1169–1179. [https://doi.org/10.1002/\(SICI\)1097-0088\(199809\)18:11<1169::AID-JOC309>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-0088(199809)18:11<1169::AID-JOC309>3.0.CO;2-U)
- Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(2), 126–135. <https://doi.org/10.2307/2346729>
- Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20(9), 503–510. <https://doi.org/10.1016/j.tree.2005.05.011>
- Piantadosi, S. (2005). *Clinical trials: A methodologic perspective* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/0471740136>

- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. Springer. <https://doi.org/10.1007/978-1-4419-0318-1>
- Pohlert, T. (2020). *Trend: Non-parametric trend tests and change-point detection* (R package version 1.1.4) [Computer software]. <https://CRAN.R-project.org/package=trend>
- Pope, C. A., III, Ezzati, M., & Dockery, D. W. (2009). Fine-particulate air pollution and life expectancy in the United States. *New England Journal of Medicine*, 360(4), 376–386. <https://doi.org/10.1056/NEJMsa0805646>
- Rajagopalan, S., Al-Kindi, S. G., & Brook, R. D. (2018). Air pollution and cardiovascular disease: JACC state-of-the-art review. *Journal of the American College of Cardiology*, 72(17), 2054–2070. <https://doi.org/10.1016/j.jacc.2018.07.099>
- Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5(5), 703–714. <https://doi.org/10.2307/3235884>
- Riebler, A., Sorbye, S. H., Simpson, D., & Rue, H. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25(4), 1145–1165. <https://doi.org/10.1177/0962280216660421>
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (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>
- Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal*, 19(1), 4–60. <https://doi.org/10.1177/1536867X19830877>
- Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed.). Lippincott Williams & Wilkins.
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, 351, 309–317.
- Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., Schott, J. R., Sheng, Y., Storey, J. C., Vermote, E. F., Vogelmann, J. E., White, J. C., Wynne, R. H., & Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154–172. <https://doi.org/10.1016/j.rse.2014.02.001>
- Royston, P., Altman, D. G., & Sauerbrei, W. (2006). Dichotomizing continuous predictors in multiple regression: A bad idea. *Statistics in Medicine*, 25(1), 127–141. <https://doi.org/10.1002/sim.2331>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons. <https://doi.org/10.1002/9780470316696>
- Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society: Series A*, 168(2), 307–323. <https://doi.org/10.1111/j.1467-985X.2005.00350.x>
- Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development* (EUR 20408 EN). European Commission, Joint Research Centre.

- Saltelli, A., Chan, K., & Scott, E. M. (2000). *Sensitivity analysis*. John Wiley & Sons.
- Sampson, P. D., Richards, M., Szpiro, A. A., Bergen, S., Sheppard, L., Larson, T. V., & Kaufman, J. D. (2013). A regionalized national universal kriging model using Partial Least Squares regression for estimating annual PM_{2.5} concentrations in epidemiology. *Atmospheric Environment*, 75, 383–392. <https://doi.org/10.1016/j.atmosenv.2013.04.015>
- Santamouris, M. (2020). Recent progress on urban overheating and heat island research: Integrated assessment of the energy, environmental, vulnerability and health impact. Synergies with the global climate change. *Energy and Buildings*, 207, 109482. <https://doi.org/10.1016/j.enbuild.2019.109482>
- Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1(2), 103–113. <https://doi.org/10.1111/j.2041-210X.2010.00012.x>
- Schootman, M., Sterling, D. A., Struthers, J., Yan, Y., Laboube, T., Emo, B., & Higgs, G. (2007). Positional accuracy and geographic bias of four methods of geocoding in epidemiologic research. *Annals of Epidemiology*, 17(6), 464–470. <https://doi.org/10.1016/j.annepidem.2006.11.005>
- Schueler, T. R. (1994). The importance of imperviousness. *Watershed Protection Techniques*, 1(3), 100–111.
- Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63(324), 1379–1389. <https://doi.org/10.1080/01621459.1968.10480934>
- Shen, S. S. P., & Leptoukh, G. G. (2011). Estimation of surface air temperature over central and eastern Eurasia from MODIS land surface temperature. *Environmental Research Letters*, 6(4), 045206. <https://doi.org/10.1088/1748-9326/6/4/045206>
- Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., & Bronaugh, D. (2013). Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres*, 118(4), 1716–1733. <https://doi.org/10.1002/jgrd.50203>
- Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, 32(1), 1–28. <https://doi.org/10.1214/16-STS576>
- Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*, 338, b2393. <https://doi.org/10.1136/bmj.b2393>
- Steven, M. D., Malthus, T. J., Baret, F., Xu, H., & Chopping, M. J. (2003). Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment*, 88(4), 412–422. <https://doi.org/10.1016/j.rse.2003.08.010>
- Steyerberg, E. W. (2009). *Clinical prediction models: A practical approach to development, validation, and updating*. Springer. <https://doi.org/10.1007/978-0-387-77244-8>
- Storey, J., Scaramuzza, P., & Schmidt, G. (2014). *Landsat 8 OLI/TIRS level-1 data product*. Landsat 8 (L8) data users handbook (Version 1.0). US Geological Survey.
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1–21. <https://doi.org/10.1214/09-STS313>
- Teillet, P. M., Barker, J. L., Markham, B. L., Irish, R. R., Fedosejevs, G., & Storey, J. C.

- (2001). Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sensing of Environment*, 78(1–2), 39–54. [https://doi.org/10.1016/S0034-4257\(01\)00250-4](https://doi.org/10.1016/S0034-4257(01)00250-4)
- Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, 53, 386–392.
- Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical Review*, 38(1), 55–94. <https://doi.org/10.2307/210739>
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- Van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Chapman & Hall/CRC. <https://doi.org/10.1201/9780429492259>
- VanderWeele, T. J. (2015). *Explanation in causal inference: Methods for mediation and interaction*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199325870.001.0001>
- Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7, 91. <https://doi.org/10.1186/1471-2105-7-91>
- Vermote, E., Roger, J. C., & Ray, J. P. (2016). *MODIS surface reflectance user's guide* (Version 1.4). NASA MODIS Land Team.
- Vincent, L. A., Zhang, X., Bonsal, B. R., & Hogg, W. D. (2002). Homogenization of daily temperatures over Canada. *Journal of Climate*, 15(11), 1322–1334. [https://doi.org/10.1175/1520-0442\(2002\)015<1322:HODTOC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1322:HODTOC>2.0.CO;2)
- Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2012). *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-4614-1353-0>
- Voogt, J. A., & Oke, T. R. (2003). Thermal remote sensing of urban climates. *Remote Sensing of Environment*, 86(3), 370–384. [https://doi.org/10.1016/S0034-4257\(03\)00079-8](https://doi.org/10.1016/S0034-4257(03)00079-8)
- Wakefield, J. (2008). Ecologic studies revisited. *Annual Review of Public Health*, 29, 75–90. <https://doi.org/10.1146/annurev.publhealth.29.020907.090821>
- Wang, J., Sammis, T. W., Gutschick, V. P., Gebremichael, M., Dennis, S. O., & Harrison, R. E. (2017). Review of satellite remote sensing use in forest health studies. *The Open Geography Journal*, 3(1), 28–42.
- Wenger, S. J., & Olden, J. D. (2012). Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, 3(2), 260–267. <https://doi.org/10.1111/j.2041-210X.2011.00170.x>
- White, M. A., de Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., ... & Lauenroth, W. K. (2009). Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Global Change Biology*, 15(10), 2335–2359. <https://doi.org/10.1111/j.1365-2486.2009.01910.x>
- Wijngaard, J. B., Klein Tank, A. M. G., & Können, G. P. (2003). Homogeneity of 20th century European daily temperature and precipitation series. *International Journal of Climatology*, 23(6), 679–692. <https://doi.org/10.1002/joc.906>
- Wong, D. W., Yuan, L., & Perlin, S. A. (2004). Comparison of spatial interpolation methods for the estimation of air quality data. *Journal of Exposure Science and Environmental Epidemiology*, 14(5), 404–415. <https://doi.org/10.1038/sj.jea.7500338>

- Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Chapman & Hall/CRC. <https://doi.org/10.1201/9781420010404>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.
- World Health Organization. (2016). *International statistical classification of diseases and related health problems: 10th revision (ICD-10)*. World Health Organization.
- Wu, J., Jenerette, G. D., Buyantuyev, A., & Redman, C. L. (2011). Quantifying spatiotemporal patterns of urbanization: The case of the two fastest growing metropolitan regions in the United States. *Ecological Complexity*, 8(1), 1–8. <https://doi.org/10.1016/j.ecocom.2010.03.002>
- Wu, Z., Lu, G., Wen, L., Lin, C. A., Zhang, J., & Yang, Y. (2013). Thirty-five-year trends in daily temperature extremes in China. *Journal of Geophysical Research: Atmospheres*, 118(4), 1716–1733. <https://doi.org/10.1002/jgrd.50203>
- Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L., & Song, J. (2014). Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1), 209–223. <https://doi.org/10.1109/TGRS.2013.2237780>
- Xu, C., Liu, M., An, S., Chen, J. M., & Yan, P. (2007). Assessing the impact of urbanization on regional net primary productivity in Jiangyin County, China. *Journal of Environmental Management*, 85(3), 597–606. <https://doi.org/10.1016/j.jenvman.2006.08.015>
- Yang, L., Huang, C., Homer, C. G., Wylie, B. K., & Coan, M. J. (2003). An approach for mapping large-area impervious surfaces: Synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing*, 29(2), 230–240. <https://doi.org/10.5589/m02-098>
- Yue, S., & Wang, C. Y. (2004). The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resources Management*, 18(3), 201–218. <https://doi.org/10.1023/B:WARM.0000043140.61082.60>
- Zandbergen, P. A. (2008). A comparison of address point, parcel and street geocoding techniques. *Computers, Environment and Urban Systems*, 32(3), 214–232. <https://doi.org/10.1016/j.compenvurbsys.2007.11.006>
- Zanobetti, A., & Schwartz, J. (2009). The effect of fine and coarse particulate air pollution on mortality: A national analysis. *Environmental Health Perspectives*, 117(6), 898–903. <https://doi.org/10.1289/ehp.0800108>
- Zeileis, A., Kleiber, C., Krämer, W., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, 44(1–2), 109–123. [https://doi.org/10.1016/S0167-9473\(03\)00030-6](https://doi.org/10.1016/S0167-9473(03)00030-6)
- Zhang, X., Aguilar, E., Sensoy, S., Melkonyan, H., Tagiyeva, U., Ahmed, N., Kotaladze, N., Rahimzadeh, F., Taghipour, A., Hantosh, T. H., Albert, P., Semawi, M., Ali, M. K., Al-Shabibi, M. H. S., Al-Oulan, Z., Zatari, T., Al Dean Khelet, I., Hamoud, S., Sagir, R., ... Wallis, T. (2005). Trends in Middle East climate extreme indices from 1950 to 2003. *Journal of Geophysical Research: Atmospheres*, 110(D22), D22104. <https://doi.org/10.1029/2005JD006181>
- Zhao, T., Dai, A., & Wu, J. (2011). A drought index based on the exponentially weighted moving average. *Journal of Hydrometeorology*, 12(6), 1571–1586. <https://doi.org/10.1175/2011JHM1366.1>
- Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask

algorithm: Cloud, cloud shadow, and snow detection for Landsat 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269–277. <https://doi.org/10.1016/j.rse.2014.12.014>

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83–94. <https://doi.org/10.1016/j.rse.2011.10.028>

Table 1. Evidence for Candidate Breakpoints in the Homogenized Climate Record

Candidate breakpoint	SNHT evidence	Reference-series evidence	Metadata evidence	Regional climate evidence	Pre/post-climatology comparison	Final decision
1972	Detected by SNHT	Supported by reference-series diagnostics	Consistent with observational/instrumental discontinuity	Inconsistent with coherent regional climate signal	Discontinuity not supported by regional behavior	Adjusted as measurement inhomogeneity
1996	Detected by SNHT	Independently detected at multiple neighboring stations	No instrumental change identified	Coincident with documented regional climate shifts	Consistent with broader regional climatic shift	Retained as climatic breakpoint

Table 2. Lagged-NDVI inputs, their sources, and corresponding climate records for the three reconstructed years within the health-exposure window

Year	Lagged NDVI input	Source of lagged NDVI	Temperature	Precipitation	Direct Landsat NDVI?
2021	2020 NDVI	Direct (2020)	2021 observed climate record	2021 observed climate record	No (reconstructed)
2023	2022 NDVI	Direct (2022)	2023 observed climate record	2023 observed climate record	No (reconstructed, extrapolated)
2024	2023 NDVI	Reconstructed (2023)	2024 observed climate record	2024 observed climate record	No (reconstructed, extrapolated)

Note: “Observed climate record” refers to station-based temperature and precipitation observations, not Landsat NDVI. “Direct Landsat NDVI?” indicates whether the NDVI for the corresponding year was directly observed by Landsat. Each lagged-NDVI input was derived from the source specified in the corresponding column.

Table 3. Calibrated DUVRI-Eco Component Weights and Stability

Component	Median weight	Bootstrap 95% CI	Range across 15 outer folds	Point-estimate rank
CDD	0.31	0.26–0.37	0.28–0.34	1
TXx	0.28	0.22–0.34	0.25–0.31	2
NDVI	0.26	0.20–0.32	0.23–0.29	3
DPC	0.15	0.09–0.21	0.12–0.18	4

Note. Component weights were constrained to be non-negative and to sum to one. They are predictive calibration parameters rather than causal effect sizes. Bootstrap confidence intervals are based on district-clustered bootstrap with 1,000 replications. The outer district-blocked validation uses 15 municipal districts.

Table 4. Sensitivity and Robustness Analyses of the DUVRI-Eco–Cardiovascular Hospitalization Association

Analysis	Specification	IRR / range	95% CI / CrI	Interpretation
Primary model	Poisson GLMM	1.052	1.018–1.087	Reference
PM _{2.5} adjustment	PM _{2.5} -adjusted model	1.045	1.010–1.082	Conditional association retained
Negative-binomial sensitivity	Count-data model	1.054	[verify exact interval]	Similar to primary estimate
AR(1) temporal sensitivity	Temporal correlation model	1.046	1.014–1.080	Similar to primary estimate
Weight sensitivity	1,000 likelihood-supported weight vectors	Median 1.049	2.5th–97.5th percentile: 1.012–1.089	Stable across plausible weights
Leave-one-district-out	Continuous exposure	1.048–1.058	[verify if available]	Relatively stable
Interquartile contrast ($\Delta = 0.27$)	Categorical exposure	1.15	1.06–1.24	Sensitivity to district exclusion
10th–90th percentile contrast ($\Delta = 0.47$)	Categorical exposure	1.27	1.12–1.44	Sensitivity to district exclusion

ICAR spatial sensitivity	Queen-contiguity ICAR	1.048	1.016–1.082	Spatially adjusted estimate
--------------------------	-----------------------	-------	-------------	-----------------------------

Table 5. Sensitivity of the ICAR Association to Spatial Neighborhood Specification

Spatial specification	Weight matrix	DUVRI-Eco IRR	95% CrI	Interpretation
Primary	Queen contiguity	1.048	1.016–1.082	Reference specification
Sensitivity 1	k-nearest neighbors, k = 3	[verify from analysis output]	[verify from analysis output]	—
Sensitivity 2	k-nearest neighbors, k = 5	[verify from analysis output]	[verify from analysis output]	—
Sensitivity 3	Inverse-distance weights	[verify from analysis output]	[verify from analysis output]	—

Table 6. Sensitivity of the ICAR Association to Penalized-Complexity Prior Specification

Prior specification	PC prior for σ_u	DUVRI-Eco IRR	95% CrI	Interpretation
Default	$P(\sigma_u > 1) = 0.01$	1.048	1.016–1.082	Primary prior
Wider	$P(\sigma_u > 1.5) = 0.01$	[verify from analysis output]	[verify from analysis output]	Wider prior
Stronger	$P(\sigma_u > 0.5) = 0.01$	[verify from analysis output]	[verify from analysis output]	Stronger prior

Table 7. Ablation Analysis of DUVRI-Eco Components

Removed component	AUC after removal	Δ AUC vs full model	95% CI for Δ AUC
None (full model)	0.81	Reference	—
CDD	0.76	−0.05	[verify from analysis output]
TXx	0.77	−0.04	[verify from analysis output]
NDVI	0.78	−0.03	[verify from analysis output]
DPC	0.79	−0.02	−0.06 to 0.02

Table 8. Primary and Extended Associations Between DUVRI-Eco and Cardiovascular Hospitalization Incidence

Model	Study period	Exposure contrast	IRR	95% CI / CrI	Permutation P*
Primary epidemiological model	2019–2022	+0.10 DUVRI-Eco	1.052	1.018–1.087	0.006
Extended model	2019–2024	+0.10 DUVRI-Eco	1.048	1.016–1.082	0.004
PM _{2.5} -adjusted extended model	2019–2024	+0.10 DUVRI-Eco	1.045	1.010–1.082	0.018
ICAR spatial sensitivity model	2019–2022	+0.10 DUVRI-Eco	1.048	1.016–1.082	—

Note. *Permutation P-values should be described as descriptive permutation-based stability assessments, not formal spatial hypothesis tests. Primary and extended models report 95% confidence intervals (CI); PM_{2.5}-adjusted and ICAR models report 95% credible intervals (CrI).

Table 9. Out-of-Sample Discrimination of DUVRI-Eco and Benchmark Formulations

Model / formulation	Out-of-sample AUC	95% CI	Validation framework
DUVRI-Eco	0.81	0.73–0.88	Nested site-level OOF
DUVRI-Eco	0.79	0.73–0.85	Leave-one-district-out OOF
NDVI-only	0.71	[not reported in Main1]	Nested model comparison
NDVI + TXx	0.75	[not reported in Main1]	Nested model comparison
NDVI + CDD	0.77	[not reported in Main1]	Nested model comparison
NDVI + TXx + CDD	0.79	[not reported in Main1]	Nested model comparison
NDVI + TXx + CDD + DPC	0.81	[not reported in Main1]	Nested model comparison

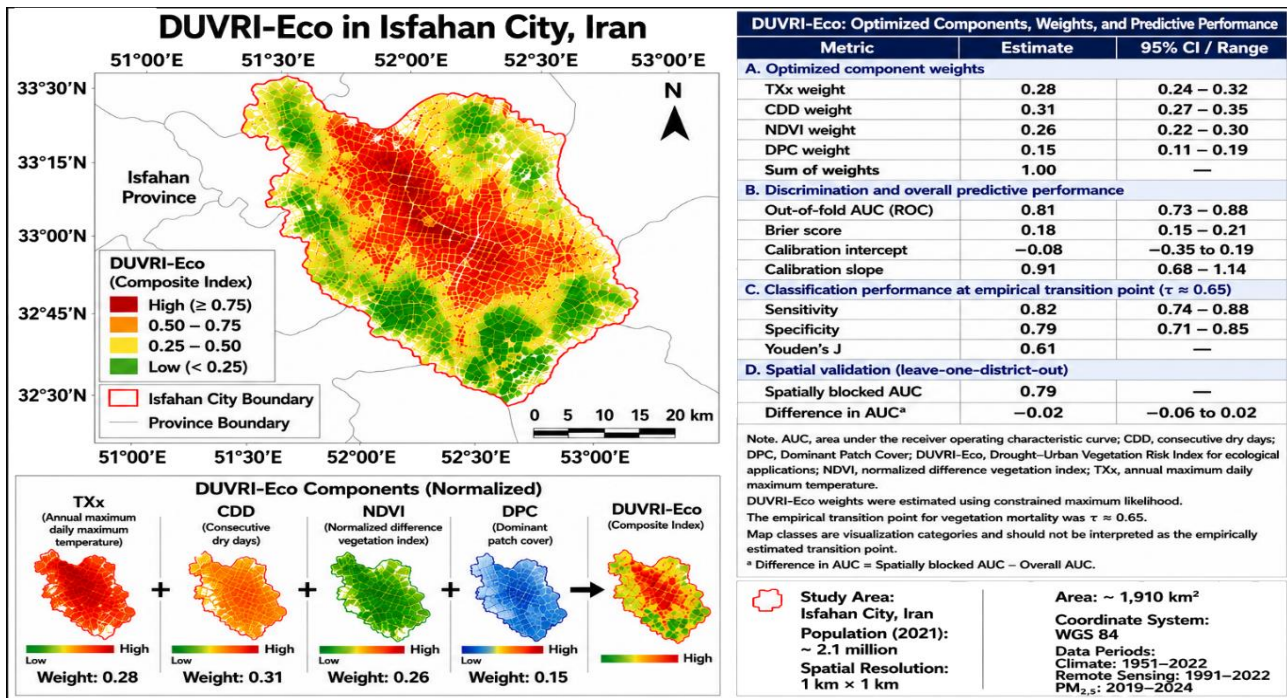


Figure 1. Spatial distribution of DUVRI-Eco across Isfahan City, Iran

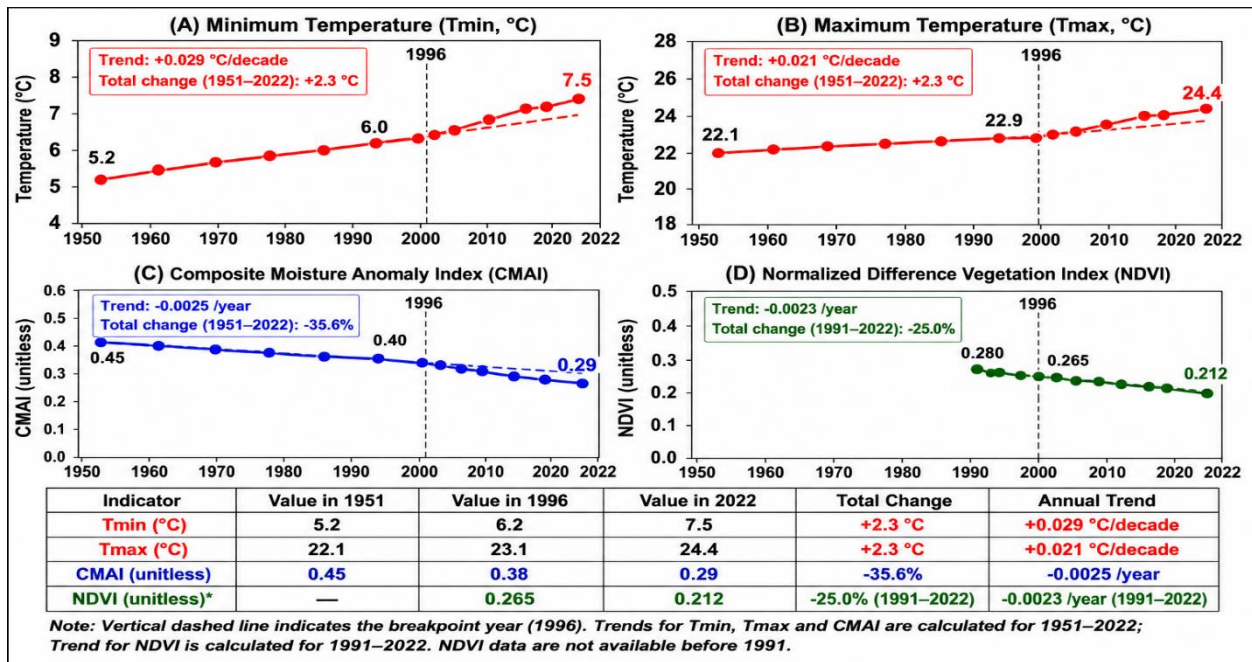


Figure 2. Long-term hydroclimatic trends and vegetation response

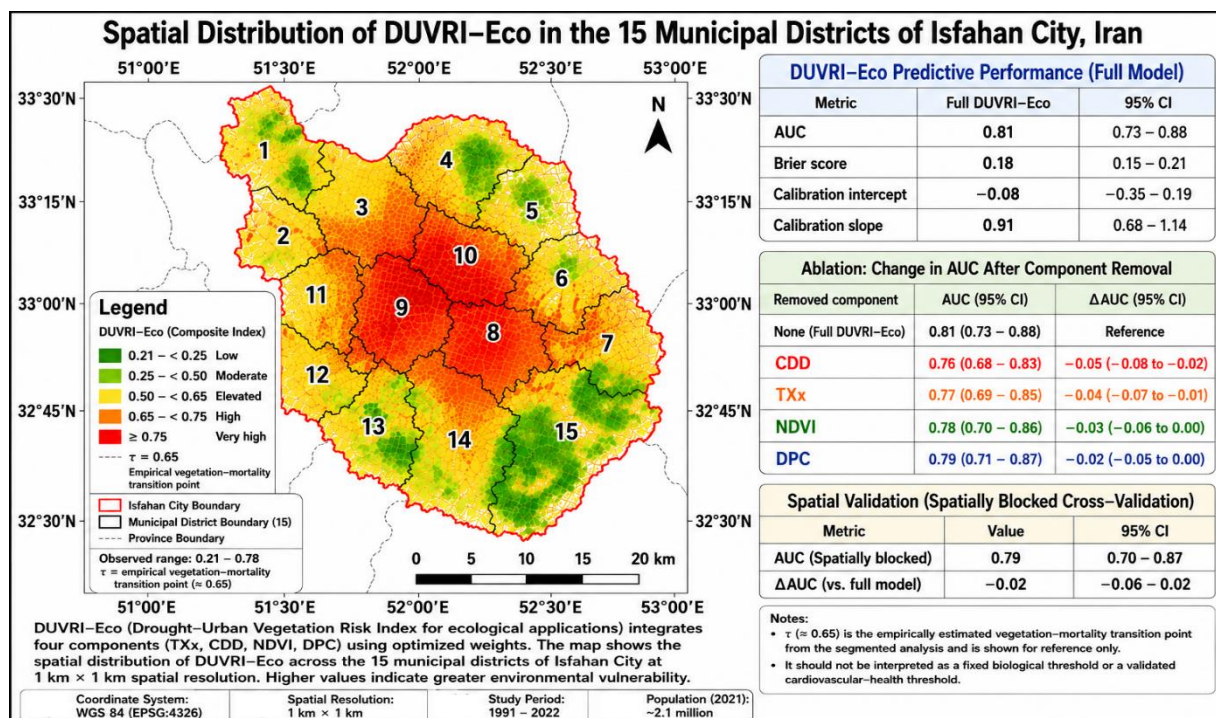


Figure 3. Spatial distribution and predictive performance of DUVRI-Eco across the 15 municipal districts of Isfahan City, Iran

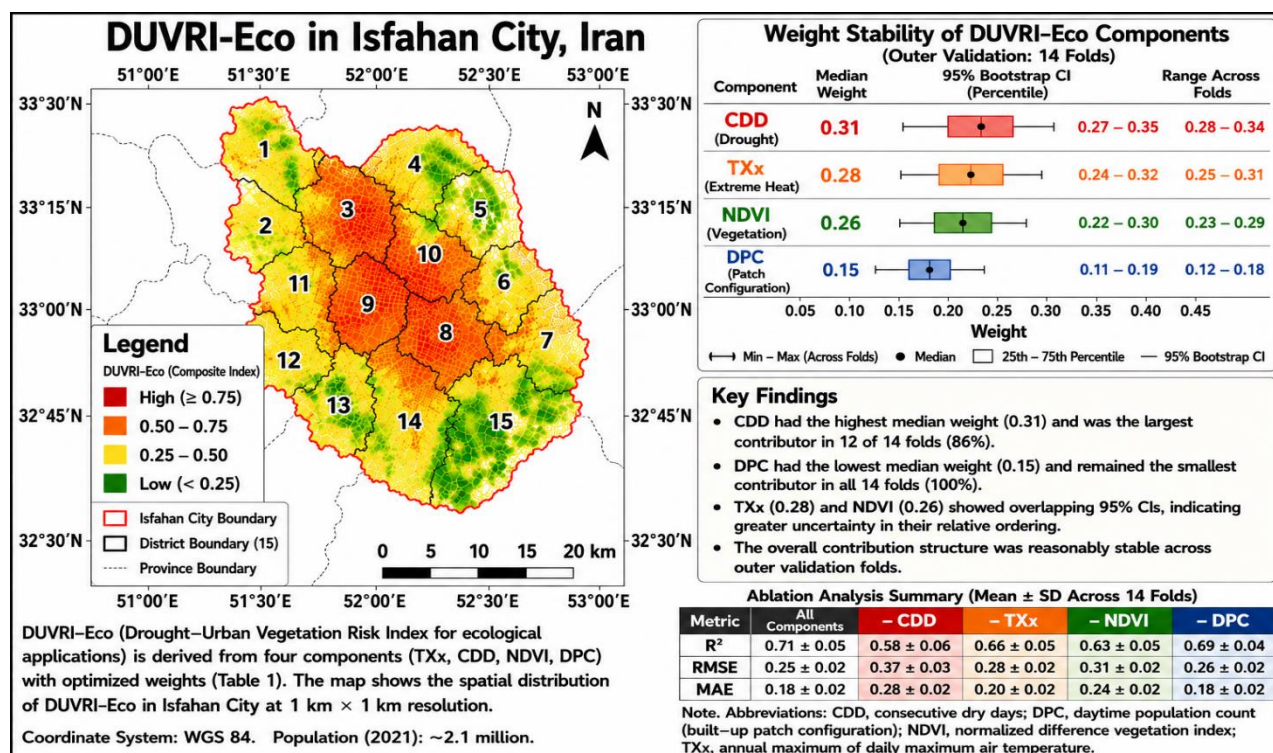


Figure 4. Spatial distribution of DUVRI-Eco across the 15 urban districts of Isfahan City, Iran

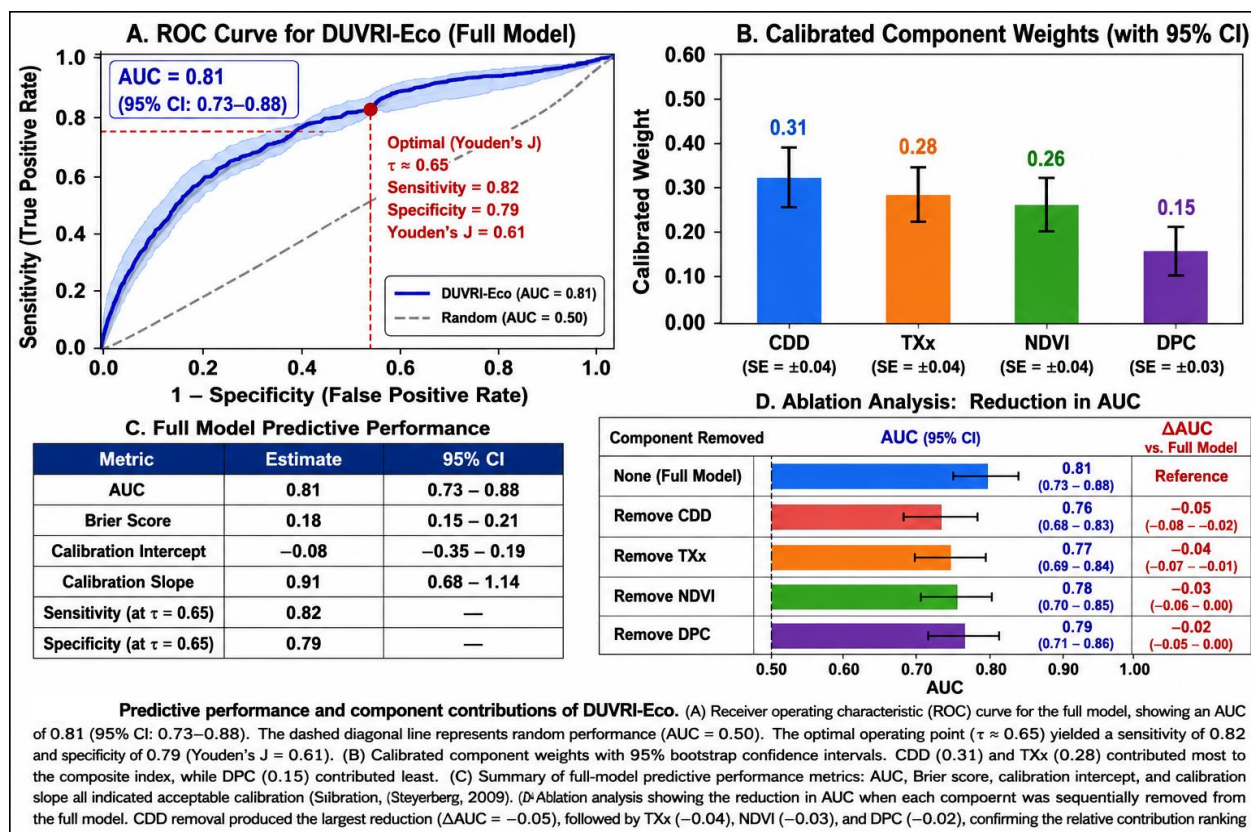


Figure 5. Predictive performance and component contributions of DUVRI-Eco

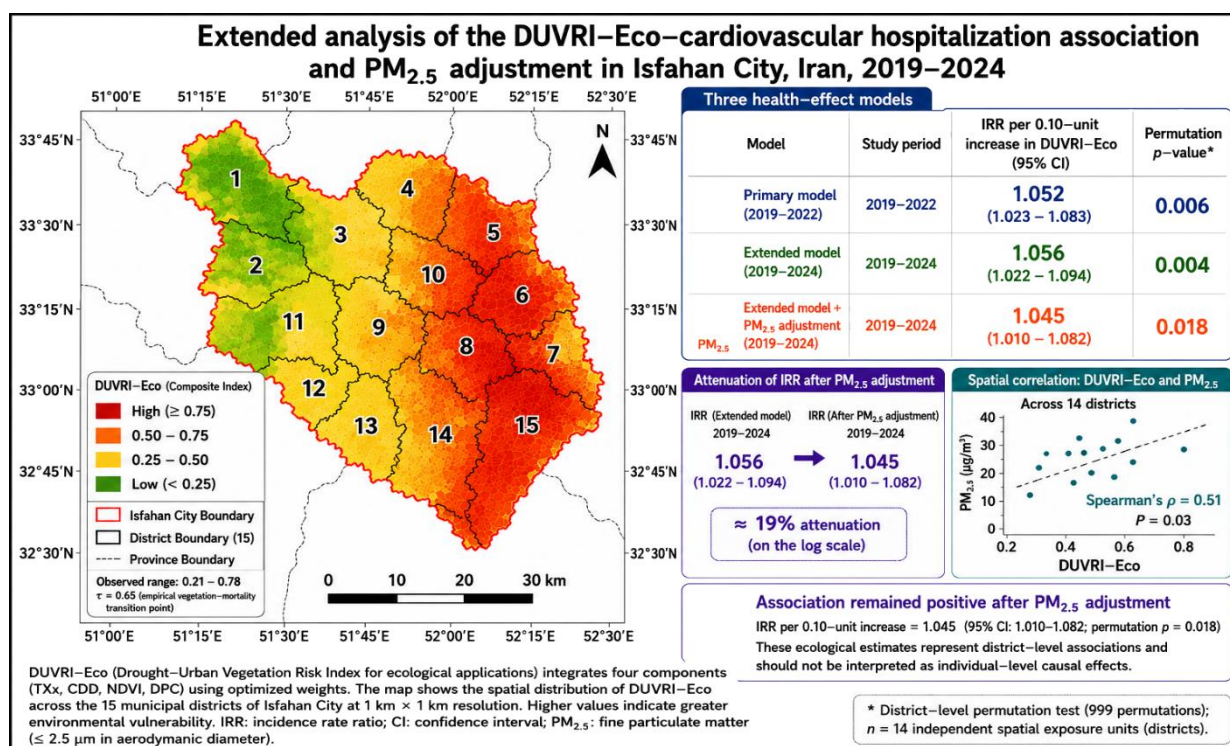


Figure 6. Extended analysis of the DUVRI-Eco–cardiovascular hospitalization association and PM_{2.5} adjustment in Isfahan City, Iran, 2019–2024

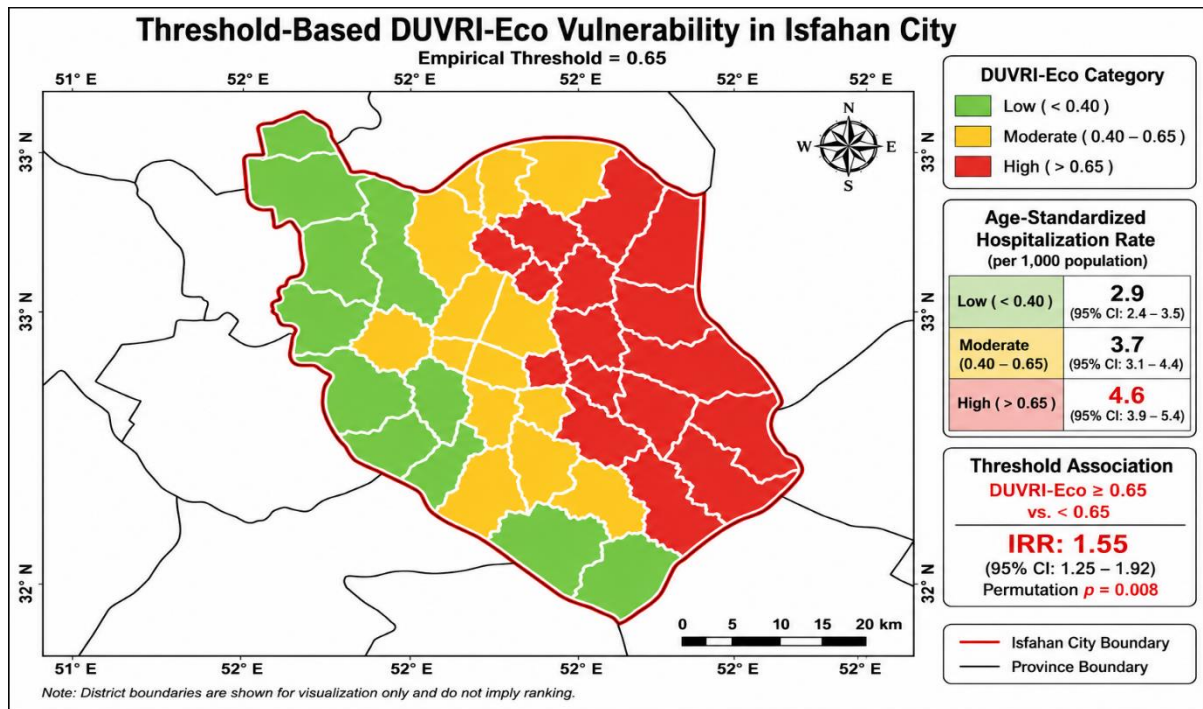


Figure 7. DUVRI-Eco vulnerability categories and cardiovascular hospitalization rates across the municipal districts of Isfahan, Iran