Theoretical and Applied Climatology (2026) 157:466
https://doi.org/10.1007/s00704-026-06389-3

RESEARCH

Spatial and severity-dependent controls of thermal comfort across Türkiye: a SHAP-based decomposition of UTCI drivers

Doğukan Doğu Yavaşlı[1]

Received: 7 April 2026 / Accepted: 12 June 2026

Abstract

Outdoor thermal stress is intensifying across Türkiye, but its meteorological controls remain poorly resolved across contrasting climate regimes. We use explainable machine learning to decompose model-attributed contributions to daily maximum Universal Thermal Climate Index (UTCI) variability, rather than to independently predict a derived thermal index. ERA5 and ERA5-HEAT data for 1950–2025 were used to assemble 826,799 warm-season grid-day observations with UTCI > 26 °C at 0.25° × 0.25° resolution. A Random Forest model was fitted using seven predictors: 2-m air temperature (T2M), relative humidity, vapor pressure deficit (VPD), wind speed, surface solar radiation downwards, surface thermal radiation downwards (STRD), and specific humidity. Model performance was treated as an internal consistency check for attribution ($R^2 = 0.936$, RMSE = 1.80 °C). Spatial and temporal blocked validation confirmed stable SHAP rankings, with T2M and VPD ranked first and second in all 12 folds (Kendall's W = 0.963). Humidity-correlation, perturbation-mode, and ablation tests showed that VPD represents a robust moisture-demand dimension rather than an independent humidity effect. Severity-stratified SHAP showed that VPD attribution share increased from 18.7% under moderate heat stress to 28.7% under very strong heat stress, while wind-speed and STRD shares declined. Pixel-wise SHAP identified VPD dominance across the semi-arid interior, wind sensitivity along the Black Sea and Marmara margins, and localized STRD dominance along parts of the Mediterranean coast. Very strong heat grid-days increased significantly (+5.0 grid-days yr^{-1}), whereas extreme heat showed a non-significant increase (+0.62 grid-days yr^{-1}). These regimes can inform regionally differentiated heat-risk adaptation.

1 Introduction

The Eastern Mediterranean and Middle East (EMME) is a globally prominent climate change hotspot, warming approximately 20% faster than the global mean and exhibiting a six- to eight-fold increase in heat-wave frequency, intensity, and duration since the 1960 s (Giorgi 2006; Lionello and Scarascia 2018; Lelieveld et al. 2016; Kuglitsch et al. 2010; Zittis et al. 2022 a; Cos et al. 2022). Projected end-of-century summer warming of 3.5–7 °C under high-emission scenarios (Lelieveld et al. 2012; Zittis et al. 2016) sharpens the public-health, labour-productivity, urban-habitability, and ecosystem-resilience implications of this trajectory (Mora et al. 2017; Raymond et al. 2020; Ahmadalipour and Moradkhani 2018).

The physiological consequences of prolonged heat exposure, which can lead to heat exhaustion, cardiovascular strain, and even death from heatstroke, are not caused by air temperature alone. Instead, they are caused by a complicated mix of weather factors, such as humidity, wind speed, and radiation (Parsons 2007; Fiala et al. 2012). Sherwood and Huber (2010) showed theoretically that atmospheric moisture is very important for determining the upper limits of human thermoregulatory capacity by setting a wet-bulb temperature limit of 35 °C for human survivability. Vecellio et al. (2022) showed empirically that actual critical Tw thresholds are much lower, around 30.6 °C in humid environments and 25–28 °C in hot-dry conditions. Recent research by Vecellio et al. (2023) integrated empirically established thresholds with CMIP6 projections, indicating a significantly increased global risk from both humid and dry heat stress. The relative significance of specific meteorological variables in influencing thermal stress fluctuates with

✉ Doğukan Doğu Yavaşlı
dogukan.yavasli@ahievran.edu.tr

[1] Department of Geography, Kırşehir Ahi Evran University, Kırşehir, Türkiye

Published online: 10 July 2026

the intensity of heat exposure and the existing atmospheric conditions. This complexity necessitates analytical methods that can accommodate non-linear, interactive, and spatially diverse relationships among climate drivers.

The Universal Thermal Climate Index (UTCI) has emerged as the most physiologically robust metric for assessing outdoor thermal comfort, integrating the thermodynamic effects of air temperature, mean radiant temperature, wind speed, and humidity through an advanced multi-node human thermoregulation model (Jendritzky et al. 2012; Fiala et al. 2012; Bröde et al. 2012). Unlike simpler indices such as the Heat Index or Wet Bulb Globe Temperature (WBGT), the UTCI captures the full spectrum of human thermal stress from extreme cold to extreme heat through an adaptive clothing model and multi-segment physiological simulation (Havenith et al. 2012 Błażejczyk et al., 2012). Comparative evaluations have demonstrated UTCI's superior sensitivity to wind speed relative to other indices such as PET (Provençal et al. 2016) and its higher discriminative power for identifying hazardous thermal conditions (Błażejczyk et al. 2012). Since its development within COST Action 730 (Błażejczyk et al. 2013), UTCI has been applied across diverse bioclimatological domains spanning urban climate assessment, tourism climatology, epidemiology, and climate change impact evaluation (Błażejczyk and Kuchcik 2021; Potchter et al. 2018).

The ERA5-HEAT dataset, developed by the Copernicus Climate Change Service, provides the first global gridded UTCI estimates derived from ERA5 reanalysis at 0.25° × 0.25° spatial resolution and hourly temporal resolution, enabling long-term, spatially continuous thermal comfort assessments at unprecedented scale (Di Napoli et al. 2021). ERA5-HEAT UTCI has been validated against European mortality data across 21 cities (Urban et al. 2021), applied to assess spatiotemporal UTCI changes across the Middle East and North Africa (Hamed et al. 2023), and used in operational thermal health hazard forecasting (Pappenberger et al. 2015). For the EMME region, Di Napoli et al. (2018) identified two distinct European bioclimatic zones using 38 years of reanalysis-based UTCI: a heat-stress-dominated southern and Mediterranean zone and a neutral-to-cold northern zone. Within Türkiye specifically, thermal comfort studies have primarily focused on urban-scale analyses (Yılmaz et al. 2024), while comprehensive national-scale UTCI assessments leveraging the full temporal depth of ERA5-HEAT remain limited.

Even though there are more UTCI datasets and studies, most thermal comfort studies still use traditional statistical methods like correlation analyses, regression models, or index-based trend classifications. These methods don't work well for climate drivers that have nonlinear, interactive, and threshold-dependent relationships (Nazarian et al. 2022; Potchter et al. 2018). The assumption that heat stress is driven uniformly by temperature across all regions and severity levels represents a critical oversimplification. In reality, the transition from moderate to extreme heat stress may be governed by fundamentally different atmospheric mechanisms depending on local topography, maritime influence, and land-surface characteristics. Identifying these spatially and severity-dependent transitions requires analytical frameworks capable of decomposing complex, high-dimensional interactions without imposing a priori parametric assumptions.

Machine learning (ML) methods, particularly ensemble tree-based models such as Random Forests (Breiman 2001), have demonstrated exceptional predictive performance in environmental and climate sciences owing to their capacity to model nonlinear relationships and high-order feature interactions (Reichstein et al. 2019). In the thermal comfort domain, ML approaches have been applied for outdoor thermal perception prediction (Guo et al. 2024; Kim et al. 2018) and high-resolution UTCI downscaling (Briegel et al. 2024), consistently outperforming traditional regression methods. The intrinsic opacity of machine learning models, often referred to as the "black-box" problem, has traditionally limited their applicability in scientific contexts that necessitate mechanistic comprehension. The advent of Explainable Artificial Intelligence (XAI) techniques, most notably SHapley Additive exPlanations (SHAP; Lundberg and Lee 2017), has transformed this landscape by providing theoretically grounded, model-agnostic explanations of individual predictions. SHAP values, rooted in cooperative game theory, quantify the marginal contribution of each input feature to a given prediction, enabling researchers to decompose model outputs into interpretable, additive feature attributions. The TreeExplainer algorithm (Lundberg et al. 2020) enables exact, computationally efficient SHAP value computation for tree-based ensembles. Recent methodological evaluations have confirmed SHAP as a robust and faithful XAI method for climate science applications (Bommer et al. 2024; Flora et al. 2024), while Jiang et al. (2024) have demonstrated how interpretable ML can advance process understanding in the geosciences through spatial SHAP analysis.

A growing number of studies have applied SHAP and related XAI methods to environmental phenomena, including precipitation estimation (He et al. 2023), urban heat island analysis (Hong et al. 2025), and environmental hazard mapping (Wang et al. 2024; Gholami et al. 2024). In the specific domain of outdoor thermal comfort, Guo et al. (2024) used SHAP to reveal that the dominant UTCI driver shifts between mean radiant temperature (in unshaded spaces) and air temperature (in shaded spaces), demonstrating condition-dependent driver hierarchies. However, no study has yet applied XAI to systematically decompose how UTCI driver importance shifts across thermal stress severity categories at the regional climatological scale, a gap that limits both scientific understanding and the design of severity-specific heat mitigation strategies.

The geographic variety of thermal comfort drivers may be studied in an excellent natural laboratory in Turkey, which is located at the crossroads of three main climatic zones: Mediterranean, continental, and oceanic. The country's seven official geographical regions span humid Black Sea coasts with persistent cloud cover and moderate temperatures, semi-arid Central and Eastern Anatolian plateaus characterized by extreme continental temperature amplitudes and atmospheric desiccation, and the thermally intense Mediterranean and Aegean coastlines where radiative forcing and maritime moisture interact in complex ways (Özturk et al. 2015). This climatic diversity, compressed within a single national territory, provides an opportunity to systematically evaluate how the dominant mechanisms of heat stress shift across contrasting environments.

This study addresses the following research objectives: First, we apply SHAP analysis not only at global and regional aggregates but at the individual 0.25° × 0.25° grid-cell level, producing a pixel-wise map of dominant secondary climate drivers for thermal comfort across the entirety of Türkiye. Second, by independently analyzing four UTCI severity bands (moderate, strong, very strong, and extreme heat stress), we examine how climate driver importance restructures as heat stress intensifies, revealing nonlinear transitions in moisture-related and radiation-related drivers that are invisible to analyses aggregating across severity levels. Third, the extended analysis period (1950–2025) enables assessment of long-term trends in severe heat event frequency, providing temporal context for the observed spatial driver patterns. Because UTCI is itself a meteorologically derived thermal index, our aim is not to treat the machine-learning model as a purely independent predictive exercise, but rather to use it as an interpretable framework for decomposing the relative contributions of related climate drivers across space and heat-stress severity levels.

The primary research question guiding this study is: at severe heat stress thresholds (UTCI > 38 °C), how does the relative importance of individual climate drivers, particularly humidity, wind speed, and radiation, change compared to moderate stress conditions, and do the dominant secondary drivers exhibit systematic spatial patterns corresponding to known climatological gradients? In this context, "importance" refers to model-attributed contribution within an explainable machine-learning framework rather than to independent causal effect size in a strict physical sense.

2 Materials and methods

2.1 Study area

The study focuses on Türkiye, a transcontinental country situated predominantly in the Anatolian Peninsula of southwestern Asia, with a smaller portion (Eastern Thrace) extending into southeastern Europe (Fig. 1). Türkiye is strategically located within the Eastern Mediterranean basin (approximately 26°E-45°E, 36°N-42°N), a region that has

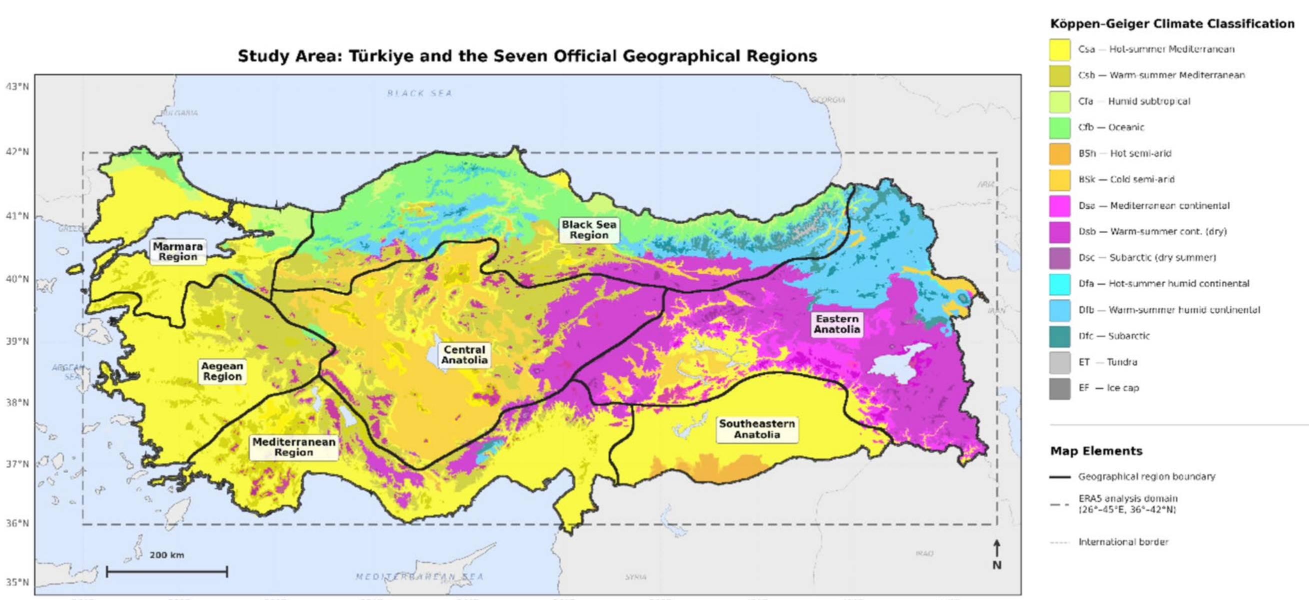


Fig. 1 Study area map of Türkiye and its seven official geographical regions. The Köppen–Geiger climate classification background is reproduced from Taşoğlu et al. (2024), with additional overlays showing the official regional boundaries and the ERA5 analysis domain (26°–45°E, 36°–42°N)

been consistently identified as one of the most prominent climate change hotspots globally, with observed warming rates approximately 20% above the global mean (Lionello and Scarascia 2018; Zittis et al. 2022 a).

Türkiye exhibits extraordinary topographical and climatological complexity within its approximately 783,562 km² territory. Elevations range from sea level along the extensive coastlines to over 5,000 m at Mount Ağrı (Ararat) in the east. This topographic heterogeneity generates a mosaic of distinct climate regimes:

The Black Sea Region is characterized by a humid oceanic climate (Köppen Cfb/Cfa) with persistent cloud cover, year-round precipitation, and moderate temperature amplitudes; the Pontic Mountains create an orographic barrier isolating this region from the continental interior. The Marmara Region exhibits a transitional climate between Mediterranean and oceanic influences, moderated by the Sea of Marmara and the proximity to both the Black Sea and the Aegean. The Aegean Region has a Mediterranean climate (Köppen Csa) with hot, dry summers and mild, wet winters, characterized by strong land-sea breeze circulation patterns and significant insolation during the warm season. The Mediterranean Region displays a hot-summer Mediterranean climate (Köppen Csa) with the highest summer temperatures among coastal regions; the Taurus Mountains create a sharp climate gradient between the coastal strip and the interior plateau. Central Anatolia is dominated by a semi-arid continental steppe climate (Köppen BSk) with extreme temperature amplitudes, low humidity, and limited precipitation across the extensive plateau (average elevation approximately 1,000 m). Eastern Anatolia experiences a severe continental climate with cold winters and warm summers, influenced by high elevation (average approximately 1,500-2,000 m), significant diurnal temperature ranges, and very low atmospheric moisture content. Southeastern Anatolia has a semi-arid to arid climate (Köppen BSh/BWh) with the highest maximum temperatures in Türkiye, frequently exceeding 45 °C during summer, and is characterized by extreme atmospheric desiccation and intense solar radiation.

To capture this spatial heterogeneity with maximum precision, Türkiye was partitioned into seven official geographical regions following the classification established by the First Turkish Geography Congress of 1941 and maintained by the Turkish Statistical Institute. A ray-casting point-in-polygon algorithm was implemented to allocate each 0.25° × 0.25° grid point to its precise geographic region, preventing the boundary-bleeding artifacts inherent in bounding-box approaches where grid cells near irregular boundaries may be misclassified. This polygon-based allocation is used consistently in the regional models and in the Leave-One-Region-Out cross-validation experiment.

2.2 Climate and thermal stress data

Long-term historical climate data spanning 75 years (1950–2025) were acquired from two complementary European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis products. The target variable for thermal comfort evaluation, the Universal Thermal Climate Index (UTCI), was obtained from the ERA5-HEAT (Human thErmAl comforT) reanalysis dataset (Di Napoli et al. 2021), available through the Copernicus Climate Data Store (CDS). ERA5-HEAT provides hourly global UTCI estimates at 0.25° × 0.25° spatial resolution, computed from ERA5 atmospheric fields using the UTCI operational procedure (Bröde et al. 2012). UTCI values represent the equivalent ambient temperature of a reference environment that would elicit the same physiological response as the actual environment, accounting for air temperature, mean radiant temperature, wind speed, and humidity. For each grid point, the daily maximum UTCI was calculated from the hourly values to capture peak thermal stress conditions.

Seven fundamental climate predictor variables were extracted from the ERA5-Land reanalysis (Hersbach et al. 2020; Soci et al. 2024) at its native 0.1° × 0.1° spatial resolution and subsequently regridded to the ERA5-HEAT 0.25° × 0.25° grid using nearest-neighbor interpolation. The ERA5-HEAT lat/lon grid was treated as the reference grid; each ERA5-Land cell was matched to its nearest ERA5-HEAT cell centre. No spatial smoothing was applied during the regridding step. Because both ERA5-HEAT and ERA5-Land are derived from the same underlying ERA5 atmospheric reanalysis, no inter-product bias correction was applied. The predictor variables and their derivations are detailed in Table 1.

Relative Humidity (RH) was computed from the 2-meter air temperature (T2M) and 2-meter dew point temperature (D2M) using the Magnus-Tetens approximation:

$$e_s = 6.112 \times \exp\left(\frac{17.67 \times T_{2m}}{T_{2m} + 243.5}\right)$$

$$e_a = 6.112 \times \exp\left(\frac{17.67 \times T_d}{T_d + 243.5}\right)$$

$$RH = 100 \times \frac{e_a}{e_s}$$

where e_s is the saturation vapor pressure (hPa) and e_a is the actual vapor pressure (hPa).

Vapor Pressure Deficit (VPD), representing the atmospheric moisture demand and a critical determinant of human evaporative cooling capacity, was calculated as:

Table 1 Climate predictor variables used in the machine learning framework

Variable	Abbreviation	Unit	Source	Temporal Aggregation
2-meter Air Temperature	T2M	°C	ERA5-Land (t2m)	Daily maximum
Relative Humidity	RH	%	Derived from T2M and D2M	Daily mean
Vapor Pressure Deficit	VPD	kPa	Derived from T2M and D2M	Daily mean
Wind Speed	Wind	m s^{-1}	Derived from U10 and V10	Daily mean
Surface Solar Radiation Downwards	SSRD	MJ/m²	ERA5-Land (ssrd)	Daily sum
Surface Thermal Radiation Downwards	STRD	MJ/m²	ERA5-Land (strd)	Daily sum
Specific Humidity	Q	g/kg	Derived from D2M and surface pressure	Daily mean

$$VPD = \frac{e_s - e_a}{10}\ [\text{kPa}]$$

using saturation and actual vapor pressures derived via the Magnus formula:

$$e_s = 0.6108 \times \exp\left(\frac{17.27 \times T_{2m}}{T_{2m} + 237.3}\right)$$

$$e_a = 0.6108 \times \exp\left(\frac{17.27 \times T_d}{T_d + 237.3}\right)$$

Saturation and actual vapour pressure were computed using the FAO Penman-Monteith form of the Magnus equation (Allen et al. 1998), applied consistently for all derived humidity variables (RH, VPD, and q).

Wind Speed was computed as the resultant of the 10-meter eastward (U10) and northward (V10) wind components:

$$\text{Wind} = \sqrt{U_{10}^2 + V_{10}^2}$$

The temporal domain was restricted to the extended warm season (April-October) to focus on heat stress conditions, as thermal discomfort from heat in the study region is negligible during the November-March period.

2.3 Data preprocessing and stratification

The preprocessing pipeline consisted of five sequential stages. First, data were restricted to the extended warm season months (April through October, inclusive) to isolate the period relevant for heat stress assessment. Second, ERA5-Land variables at 0.1° × 0.1° resolution were regridded to the ERA5-HEAT 0.25° × 0.25° grid using nearest-neighbor interpolation to ensure spatial consistency between predictor and target variables. Third, for each 0.25° × 0.25° grid cell, hourly UTCI values were aggregated to daily maximum UTCI (representing peak thermal stress), while predictor variables were aggregated as daily maxima (T2M), daily means (RH, VPD, Wind, Q), or daily sums converted to MJ/m² (SSRD, STRD). Fourth, to strictly analyze heat stress mechanisms rather than general thermal conditions, the dataset was filtered to retain only grid-day observations where the daily maximum UTCI exceeded 26.0 °C, corresponding to the onset of "moderate heat stress" as defined by the UTCI assessment scale (Bröde et al. 2012); this threshold-based filtering ensures that the machine learning model and subsequent SHAP analysis are trained exclusively on physiologically relevant heat stress conditions. Fifth, the continuous UTCI values were classified into four established heat stress severity bands following the standard UTCI assessment scale: moderate heat stress (26–32 °C), strong heat stress (32–38 °C), very strong heat stress (38–46 °C), and extreme heat stress (>46 °C). The final preprocessed dataset comprised 826,799 grid-day observations across the 75-year study period, stored as an optimized Apache Parquet feature matrix for computational efficiency.

After temporal filtering, daily aggregation, and feature derivation, the preprocessed feature matrix contained no missing values. The standard listwise call in the preprocessing pipeline therefore retained all 826,799 rows; no rows were removed at this step. We report this explicitly because the absence of missing data is a non-trivial property of the analysis dataset and removes one potential source of selection bias.

2.4 Machine learning framework

2.4.1 Model selection and architecture

A Random Forest (RF) Regressor (Breiman 2001) was selected as the predictive model for several methodologically motivated reasons: (i) RF models handle nonlinear relationships and high-order feature interactions without requiring explicit specification; (ii) they are robust to multicollinearity among predictors, which is expected given the thermodynamic interdependencies among climate variables; (iii) they provide built-in out-of-bag error estimation; and (iv) the TreeExplainer algorithm enables exact, computationally efficient SHAP value computation for tree-based ensembles (Lundberg et al. 2020). Because the target variable (UTCI) is itself derived from meteorological inputs conceptually related to the predictor set, the RF model is used here primarily as a stable, interpretable decomposition framework rather than as an end in itself for independent prediction.

We make this framing explicit because the structural relationship between UTCI and our predictor set has a quantitative consequence: a high test R^2 should not be interpreted as evidence of an independent prediction task in which the model has discovered a previously hidden mapping. UTCI is computed hourly from temperature, mean radiant temperature, humidity, and wind via the multi-node UTCI-Fiala operational procedure (Fiala et al. 2012; Bröde et al. 2012). Our model is trained on daily aggregates of variables that overlap conceptually with UTCI inputs, so a substantial part of the model's R^2 is structurally guaranteed. The scientific value of the framework therefore resides in the attribution-decomposition of UTCI variance into per-predictor SHAP contributions, not in the predictive skill metric itself. Throughout the Results section, R^2 and RMSE are reported as evidence that the fitted model is sufficiently stable to support attribution-focused analysis (Sect. 3.2) and not as the principal contribution of the study.

The RF model was configured with the following hyperparameters (Table 2), selected through preliminary cross-validation experiments.

2.4.2 Training strategy

Three categories of models were independently trained and evaluated. A global model was trained on the entire Türkiye dataset ($n = 826{,}799$) to capture country-wide thermal comfort dynamics. Seven regional models were trained on geographically subsetted data, with each grid point allocated to its region using the polygon-based classification described in Sect. 2.1; regional sample sizes ranged from 59,139 (Marmara) to 161,403 (Central Anatolia). Severity-stratified driver importance was obtained by evaluating the SHAP attributions of the single global model on the subset of the test partition belonging to each UTCI severity band, holding the model, training data and SHAP procedure constant across bands. This single-model conditional approach removes the response-range scaling artefact that would otherwise be introduced by training separate models per band (see Sect. 3.4 and Discussion 4.3).

For each model, the data were split into training (80%) and testing (20%) sets using stratified random sampling with a fixed random seed (42) for reproducibility. Model performance was evaluated using the Coefficient of Determination (R^2) and Root Mean Square Error (RMSE) computed on the held-out test set.

2.4.2.1 Sensitivity to blocked cross-validation Random train/test splitting under spatial-temporal autocorrelation has been shown to overstate held-out skill in gridded reanalysis applications because nearby grid-day observations are not independent. We therefore replicated the analysis under two structurally different blocking strategies and reported (i) held-out R^2 for each fold and (ii) Kendall's coefficient of concordance W on the SHAP-derived feature rankings across all folds. The blocked-CV results are summarised in Sect. 3.2 and reported per-fold in Supplementary Tables S1–S2 and Supplementary Figure S2.

Leave-One-Region-Out (LORO) spatial blocking. Using the TopoJSON polygon boundaries from Sect. 2.1, the data were partitioned into seven non-overlapping regional folds. For each fold, the model was retrained on the six in-set regions ($n \approx 660{,}000$–$770{,}000$) and tested on the held-out region ($n \approx 59{,}000$–$161{,}000$). The SHAP TreeExplainer (tree_path_dependent mode, see Sect. 2.5.1) was then applied to a fixed 1,500-row random sample of the held-out region.

Contiguous temporal blocking (5 folds). Five contiguous 15-year blocks (1950–1964, 1965–1979, 1980–1994, 1995–2009, 2010–2025) were defined. For each fold, the model was retrained on four blocks and tested on the held-out block ($n \approx 160{,}000$–$175{,}000$). The same SHAP procedure was applied to a fixed 1,500-row random sample of the held-out block.

For both blocking strategies, model hyperparameters, random seed, and the SHAP sampling procedure were held identical to the global model. The held-out R^2 values and the SHAP ranking concordance metrics are reported in Sect. 3.2.

Table 2 Random Forest Regressor hyperparameter configuration

Parameter	Value	Justification
Number of estimators	500	Sufficient for convergence with diminishing returns beyond this threshold
Maximum depth	12	Balances model complexity against overfitting risk
Minimum samples per leaf	20	Prevents overly specific terminal nodes
Minimum samples per split	40	Ensures statistically meaningful splits
Maximum features per split	sqrt(p)	Standard heuristic for regression forests
Random state	42	Ensures reproducibility

2.5 Explainable AI (XAI) framework

2.5.1 SHAP value computation

To overcome the opacity of the RF model and extract mechanistic insights, SHapley Additive exPlanations (SHAP; Lundberg and Lee 2017) were integrated at multiple analytical scales. SHAP values are grounded in Shapley values from cooperative game theory, where each feature is treated as a "player" in a coalition, and its contribution to the prediction is computed as the weighted average of its marginal contributions across all possible feature coalitions. For a prediction $f(x)$, the SHAP value ϕ_j for feature j satisfies the additive decomposition:

$$f(x) = \phi_0 + \sum{}_{j=1}^{p} \phi_j(x)$$

where ϕ_0 is the expected model output (base value) and $\phi_j(x)$ represents the contribution of feature j to the deviation from the base value for instance x.

The TreeExplainer algorithm (Lundberg et al. 2020) was used to compute exact SHAP values in polynomial time for tree-based models. Feature importance was quantified as the mean absolute SHAP value $((|\acute{\phi}|))$ across all instances, representing the average magnitude of each feature's contribution to the predicted UTCI, expressed in degrees Celsius. These attributions should be interpreted as model-based contribution magnitudes within the fitted predictor space, not as direct estimates of independent physical effect size.

SHAP perturbation mode. TreeExplainer supports two complementary perturbation strategies for handling "missing" features during the Shapley computation: `tree_path_dependent` (the default), which uses the conditional expectation of the tree path given the joint distribution actually observed in the training data, and `interventional`, which uses an independent background distribution and therefore breaks correlations among features. The two modes answer different questions: `tree_path_dependent` returns the SHAP attribution under the empirical climate joint distribution in Türkiye (which is the appropriate framing for our sensitivity-decomposition objective), while `interventional` returns the marginal counterfactual attribution under feature independence. We report results from `tree_path_dependent` as the primary analysis and use `interventional` as a robustness check against the possibility that observed feature correlations inflate certain attributions. The two modes return a perfect Kendall rank concordance (τ = +1.000, $p = 4 \times 10^{-4}$), with magnitudes shifting in directions consistent with the documented correlation structure (Sect. 3.3.1).

2.5.2 Multi-scale SHAP analysis

SHAP analysis was conducted at four complementary scales. At the global scale, SHAP summary plots (beeswarm plots) for the Whole Türkiye model were generated, revealing overall feature importance and the direction of feature effects. At the regional scale, independent SHAP analyses for each of the seven geographic regions identified the top three climate drivers per region and their relative rankings. At the severity-stratified scale, SHAP importance matrices were computed by subsetting the single global model's SHAP values according to UTCI severity band and visualized as heatmaps to reveal how driver importance shifts with intensifying heat stress. At the pixel-wise scale, for every distinct geographic grid cell in the dataset, the mean climate conditions during heat events were computed and SHAP values were calculated using the global model; the dominant secondary driver (the highest-importance feature excluding T2M) was identified for each pixel, producing a high-resolution 0.25° × 0.25° spatial map of driver dominance across Türkiye.

Justification of mean-condition pixel-wise SHAP. Computing SHAP at the pixel's mean heat-event meteorology anchors the attribution at a single, climatologically representative reference state per pixel. This choice has three properties: (i) it is consistent across pixels (each pixel is summarized at its own typical event), (ii) it is computationally tractable at the 0.25° × 0.25° resolution across the entire Türkiye bounding box, and (iii) it preserves the climatological structure of the underlying data without introducing day-level noise. We additionally provide a per-day SHAP aggregation (Sect. 3.6.2) that recomputes SHAP at 20 randomly sampled heat-event days per pixel, reports the modal driver and its stability, and quantifies the agreement between the mean-condition and per-day classifications. This robustness check directly addresses the concern that mean-condition SHAP may suppress day-to-day variability.

2.5.3 SHAP dependence analysis

SHAP dependence plots were generated for key features (T2M and RH) to examine nonlinear interaction effects. Each dependence plot displays the SHAP value of a feature against its actual value, with points colored by a secondary interacting feature, revealing how the effect of one climate driver on UTCI is modulated by another.

2.6 Trend analysis

Annual counts of severity-band grid-days were computed for each of the three severity bands (Strong Heat 32–38 °C, Very Strong Heat 38–46 °C, Extreme Heat > 46 °C) over the 75-year window. Two weightings were reported: (i) raw grid-day counts (used in Sect. 3.1) and (ii) cell-area-weighted counts via Σ cos(latitude), reported as an area-weighted sensitivity. Trend significance was evaluated using the classical Mann-Kendall test (Mann 1945; Kendall 1948) and the Hamed-Rao modification (Hamed and Rao 1998) that adjusts the variance of S for first-order serial autocorrelation. Theil-Sen non-parametric slopes (Sen 1968) with rank-based 95% confidence intervals were used as the companion slope estimator. All trend statistics were computed using the `pymannkendall` Python package (Hussain and Mahmud, 2019).

3 Results

3.1 Temporal trends of severity classes (1950–2025)

The 75-year longitudinal analysis reveals a pronounced intensification of severe heat stress events across Türkiye (Fig. 2). We report trends for all three heat-stress categories (Strong (32–38 °C), Very Strong (38–46 °C), and Extreme (>46 °C)) using Theil-Sen slopes as the rank-based companion to the Mann-Kendall test, and additionally using the Hamed-Rao modification (Sect. 2.6) that adjusts the variance for first-order serial autocorrelation. Both raw grid-day counts, and cell-area-weighted (Σ cos latitude) counts were evaluated; the area-weighted slopes are approximately 20–25% smaller in magnitude but do not alter trend directions or significance verdicts.

Annual counts of Strong Heat (32–38 °C) grid-days exhibit the largest absolute trend (Theil-Sen slope + 6.68 grid-days yr^{-1}, 95% CI [+2.70, +11.14]; classical Mann-Kendall $\tau = +0.256$, $p = 1.1 \times 10^{-3}$; Hamed-Rao modified MK $p = 1.1 \times 10^{-3}$), indicating that the warming signal is not confined to the upper-tail categories. Very Strong Heat (38–46 °C) grid-days exhibit a significant upward trend (Theil-Sen slope + 5.00 grid-days yr^{-1}, 95% CI [+2.96, +7.15]; classical MK $\tau = +0.336$, $p = 1.7 \times 10^{-5}$; Hamed-Rao modified MK $p = 0.012$); the very strong trend remains significant at $\alpha = 0.05$ under the autocorrelation-robust test, although the modified-MK p-value is more conservative than the classical one. By the most recent decade (2015–2025), very strong heat grid-days regularly exceed 1,000 per year, compared to a baseline of approximately 500–700 during 1950–1970. Extreme Heat (>46 °C) grid-days, while less frequent in absolute terms, show a positive Theil-Sen slope of + 0.62 grid-days yr^{-1}, with a 95% CI [−0.06, + 1.39] that includes zero; the trend does not reach conventional

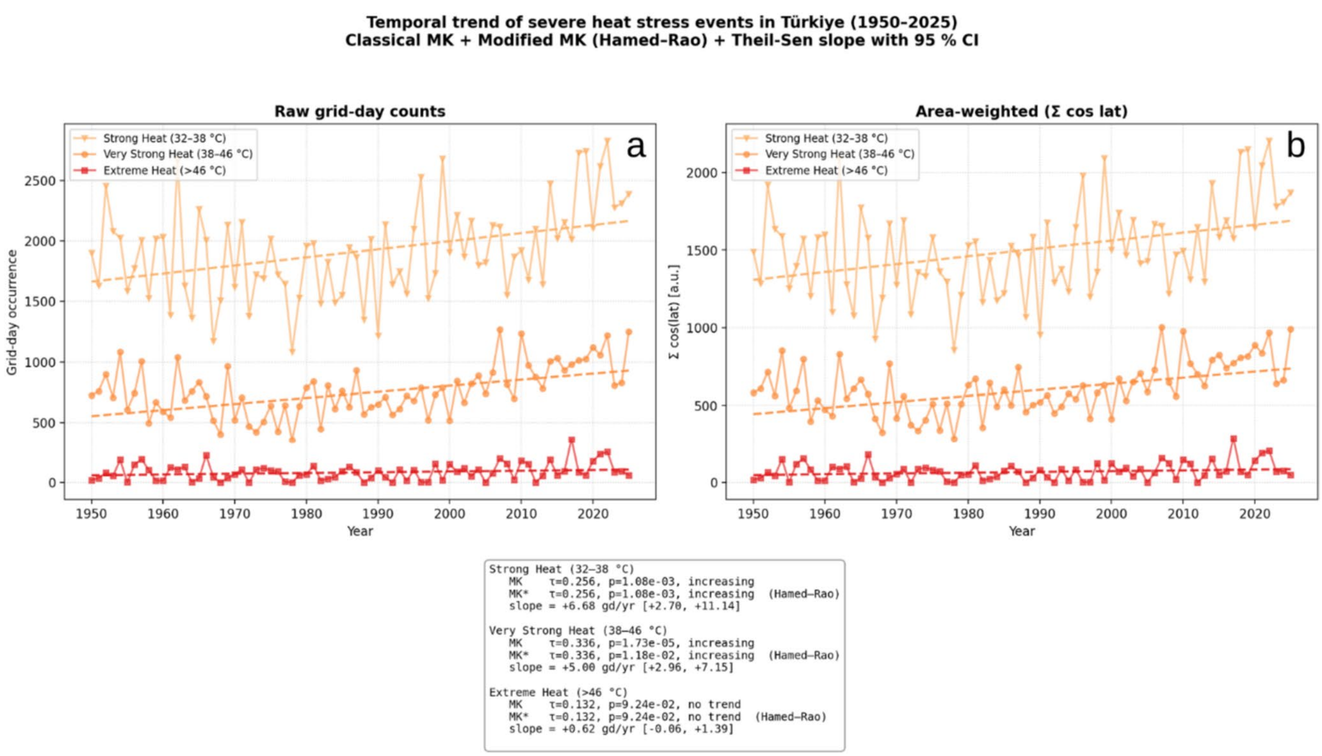


Fig. 2 Temporal trend of severe heat stress events in Türkiye (1950–2025). Two panels: (**a**) raw grid-day counts; (**b**) area-weighted (Σ cos lat) counts. Three series per panel: Strong Heat (yellow triangles), very strong heat (orange circles), Extreme Heat (red squares). Dashed lines are Theil-Sen non-parametric trend fits. Annotation block reports the classical Mann-Kendall test (τ, p), the Hamed-Rao modified MK (τ, p), and the Theil-Sen slope with 95% CI for every series

significance under either the classical or the Hamed-Rao modified Mann-Kendall test ($p = 0.092$). These events were exceedingly rare during 1950–1970 (typically 0–50 grid-days/year) but have become more common since the mid-1990s, with individual years recording 100–350 extreme heat grid-days. These trends are consistent with broader Eastern Mediterranean warming patterns documented in the IPCC AR6 (Ali et al. 2022), long-term UTCI trend analyses for the MENA region showing increases of 0.1–0.7 °C per decade (Hamed et al. 2023), and regional projections of further intensification of summer heat extremes (Zittis et al. 2016; Ntoumos et al. 2022).

3.2 Global model performance and validation

The Random Forest model trained on the full Türkiye dataset ($n = 826{,}799$) achieved high internal consistency, with $R^2 = 0.936$ and RMSE = 1.80 °C on the held-out test set (20% of data; Fig. 3). As discussed in Sect. 2.4.1, because UTCI is itself derived from meteorological variables that overlap with the predictor set, this skill metric is partly structural and is interpreted as evidence that the fitted framework is sufficiently stable to support attribution-focused analysis, not as the principal contribution of the study. The observed-versus-predicted density scatter plot reveals tight clustering along the 1:1 diagonal across the UTCI spectrum, from moderate heat stress, with negligible systematic bias.

The hexbin density visualization demonstrates that the highest prediction densities are concentrated in the 26–35 °C range, corresponding to the most frequently

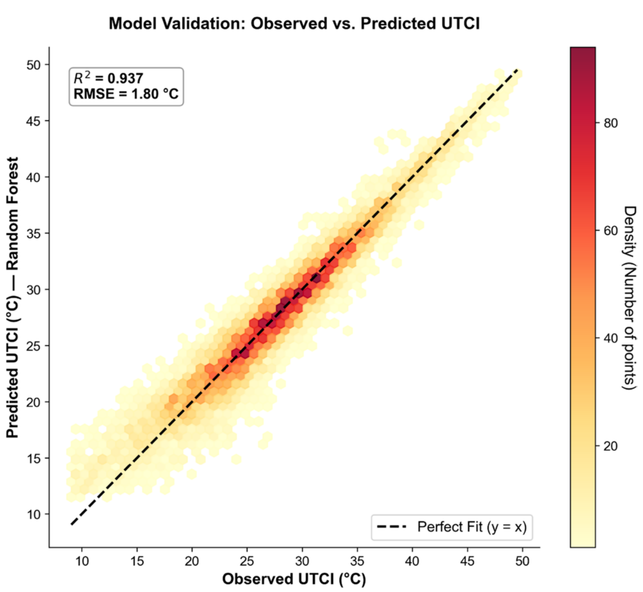


Fig. 3 Model validation scatter plot: Observed versus predicted UTCI for the global (Whole Türkiye) random forest model. Color intensity (hexbin density) represents the number of data points in each cell. The dashed black line indicates perfect prediction (y = x). R^2 = 0.936 and RMSE = 1.80 °C on the held-out test set ($n \approx 165{,}000$)

observed UTCI conditions in the dataset. Model performance remains robust at the extreme tails of the distribution (UTCI > 40 °C), where accurate predictions are most consequential for public health applications. The slight dispersion at the highest UTCI values (> 45 °C) reflects both the inherent stochasticity of extreme events and the smaller sample sizes in this severity band, but the overall predictive envelope remains well-calibrated.

A direct test of whether the global R^2 is inflated by a small subset of well-fit pixels was conducted by computing per-pixel R^2 across 1,280 Türkiye grid cells with ≥ 5 heat-event days in the held-out test partition. The median per-pixel R^2 is 0.935, essentially identical to the global R^2 of 0.936; the 10th and 90th percentiles are 0.860 and 0.968 respectively, and 76% of pixels achieve $R^2 > 0.90$. The global skill metric is therefore representative of the typical pixel rather than driven by averaging across a heterogeneous skill landscape.

The residual variance at the pixel level correlates weakly but significantly with the climatological mean SSRD at the pixel (Spearman $\rho = -0.10$, $p = 3 \times 10^{-4}$): pixels with persistently low SSRD (predominantly the cloudy Black Sea coastline) carry the largest residuals, consistent with daily-aggregated SSRD and STRD predictors losing the sub-daily cloud-cover variability that the hourly UTCI calculation does sample. The 6.4% of variance left unexplained by the global model is therefore consistent with a temporal aggregation gap between hourly-computed UTCI and daily-aggregated predictors rather than with missing physical inputs. This finding is reported in Supplementary Figure S1 and is discussed further in Sect. 4.7.1.

The 80/20 random split that produced the $R^2 = 0.936$ estimate above was supplemented by Leave-One-Region-Out spatial blocking (7 folds) and contiguous 15-year temporal blocking (5 folds) as described in Sect. 2.4.2.1. Held-out R^2 ranged from 0.86 (Black Sea, the most climatologically heterogeneous fold) to 0.96 (Southeastern Anatolia) across the seven spatial folds, and from 0.92 to 0.94 across the five temporal folds. The temporal range is essentially identical to the random-split R^2 of 0.936, indicating that the random split does not appreciably inflate the skill estimate. More importantly for attribution analysis, the SHAP feature ranking returned a Kendall's coefficient of concordance $W = 0.943$ (spatial folds; $p = 5.4 \times 10^{-7}$), $W = 1.000$ (temporal folds; $p = 3.9 \times 10^{-5}$, identical top-7 ranking in every temporal block), and $W = 0.963$ across all twelve folds combined ($p = 5.5 \times 10^{-13}$). T2M is rank 1 in 12 of 12 folds; VPD is rank 2 in 12 of 12 folds. The headline ranking T2M > VPD > Wind > STRD > RH > Q > SSRD is therefore robust to the blocking strategy. Per-fold metrics are reported in Supplementary Tables S1 (spatial) and S2 (temporal), with the rank heatmap and pairwise Kendall τ matrix shown in Supplementary Figure S2.

3.3 Regional model performance and driver rankings

Regional models demonstrated consistently high predictive accuracy across all seven geographic regions, with R^2 values ranging from 0.882 to 0.964 (Fig. 4; Table 3). The inter-regional variation in model performance provides meaningful insights into the climatological complexity of each region.

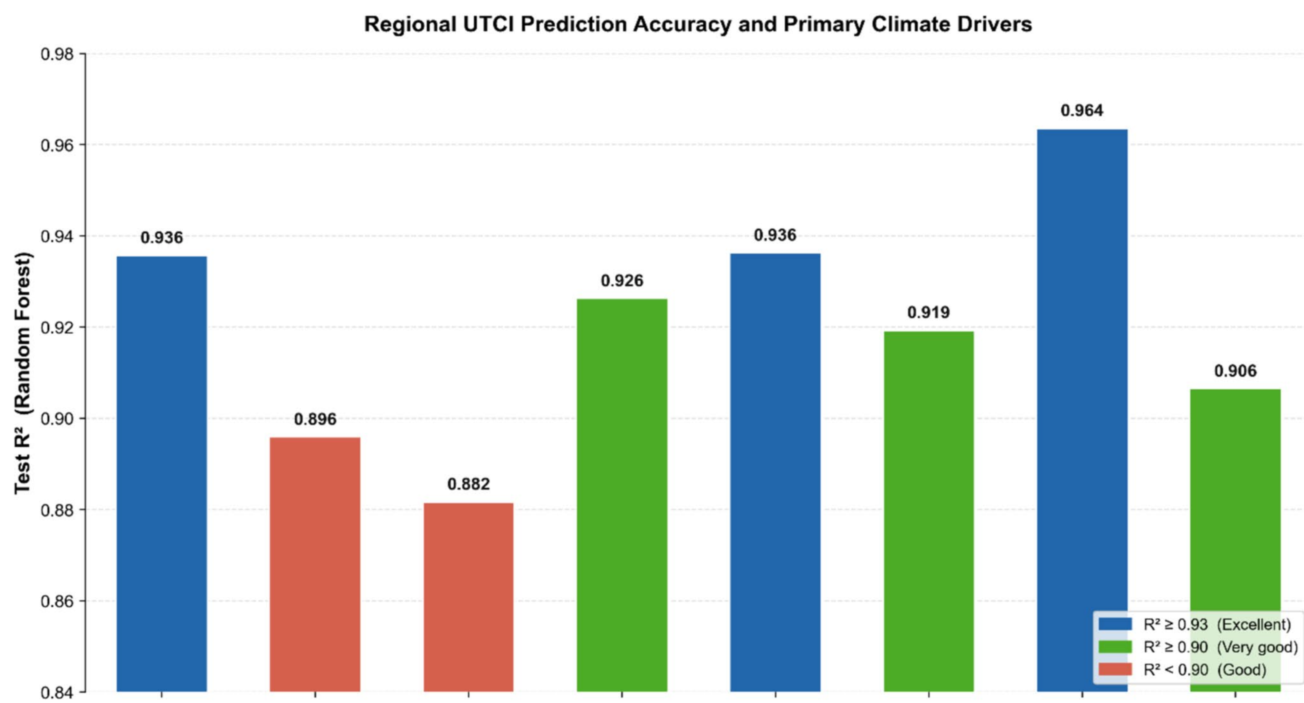
Southeastern Anatolia achieved the highest model accuracy ($R^2 = 0.964$), reflecting the region's relatively homogeneous semi-arid climate where thermal dynamics are governed by a consistent set of drivers (T2M, VPD, RH). The limited cloud cover and moisture variability in this region produce highly predictable UTCI responses to the input climate variables.

The Black Sea Region yielded the lowest, yet still robust, model accuracy ($R^2 = 0.882$). This reduced performance is attributable to the region's exceptional climatological complexity: persistent cloud cover creates highly variable radiation conditions, orographic effects modulate wind patterns unpredictably, and the persistent maritime moisture supply generates complex humidity-temperature interactions that challenge even nonlinear models. The emergence of Specific Humidity (Q) as the third-ranked driver which is unique among all regions, reflects the fundamental role of absolute atmospheric moisture content in determining thermal comfort along this persistently humid coastline.

The Marmara Region is the only region where Wind Speed displaces VPD as the second-most-important driver, reflecting the strong influence of the Bosphorus and Dardanelles

Table 3 Regional model performance metrics and SHAP-derived climate driver rankings

Region	*N* Samples	R^2	1 st Driver	2nd Driver	3rd Driver
Whole Türkiye	826,799	0.936	T2M	VPD	Wind
Southeastern Anatolia	65,274	0.964	T2M	VPD	RH
Aegean Region	75,527	0.936	T2M	VPD	Wind
Marmara Region	59,139	0.926	T2M	Wind	VPD
Central Anatolia	161,403	0.919	T2M	VPD	Wind
Mediterranean Region	76,708	0.906	T2M	VPD	STRD
Eastern Anatolia	133,709	0.896	T2M	VPD	Wind
Black Sea Region	101,529	0.882	T2M	VPD	Q



	Whole Türkiye	Eastern Anatolia	Black Sea Region	Marmara Region	Aegean Region	Central Anatolia	Southeastern Anatolia	Mediterranean Region
1st Driver	T2M	T2M	T2M	T2M	T2M	T2M	T2M	T2M
2nd Driver	VPD	VPD	VPD	Wind	VPD	VPD	VPD	VPD
3rd Driver	Wind	Wind	Q	VPD	Wind	Wind	RH	STRD

Fig. 4 Regional UTCI prediction accuracy (R^2) and SHAP-derived primary, secondary, and tertiary climate drivers for each of Türkiye's seven geographical regions. Bar colors indicate performance tiers: blue ($R^2 \geq 0.93$), green ($R^2 \geq 0.90$), and red ($R^2 < 0.90$). The table below the chart lists the top three climate drivers ranked by mean absolute SHAP value for each region

strait winds, combined with the moderating effects of the Sea of Marmara on the regional thermal environment.

The Mediterranean Region is distinguished by STRD (Thermal Radiation) as the third-ranked driver, consistent with the intense surface heating, clear-sky conditions, and high longwave radiative forcing that characterize the southern Turkish coast during summer.

3.3.1 Robustness of feature ranking to SHAP perturbation mode

A complementary test of ranking robustness was performed by recomputing SHAP under both `tree_path_dependent` (the default mode used throughout this manuscript) and `interventional` perturbation modes on the same global-model test sample (Sect. 2.5.1). The two modes give a perfect Kendall rank concordance ($\tau = +1.000$, $p = 4 \times 10^{-4}$): every feature retains its rank under both modes. The absolute magnitudes shift in interpretable directions: the top three features (T2M, VPD, Wind) gain attribution under interventional (+4% to +9%), while STRD, RH, Q, and SSRD lose attribution (−15% to −38%). These shifts are consistent with the lower-ranked features sharing attribution along correlation channels (notably the T2M↔VPD correlation r=+0.901, the RH↔VPD correlation $r = -0.807$, and the STRD↔Q correlation r=+0.759) and lend further weight to the conclusion that the ranking hierarchy is mathematically robust to feature-correlation handling. The full comparison is provided in Supplementary Tables S3a and S3b and Figure S3.

3.4 Severity-stratified driver importance

The SHAP-derived feature importance matrix, stratified across the four UTCI severity bands, shows that attribution changes systematically of climate driver importance as heat stress intensifies (Fig. 5). These SHAP values quantify model-attributed importance, namely the contribution of each predictor to the fitted Random Forest's predictions, rather than direct physical effect sizes.

Severity-stratified importance was computed by evaluating the SHAP attributions of the single global model on the subset of the test partition belonging to each UTCI severity band, holding the model, training data, and SHAP procedure constant across bands. For each band we report both the absolute mean |SHAP| (°C) and the normalized SHAP share (the percentage of total |SHAP| within the band). The share metric is invariant to the band-to-band differences in target range and is therefore the appropriate basis for cross-band comparison; both metrics are shown in Fig. 5 and tabulated in Supplementary Tables S4a and S4b.

Three patterns emerge from this severity-stratified decomposition. First, the moisture-demand dimension (represented by VPD) intensifies as heat stress increases: its share of attribution grows monotonically from 18.7% at moderate heat stress (26–32 °C) to 28.7% at very strong heat stress (38–46 °C), a gain of 10.0% points, while its absolute mean |SHAP| grows from 0.93 to 4.17 °C across the same range. Second, the wind-speed share collapses from 17.1% at moderate heat stress to 4.2% at very strong heat stress and 3.4% at extreme heat, a relative reduction

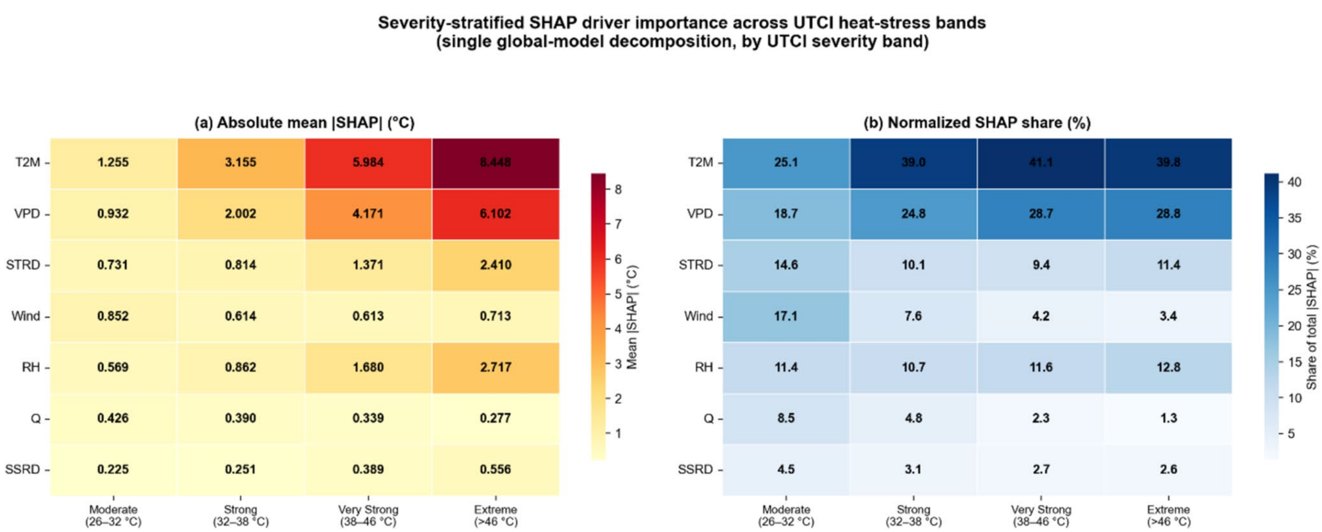


Fig. 5 Severity-stratified SHAP driver importance across UTCI heat-stress bands using the single global-model decomposition. Panel (**a**) shows absolute mean |SHAP| values (°C), while panel (**b**) shows normalized SHAP shares (% of total |SHAP| within each severity band). T2M and VPD increase both in absolute magnitude and attribution share from moderate to very strong heat stress, whereas wind speed collapses in share. STRD increases in absolute magnitude but decreases in relative share, indicating that radiative attribution does not intensify proportionally with severity

of approximately 80%. Third, the radiative contribution (STRD) does not intensify in relative terms: although its absolute mean |SHAP| grows from 0.73 to 1.37 °C, its share of attribution declines slightly from 14.6% to 9.4% between moderate and very strong heat stress, because the contributions of T2M and VPD grow faster. Relative humidity (RH) maintains a roughly constant share (11–13%) and consequently rises to the third-ranked position at strong, very strong, and extreme heat stress, displacing wind speed. Air temperature (T2M) remains the dominant driver throughout, its share rising from 25.1% to 41.1% between moderate and very strong heat stress. Taken together, these patterns indicate that the transition into severe heat stress is characterized by an enlarged moisture-demand contribution and a collapsed wind contribution, while the radiative contribution remains a steady rather than escalating term.

The collapse of the wind-speed contribution at higher severities has a clear thermophysiological interpretation. At moderate heat stress, convective and evaporative cooling via wind provides meaningful mitigation; under extreme conditions where air temperature approaches or exceeds skin temperature (~35 °C), however, ventilation may no longer be uniformly beneficial: its physiological effect depends on the local humidity regime, the exposure conditions, and the timing of the heat event, and at sufficiently extreme combinations of dry-bulb temperature and humidity, sensible-heat exchange driven by wind can transition from a cooling to a heating mechanism. This finding aligns with the experimental work of Bröde and Kampmann (2023), who demonstrated that wind reduces physiological strain below 35 °C but increases it in hot-dry conditions above this threshold.

The intensification of the VPD share at very strong heat stress demonstrates that entry into physiologically dangerous thermal zones is associated with an enlarged moisture-demand contribution, not by raw temperature increments alone. The relative contribution of the radiative cluster (SSRD + STRD jointly; see Sect. 3.4.2) decreases from moderate to very strong heat stress (19.1% to 12.1%; Supplementary Table S4b). Thus, radiation remains a contributing term, but it does not show the severity-amplifying pattern observed for the moisture-demand dimension.

3.4.1 Feature ablation and the moisture-demand dimension

To test whether the apparent dominance of VPD is robust to the documented inter-correlations among moisture predictors (T2M ↔ VPD r = +0.901; RH ↔ VPD r = −0.807; STRD ↔ Q r = +0.759), we trained four reduced-feature variants of the full global model (hereafter M0) differing only in which moisture-related predictors are included:

- M1: full set without RH (n_features = 6).
- M2: full set without Q (n_features = 6).
- M3: full set without both RH and Q (only VPD as moisture proxy, n_features = 5).
- M4: full set without VPD (with RH and Q both retained, n_features = 6).

Test R^2 ranged narrowly across all five models: 0.9357 (M0), 0.9369 (M1), 0.9368 (M2), 0.9373 (M3), and 0.9369 (M4). Removing either RH or Q individually, or both jointly, does not degrade predictive skill, and M3 (with only VPD as the moisture predictor) achieves the highest R^2 of the five models. RH and Q therefore carry no independent predictive content once VPD is included; they are statistically redundant.

The SHAP ranking under each ablation is consistent with this conclusion: VPD retains rank 2 in M0, M1, M2, and M3 (every model where VPD is present). When VPD is removed (M4), RH inherits rank 2, with its absolute SHAP magnitude rising by 102% relative to M0. The "moisture-demand signal" is therefore transferable between labels: drop VPD → RH inherits; drop RH and Q → VPD absorbs (its share grows from 9.7% in M0 to 24.8% in M3, a + 156% relative increase).

We interpret these results together as evidence that the model is fitting a single moisture-demand dimension, of which VPD is the most informative single proxy but for which RH and Q are interchangeable representations. Accordingly, the manuscript refers to the second-ranked driver throughout as the "moisture-demand dimension" (with VPD as the representative variable) rather than as "VPD dominance" per se. The ablation summary (R^2, RMSE, and top-5 SHAP rankings for each model) is reported in Supplementary Table S5, with diagnostic plots in Supplementary Figure S4.

3.4.2 Radiation cluster reporting

Because SSRD and STRD are both radiative inputs into the UTCI computation and share a common atmospheric-radiative state, we additionally report the combined "Radiation" cluster (SSRD + STRD) as a single dimension alongside the individual rankings. Under the global model, the cluster's mean |SHAP| is 1.175 °C (15.2% of total |SHAP|, ranking 3rd) under `tree_path_dependent` mode, and 0.890 °C (11.8%, ranking 4th) under `interventional` mode. Across the 12 blocked-CV folds (Sect. 2.4.2.1), the Radiation cluster outranks Wind Speed in 11 of 12 folds; the single exception is the Marmara fold, where Wind Speed displaces Radiation in third place (consistent with the regional analysis in Sect. 3.3, which identifies Wind Speed as the second-most-important driver in Marmara). The radiation-cluster framing is more parsimonious for cross-band and cross-region

SHAP attribution comparison; however, as emphasized in Sect. 4.4, SSRD (shortwave) and STRD (longwave) correspond to physically distinct adaptation pathways, so the cluster should not be read as a single intervention target. The clustered rankings are summarized in Supplementary Table S6.

3.5 Temperature versus humidity dynamics

The relationship between air temperature and humidity contributions to thermal stress was further examined through fine-resolution SHAP analysis across 2 °C UTCI bins (Fig. 6). The absolute SHAP contributions of both T2M and RH increase with UTCI intensity, but T2M contributions escalate far more steeply. At the lowest bin (26–30 °C), T2M contributes 1.10 °C while RH contributes 0.55 °C. At the highest bin (46–50 °C), T2M reaches 8.32 °C while RH reaches 2.62 °C.

The RH/T2M SHAP dominance ratio analysis (Fig. 6, right panel) reveals that temperature remains consistently dominant over relative humidity across all UTCI bins, with the ratio never exceeding 1.0. The ratio is highest at the lowest UTCI bin (0.50 at 26–30 °C) and decreases to approximately 0.26–0.32 at higher stress levels. The correct conclusion is therefore not humidity dominance, but persistent temperature primacy accompanied by a severity-dependent increase in moisture-related influence. This pattern is consistent with the finding of Simpson et al. (2023), who showed that commonly used thermal indices disagree about the relative importance of moisture precisely because this importance varies systematically across the intensity spectrum of thermal stress.

3.6 High-resolution spatial heterogeneity of climate drivers

3.6.1 Regional choropleth map

The regional-scale cartographic visualization (Fig. 7) provides a synoptic overview of the spatial distribution of tertiary climate drivers across Türkiye's seven geographic regions. With T2M universally ranking as the primary driver and VPD as the secondary driver in six of seven regions, the tertiary driver reveals the most informative spatial differentiation. Central Anatolia, the Aegean, Eastern Anatolia, and Marmara share wind as their tertiary driver, although the processes underpinning this pattern differ, reflecting convective cooling over the arid plateau and sea-breeze influences along the coasts. The Mediterranean coastline is uniquely characterized by thermal radiation (STRD) as its tertiary driver, reflecting intense surface heating, clear-sky longwave emission, and the thermal amplification effect of the Taurus Mountain barrier. Specific humidity (Q) emerges as the tertiary driver only in the Black Sea region, where persistent oceanic moisture supply and orographic precipitation maintain high absolute humidity levels throughout the warm season. In Southeastern Anatolia, relative humidity (RH) serves as the tertiary driver, reflecting the critical role of even modest moisture fluctuations in modulating extreme heat stress where baseline humidity is very low.

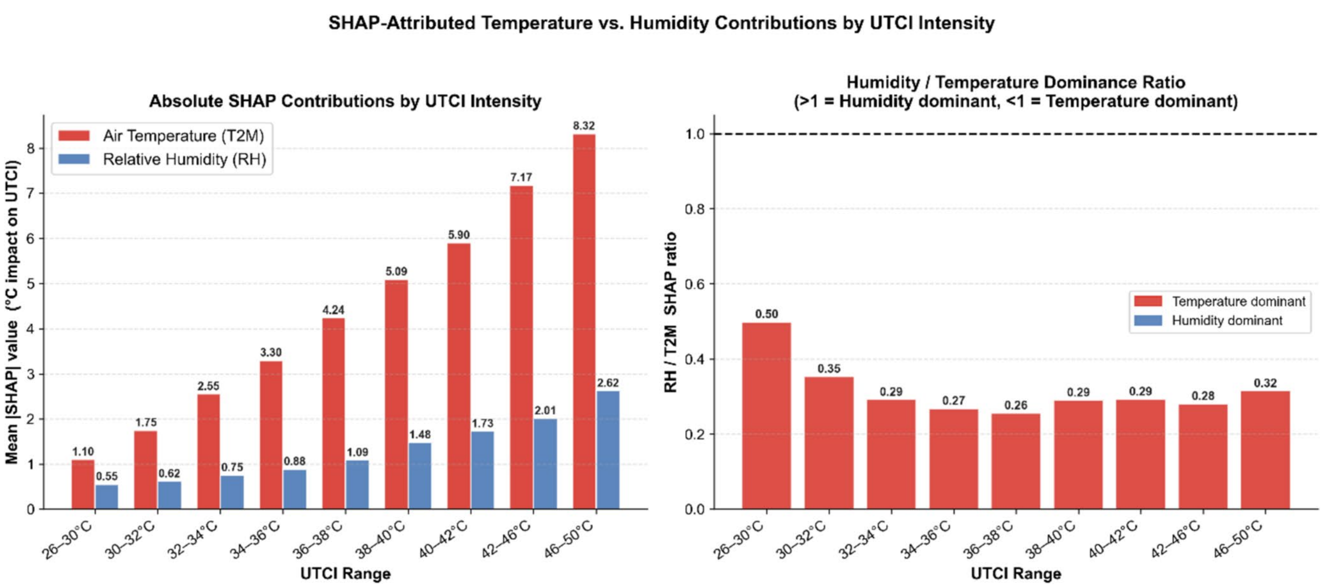


Fig. 6 Model-attributed temperature versus humidity SHAP contributions across UTCI heat-stress bands. Left panel: absolute SHAP contributions of T2M (red) and RH (blue) by 2 °C UTCI bins. Right panel: RH/T2M SHAP dominance ratio (values > 1 would indicate humidity dominance). Temperature remains dominant across all severity levels, with the ratio peaking at 0.50 in the 26–30 °C range and declining at higher stress levels

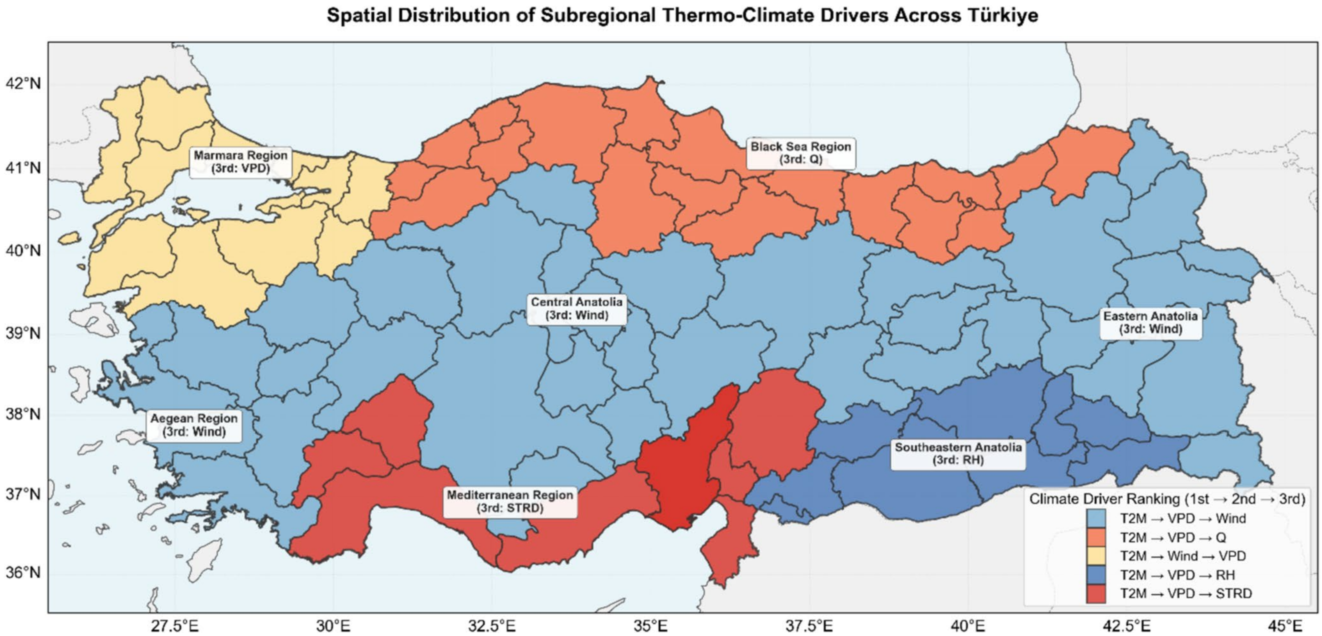


Fig. 7 Spatial distribution of subregional thermo-climate driver dominance across Türkiye (TopoJSON polygon boundaries). Regions are colored according to their top three climate driver combinations. T2M and VPD are the first and second drivers in six of seven regions; the tertiary driver captures the most informative regional differentiation

3.6.2 Pixel-wise SHAP dominance map

The 0.25° × 0.25° pixel-wise map of dominant secondary climate drivers, computed by applying SHAP analysis to the mean climate conditions of every geographic grid cell during heat events (Fig. 8), bypasses regional aggregation entirely, revealing intra-regional heterogeneity that is invisible in choropleth representations.

The pixel-wise map exposes three fundamental spatial regimes. The vast semi-arid interiors, spanning Central

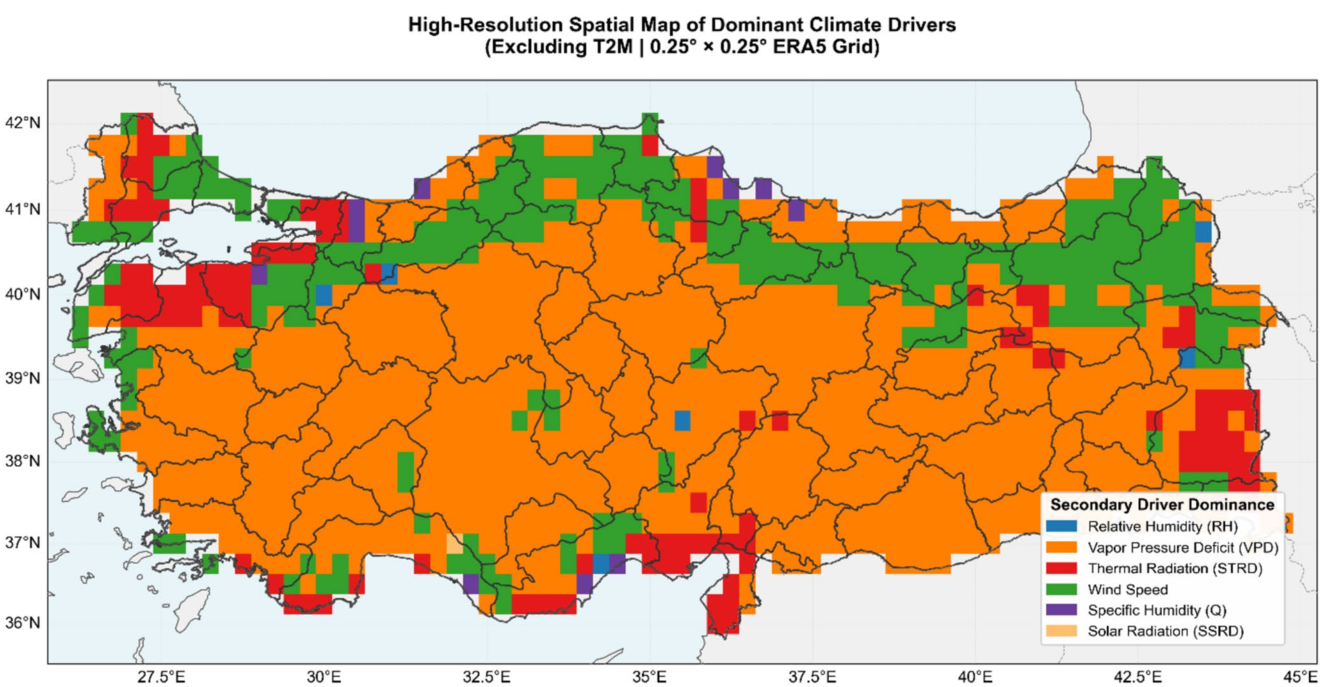


Fig. 8 High-resolution (0.25° × 0.25°) spatial map of dominant secondary climate drivers for thermal comfort across Türkiye, excluding T2M. Each pixel is colored by the climate variable with the highest mean absolute SHAP value (after T2M) based on mean heat-event conditions. Orange: VPD; Green: Wind Speed; Red: STRD; Blue: RH; Purple: Q; Yellow: SSRD

Anatolia, Eastern Anatolia, Southeastern Anatolia, and the inland portions of the Aegean and Mediterranean regions, show VPD as the dominant secondary driver. This spatial coherence reflects the thermodynamic reality that in arid and semi-arid environments, where air temperatures routinely exceed skin temperature, the atmospheric moisture deficit (VPD) determines the rate of evaporative cooling from the human body. The geographic extent of VPD dominance encompasses approximately 70% of Türkiye's land area (computed over the Türkiye-masked grid cells with at least 10 heat-event days), underscoring the pervasive influence of atmospheric desiccation on thermal comfort across the Anatolian landmass.

A distinct band of wind-dominated pixels lines the Black Sea coast and extends westward through the Marmara region. The XAI framework has autonomously identified that in environments where humidity is persistently high and temperature amplitudes are moderate (oceanic climate), sensible heat loss via convective wind cooling becomes the critical mitigating variable. The Pontic Mountains create a sharp transition: pixels immediately inland from the mountain crest shift abruptly from wind to VPD dominance, demonstrating the orographic control on climate driver regimes.

Along the Aegean and Mediterranean coastlines, STRD emerges as the dominant secondary driver in discrete pixel clusters. Because STRD here is downward longwave radiation reaching the surface (i.e. the atmospheric thermal emission term, not surface upwelling), this pattern reflects the broader coastal radiative-load environment — warm lower-tropospheric air, atmospheric humidity, cloud-cover state, and land–sea thermal contrasts that jointly raise the downwelling longwave flux during the warm season. The patchy spatial distribution of STRD-dominant pixels therefore indicates localized longwave-radiative sensitivity, which should be interpreted alongside the radiation-cluster reading in Sect. 3.4.2 rather than as direct evidence of shortwave shading control.

Isolated pixels dominated by RH, Q, or SSRD appear in transitional zones and at higher elevations, representing microclimatic niches where specific humidity gradients or solar radiation variability locally override the dominant VPD signal.

Robustness of pixel-wise SHAP via per-day aggregation. To address the concern that mean-condition SHAP may suppress day-to-day driver variability, we additionally computed per-day SHAP attributions on a random sample of 20 heat-event days per pixel for the 1,280 Türkiye grid cells (25,600 day-level SHAP evaluations). For each pixel we recorded (i) the per-day modal dominant driver (the most frequent argmax across the 20 days), (ii) the modal-driver stability percentage (=modal count/20), and (iii) the average first-vs-second |SHAP| margin gap. The per-day modal dominant driver agrees with the mean-condition Fig. 8 assignment in 79.5% of pixels, the median per-day stability is 60%, and 75.8% of pixels achieve stability ≥ 50% (32.5% achieve stability ≥ 70%). The 20.5% of pixels where the two methods disagree concentrate in two interpretable structures: STRD-mean-cond pixels with marginally elevated mean STRD reclassify as VPD-dominant when averaged across days (consistent with Sect. 3.6.4 showing that STRD-dominant pixels have only modestly elevated mean STRD), and Wind/VPD oscillations along the Black Sea boundary. The mean-condition map (Fig. 8) is therefore retained as the headline visualisation, with the per-day modal map and the stability-percentage map provided in Supplementary Figures S6 and S7 respectively, and the first-vs-second |SHAP| margin gap histogram shown in Supplementary Figure S8.

3.6.3 Regional SHAP summary analysis

Individual SHAP summary (beeswarm) plots for the national model and seven regional models are presented as a composite panel in Fig. 9. T2M dominates in every region, but the secondary and tertiary driver hierarchies diverge in ways that are consistent with the regional climatological characteristics described in Sect. 2.1.

At the national scale (Fig. 9a), VPD ranks second and Wind Speed third, with STRD and RH contributing at lower magnitudes. The Black Sea Region (Fig. 9d) presents the most distinctive profile: Specific Humidity (Q) rises to third position, unique among all regions, reflecting the persistent oceanic moisture supply, while wind speed exhibits particularly strong bidirectional effects (up to −4 °C), demonstrating the critical cooling role of maritime ventilation. The relatively compressed SHAP range across all features reflects the temperature-moderating influence of the Black Sea. Southeastern Anatolia (Fig. 9g) shows the widest T2M SHAP range (approximately −6 to +8 °C), reflecting the extreme temperature amplitudes of this semi-arid zone; RH appears as the third driver, also unique, indicating that even modest humidity fluctuations in this inherently dry environment significantly modulate thermal comfort, while wind speed clusters near zero, confirming that ventilation plays a negligible role.

The Mediterranean Region (Fig. 9h) is distinguished by STRD's elevated importance (third position), consistent with intense longwave radiative forcing from sun-heated limestone terrain. Wind speed shows a bimodal pattern in which high wind values mitigate UTCI while calm conditions amplify it, consistent with the importance of sea breeze circulation for coastal thermal comfort. The Marmara Region (Fig. 9b) is the only region where Wind Speed displaces VPD as the second-most-important driver, reflecting the unique influence of the Bosphorus and Dardanelles

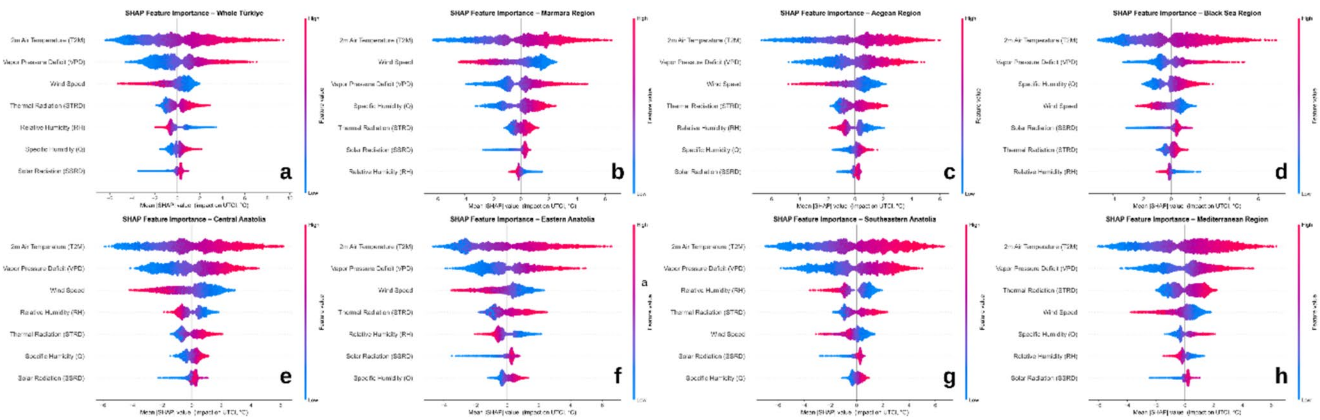


Fig. 9 Composite SHAP beeswarm plots for the national model and seven regional models. (**a**) Whole Türkiye, (**b**) Marmara Region, (**c**) Aegean Region, (**d**) Black Sea Region, (**e**) Central Anatolia, (**f**) Eastern Anatolia, (**g**) Southeastern Anatolia, (**h**) Mediterranean Region. Each point represents a single prediction; horizontal position indicates the SHAP value (impact on UTCI in °C), and color indicates the feature value (red/pink = high, blue = low). Features are ordered by mean absolute SHAP value within each panel

strait winds and the Sea of Marmara's thermal moderation. Central (Fig. 9e) and Eastern Anatolia (Fig. 9f) share structurally similar profiles dominated by T2M and VPD, though Eastern Anatolia exhibits a compressed overall SHAP range reflecting altitude-moderated temperature extremes. The Aegean profile (Fig. 9c) closely resembles the national average, with T2M, VPD, and Wind Speed as the top three drivers.

3.6.4 Physical cross-checks of dominant-driver classifications

To assess whether the pixel-wise SHAP dominance pattern is consistent with the underlying climatological structure, each dominant-driver class was cross-checked against three independent contexts: the climatological value of the dominant driver at the pixel, elevation, and boundary distance. Boundary distance is used here as a peripheral/coastal-context proxy rather than as a strict distance-to-coast metric. The full results are reported in Supplementary Table S7 and Supplementary Figure S5.

VPD-dominant pixels (n = 896; 70.0% of Türkiye-masked grid cells) exhibit substantially higher climatological VPD than the rest of Türkiye (2.64 vs. 1.86 kPa; Welch t = +29.2) and have the largest mean boundary distance (124.8 km). This supports the interpretation that the SHAP-derived moisture-demand regime corresponds to the empirically driest and most interior parts of the country.

Wind-dominant pixels (n = 255; 19.9%) have a lower mean wind speed than non-wind pixels (1.63 vs. 1.90 m s^{-1}; Welch t = −5.2). This is informative rather than contradictory: SHAP measures marginal model sensitivity, not climatological magnitude. These pixels occupy elevated coastal-ridge and peripheral settings, consistent with the Black Sea and Marmara wind-sensitivity regime.

STRD-dominant pixels (n = 111; 8.7%) have modestly higher STRD than the rest of Türkiye (362 vs. 351 W m^{-2}, mean flux; Welch t = +4.3) and the smallest mean boundary distance, consistent with a localized coastal radiative-load regime. Because STRD represents downward longwave radiation, this pattern should be interpreted cautiously and together with the radiation-cluster analysis rather than as direct evidence of shortwave exposure alone.

These results are interpreted as climatological consistency checks, not as causal proof. They support the physical plausibility of the SHAP-derived spatial regimes while preserving the distinction between model attribution and independent physical causality.

3.7 SHAP dependence interactions

The SHAP dependence plots for T2M and RH reveal the nature of the temperature-humidity interaction (Fig. 10). T2M Dependence (Fig. 10a). The T2M SHAP value increases approximately linearly with air temperature, transitioning from strongly negative contributions at low temperatures (~0 °C, SHAP ≈ −5 °C) to strongly positive contributions at high temperatures (~45 °C, SHAP ≈ +10 °C). The color gradient (representing RH) reveals a humidity interaction: at temperatures above ~30 °C, points with lower RH (blue in the figure) tend to cluster at higher SHAP values than those with higher RH (pink/red). This suggests that dry heat conditions at high temperatures lead to somewhat greater T2M SHAP contributions, consistent with the VPD amplification mechanism identified in the severity-stratified analysis.

RH Dependence (Fig. 10b). The RH dependence plot reveals a strongly nonlinear, monotonically decreasing relationship. At low RH values (<20%), SHAP values are strongly positive (up to +3.5 °C), while the transition to negative SHAP values occurs at approximately

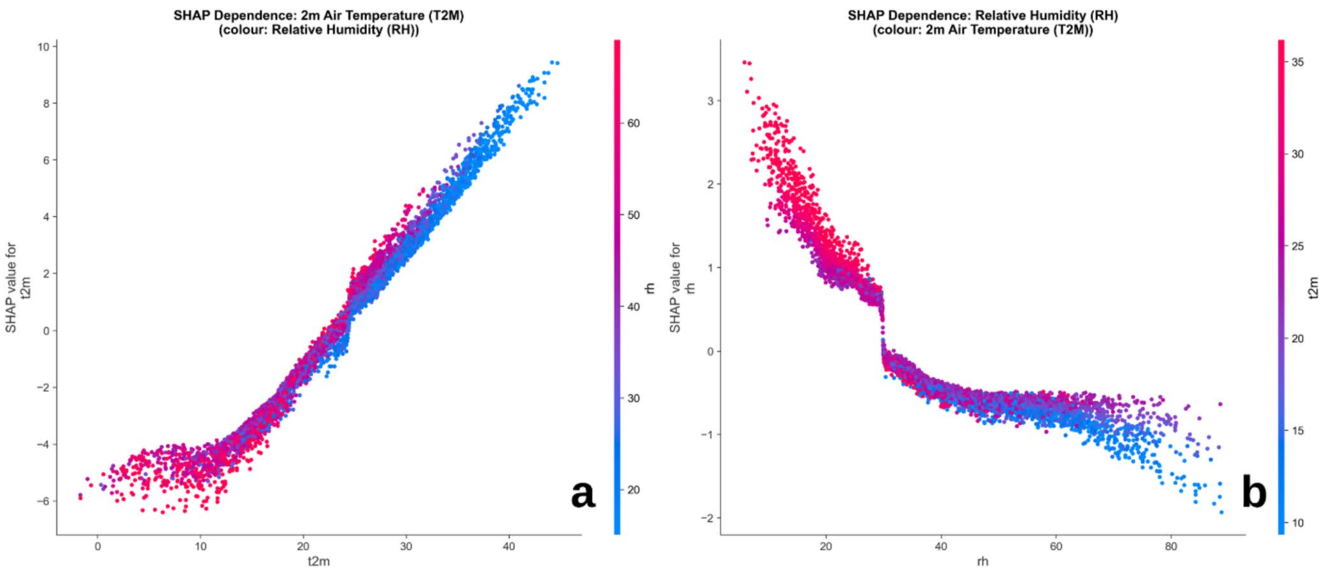


Fig. 10 SHAP dependence plots showing model-attributed contributions of individual climate drivers to predicted UTCI. (**a**) T2M SHAP values versus air temperature, colored by relative humidity. (**b**) RH SHAP values versus relative humidity, colored by 2-meter air temperature

RH ≈ 28–30%, beyond which increasing humidity is associated with decreasing UTCI contributions. This pattern requires careful interpretation: it should not be read as direct physical evidence that higher humidity reduces heat stress. Rather, it reflects the covariance structure within the filtered dataset and the fitted model. In Türkiye's warm-season heat-stress sample, the most extreme UTCI values often coincide with very high temperature and low-to-moderate RH in dry inland settings, whereas higher RH more commonly occurs in maritime environments where temperature is partly moderated. The color gradient confirms this: the strongest positive RH SHAP values occur at high temperatures (red/pink points at low RH), while the most negative values occur at low temperatures (blue points at high RH). The RH dependence therefore captures the conditional statistical association between moisture regime and thermal stress within this domain, not a universal physical relationship.

4 Discussion

4.1 Principal findings and scientific contribution

The main contribution of this study is not that heat stress is intensifying in the Eastern Mediterranean, since that broad conclusion is already well established in the regional literature (Lelieveld et al. 2016; Zittis et al. 2022 a). Nor does the primary contribution lie in the high predictive skill of the Random Forest model, which is partly expected because UTCI is itself derived from meteorological inputs that conceptually overlap with the predictor set. Rather, the scientific value of this work resides in the systematic decomposition of thermal stress variability into spatially differentiated and severity-aware driver structures using explainable AI.

By integrating SHAP analysis at global, regional, severity-stratified, and pixel-wise scales, complemented by blocked spatial-temporal cross-validation, feature ablation, per-day pixel-wise stability analysis, and physical cross-checks, we have decomposed the model-attributed contributions of radiation, moisture, wind, and temperature to predicted UTCI, revealing patterns that are not accessible through conventional correlation or regression approaches. To our knowledge, this represents the first study to apply XAI to systematically decompose UTCI driver importance across thermal stress severity categories at the regional climatological scale, thereby addressing a gap at the intersection of UTCI-based bioclimatology, explainable AI methodology, and severity-stratified thermal stress analysis. The novelty lies less in the prediction of a derived thermal index per se than in the spatially explicit and severity-aware attribution framework used to identify where different classes of meteorological controls become more or less prominent. The central finding is that thermal comfort drivers are neither spatially uniform nor severity-invariant. The pixel-wise SHAP map (Fig. 8) provides the clearest spatial evidence: the dominant secondary climate driver shifts across Türkiye's climatic gradient, with VPD governing much of the semi-arid interior (~70% of Türkiye-masked grid cells), wind speed becoming comparatively more influential along the northern maritime periphery, and a localized radiative-load regime appearing along parts of the southern coastline. These spatial patterns are coherent and consistent with known climatological gradients. They survive the per-day pixel-wise stability check (79.5% mean-condition/per-day modal agreement; Sect. 3.6.2), are corroborated by physical cross-checks against climatological values, elevation, and boundary distance (Sect. 3.6.4), and are stationary across spatial and temporal blocked cross-validation folds (Kendall's W = 0.943/1.000/0.963; Sect. 3.2). They align with the underlying thermophysiological mechanisms by which the human body exchanges heat with its environment under different atmospheric regimes. This insight is essential if heat-risk assessment is to move beyond trend description toward process-aware and region-specific adaptation planning.

4.2 Mechanistic interpretation of spatial driver patterns

The spatial coherence of the VPD-dominated interior can be interpreted through the lens of human heat balance physiology. A critical distinction must be drawn: high VPD does not indicate "humid heat" but rather atmospheric dryness and high evaporative demand. In semi-arid environments where air temperatures frequently exceed skin temperature (~35 °C), dry-bulb convective heat exchange transitions from a cooling mechanism to a heating mechanism, and the body becomes almost entirely dependent on evaporative cooling (sweat evaporation) for thermoregulation (Parsons 2007). The rate of sweat evaporation is directly proportional to VPD: higher VPD accelerates evaporation, but the extremely high VPD values observed in Central, Eastern, and Southeastern Anatolia (often exceeding 3–4 kPa) indicate conditions where, despite high evaporative demand, the rate of metabolic heat production may exceed the rate of dissipation under extreme conditions. The physical correspondence is supported by the Sect. 3.6.4 cross-checks: VPD-dominant pixels coincide with the climatologically driest pixels (Welch t = +29.2 vs. the rest of Türkiye) and lie more than twice as far inland on average (124.8 vs. 59.4 km from the boundary), ruling out the alternative that the SHAP-derived VPD dominance is a model-internal artefact disconnected from the climatology. The cross-check is a consistency check, not a causal inference, but it strengthens the physical reading of the SHAP pattern. This finding is consistent with recent work demonstrating rising VPD as a fundamental constraint on biological evaporative cooling (Yuan et al. 2019) and the empirical evidence that dry heat stress thresholds are substantially lower than humid heat thresholds (Vecellio et al. 2022). The conflation of humidity's role in hot-humid coastal stress with VPD's role in dry inland stress should be carefully avoided; they are related through moisture physics but not interchangeable.

The wind-dominated regime along the Black Sea coast and Marmara region reflects a fundamentally different thermodynamic pathway. An important subtlety in the SHAP attribution should be noted before discussing the mechanism: the Wind-dominant pixels identified by the model have a lower mean wind speed than the rest of Türkiye (1.63 vs. 1.90 m s^{-1}; Sect. 3.6.4), yet SHAP nonetheless assigns these pixels Wind as their dominant secondary driver. This is not a contradiction. SHAP measures the marginal sensitivity of the predicted UTCI to a feature, not the climatological magnitude of the feature. The Black Sea/Marmara coastal-ridge belt is the region where modest wind variations produce substantial UTCI variations, temperatures are close to thresholds where convective cooling matters most, and humidity is high enough that any breeze meaningfully alters thermal comfort. A univariate correlation analysis between mean wind speed and UTCI would not have identified these pixels, illustrating concretely why SHAP analysis was the appropriate methodological choice. In humid oceanic environments (RH typically 60–80%), evaporative cooling is inherently limited by the low VPD, and the body relies more heavily on convective heat loss via wind-driven sensible heat exchange for thermal regulation. The SHAP decomposition recovered a pattern that is consistent with established thermophysiological expectations for humid maritime environments, lending physical plausibility to the attribution results. This is consistent with the findings of Foster et al. (2022), who demonstrated that wind's cooling effectiveness depends critically on the temperature-humidity regime, and with Van der Hoeven and Wandl (2015), who found that intra-urban thermal comfort variability is mainly wind-driven during daytime.

The radiation-dominated clusters along the Mediterranean coast are consistent with the localized nature of longwave radiative heat gain from sun-heated surfaces. In the Mediterranean climate, the combination of high atmospheric temperatures, summer humidity, and clear-sky conditions raises the downwelling longwave flux that STRD measures; this elevated atmospheric thermal emission contributes to the mean radiant temperature that enters the UTCI computation. The STRD-dominant pixels should not be interpreted as locations where shortwave shading interventions alone would be the appropriate adaptation lever. The Sect. 3.6.4 cross-check, however, recommends caution: STRD-dominant pixels have only a modestly elevated mean STRD (+3% over the rest of Türkiye; Welch t = +4.3), considerably smaller than the climatological elevation seen in VPD-dominant pixels. The pattern is therefore better described as identifying pixels where STRD is least correlation-redundant with VPD and T2M, rather than as pixels with the absolutely highest radiative loading. This more cautious reading is also supported by the corrected severity-stratified analysis (Sect. 3.4), which shows that STRD's share of cross-band attribution actually decreases with increasing severity once the response-range scaling artefact of training separate models per band is removed. The patchy spatial distribution of STRD-dominant pixels suggests that terrain-scale factors such as aspect, surface material, and urbanization modulate this effect at sub-regional scales, consistent with previous observations that surface properties exert strong local control on the thermal radiation environment in Mediterranean settings (Altunkasa and Uslu 2020).

4.3 Severity-dependent driver transitions

The severity-stratified analysis provides critical insights for public health risk assessment. The finding that the VPD share of cross-band attribution grows sharply (and the wind share collapses) as conditions transition from moderate to very strong heat stress, while the STRD share declines once response-range scaling is removed (Sect. 3.4; Fig. 5) carries important implications: the atmospheric conditions associated with extreme physiological danger are qualitatively different from those producing moderate discomfort. These results should therefore not be read as a "humidity takeover" of temperature dominance, but rather a transition toward a more compound structure of severe heat in which the moisture-demand dimension (labelled by VPD; Sect. 3.4.1) becomes increasingly consequential, the wind contribution collapses, and temperature retains its primary role.

Specifically, during moderate heat stress (26–32 °C UTCI), the climate driver portfolio is relatively balanced, with wind speed contributing substantially (a 17.1% share of attribution) alongside T2M and VPD. This suggests that passive cooling strategies (ventilation, air movement) may be effective mitigation measures at this severity level, with the important caveat that effectiveness depends on the local humidity regime and exposure context. At very strong heat stress (38–46 °C UTCI), the VPD share surges (+10.0% points; Sect. 3.4) while wind speed collapses to a share of 4.2%, a 75% relative reduction from its moderate-heat share, reaching 3.4% at extreme heat (80% relative reduction). This model-attributed decline in wind importance is consistent with the experimental demonstration by Bröde and Kampmann (2023) that wind cooling reverses to heating when air temperature exceeds approximately 35 °C. At these severities, adaptation strategies that rely primarily on ventilation may become less effective, and the radiative and moisture environment (mean radiant temperature, atmospheric moisture demand) may merit relatively greater attention; specific intervention choices, however, depend on urban morphology, exposure, water availability, and public-health capacity considerations that lie outside the scope of this paper (see Sect. 4.4).

This severity-dependent transition has direct implications for heat action plans and early warning systems. Many existing European heat warning systems rely on temperature thresholds alone (Casanueva et al. 2019; Di Napoli et al. 2019); the present findings suggest that incorporating moisture-demand parameters (e.g. VPD or equivalent) could improve risk stratification, particularly for regions where these drivers gain importance at severe heat stress levels. Heat action plans should also explicitly address the misconception that ventilation universally mitigates heat stress, particularly during extreme events where air temperature exceeds skin temperature.

4.4 Implications for climate adaptation

The spatially resolved driver-regime decomposition motivates indicative, regionally differentiated lines of inquiry for climate adaptation planning in the Eastern Mediterranean (Ali et al. 2022; Nastos and Saaroni 2024; Pietrapertosa et al. 2023), but it is important to be clear about the scale and limits of these implications. The framework reported here does not include urban morphology, population exposure, land cover, green infrastructure, public-health capacity, or socioeconomic vulnerability, all of which are essential inputs to any real adaptation decision. The driver-regime patterns are best read as initial diagnostic information about which atmospheric mechanisms appear most consequential in each region, to be integrated with the operational, urban, and social context that this study does not directly address.

With that caveat stated explicitly, three indicative directions emerge from the SHAP-derived driver structure. First, in the semi-arid interior (Central, Eastern, and Southeastern Anatolia), the dominance of the moisture-demand dimension suggests that thermal comfort interventions may need to address atmospheric dryness and evaporative demand in addition to temperature exposure. Green infrastructure that enhances local evapotranspiration is one candidate pathway, contingent on water availability and exposure conditions; misting and similar evaporative interventions depend critically on water resources, ventilation, and infrastructure context that the model does not represent. Second, along the northern maritime periphery (Black Sea and Marmara), wind sensitivity emerges as a distinctive feature of the SHAP map, with the important caveat from Sect. 3.6.4 that wind-dominant pixels have a lower mean wind speed than the rest of Türkiye; the SHAP-identified pattern is therefore one of high marginal sensitivity, not of high climatological magnitude. The implication is that calm-wind periods in these regions may represent disproportionate heat-risk events, and urban-design choices that preserve ventilation pathways may yield local benefits. Third, along the southern coastline (Mediterranean and Aegean), the radiation cluster (SSRD + STRD) ranks third in the M0 model. We note carefully that STRD is downward longwave radiation, governed by atmospheric temperature, humidity, and cloud-cover state; shading and reflective surfaces directly reduce shortwave exposure (SSRD), not longwave radiation. The radiation-cluster framing therefore corresponds to a heterogeneous mix of intervention pathways: reflective surfaces and shade structures (Santamouris 2014) primarily address SSRD; broader urban heat-island mitigation that reduces ambient atmospheric loading and surface heating contributes to STRD reduction more indirectly (Bouketta et al. 2024; Cascone and Leuzzo 2023).

4.5 Temporal trends and future projections

The 75-year temporal analysis (Fig. 3) demonstrates that the spatial driver patterns documented in this study are set against a backdrop of intensifying heat stress. The statistically significant acceleration of very strong heat events (classical Mann-Kendall $\tau = 0.336$, $p = 1.7 \times 10^{-5}$; Hamed-Rao modified MK $p = 0.012$; Theil-Sen slope + 5.0 grid-days yr^{-1}, 95% CI [+2.96, +7.15]) implies that the conditions under which the moisture-demand dimension intensifies and wind sensitivity collapses (i.e., the upper tail of the heat stress distribution) will become increasingly frequent. The Strong Heat (32–38 °C) band exhibits the largest absolute trend (Theil-Sen slope = +6.68 grid-days yr^{-1}, 95% CI [+2.70, +11.14]; classical and modified MK $p = 1.1 \times 10^{-3}$), indicating that the warming signal is not confined to the upper-tail categories. The extreme heat trend (+0.62 grid-days/year), while not statistically significant at the 0.05 level ($p = 0.092$) under either the classical or the autocorrelation-robust Mann-Kendall test, and with a 95% CI of [−0.06, +1.39] that crosses zero, shows a visually consistent upward drift that may reflect the rarity and high interannual variability of events in this tail category rather than the absence of a physical trend. The nonlinear nature of the severity-dependent driver transition (Sect. 3.4) suggests a potential regime shift: as mean temperatures rise, a disproportionately large number of previously moderate-stress days will cross into the very strong and extreme stress categories where fundamentally different atmospheric mechanisms dominate. This has cascading implications for the effectiveness of current heat mitigation infrastructure, much of which was designed for moderate heat conditions. These temporal patterns are consistent with the centennial-scale analyses of Founda et al. (2022) in Athens and the UTCI trend intensification documented by Katavoutas and Founda (2019), who observed the "very strong heat stress" season elongating by approximately 4.3 days per decade over the past six decades.

4.6 Methodological considerations

The approach of computing SHAP values for every individual grid cell, rather than aggregating at the regional level, represents a methodological contribution that extends beyond the specific findings reported here. This pixel-wise SHAP mapping technique reveals intra-regional heterogeneity, such as the sharp VPD-to-wind transition at the Pontic Mountain crest, that is invisible in regional aggregations and enables climate adaptation strategies to be tailored at sub-regional scales. Similar spatial SHAP mapping approaches have been demonstrated for susceptibility modeling (Wang et al. 2024) and urban thermal comfort prediction (Zhang et al. 2025), suggesting a broader applicability of this methodology across geoscience applications.

4.7 Limitations

4.7.1 Limitations of explainable machine learning applied to a derived thermal index

UTCI is a derived thermal index computed from meteorological inputs, including air temperature, humidity, wind, and radiation, through the UTCI-Fiala operational procedure (Fiala et al. 2012; Bröde et al. 2012). This has an important implication for the present analysis: the Random Forest model is not predicting an independent target variable. Because the predictor set overlaps conceptually with the variables used to compute UTCI, part of the model skill is structural. For this reason, R^2 and RMSE are interpreted here as indicators of internal model consistency for attribution analysis, not as the main scientific result of the study.

The remaining unexplained variance should also be read in light of the temporal mismatch between the target and the predictors. UTCI is computed from hourly fields, whereas the predictors used here are daily maxima, means, or sums. The per-pixel residual diagnostic in Sect. 3.2 suggests that this daily-aggregation gap is largest in cloudier regimes, especially along the Black Sea coast, where sub-daily radiation variability is likely to be smoothed by daily aggregation. This diagnostic is reported in Supplementary Figure S1. A future analysis using predictor values at the hour of maximum UTCI would provide a useful complement, especially for radiation and wind, but it would address a different question: instantaneous peak-hour attribution rather than the event-day attribution analysed here.

A second limitation concerns correlated predictors. RH, VPD, and Q are not independent physical controls; they are related descriptions of the atmospheric moisture state. The correlation matrix, SHAP perturbation-mode comparison, and feature-ablation experiment show that the model captures a robust moisture-demand dimension, but they do not allow VPD, RH, and Q to be interpreted as fully separate causal pathways. The ablation results therefore support the interpretation of VPD as the most compact label for this moisture-demand dimension within the present predictor set, rather than as a uniquely isolated physical cause.

A third limitation is that SHAP values explain the fitted model rather than establishing causality in an experimental sense. The physical cross-checks in Sect. 3.6.4 compare SHAP-derived dominance classes with climatological values, elevation, and boundary distance. These checks make the spatial attribution patterns more physically plausible and reduce the likelihood that they are purely model-internal artefacts, but they do not prove that the identified variables are the sole physical causes of UTCI variability in those locations. This distinction is particularly important for wind: Wind-dominant pixels have lower mean wind speed than the rest of Türkiye, showing that SHAP measures marginal model sensitivity rather than climatological magnitude.

4.7.2 General data and modelling limitations

The analysis is based on ERA5, ERA5-Land, and ERA5-HEAT reanalysis products. Although these datasets provide spatially continuous, long-term coverage, they may under-represent local processes such as urban heat islands, foehn winds, fine-scale coastal circulations, and neighbourhood-scale topographic effects. The 0.25° analysis grid is therefore appropriate for synoptic-to-regional interpretation, but it is too coarse to support site-specific urban design decisions. This scale limitation is especially relevant for the adaptation discussion, which should be read as indicative rather than prescriptive.

The predictor set is limited to seven meteorological variables. Important planning-relevant variables, including urban geometry, land cover, surface materials, anthropogenic heat, population exposure, public-health capacity, and socioeconomic vulnerability, are not represented. The adaptation pathways discussed in Sect. 4.4 therefore identify plausible process linkages rather than operational recommendations. Any practical implementation would require integration with urban morphology, exposure, vulnerability, water availability, and local public-health information.

The primary model is a Random Forest. This choice is appropriate for nonlinear attribution with TreeExplainer, but other model classes, including gradient boosting and regularized regression approaches, could yield different fitted response surfaces. The main SHAP hierarchy is robust across spatial and temporal blocking within the Random Forest framework, but future work should test whether the same hierarchy is preserved across alternative model architectures.

The original random 80/20 train-test split is retained for comparability with the baseline analysis, but its interpretation is limited to internal model stability. Spatial-temporal autocorrelation was explicitly addressed through Leave-One-Region-Out spatial validation and contiguous temporal-block validation (Sect. 3.2). These blocked experiments show that the SHAP ranking is highly stable across folds, but they do not remove all forms of dependence inherent in gridded reanalysis data. The remaining limitation is therefore the scale of the blocking: spatial folds were defined by official geographical regions, and finer-scale spatial blocking or nested resampling could provide an additional test of local generalization.

Finally, the trend analysis was strengthened using classical and Hamed-Rao modified Mann-Kendall tests, Theil-Sen slopes with 95% confidence intervals, and an area-weighting sensitivity. Even so, the trend analysis is restricted to monotonic change. It does not test for step changes, nonlinear trend forms, changing variance, or shifts in the timing and persistence of heat-stress events. These temporal structures remain important topics for future work.

5 Conclusions

This study presents a comprehensive, multi-scale explainable machine learning framework for analyzing the spatial and severity-dependent controls of thermal comfort across Türkiye, an Eastern Mediterranean climate change hotspot. The principal conclusions are:

1. The selected predictor set provides a sufficiently stable Random Forest model (R^2 = 0.936; per-pixel R^2 median = 0.937; Kendall's W = 0.963 across 12 blocked-CV folds) to support attribution-focused XAI analysis of daily maximum UTCI across 826,799 grid-day observations spanning 75 years (1950–2025). The predictive skill is partly structural because UTCI is itself derived from related meteorological variables; the contribution of this work lies in the attribution decomposition, not in the prediction skill per se.
2. The severity-stratified analysis under a single-model conditional decomposition reveals a clear transition: as heat stress intensifies from moderate (26–32 °C) to very strong (38–46 °C), the moisture-demand-dimension share grows by approximately 10% points and the wind share collapses by approximately 13% points, while the radiative-cluster share declines and does not show the monotonic escalation observed for the moisture-demand dimension.
3. The pixel-wise (0.25° × 0.25°) SHAP map reveals three distinct spatial regimes of secondary driver dominance: a moisture-demand-dimension regime across the semi-arid interior (~70% of Türkiye-masked grid cells), wind sensitivity along the northern maritime coast, and a localized radiative-load regime along parts of the southern coastline, with these patterns surviving per-day robustness checks (79.5% mean-condition/per-day

modal agreement) and physical cross-checks against climatological values, elevation, and boundary distance under contrasting climatic settings.

4. Strong (32–38 °C), Very Strong (38–46 °C), and Extreme (>46 °C) heat events have increased at Theil-Sen slopes of +6.68, +5.00, and +0.62 grid-days yr^{-1} respectively over 1950–2025. The Strong and Very Strong trends are statistically significant under both classical and autocorrelation-robust Mann-Kendall tests; the Extreme trend does not reach conventional significance under either test, implying that the severe-heat driver regimes will become increasingly prevalent under continued warming.
5. Robust adaptation planning requires integration of the SHAP-derived driver regimes identified here with the local urban morphology, exposure, land cover, water availability, and public-health context, which our framework does not directly address.

These findings provide a data-driven, spatially resolved foundation for climate adaptation planning in the Eastern Mediterranean and demonstrate the potential of XAI methodologies for extracting mechanistic insights from complex environmental datasets.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00704-026-06389-3.

Acknowledgements The author gratefully acknowledges the Copernicus Climate Change Service (C3S) and the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing ERA5, ERA5-Land, and ERA5-HEAT data through the Climate Data Store. The author thanks the editor and anonymous reviewers for their constructive comments. Generative AI tools were used during revision for language polishing and consistency checking under the author's supervision; all scientific content, analyses, interpretations, and conclusions were reviewed and verified by the author.

Author contributions D.D.Y. conceived and designed the study, conducted all analyses, prepared all figures and tables, wrote the manuscript, and reviewed and approved the final version.

Funding Not applicable. The authors received no specific funding for this study.

Data availability ERA5 and ERA5-HEAT data are publicly available through the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/). The boundary data for Türkiye's geographic regions were obtained from official geographic databases. All analysis code is available upon request.

Code availability The Python scripts used for data analysis and figure preparation are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors declare no competing interests.

References

Ahmadalipour A, Moradkhani H (2018) Escalating heat-stress mortality risk due to global warming in the Middle East and North Africa (MENA). Environ Int 117:215–225. https://doi.org/10.1016/j.envint.2018.05.014

Ali E, Cramer W, Carnicer J, Georgopoulou E, Hilmi NJM, Le Cozannet G, Lionello P (2022) Cross-Chapter Paper 4: Mediterranean Region. In: Pörtner HO, Roberts DC, Tignor M et al (eds) Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp 2233–2272. https://doi.org/10.1017/9781009325844.021

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper No. 56. Food and Agriculture Organization of the United Nations, Rome, Italy

Altunkasa MF, Uslu C (2020) Upgrading New Zealand's existing office buildings – an assessment of life cycle impacts and its influence on 2050 climate change mitigation target. Sustain Cities Soc 57:102134. https://doi.org/10.1016/j.scs.2020.102134

Błażejczyk K, Kuchcik M (2021) UTCI-10 years of applications. Int J Biometeorol 65:1631–1635. https://doi.org/10.1007/s00484-021-02109-3

Błażejczyk K, Epstein Y, Jendritzky G, Staiger H, Tinz B (2012) Comparison of UTCI to selected thermal indices. Int J Biometeorol 56:515–535. https://doi.org/10.1007/s00484-011-0455-5

Błażejczyk K, Jendritzky G, Bröde P, Fiala D, Havenith G, Epstein Y, Psikuta A, Kampmann B (2013) An introduction to the Universal Thermal Climate Index (UTCI). Geogr Pol 86:5–10. https://doi.org/10.7163/GPol.2013.1

Bommer PL, Kretschmer M, Hedström A, Bareeva D, Höhne MM-C (2024) Finding the right XAI method-A guide for the evaluation and ranking of explainable AI methods in climate science. Artif Intell Earth Syst 3:e230074. https://doi.org/10.1175/AIES-D-23-0074.1

Bouketta S, Bouchahm Y, Djekic J (2024) Coupling of different nature-based solutions for pedestrian thermal comfort in a Mediterranean climate. Build Environ 255:111413. https://doi.org/10.1016/j.buildenv.2024.111413

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Briegel F, Wehrle J, Schindler D, Christen A (2024) High-resolution multi-scaling of outdoor human thermal comfort and its intra-urban variability based on machine learning. Geosci Model Dev 17:1667–1688. https://doi.org/10.5194/gmd-17-1667-2024

Bröde P, Kampmann B (2023) Temperature–humidity-dependent wind effects on physiological heat strain of moderately exercising individuals reproduced by the Universal Thermal Climate Index (UTCI). Biology (Basel) 12:802. https://doi.org/10.3390/biology12060802

Bröde P, Fiala D, Błażejczyk K, Holmér I, Jendritzky G, Kampmann B, Tinz B, Havenith G (2012) Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). Int J Biometeorol 56:481–494. https://doi.org/10.1007/s00484-011-0454-6

Casanueva A, Burgstall A, Kotlarski S, Messeri A, Morabito M, Schwierz C, Töglhofer C (2019) Overview of existing heat-health warning systems in Europe. Int J Environ Res Public Health 16:2657. https://doi.org/10.3390/ijerph16152657

Cascone S, Leuzzo A (2023) Thermal comfort in the built environment: a digital workflow for the comparison of different green infrastructure strategies. Atmosphere (Basel) 14:685. https://doi.org/10.3390/atmos14040685

Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, New York, pp 785–794. https://doi.org/10.1145/2939672.2939785

Cos J, Doblas-Reyes F, Jury M, Marcos R, Bretonnière P-A, Samsó M (2022) The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. Earth Syst Dyn 13:321–340. https://doi.org/10.5194/esd-13-321-2022

Di Napoli C, Pappenberger F, Cloke HL (2018) Assessing heat-related health risk in Europe via the Universal Thermal Climate Index (UTCI). Int J Biometeorol 62:1155–1165. https://doi.org/10.1007/s00484-018-1518-2

Di Napoli C, Pappenberger F, Cloke HL (2019) Verification of heat stress thresholds for a health-based heat-wave definition. J Appl Meteorol Climatol 58:1177–1194. https://doi.org/10.1175/JAMC-D-18-0240.1

Di Napoli C, Barnard C, Prudhomme C, Cloke HL, Pappenberger F (2021) ERA5-HEAT: A global gridded historical dataset of human thermal comfort indices from climate reanalysis. Geosci Data J 8:2–10. https://doi.org/10.1002/gdj3.137

Fiala D, Havenith G, Bröde P, Kampmann B, Jendritzky G (2012) UTCI-Fiala multi-node model of human heat transfer and temperature regulation. Int J Biometeorol 56:429–441. https://doi.org/10.1007/s00484-011-0451-6

Flora ML, Potvin CK, McGovern A, Handler S (2024) A machine learning explainability tutorial for atmospheric sciences. Artif Intell Earth Syst 3:e230018. https://doi.org/10.1175/AIES-D-23-0018.1

Foster J, Smallcombe JW, Hodder S, Jay O, Flouris AD, Havenith G (2022) Quantifying the impact of heat on human physical work capacity; part II: the observed interaction of air velocity with temperature, humidity, sweat rate, and clothing. Int J Biometeorol 66:507–520. https://doi.org/10.1007/s00484-021-02231-2

Founda D, Katavoutas G, Pierros F, Mihalopoulos N (2022) Centennial changes in heat waves characteristics in Athens (Greece) from multiple definitions based on climatic and bioclimatic indices. Glob Planet Change 212:103807. https://doi.org/10.1016/j.gloplacha.2022.103807

Gholami H, Falah F, Azarafza M, Mohammadi P, Jafari A, Lee S, Tien Bui D (2024) An interpretable model based on machine learning and SHAP interpretation technique for mapping wind erosion hazard. Environ Sci Pollut Res 31:64628–64643. https://doi.org/10.1007/s11356-024-32450-8

Giorgi F (2006) Climate change hot-spots. Geophys Res Lett 33:L08707. https://doi.org/10.1029/2006GL025734

Guo R, Yang B, Guo Y, Li H, Li Z, Zhou B, Hong B, Wang F (2024) Machine learning-based prediction of outdoor thermal comfort: combining Bayesian optimization and the SHAP model. Build Environ 254:111301. https://doi.org/10.1016/j.buildenv.2024.111301

Hamed KH, Rao AR (1998) A modified Mann–Kendall trend test for autocorrelated data. J Hydrol 204(1–4):182–196. https://doi.org/10.1016/S0022-1694(97)00125-X

Hamed MM, Nashwan MS, Shahid S (2023) Spatiotemporal changes in Universal Thermal Climate Index in the Middle East and North Africa. Atmos Res 298:107089. https://doi.org/10.1016/j.atmosres.2023.107089

Havenith G, Fiala D, Błażejczyk K, Richards M, Bröde P, Holmér I, Rintamäki H, Benshabat Y, Jendritzky G (2012) The UTCI-clothing model. Int J Biometeorol 56:461–470. https://doi.org/10.1007/s00484-011-0458-2

He Z, Yang Y, Fang R, Zhou S, Zhao W, Bai Y, Ma Y, Guan Y, Tang X, Wang B (2023) Integration of shapley additive explanations with random forest model for quantitative precipitation estimation of mesoscale convective systems. Front Environ Sci 10:1057081. https://doi.org/10.3389/fenvs.2022.1057081

Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut J-N (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146:1999–2049. https://doi.org/10.1002/qj.3803

Hong T, Yim SHL, Heo Y (2025) Interpreting complex relationships between urban and meteorological factors and street-level urban heat islands: application of random forest and SHAP method. Sustain Cities Soc 126:106353. https://doi.org/10.1016/j.scs.2025.106353

Hussain M, Mahmud I (2019) pyMannKendall: a python package for non parametric Mann Kendall family of trend tests. J Open Source Softw 4(39):1556. https://doi.org/10.21105/joss.01556

Jendritzky G, de Dear R, Havenith G (2012) UTCI-why another thermal index? Int J Biometeorol 56:421–428. https://doi.org/10.1007/s00484-011-0455-4

Jiang Z, Bevacqua E, Zscheischler J (2024) How interpretable machine learning can benefit process understanding in the geosciences. Earths Future 12:e2024EF004540. https://doi.org/10.1029/2024EF004540

Katavoutas G, Founda D (2019) Intensification of thermal risk in Mediterranean climates: evidence from the comparison of rational and simple indices. Int J Biometeorol 63:1251–1264. https://doi.org/10.1007/s00484-019-01724-4

Kendall MG (1948) Rank correlation methods, 1st edn. Charles Griffin, London

Kim J, Zhou Y, Schiavon S, Raftery P, Brager G (2018) Personal comfort models: predicting individuals' thermal preference using occupant heating and cooling behavior and machine learning. Build Environ 129:96–106. https://doi.org/10.1016/j.buildenv.2017.12.024

Kuglitsch FG, Toreti A, Xoplaki E, Della-Marta PM, Zerefos CS, Türkeş M, Luterbacher J (2010) Heat wave changes in the eastern Mediterranean since 1960. Geophys Res Lett 37:L04802. https://doi.org/10.1029/2009GL041888

Lelieveld J, Hadjinicolaou P, Kostopoulou E, Chenoweth J, El Maayar M, Giannakopoulos C, Hannides C, Lange MA, Tanarhte M, Tyrlis E, Xoplaki E (2012) Climate change and impacts in the Eastern Mediterranean and the Middle East. Clim Change 114:667–687. https://doi.org/10.1007/s10584-012-0418-4

Lelieveld J, Proestos Y, Hadjinicolaou P, Tanarhte M, Tyrlis E, Zittis G (2016) Strongly increasing heat extremes in the Middle East and North Africa (MENA) in the 21st century. Clim Change 137:245–260. https://doi.org/10.1007/s10584-016-1665-6

Lionello P, Scarascia L (2018) The relation between climate change in the Mediterranean region and global warming. Reg Environ Change 18:1481–1493. https://doi.org/10.1007/s10113-018-1290-1

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 30:4765–4774. https://doi.org/10.48550/arXiv.1705.07874

Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, Katz R, Himmelfarb J, Bansal N, Lee SI (2020) From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2:56–67. https://doi.org/10.1038/s42256-019-0138-9

Mann HB (1945) Nonparametric tests against trend. Econometrica 13(3):245–259. https://doi.org/10.2307/1907187

Mora C, Dousset B, Caldwell IR, Powell FE, Geronimo RC, Bielber CR, Counsell BW, Dietrich BS, Johnston EW, Louis LV, Lucas MP, McKenzie MM, Shea AG, Tseng H, Giambelluca TW, Leon LR, Hawkins E, Trauernicht C (2017) Global risk of deadly heat. Nat Clim Change 7:501–506. https://doi.org/10.1038/nclimate3322

Nastos P, Saaroni H (2024) An assessment of CMIP6 climate signals and biases in temperature, precipitation and soil moisture over Europe. Int J Climatol 44:3169–3190. https://doi.org/10.1002/joc.8169

Nazarian N, Krayenhoff ES, Bechtel B, Hondula DM, Paolini R, Vanos J, Vantrepotte V, Middel A, Talaei S, Cheung T, Foster J, Lee JKW, Chow WTL (2022) Integrated assessment of urban overheating impacts on human life. Earths Future 10:e2022EF002682. https://doi.org/10.1029/2022EF002682

Ntoumos A, Hadjinicolaou P, Zittis G, Lelieveld J (2022) Projected air temperature extremes and maximum heat conditions over the MENA region. Earth Syst Environ 6:343–356. https://doi.org/10.1007/s41748-022-00246-2

Özturk T, Ceber ZP, Türkeş M, Kurnaz ML (2015) Projections of climate change in the Mediterranean Basin by using downscaled global climate model outputs. Int J Climatol 35:4276–4292. https://doi.org/10.1002/joc.4285

Pappenberger F, Jendritzky G, Staiger H, Dutra E, Di Giuseppe F, Richardson DS, Cloke HL (2015) Global forecasting of thermal health hazards. Int J Biometeorol 59:311–323. https://doi.org/10.1007/s00484-014-0843-3

Parsons K (2007) Human Thermal Environments: The Effects of Hot, Moderate, and Cold Environments on Human Health, Comfort, and Performance, 3rd edn. CRC Press, Boca Raton

Pietrapertosa F, Olazabal M, Simões SG, Salvia M, Fokaides PA, Kyriakopoulos G, Spyridaki N, Zaganczyk K, Kurnik B, Swiderska M, Ruggiero S, Di Munno R, Manfren M, De Felice M, Musco A, Damm A, Vollmer D, Schaub S, Smigaj M, Magdziak M, Tikkanen J, Marin J, Sanz MJ, Lo Presti V, Fonti V, Gagliano A, Gori A, Pacheco MM, Sousedik B, Berry PM (2023) Adaptation to climate change in cities of Mediterranean Europe. Cities 140:104459. https://doi.org/10.1016/j.cities.2023.104459

Potchter O, Cohen P, Lin TP, Matzarakis A (2018) Outdoor human thermal perception in various climates: a comprehensive review of approaches, methods and quantification. Sci Total Environ 631:390–406. https://doi.org/10.1016/j.scitotenv.2018.03.184

Provençal S, Bergeron O, Leduc R, Barrette N (2016) Thermal comfort in Quebec City, Canada: sensitivity analysis of the UTCI and other popular thermal comfort indices in a mid-latitude continental city. Int J Biometeorol 60:591–603. https://doi.org/10.1007/s00484-015-1066-8

Raymond C, Matthews T, Horton RM (2020) The emergence of heat and humidity too severe for human tolerance. Sci Adv 6:eaaw1838. https://doi.org/10.1126/sciadv.aaw1838

Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat (2019) Deep learning and process understanding for data-driven Earth system science. Nature 566:195–204. https://doi.org/10.1038/s41586-019-0912-1

Santamouris M (2014) Cooling the cities-a review of reflective and green roof mitigation technologies. Sol Energy 103:682–703. https://doi.org/10.1016/j.solener.2013.12.006

Sen PK (1968) Estimates of the regression coefficient based on Kendall's tau. J Am Stat Assoc 63(324):1379–1389. https://doi.org/10.1080/01621459.1968.10480934

Sherwood SC, Huber M (2010) An adaptability limit to climate change due to heat stress. Proc Natl Acad Sci U S A 107:9552–9555. https://doi.org/10.1073/pnas.0913352107

Simpson CH, Brousse O, Ebi KL, Heaviside C (2023) Commonly used indices disagree about the effect of moisture on heat stress. npj Clim Atmos Sci 6:78. https://doi.org/10.1038/s41612-023-00408-x

Soci C, Hersbach H, Simmons A, Berrisford P, Poli P, Dee D, Peubey C, Radu R, Rayner N, Oki T (2024) The ERA5 global reanalysis from 1940 to 2022. Q J R Meteorol Soc. https://doi.org/10.1002/qj.4803

Taşoğlu E, Öztürk MZ, Yazıcı Ö (2024) High resolution Köppen-Geiger climate zones of Türkiye. Int J Climatol 44:5248–5265. https://doi.org/10.1002/joc.8635

Urban A, Di Napoli C, Cloke HL, Kyselý J, Pappenberger F, Sera F, Gasparrini A (2021) Evaluation of the ERA5 reanalysis-based universal thermal climate index on mortality data in Europe. Environ Res 198:111227. https://doi.org/10.1016/j.envres.2021.111227

Van der Hoeven F, Wandl A (2015) Temporal and spatial variability of urban heat island and thermal comfort within the Rotterdam agglomeration. Build Environ 83:91–103. https://doi.org/10.1016/j.buildenv.2014.05.015

Vecellio DJ, Wolf ST, Cottle RM, Kenney WL (2022) Evaluating the 35°C wet-bulb temperature adaptability threshold for young, healthy subjects (PSU HEAT project). J Appl Physiol 132:340–345. https://doi.org/10.1152/japplphysiol.00790.2021

Vecellio DJ, Kong Q, Kenney WL, Huber M (2023) Greatly enhanced risk to humans as a consequence of empirically determined lower moist heat stress tolerance. Proc Natl Acad Sci U S A 120:e2305427120. https://doi.org/10.1073/pnas.2305427120

Wang N, Zhang H, Dahal A, Cheng W, Zhao M, Lombardo L (2024) On the use of explainable AI for susceptibility modeling: examining the spatial pattern of SHAP values. Geosci Front 15:101800. https://doi.org/10.1016/j.gsf.2024.101800

Yılmaz M, Kara Y, Toros H, İncecik S (2024) Analysis of the summer thermal comfort indices in İstanbul. Int J Biometeorol 68:1327–1342. https://doi.org/10.1007/s00484-024-02633-3

Yuan W, Zheng Y, Piao S, Ciais P, Lombardozzi D, Wang Y, Ryu Y, Chen G, Dong W, Hu Z, Jain AK, Jiang C, Kato E, Li S, Lienert S, Liu S, Nabel JEMS, Qin Z, Quine T, Sitch S, Smith WK, Wang F, Wu C, Xiao Z, Yang S (2019) Increased atmospheric vapor pressure deficit reduces global vegetation growth. Sci Adv 5:eaax1396. https://doi.org/10.1126/sciadv.aax1396

Zhang X, Chen T, Zhang S, Yuan C, Tan PY (2025) Toward healthy neighborhood thermal environments: interpretation of thermal comfort and rapid identification of hotspot areas using image-based metrics and explainable machine learning. Sustain Cities Soc 120:106648. https://doi.org/10.1016/j.scs.2024.106648

Zittis G, Hadjinicolaou P, Fnais M, Lelieveld J (2016) Projected changes in heat wave characteristics in the eastern Mediterranean and the Middle East. Reg Environ Change 16:1863–1876. https://doi.org/10.1007/s10113-016-0995-3

Zittis G, Almazroui M, Alpert P, Ciais P, Cramer W, Dahdal Y, Fnais M, Francis D, Hadjinicolaou P, Howari F, Jrrar A, Klingmüller K, Lelieveld J, Lelieveld J, Lionello P, Losic L, Luterbacher J, Manzanas R, Mariotti A, Michaelides S, Pal JS, Perkins-Kirkpatrick SE, Scoccimarro E, Solmon F, Torma C, Ztitis G (2022) Climate change and weather extremes in the Eastern Mediterranean and Middle East. Rev Geophys 60:e2021RG000762. https://doi.org/10.1029/2021RG000762

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.