ISPRS Journal of Photogrammetry and Remote Sensing 233 (2026) 209–230

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Satellite-based heat Index estimatioN modEl (SHINE): An integrated machine learning approach for the conterminous United States

Seyed Babak Haji Seyed Asadollah, Giorgos Mountrakis *, Stephen B. Shaw

Department of Environmental Resources Engineering, State University of New York, College of Environmental Science and Forestry, 1 Forestry Drive, Syracuse, NY 13210, United States

ARTICLE INFO

Keywords:
Heat index
Air temperature
Land surface temperature
MERRA-2
Extreme heat events

ABSTRACT

The accelerating frequency, duration and intensity of extreme heat events demand accurate, spatially complete heat exposure metrics. Here, a modeling approach is presented for estimating the daily-maximum Heat Index (HI) at 1 km spatial resolution. Our study area covered the conterminous United States (CONUS) during the warm season (May to September) between 2003 and 2023. More than 4.6 million observations from approximately 2000 weather stations were paired with weather-related, geographical, land cover and historical climatic factors to develop the proposed Satellite-based Heat Index estimatioN modEl (SHINE). Selected explanatory variables at daily temporal intervals included reanalysis products from Modern-Era Retrospective analysis for Research and Applications (MERRA) and direct satellite products from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor.

The most influential variables for HI estimation were the MERRA surface layer height and specific humidity products and the dual-pass MODIS daily land surface temperatures. These were followed by land cover products capturing water and forest presence, historical norms of wind speed and maximum temperature, elevation information and the corresponding day of year. An Extreme Gradient Boosting (XGBoost) regressor trained with spatial cross-validation explained 93 % of the variance ($R^2 = 0.93$) and attained a Root Mean Square Error (RMSE) of 1.9°C and a Mean Absolute Error (MAE) of 1.4°C. Comparison of alternative configurations showed that while a MERRA-only model provided slightly higher accuracy (RMSE of 1.8°C), its coarse resolution failed to capture fine-scale heat variations. Conversely, a MODIS-only model offered kilometer-scale spatial resolution but with higher estimation errors (RMSE of 2.9°C). Integrating both MERRA and MODIS sources enabled SHINE to maintain spatial detail and preserved accuracy, underscoring the complementary strengths of reanalysis and satellite products. SHINE also demonstrated resistance to missing MODIS LST observations due to clouds as the additional RMSE error was approximately 0.5°C in the worst case of missing both morning and afternoon MODIS land surface temperature observations. Spatial error analysis revealed <1.7°C RMSE in arid and Mediterranean zones but larger, more heterogeneous errors in the humid Midwest and High Plains. From the policy perspective and considering the HI operational range for public-health heat effects, the proposed SHINE approach outperformed typically used proxies, such as land surface and air temperature. The resulting 1 km daily HI estimations can potentially be used as the foundation of the first wall-to-wall, multi-decadal, high resolution heat dataset for CONUS and offer actionable information for public-health heat studies, energy-demand forecasting and environmental-justice implications.

1. Introduction

Extreme heat events (EHEs) are temporal periods in which the average temperature exceeds the historical meteorological norms for a certain region (Sulikowska and Wypych, 2021). The prolonged duration of these events, which could extend from few days to weeks, are known as heatwaves (Chapman et al., 2019). Due to climate change, EHEs are expected to increase in terms of length, intensity and frequency (Trancoso et al., 2020), with considerable consequences on human health, social and industrial activities as well as agricultural productivity (Nukala, 2023; Wang et al., 2022a). A recent survey by Ballester et al. (2023) estimated approximately 60,000 heat-related deaths in

* Corresponding author.
E-mail address: gmountrakis@esf.edu (G. Mountrakis).

https://doi.org/10.1016/j.isprsjprs.2026.01.018
Received 22 September 2025; Received in revised form 6 January 2026; Accepted 10 January 2026
Available online 23 January 2026
0924-2716/Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS).

Europe during the summer of 2022. Similar surges in death rates caused by heat were also reported in the US and China (Chen et al., 2017; Sheridan et al., 2021).

The development of a robust EHE detection model that could identify elevated heat exposure at sites with no surface-based meteorological monitoring is critical for improving the response to EHEs. (McElroy et al., 2020). The foundation for designing such systems is the explicit quantification of heat exposure at a high spatial resolution. While some studies solely considered the daily maximum air temperature (Tmax) as criteria EHE occurrence (Asadollah et al., 2021; Espinosa and Portela, 2025; Nouri et al., 2022), it has been shown that inclusion of other additional factors enhances the exposure rate measurement. Relative Humidity (RH) is evidently another influential component, which, in conjunction with Tmax, can amplify the heat stress on human body and contributes to higher heat-related illnesses (Coffel et al., 2017; Min et al., 2025; Sobolewski et al., 2021).

Developing heat detection indices that incorporate multiple climate components has proven to be an effective solution for overcoming the limitations of single-criterion heat exposure detection. Indices such as the Universal Thermal climate Index and Canada's Humidex (Bröde et al., 2012; Ho et al., 2016) enable a more comprehensive heat exposure measurement. The heat scale provided by Steadman is one of the most popular heat measurement criteria which operates as the foundation of heat vulnerability assessment systems in U.S. (Steadman, 1984) as well as Europe (Michelozzi et al., 2009), Asia (Burkart et al., 2011) and Central and South America (Bell et al., 2008). Steadman's index estimates the severity of heat exposure considering a diverse set of parameters including air temperature, vapor pressure, body size, clothing coverage, effective wind speed and etc. The impact of all these factors is translated into a single aggregated heat estimator called Heat Index (HI) with the same units as air temperature (Steadman, 1979). Steadman's equation was comprehensive and accurate; however, it was very complex as it required large number of variables. Rothfusz developed a modified HI equation which was simultaneously suitable for quick measurement and also operational for weather services (Rothfusz and Headquarters, 1990). Since the Rothfusz equation was obtained through iterative regression analysis on Steadman's original data, it is better tuned, making it reliable for use in nationwide weather forecasting. This equation is also utilized by the U.S. National Weather Service (NWS) to measure health threats due to heat-related disorders (Center, 2012).

HI has been shown to be crucial in identifying EHE and critical thresholds, highlighting its role in understanding heat-health impacts, and thus enhancing the effectiveness of required health interventions (Cristo et al., 2007; Petitti et al., 2016). Based on studies from Perera et al. (2022), HI is a valuable instrument for detection of thermal hotspots in urban areas as micro climate environments. Khatana et al. (2022) demonstrated that the HI is a more sensitive and a locally adaptable heat exposure metric capturing heat stress on humans more accurately than temperature-based alternatives. In a health-related study, Benmarhnia et al. (2019) replaced the heat detection mechanism of the New York City heat emergency plan between 2004 and 2010 with HI and estimated that this change could lead to a reduction of approximately 0.80 fewer heat-related illnesses per day, translating to a significant public health benefit. This adjustment could also lead to about 50 avoided cases of heat-related illnesses during the 2009–2010 summers. By evaluating 1617 US counties over 22 states and over 50 million hospital records from 2003 to 2012, a strong relationship was established between HI levels and heat-related hospitalizations (Vaidyanathan et al., 2019). Their findings further support the claim that HI metrics better align with heat related health impacts.

HI estimation models should be developed by considering an extensive range of historical events (Gobbi et al., 2017) and using a diverse range of meteorological and geographical conditions to capture the effects of weather and land variation on heat anomalies (Sangelantoni et al., 2023). Satellite-based Remote Sensing (RS) observations have been proven to be a valuable source as they benefit from observational uniformity and span wide geographical areas and temporal extents. RS has been widely utilized for heat vulnerability assessment and urban heat analysis (Liu et al., 2023; Schaefer et al., 2021; Wu et al., 2022). To effectively integrate the complex, high-dimensional data derived from remote sensing and meteorological sources, machine learning (ML) techniques offer significant advantages. ML models are capable of capturing nonlinear relationships and intricate interactions among variables that traditional statistical models may overlook (Chowdhury et al., 2023; Kheyruri et al., 2025). This is particularly important in the context of heat estimation and EHE detection, where the spatiotemporal dynamics are highly heterogeneous and influenced by multifactorial drivers (Cebrián et al., 2022; Jiang et al., 2023). Furthermore, ML approaches have demonstrated superior performance in estimation accuracy and generalization when trained on a large and diverse dataset, making them well-suited for applications requiring high-resolution spatiotemporal estimations across diverse geographic and climatic contexts (Chinchanikar and Shaikh, 2022; Rodrigues et al., 2021).

Combination of RS and ML have been widely used for Land Surface Temperature (LST) estimations (Benali et al., 2012; Li et al., 2023b). Landsat imagery, with its high spatial resolution and thermal infrared bands, provides reliable input data for estimating LST, making it particularly valuable for analyzing localized thermal patterns and urban heat dynamics (Ermida et al., 2020; Parastatidis et al., 2017). However, Moderate Resolution Imaging Spectroradiometer (MODIS) is often preferred over Landsat for daily heat estimations due to its higher temporal resolution, providing near-daily observations that enable continuous monitoring of temperature dynamics. For example, Wasif Ali et al. (2022) incorporated day and night LST MODIS products to examine the trends in maximum and minimum LST over Islamabad Capital Territory, Pakistan, using linear regression and the non-parametric Mann–Kendall test. In another study, Na et al. (2024) estimated daily mean, maximum, and minimum LSTs at the global scale using MODIS products as the input for eight different ML algorithms. However, as Krehbiel and Henebry (2016) point out, satellite-derived LST data measure something qualitatively different from weather station observations of the air temperature at the nominal height of 2 m as the spaceborne sensors sense the radiometric skin temperature at the land surface. Considering the fact that air temperature is essential to estimate the Heat Index, satellite LST data may not be a reasonable proxy for HI estimations.

Table 1 summarizes existing studies that have utilized RS products and statistical approaches to estimate daily maximum air temperature and relative humidity. While these studies do not explicitly estimate HI, they can indirectly be utilized for that purpose. The majority of the conducted studies concentrated on broader differences between statistics of air temperature (such as minimum, mean or maximum), and only a handful used Tmax to characterize EHEs. For example, with a stronger focus on capturing the spatial variations of heat exposure rate, Alqasemi et al. (2022) used both day- and night-time MODIS LSTs to estimate monthly maximum and minimum air temperatures using ordinary linear regression model over the United Arab Emirates. Chung et al. (2020) used four daily MODIS LSTs as inputs for a Long Short Term Memory (LSTM) deep learning algorithm to estimate daily Tmax during heat waves and daily Tmin during cold waves. Moreover, studies such as Suthar et al. (2023), estimated Tmax using MODIS LSTs and used that estimation to detect EHE occurrences.

Estimation of relative humidity (RH), another essential component beyond the Tair for HI calculation, using MODIS products is less common (see Table 1). Most studies incorporated MODIS-derived moisture variables, such as precipitable water vapor (PWV) or atmospheric moisture (AM), with the sole exception of Li and Zha (2018). Notably, Recondo et al. (2013) was the only study to develop models for both Tmax and RH, providing the necessary components for estimating apparent temperature. Among the reviewed literature, the study by Ho et al. (2016) stands out as the most comprehensive for HI estimation, as

Table 1
Summary of statistical approaches for daily maximum air temperature and relative humidity estimation.

Tmax estimation	Predictor variables	Method	Location	Time	RMSE	R^2
Joy et al. (2025)	LSTs (Ad, An), NDVI, Land Cover, Elevation, Lat, Lon, Julian day, Distance to ocean	XGBoost, ANN, GAM, MLR	Mainland India	2010–2022	1.7°C	0.90
Zhang et al. (2024)	LSTs (Td, Tn), NDVI, Land Cover, Elevation, Lat, Lon, Julian day, Daylength	RF	Iran	2010–2020	2.4°C	0.94
He et al. (2023)	LSTs (Td, Tn), NDVI, Land Cover, Elevation, Month, Lat, Lon	BPNN	Loess Plateau, China	2003–2012	1.2°C	0.94
Suthar et al. (2023),	LSTs (Ad, An, Td, Tn), NDVI, Aerosol optical depth, Surface layer height, Surface pressure, Water vapor, Solar radiation, Wind speed, RH, Lat, Lon	MLR, SVR, RF	Rajasthan and Karnataka, India	2013–2022	0.7–1.3°C	0.90–0.92
Alqasemi et al. (2022)	LSTs (Ad, An)	SLR	United Arab Emirates	2003–2019	1.7°C	0.94
Zheng et al. (2022)	LSTs (Td), NDVI, Land Cover, Elevation, Lat, Lon, Julian day, Slope, Aspect	SLR, SVR, RF, GBR, AdaBoost, XGBoost, LightGBM, MLP	Qinghai–Tibet Plateau, China	2003–2018	2.1°C	0.94
Zheng et al. (2022)	LSTs (Ad, An, Td, Tn), Solar and net Radiation, EVI, LAI, Albedo, Elevation, Lat, Lon, Julian day, Delta LST	HGB, RF, ET, DBN	Eurasian continent	2003–2018	1.7–2.2°C	0.98
Zeng et al. (2021)	LSTs (Ad), SRs, NDVI, Land cover, Snow Cover, Solar Radiation, Elevation, Julian day, View time, Climate zone, Distance to ocean	RF	Global (Excluding polar zones)	2009–2018	2.6°C	0.95
Serra et al. (2020)	LSTs (Td, Tn), NDVI, NDBI, Elevation, Lat, Lon, Slope, Aspect, Julian day	MLR	Barcelona, Spain	2015	2°C	0.92
Hough et al. (2020)	LSTs (Ad, An, Td, Tn), Ta, NDVI, Elevation, Land Cover, Lat, Lon, Slope, Aspect, Julian day, Climate region	XGBoost, RF, GAM	Continental France	2000–2016	1.8°C	0.95
Chung et al. (2020)	LSTs (Ad, An, Td, Tn)	LSTM	South Korea	2008–2018	2.2°C	0.24
Phan et al. (2019)	LSTs (Ad, An, Td, Tn)	SLR, MLR	Vietnam	2004–2013	0.9–1.1°C	0.86–0.94
Yoo et al. (2018)	LSTs (Ad, An, Td, Tn), NDVI, Elevation, Aspect, Land Cover, Solar radiation, Lat, Lon	RF	Los Angeles, USA Seoul, South Korea	2006–2016	1.0–1.7°C	0.73–0.85
Kitsara et al. (2018)	LSTs (Ad, An), NDVI	SLR	Greece	2002–2004	5.9°C	0.42
Yang et al. (2017)	LSTs (Ad, An), Solar Zenith Angle, acquisition time, NDVI, Elevation, Lat, Lon, Distance to ocean, Julian day	MLR	Northeast China	2002–2016	4.6°C	0.90
Janatian et al. (2017)	LSTs (Td, Tn), Solar Zenith Angle, NDVI, Elevation, Lat, Julian day	MLR	Iran	2000–2004	3.0–3.6°C	0.84–0.90
Cai et al. (2017)	LST (Td), Land Cover	OLS-LR	Yangtze River, China	2008–2010	2.5°C	0.87
Chen et al. (2016)	LSTs (Ad, An, Td, Tn), EVI	SPR	Beijing urban area	2009–2010	1.6–2.3°C	0.95–0.98
Lin et al. (2016)	LSTs (Ad, An, Td, Tn), NDVI, Humidity, TRI, Elevation	SLR, MLR	China	2000–2010	3.3°C	0.90
Noi et al. (2016)	LSTs (Ad, An, Td, Tn), NDVI, Elevation, Lat, Long, day-length in hours, Julian day, Solar zenith angles	Multivariate linear-regression framework	Northern Vietnam	2003–2013	1.4°C	0.88
RH estimation	**Predictor variables**	**Method**	**Location**	**Time**	**RMSE**	**R^2**
Liao et al. (2020)	PWV (Td), Cloud (MOD06), Atm (Td)	SLR	U.S. Pacific Coast, Great Lakes Area	2009–2011	15.3-17.0 %	–-
Ramírez-Beltrán et al. (2019)	PWV (Td, Ad), LSTs (Ad, An, Td, Tn), NDVI, Albedo, Elevation, Soil classes	MLR	Meso-America & Caribbean region	2010–2012	9.9 %	0.58
Li and Zha (2018)	Red, Green, Blue SRs (MODIS), EVI, Landcover, Distance to water, road and city, Elevation	RF	China	2009	–-	0.70
Ho et al. (2016)	LST (Landsat), PWV (Td), NDWI, Distance to ocean, Solar radiation, Sky view factor	RF	Greater Toronto	2002–2009	9.0 %	0.87
Adab et al. (2013)	PWV (Td), Elevation	PLR	Iran	2012	4.8 %	–-
Lin et al. (2013)	AT (ad), AM (Ad), NDVI	LR	East Africa		–-	0.17–0.39
Recondo et al. (2013)	PWV (Td), LSTs (Td, Tn), NDVI, Elevation, Lon, Distance to ocean, Julian day	MLR	Peninsular Spain	2010	8-11 %	0.39–0.49
Peng et al. (2006)	PWV (Td), AT (Td), Elevation	Quadratic Regression	Peninsular Malaysia	2000–2004	2.9 %	0.68

Notes: Ad, An, Td, Tn refers to Aqua Day, Aqua Night, Terra Day, and Terra Night, respectively. LST: Land Surface Temperature (MOD11A1/MYD11A1), PWV: Precipitable Water Vapor (MOD05/MYD05), AM: Atmospheric Moisture (MOD07/MYD07), AT: Atmospheric Temperature (MOD07/MYD07), NDVI: Normalized Difference Vegetation Index, EVI: Enhanced Vegetation Index, NDWI: Normalized Difference Water Index, TRI: Terrain Relief Index.
AdaBoost: Adaptive Boosting, ANN: Artificial Neural Network, BPNN: Backpropagation Neural Network, DBN: Deep Belief Network, DT: Decision Tree, ET: Extremely Randomized Trees, FFNN: Feedforward Neural Network, GAM: Generalized Additive Model, GLM: Generalized Linear Model, GBR: Gradient Boosting Regressor, HGB: Histogram-based Gradient Boosting, LightGBM: Light Gradient Boosting Machine, LSTM: Long Short-Term Memory, MLP: Multi-Layer Perceptron, MLR: Multiple Linear Regression, OLS-LR: Ordinary Least Squares Linear Regression, PLR: Polynomial Linear Regression, PLS: Partial Least Squares, RF: Random Forest, SLR: Simple Linear Regression, SPR: Semi-Physical Regression, SVR: Support Vector Regression, XGBoost: Extreme Gradient Boosting.

it estimated HUMIDEX, a Canadian index of apparent temperature, using a combination of Landsat-derived LST, MODIS PWV, and urbanization-related auxiliary data. Furthermore, aside from a few studies (Ho et al., 2016; Li and Zha, 2018), the table reveals that the application of ML models as RH estimators remains significantly underrepresented in the literature.

Although the HI is widely recognized as an effective assessor of heat-health conditions, very few studies have directly estimated HI utilizing RS observations. However, HI estimation efforts using the combination of RS products and ML algorithms, especially in large geographic areas,

are scarce. Pede and Mountrakis (2022) estimated the daily HI across Conterminous Unites States (CONUS) over the summer of 2012, using climatical, topographical, geographical and land cover variables obtained from different RS sources. Using a Random Forest ML model on a dataset comprising data from 1395 CONUS stations over a 5-month summer period, HI model accuracy had an RMSE of 2.4°C and an R^2 of 0.83. It should be noted though that this study used only clear sky observations thus limiting the applicability of the model. The majority of conducted studies incorporated short temporal lengths, which limits the modeling ability to capture the year-to-year variability. Furthermore, the concentration on a specific region or use of a small number of weather stations limits the geographical diversity and regional variation. In addition, cloud coverage is a significant concern that greatly affects the availability of satellite products, often reducing the number of usable samples for estimation and, consequently, impacting the validity and applicability of the resulting models. Based on existing literature, nearly all studies focusing on Tmax and RH estimations (Table 1) considered only clear-sky conditions, thereby entirely overlooking the influence of cloud contamination. This would cause serious issues in the estimation over regions with frequent cloud cover or when cloud cover leads to significant reductions in the observed LST values, especially during midday when the heat intensity is highest (Østby et al., 2014; Zeng et al., 2018).

Based on these gaps, our study aims to develop an HI estimation framework which takes advantage of RS products to build an accurate and spatially detailed wall-to-wall HI estimation framework at the 1 km^2 spatial scale. Our approach employs both MODIS and Modern-Era Retrospective analysis for Research and Applications (MERRA-2) as primary sources of dynamic daily climate data. Studies such as Grotjahn and Huynh (2018) and Vargas Zeppetello et al. (2022) demonstrated an ability of reanalysis products (e.g. MERRA-2 and ERA) to provide accurate assessment of Tmax, RH and HI on both continental and global levels. However, these studies also highlighted the limitation of the coarse spatial resolution (~50 km^2) of reanalysis products, which significantly constraints their applicability on fine scale heat exposure assessments and overlooks critical local temperature anomalies. Additionally, reanalysis products are produced based on assimilation of multiple sources into physically-based atmospheric models. Such process models may introduce systematic biases and uncertainties in high humidity conditions (Gupta et al., 2020). Thus, the application of downscaling methods solely to MERRA output could potentially amplify biases and cause inter-variable dependencies, reducing the accuracy of estimation in multivariate heat metrics like HI (Adinolfi et al., 2023; Wang et al., 2021).

To address these limitations, we propose a model called Satellite-based Heat Index estimatioN modEl (SHINE). By integrating MERRA-2 climate products with higher resolution MODIS-derived climatic products and land cover topographic information, SHINE aims to achieve significant spatial resolution improvements with consistent high estimation accuracy. To address spatial coverage limitations from prior studies here we utilize over 2000 stations across CONUS. This empowers our estimation model to capture a wide range of climatical and geographical variabilities and enables large scale spatial validation. This study also utilizes daily data for over two decades of continuous temporal coverage. Exposure to such long-term dataset enables the model to consider historical heat trends and interannual variability, potentially enhancing the reliability of proposed framework. Lastly, the presented model development and accuracy assessment considers different cloud coverage scenarios to evaluate robustness and applicability under both full and partial cloud contamination conditions. Detailed error assessment is also provided as the overarching goal is to operationalize this model and create a series of historical HI-based heatmaps across the continental U.S.

2. Data and applied methodology

2.1. Study area and heat index estimation (target variable)

The Conterminous United States (CONUS) is a popular study area due to its geographical and climatic diversity as well as availability of long term publicly available datasets. To establish a reference dataset for the heat estimations, this study utilized NOAA's Local Climatological Data (LCD), which included surface observations collected from the Automated Surface Observing System (ASOS) and the Automated Weather Observing System (AWOS) stations located across CONUS. The LCD data were obtained annually from 2003 to 2023 for five months starting May 1st to September 30th. Selection of this period ensured that the estimation model would capture year-to-year variability and also aligned with satellite data availability (MODIS products). Fig. 1 shows the study area, the geographic distribution of the included active weather stations and their observational annual frequency. An active station was defined as a station having at least 95 percent of valid data within each year. The stations on the spatial distribution map are the union of stations over considered years. As depicted in Fig. 1, the distribution of stations is denser in the eastern part of the US with limited distribution in sparsely populated areas.

For each day, the reference HI value for each station was estimated as follows: for each hourly observation of air Temperature (T) and Relative Humidity (RH) the HI was calculated using the approach of Lu and Romps (2022). Unlike the traditional NWS calculation approach, which was based on Steadman's approach, the Lu and Romps method re-solves Steadman's full human heat-balance equation for all probable temperature and humidity pairs. This enables the physical validation of calculated HI from cold conditions to extreme heat and humidity situations. Their physics-based approach replaced Steadman's original simplified linear long-wave radiation term with the nonlinear Stefan-Boltzmann law. Additionally, they adopted a novel saturation-vapor-pressure equation, while all other parameters remained unchanged.

Steadman's original method was based on seven major constraints, so in extremely cold or hot-humid conditions the solution no longer considered body's core at the life-critical 36.8°C. Lu and Romps added a new eighth constraint which fixed the core temperature at around 36.8°C and simultaneously freed one parameter that Steadman had treated as fixed, turning it into the extra unknown needed to keep the system solvable. The freed parameter was depended on different climate regions, ranging from very cold to extreme hot-humid. By shifting from one control parameter to another based on the governing climate condition, the extended model always has eight equations in eight unknowns, guaranteeing a physically consistent solution for every temperature–humidity pair while preserving the core-temperature setpoint that Steadman assumed (Lu and Romps, 2022). Through reassessing historical US heat waves from 1984 to 2020, application of this extended method showed that the traditional NWS approximation underestimated the true heat index by up to 10°C (20°F) in extreme cases (Romps and Lu, 2022). Using this extended approach, the highest value from all measured HIs within a day was selected to represent the maximum daily HI value.

Ground-based observational HI data were generated for the summer months (May to September) from 2003 to 2023 at active stations for each year leading to a total of 5,371,371 samples. From these samples, 120,023 samples had no observation for Tmax which led to no meaningful HI value and consequently were removed from the dataset. Also a total of 624 samples had unusually high HI estimations (above ~70°C) and were excluded based on justifications provided by Lu and Romps (2022). Furthermore, 636,484 samples had RH value of zero, possibly due to measurement error, mistakes in data entry or stations instrumental issues, and they were removed from the dataset. Applying all these removal criteria resulted in a final dataset with 4,614,240 valid daily samples.

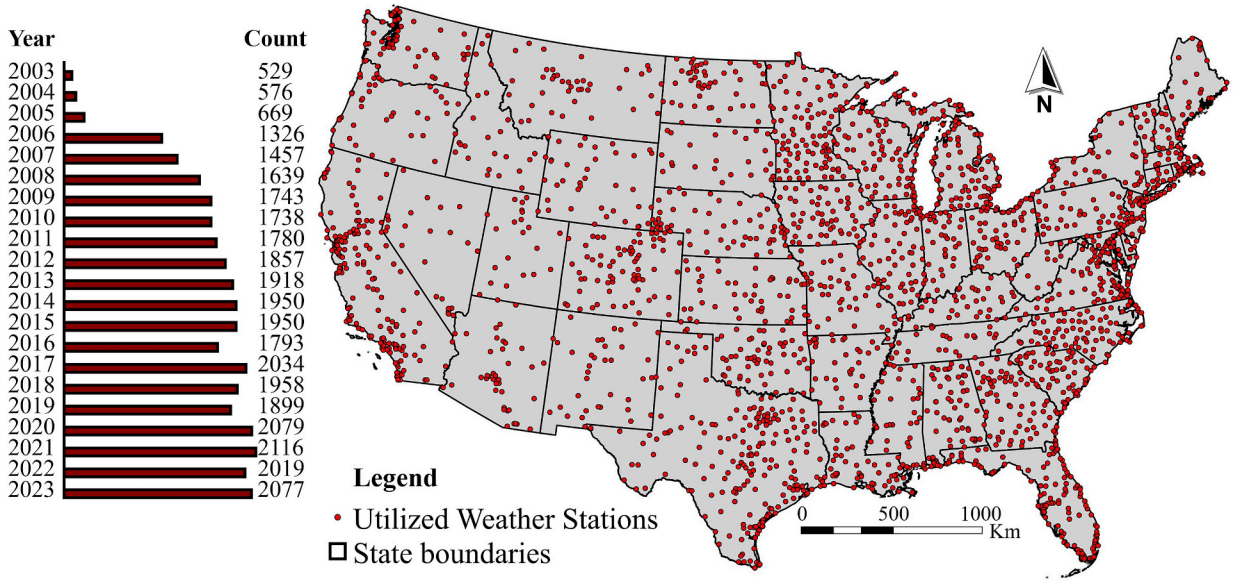

Fig. 1. Spatial and temporal distribution of the utilized ground weather stations.

2.2. SHINE explanatory variables

The main objective of this study is to provide a high resolution wall-to-wall CONUS HI estimation model. Variable generation initially included coarse resolution (~50 km^2) Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) products, such as temperature, humidity and wind information. MERRA-2 products integrate RS products and ground-based observations into the GEOS physical atmospheric model (Dorji et al., 2025; Jordan et al., 2010). To downscale the MERRA-2 variables and improve model accuracy additional gridded products are included. The Moderate Resolution Imaging Spectroradiometer (MODIS) and the Shuttle Radar Topography Mission (SRTM) are two primary sensors-based sources providing gridded observational data for climate and topography, respectively. Additionally, the Landsat-based National Land Cover Dataset (NLCD) is chosen for land cover land use (LULC) representations (Singh et al., 2012; Wan et al., 2019). Finally, long term climatic averages are included to provide general contextual information to the model using the TerraClimate gridded climate products (Abatzoglou et al., 2018; Black, 2010).

Table 2 shows the list of tested independent variables, their description, spatial and temporal resolution, aggregation approaches and RS source used for HI estimation. Note that the final model utilized a subset of these variables (see Results section).

2.2.1. Merra-based land surface forcing

Land Surface Forcin (LSF) products from NASA's MERRA version 2 were also included as explanatory variables. Data provided by MERRA-2 have been widely used in large- and local-scale heatwave detection over US, Africa and other regions (Coughlan de Perez et al., 2023; Engdaw et al., 2022; Yang et al., 2019). This product is delivered with an hourly resolution with a global pixel size of nearly 50 km. LSF variables include instantaneous near surface variables, namely the air temperature, specific humidity, wind speed, surface pressure and surface layer height. These variables have been dynamically generated by the Goddard Earth Observing System (GEOS) atmospheric model while continuously assimilating satellite and in-situ observations since 1980.

2.2.2. Land surface temperature (LST)

LST products from MODIS's MOD11A1 (Terra) and its Aqua twin (MYD11A1) were utilized. These products provide daily earth surface temperature estimations in pixels with 1 km resolution. For each single day four observations were available, including Terra daytime (10:30 a.m. local time), Aqua noontime (1:30p.m.), Terra nighttime (10:30p.m.) and Aqua nighttime (1:30 a.m.). These products are derived from the sensor's thermal-infrared bands after cloud-screening and atmospheric-water–vapor correction. As table 1 shows, this specific set of MODIS products has been extensively used for both maximum air temperature and relative humidity estimations (Agathangelidis et al., 2022; Albright et al., 2011). While some studies used surface skin temperature from MODIS atmospheric profile products (MOD/MYD07), the selected products (MOD/MYD11A1) are closer to observational ground temperature (Wang et al., 2007; Zhang et al., 2025). Invalid observations due to cloud presence were kept in the dataset and their accuracy impact was assessed.

2.2.3. Precipitable water vapor (PWV)

PWV products were obtained from the Terra and Aqua (MOD/MYD05) Total Precipitable Water product at daytime and represent the total depth of water in an atmospheric column above each pixel. While availability of LST products is affected by cloud coverage, PWV observations need a clear line of sight through the column of air, not a direct view of the land surface (Gao and Kaufman, 2003). Thus it can operate under thin or broken-cloud conditions and even above low clouds, and contains far fewer missing observations making it good auxiliary variable to be used beside the LSTs for HI estimations (Gao et al., 2003). While MODIS atmospheric profile products (MOD/MYD07) offer various products for vapor content at different atmospheric levels, the utilized PWV products were a better match to in-situ ground relative humidity (Albert et al., 2005; Peng et al., 2006).

2.2.4. Surface Reflectance (SR)

SR bands 1 to 7 (ranging from visible to short-wave infrared) were also considered as a separate set of input variables. These observations were obtained from both Terra and Aqua at daytime (MOD09GA/MYD09GA) at a daily basis with 500-meter spatial resolution. Inclusion of such variables enables the model to capture better Albedo, which directly governs the surface energy balance, physically based heat-fluxes (Angelini et al., 2021; Liang, 2001). Based on table 1, vegetation indices were extensively used as input parameter for both air temperature and relative humidity estimations. To capture daily variations in vegetation indices, the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI) from bands 1–2 (red and NIR) and bands 2–7 (NIR and MIR) pairs were calculated and considered. While MODIS provides its own vegetation indices products (MOD/

Table 2
Overview of predictor variables, their description, temporal and spatial resolution, spatial and temporal aggregations, source and product name.

Variable	Description	Temporal resolution	Spatial resolution	Spatial aggregations	Temporal aggregations	Source product
TLML	Surface air temperature (°K)	Hourly	~ 50 Kilometers	—-	Average and Max Daily aggregations	MERRA-2 M2I1NXLFO (V 5.12.4)
HLML	Surface Layer Height (meters)					
PS	Surface Pressure (Pa)					
QLML	Surface Specific Humidity (Kg/Kg)					
SPEEDLML	Surface Wind Speed (m/s)					
LST_{TD}	10:30 am (Terra Day) Land Surface Temperature (LST) (°K)	Daily	1000 m	1-, 3-, 5- and 10-km	—-	MODIS MOD11A1 (V 6.1)
LST_{TN}	10:30 pm (Terra Night) LST(°K)					
LST_{AD}	01:30 pm (Aqua Noon) LST (°K)	Daily	1000 m	1-, 3-, 5- and 10-km	—-	MODIS MYD11A1 (V 6.1)
LST_{AN}	01:30 am (Aqua Night) LST (°K)					
PVW_{TD}	10:30 am (Terra Day) Precipitable Water Vapor (cm)	Daily	1000 m	1-, 3-, 5- and 10-km	—-	MODIS MOD05 (V 6.1)
PVW_{AD}	01:30 pm (Aqua Noon) Precipitable Water Vapor (cm)					MODIS MYD05 (V 6.1)
SR_{TD}	Surface Reflectance's Band 1–7 (Terra Day)	Daily	500 m	—-	—-	MODIS
$NDVI_{TD}$	Normalized Difference Vegetation Index		1000 m			MOD09GA (V 6.1)
$NDWI_{TD}$	Normalized Difference Water Index					
SR_{AD}	Surface Reflectance's Band 1–7 (Aqua Noon)	Daily	500 m	—-	—-	MODIS
$NDVI_{AD}$	Normalized Difference Vegetation Index		1000 m			MYD09GA (V 6.1)
$NDWI_{AD}$	Normalized Difference Water Index					
%Barren	Percent of barren (NLCD class 31)	Yearly	30 m	Percentage of each LULC within 1-, 3-, 5- and 10-kilometer r	—-	NLCD LULC maps
%Cropland	Percent of agriculture (NLCD classes 81–82)					
%Developed	Percent of developed (NLCD classes 21–24)					
%Forest	Percent of forest (NLCD classes 41–43)					
% Herbaceous	Percent of herbaceous (NLCD classes 71–74)					
%Shrubland	Percent of shrubland (NLCD classes 51–52)					
%Water	Percent of water (NLCD class 11)					
%Wetlands	Percent of wetland (NLCD classes 90–95)					
Elev	Elevation (meters)	Single date (2000 s)	30 m	1-, 3-, 5- and 10-km	—-	SRTM
Slope	Slope (Percentage)					
Aspect	Aspect (Degrees)					
TPI	Topographic Position Index (Unitless)					
TRI	Terrain Ruggedness Index (Unitless)					
Aet	Actual Evapotranspiration (mm/month)	Monthly	~4000 m	—-	Statistics calculated for 20 (2000–2021) 30 (1990–2021) and 50 (1960–2011) years of historical periods	TerraClimate
Pet	Potential Evapotranspiration (mm/month)					
Ppt	Precipitation (mm/month)					
Soil	Soil Moisture (mm^3/ mm^3)					
Srad	Downward shortwave radiation (W/m^2)					
Swe	Snow water equivalent (mm)					
Tmax	Max Temperature (°C)					
Tmin	Min Temperature (°C)					
Vap	Vapor pressure (kPa)					
Ws	Wind speed (m/s)					
PDSI	Palmer Drought Severity Index (Unitless)					
Julian Day	Day-Of-Year based on Julian calendar	Daily	—-	—-	—-	—-

MYD13), their 16-days temporal resolution did not align with the daily estimation goal of our study.

2.2.5. Land use land cover (LULC)

Based on table 1, variables capturing the land use/cover were considered as an input variable among the majority of studies for both Tmax and RH estimations. Here, the percentage of different LULC classes was calculated within defined neighborhoods using the National Land Cover Dataset (NLCD) database (Jin et al., 2023b,2023a). NLCD classifies every pixel into 8 main and 20 sub LULC classifications, with annual updates released every year (starting from 1992 to current) to track landscape change, canopy cover, and impervious surfaces. For the purpose of this study, corresponding sub classifications were merged under their main category and only the percentage of final eight major LULC (namely barren, cropland, developed, forest, herbaceous, shrubland, water and wetlands) were reported. Several studies justified the inclusion of cropland, developed and forest LULC in their heat estimations as they govern the high portion of heat exchange between different mediums. Inclusion of herbaceous, shrubland and wetlands LULCs can provide valuable information regarding the heat mitigation factors as they could contribute to temperature decrease (Muro et al., 2018; Shen et al., 2022). On the other hand, studies show that large barren or arid tracts act as heat sources (Aghazadeh et al., 2023; Baronian et al., 2024). Water is the only LULCs which could act as both depleting and exacerbating heat exposure (Akter et al., 2021; Sadiq Khan et al., 2020).

2.2.6. Topographic indicators

Since altitude changes directly influence the fluctuation in Tmax and RH, numerous studies have incorporated it (Chen et al., 2021; Fernández-Duque et al., 2023). Based on Table 1, beside elevation, derivatives including slope and aspect were also frequently listed among input variables. Here, elevation from Shuttle Radar Topography Mission (SRTM) mission and calculate the slope and aspect from the digital elevation mode (Van Zyl, 2001; Werner, 2001). Beside these common topographic variables, the Topographic Position Index (TPI) and the Terrain Ruggedness Index (TRI) were also included. TPI and TRI could relate to heat exposure as the terrain shape and ruggedness control sunlight exposure, shading, and local microclimates, with ridges and flat areas typically experiencing higher heat stress (Dobrowski, 2011; Jucker et al., 2018; Xu et al., 2025a).

2.2.7. Historical climate indicators

Long-term, high-resolution data on temperature, humidity and hydrometeorological conditions are essential for identifying heat stress trends and anomalies (Beniston and Stephenson, 2004; Grotjahn et al., 2016). TerraClimate is a global, monthly dataset from 1958 to present with nearly 4-kilometer spatial resolution, which blends station-based climatologies with monthly anomalies from reanalysis datasets (Abatzoglou et al., 2018). To fully capture the effect of historical hydrometeorological conditions and provide related historical context several TerraClimate variables were considered. These included atmospheric (min and max temperature, vapor pressure, wind speed and solar radiation), hydrological (precipitation, soil moisture, snow water, Palmer Drought Severity Index (PDSI)) and energy-water balance (evapotranspiration) variables. Several historical periods were also considered to evaluate the governing pattern in near-, middle- and far-past.

2.3. Explanatory variables: spatial resampling and aggregated statistics generation

To achieve an operation spatial resolution of 1 km^2, the spatial coordinate system of MODIS products was used as reference and other utilized satellite products were either down- or up-sampled to match it. The bilinear interpolation was used to down sample the MERRA-2 products (~50 km latitudinal × 60 km longitudinal) and TerraClimate (~4 km × 4 km) to match the MODIS cell size and projection. Bilinear interpolation is a widely used up sampling approach that calculates the weighted average of the four adjacent pixels to estimate a new value within that pixel (Le Bideau et al., 2019; Vandal and Nemani, 2021). Application of bilinear interpolation to resample MERRA-2 products to the 1-km grid is conducted for co-registration with other higher spatial resolution covariates. The expectation is that the coarser MERRA-2 products will capture large scale effects while the more detailed MODIS LST, NLCD land-cover fractions, and SRTM topographic variables will contribute to the distribution of the MERRA-2 coarse information into the operating 1 km reference system. Area-weighted aggregation was employed to down sample NLCD-based LULC percentages and SRTM variables (both with an original grid size of 30 m) to the reference 1 km grid size. Area-weighted aggregation is used to down sample raster data by aggregating the values of multiple higher-resolution pixels into a single lower-resolution pixel, with each pixel's value being weighted according to its area or proportion of the total area in the new pixel (Chen et al., 2024; Zhong et al., 2024).

After resampling all explanatory variables to the reference 1 km spatial resolution, several spatial and temporal statistical aggregated metrics were derived. From the spatial perspective, buffer zones of 1, 2, 3, 5, and 10 km were created and within each buffer statistical summaries such as mean, standard deviation, 10th and 90th percentiles, and the range were calculated. This process was specifically applied to MODIS- and STRM-based variables and each aggregated metric was used as a separate explanatory variable. The same buffer zones were also used as neighborhood areas to calculate the percentages of different LULC classes. Incorporation of spatial aggregation could enable capturing neighborhood characteristics and landscape heterogeneity around each station, which may outperform single-pixel value extractions (Barbosa et al., 2021; Wu et al., 2019). In terms of intraday temporal matching, our model incorporates observations of variable temporal length. For example, the MERRA-2 variables were measured on a daily-basis and aggregated into a daily mean and maximum. On the other hand MODIS provides instantaneous LST snapshots at ~10:30 a.m. and ~01:30p.m. local time. We used a mixed-timescale predictor set by design, where instantaneous MODIS LSTs provide snapshots that capture the typical rise toward daily HImax and carry information about the daytime surface energy and cloud coverage. Meanwhile, MERRA-2 variables summarize the day-integrated thermodynamic environment that governs evaporative cooling and the humidity contribution to HI. Due to the fact that fixing all predictors to a single time window can misrepresent the true HI and introduce time-mismatch noise (Menne and Williams, 2009; Zhang et al., 2011), one-to-one hour alignment between MERRA-2 and MODIS has not been carried out. For TerraClimate, the mean and standard deviation of variables were estimated for all 12 months and the 5 summer months for 20 years (2000–2020), 30 years (1990–2020), 40 (years 1980–2020), 50 years (1970–2020) and 60 years (1960–2020) as historical periods. For temporal aggregation, a 30-year historical period is considered as climatological normal based on the World Meteorological Organization (WMO). This 30-year span is long enough to smooth the year-to-year anomalies and short enough to reflect the current climate conditions. While considering 40–60 years historical intervals might reduce the weight of today's climate, it enables the model to capture multi decadal climate models and provides long-memory climate trends (Mann et al., 2021). Furthermore, the 12-month span allows for an annual climate envelope while focusing on summer months concentrates on the primary heat exposure season.

2.4. Machine learning model selection – extreme gradient boosting

This study uses the Extreme Gradient Boosting (XGBoost) machine learning (ML) model (Chen and Guestrin, 2016). XGBoost is an ensemble machine learning model which estimates a target by sequentially adding regularized decision trees that fit to the residuals of prior trees using second-order gradient information, so the weighted ensemble

progressively minimizes the chosen loss function (Li et al., 2022; Niazkar et al., 2024). As discussed in section 2.2, MODIS products are highly affected by cloud coverage. In our case, this resulted in numerous samples with missing values (e.g., LST, SR and to a lesser extent PVW MODIS data).

One advantage of XGBoost is its ability to accept the missing observations without further processing (Shaik et al., 2024; Shi et al., 2022). During tree construction XGBoost learns for every split the "default" branch that yields the lowest loss when a variable is absent, so each observation with a gap is routed automatically without prior imputation. This lets the algorithm to train directly on incomplete data, preserves any predictive signal carried by every sample, and avoids the bias or information loss that simple fill-in methods can introduce. To perform the optimization task on the XGBoost associated parameters, a one-parameter-at-a-time sweep was first applied, varying each parameter across a wide numeric range and selecting the value that maximized mean k-fold cross-validation accuracy before proceeding to the next parameter. Once all hyper-parameters had been tuned sequentially, a final model was trained with the chosen settings and its generalization ability evaluated via nested five-fold cross-validation (outer loop) and an independent hold-out test set. This two-stage workflow keeps computational cost low while still yielding an out-of-sample performance estimate free from optimization bias.

2.5. Variable selection and accuracy assessment

The list of candidate variables represented in Table 2 was nearly 200. This high number of candidate inputs necessitated the inclusion of a feature engineering phase to reduce the dataset size and its complexity. Variable selection was conducted using a Forward Feature Selection (FFS) method. Each variable was progressively selected and added by considering the accuracy variation caused by the addition of each single variable using criteria, named Root Mean Squared Error (RMSE). Considering this accuracy metric enabled the model to capture the overall magnitude of errors and has been applied in numerous studies as the core metric for determining a variable inclusion in the final optimize dataset (Balram et al., 2019; Meyer et al., 2018; Zhang et al., 2017).

The presence of multicollinearity was another concern, as some variables are direct statistical estimations of an original variable and may be highly correlated. Due to the high number of included variables and to increase the model's computation efficiency, variables which were statistical derivates from a common variable (e.g., spatial aggregations) were initially grouped as subsets. Next, the Variance Inflation Factor (VIF) analysis was applied among the features in each subset to assess multicollinearity. The VIF method quantifies how much the variance of a regression coefficient is inflated due to multicollinearity among the independent variables, helping to identify and address redundancy in predictors (Cheng et al., 2022; Thompson et al., 2017). In the case of multicollinearity, the Random Forest tree-based importance method was applied to the subset. The feature with the highest importance rate with HI, the target variable, was extracted from the subset and passed to the second feature selection level. If the subset showed no multicollinearity, all features within it would pass to the next level. In the second feature engineering step, the newly created dataset from the concatenation of aforementioned subsets was used in the FFS algorithm.

Accuracy assessment for SHINE was conducted across all 4.6 million samples using a spatial 10-fold cross-validation where each fold had samples from non-overlapping weather stations. This validation method ensured that model assessment took place on samples which were spatially independent from those used for training (Pede and Mountrakis, 2022). This process also generated a more reliable accuracy assessment in real-world scenarios where spatial autocorrelation is a critical consideration (Deppner and Cajias, 2024).. In addition, geographic distribution of errors was assessed spatially through station-level analysis carried out across the CONUS as well as within selected Köppen–Geiger climate divisions. The Köppen–Geiger system is a globally consistent climate-zone taxonomy based on long-term temperature and precipitation thresholds, and it is often preferred because it's simple, empirically derived classes align closely with vegetation and hydrological regimes while remaining readily comparable across regions and studies (Skandalos et al., 2022; Zhang and Gao, 2023). Finally, the model's estimation performance was examined under different cloud coverage scenarios to quantify how missing LST information affected the overall estimation accuracy. Fig. 2 shows a schematic overview of the methodology used in developing SHINE.

3. Results

3.1. Variable final selection considerations

The Forward Feature Selection (FFS) methodology was applied to evaluate the importance of each SHINE variable's addition to the dataset, based on RMSE accuracy metrics. The FFS procedure initiated by selecting the first variable that estimates HI with the lowest RMSE and designated it as the base model. Then, FFS iteratively introduced the remaining variables and measured their contribution to models' overall accuracy by considering the RMSE reduction. Once the second variable was determined and added, the updated RMSE was set as the new accuracy baseline, against which the importance of the remaining variables was evaluated. The process continued until a predefined termination criterion was met, namely when the inclusion of an additional variable resulted in less than a 1 % improvement in RMSE. XGBoost served as the base estimation algorithm in the SHINE FFS model and was tuned for optimal hyperparameter selection using an iterative procedure. Table 3 shows the final 10 selected variables, their spatial and temporal aggregations and corresponding error absorption metrics.

As shown in Table 3, two of products obtained from MERRA-2 have high contribution to HI estimation. It is essential to emphasize that these variables are not categorized under solely satellite observations but modeled products created from the integration of multi observational sources (including satellite and ground-based observations) into the physically-based Goddard Earth Observing System (GEOS) atmospheric model. Consequently, these variables inherently carry the model-originated complexity and uncertainty caused in assimilation processes, yet their inclusion increased overall model estimation accuracy by effectively representing critical atmospheric mixing dynamics related to heat and humidity (Budakoti and Singh, 2021; Miao et al., 2017). Beyond MERRA-2, the MODIS LST variables are the remaining satellite-based products represented at sub-daily scales. These variables not only serve as spatial fine-scaling agents that help to deal with coarser reanalysis estimates, but also provide detailed information regarding changes in surface energy balance (Wang et al., 2023b). LSTs also closely track near surface air temperature fluctuations which is essential to capture diurnal heat dynamics (Safieddine et al., 2025; Sharifnezhadazizi et al., 2019).

The STRM elevation and its spatial aggregated derivative captured the influence of topographic conditions for modeling local climate variability (de Meij and Vinuesa, 2014). Higher elevations are typically associated with lower temperature and humidity, which in turn lead to lower HI values (Navarro-Serrano et al., 2020). In addition, land cover percentages estimated from the NLCD provide the model with information regarding surface characteristics which either exacerbate or mitigate heat exposure (Zhou et al., 2014). Forested areas contribute to cooling mechanisms through shading and evapotranspiration, while proximity to water bodies can both enhance heat exposure or operate as heat mitigation factor, depending on local conditions and type of water body (Cai et al., 2022; Meili et al., 2021). Furthermore, the two TerraClimate-derived historical averages contributed to model overall accuracy by providing valuable information regarding background climate regimes and their variabilities over certain temporal windows (MacDonald et al., 2020). Finally, the inclusion of the Julian day

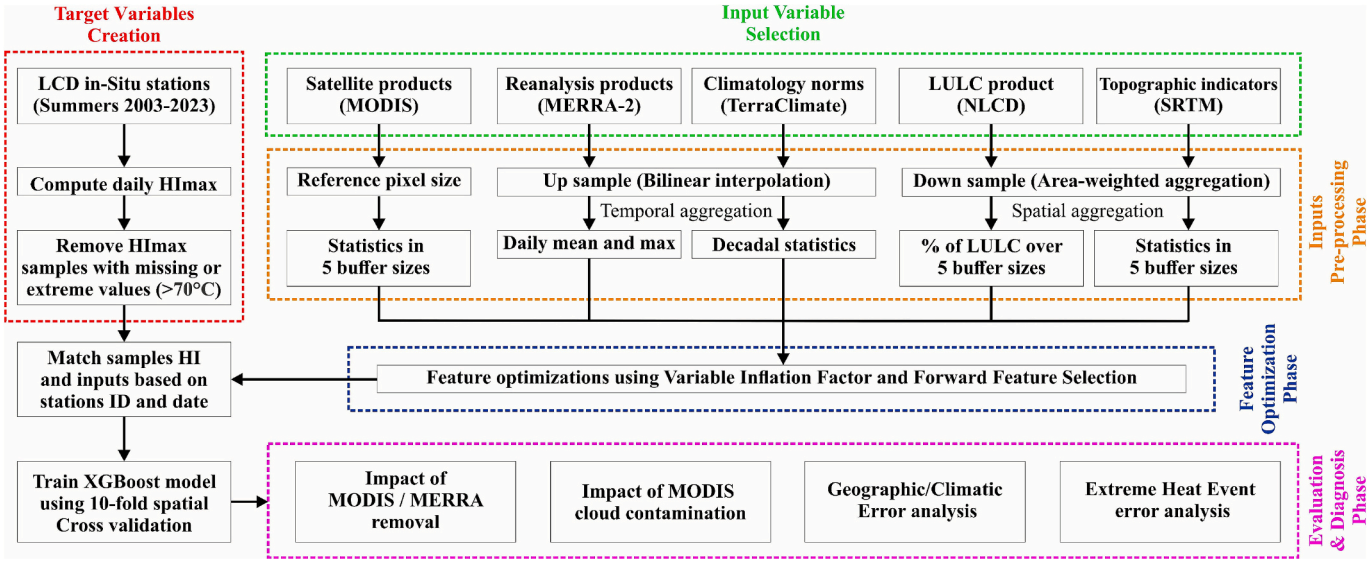

Fig. 2. Schematic overview of the methodology used to develop SHINE.

Table 3
Variable selection order in forward feature selection and variable contribution to RMSE (in °C), MAE (in °C) and R^2 sequential improvements.

Order	Added variable (source)	Spatial aggregation	Temporal aggregation	RMSE	MAE	R^2	VIF
1	Surface layer height (MERRA-2)	–	Mean daily from hourly records	2.8	2.0	0.81	3.2
2	LST of Terra at 10:30 a.m.(MODIS)	–	–	2.6	1.9	0.85	1.6
3	Average of Elevation (SRTM)	% Within a 3 km buffer	–	2.4	1.8	0.88	1.3
4	% Of Water LULC(NLCD)	% Within a 5 km buffer	–	2.2	1.7	0.91	1.1
5	Specific Humidity (MERRA-2)	–	Mean daily from hourly records	2.2	1.6	0.92	3.2
6	LST of Aqua at 01:30p.m. (MODIS)	–	–	2.2	1.5	0.93	1.5
7	Wind speed (TerraClimate)	–	Standard Deviation of summer months from 1990 to 2021	2.1	1.5	0.93	1.3
8	Julian Day	–	–	2.1	1.5	0.93	1.1
9	Maximum Temp (TerraClimate)	–	Mean of summer months from 1990 to 2021	2.0	1.5	0.93	1.7
10	% Of Forest LULC (NLCD)	% Within a 10 km buffer	–	1.9	1.4	0.93	1.2

variable provided a temporal reference for the seasonal cycle of solar forcing, enabling the model to align estimations with typical seasonal patterns of heat stress (MacDonald et al., 2020).

The obtained VIF values for incorporated variables were between 1.1 and 3.2. Being below the 5–10 range, a common threshold used to detect multicollinearity issues (O'brien, 2007; Tsagris and Pandis, 2021) it can be concluded that the predictors provided largely non-redundant information and that the selected variables can be utilized as inputs for HI estimation into the XGBoost model. Together, this set of input variables integrated dynamic atmospheric states, land-surface feedbacks, topographic modifiers, and long-term climatic context, yielding a more comprehensive representation of the contributors of HI across CONUS.

3.2. Accuracy assessment of the full SHINE model

Model error estimation was carried out using 10-fold spatial-cross validation with nearly 4.6 million samples, thus the XGBoost model was exposed to a wide range of HI values. Fig. 3 shows the Scatter-Density plot of estimated HI against their observed values. To generate this plot, first a classic scatter plot was created and then the number of points in each 0.5°C by 0.5°C cell was used to visualize the data density.

Distribution of samples formed a narrow ellipse shape with strong alignment with the optimal line (1:1 dashed line) indicating that the developed estimation model was able to successfully capture the HI variability. Considering the R^2, nearly 93.2 percent of the HI variance was effectively modeled by predictors. The estimated errors of RMSE (=1.9°C) and MAE (=1.4°C) suggest that the model offers sufficiently accurate HI estimations for a wide range of applications. Color distribution reveals that majority of samples are between 20°C and 40°C, however, HI ranges below ~0°C and above ~50°C show respectively systematic positive and negative biases, with the latter being more significant. Extreme HI observations (above 50°C) represented only 0.2 % of the dataset and are mainly recorded during midsummer heat peaks window (mid-July to early August) throughout different summers within the study period of 2003 to 2023. Mapping their spatial frequency revealed that stations which observed these extreme HI occurrences were scattered across the Central and Southeastern CONUS rather than clustered in a specific zone. This spatial analysis reduced the probability of systematic sensor glitches or climatic biases. Thus, we retained these samples as they were deemed both rare and heterogeneous in origin. While this subset of extreme HI values (>50°C) is physically plausible, several studies suggest that they be errors due to sampling biases or instrumental uncertainty under extremely hot and humid conditions (Li et al., 2023a; Yiou and Jézéquel, 2020). For example, it has been shown that ASOS and AWOS stations can exhibit dew point sensor saturation and ventilation biases under such extreme events (Ingleby et al., 2013; Lin and Hubbard, 2004; Sun et al., 2005). Because distinguishing, on a case-by-case basis, between valid extremes and instrument-affected outliers is not reliable with available metadata, we retained all extreme observations for model training but did not consider them in our interpretation as their sample size was small (<0.2 %).

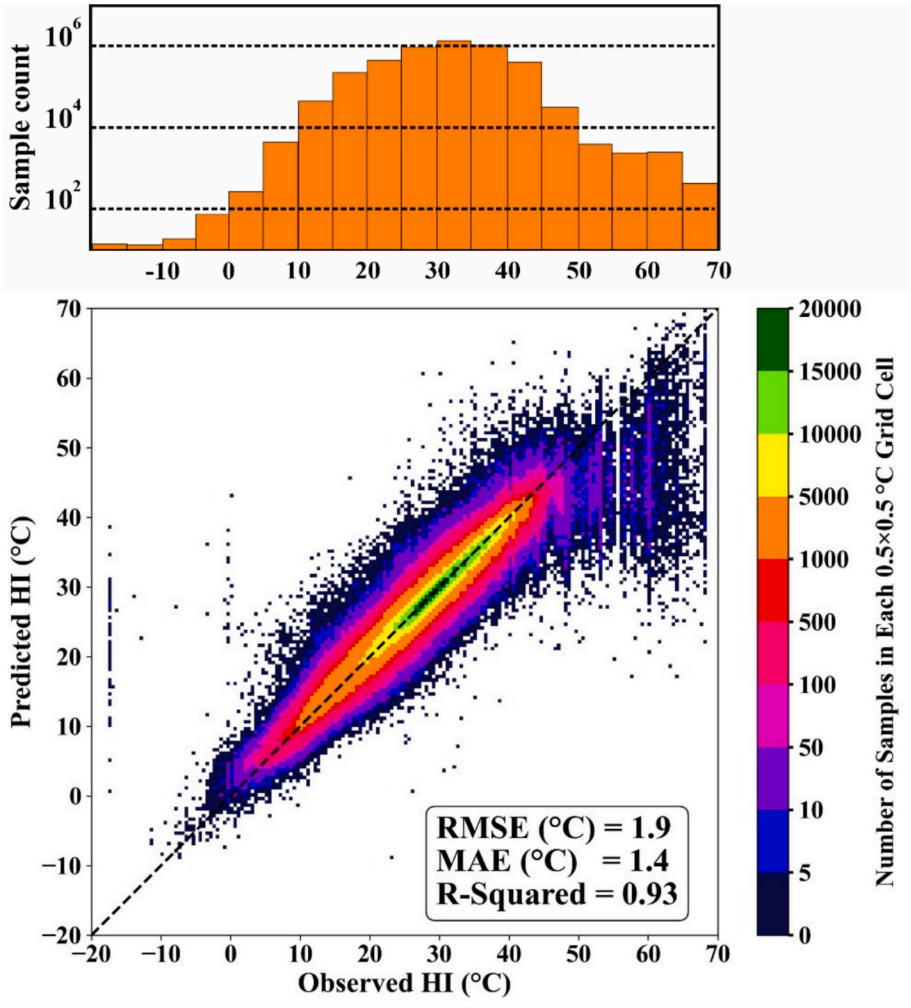

Fig. 3. Density-Scatter plot between the observed and estimated HI on ~4.6 million samples using a spatial cross validation.

3.3. Impact of MODIS/MERRA-2 variable removal

The prior results showed that the fully integrated SHINE model (combining both MODIS and MERRA-2 inputs), provided accurate estimations across the CONUS over a 20 year temporal range. To isolate the benefit of each daily data stream, here we benchmarked SHINE against two alternatives: a MODIS-only configuration that was solely dependent on direct satellite daily observations and excluded MERRA products, and a MERRA-only configuration which only utilized reanalysis daily products without MODIS. The objective was to quantify the trade-offs between satellite-based, reanalysis-based, and integrated approaches. For each configuration, a separate feature selection procedure was applied using the same FFS scheme and termination criteria described in section 3.1. List of FFS-selected features for MODIS-only and MERRA-only configurations and their corresponding contribution to error statistics are tabulated in Table 4. Additionally, Fig. 4 shows density-scatter plots of estimated versus observed HI for these two modeling configurations, with color intensity indicating the frequency of each HI pair.

Looking at Table 4, FFS results show that the MODIS-only model reaches the highest estimation accuracy by utilizing 12 predictors (including two PWV and four LST products) to explain approximately 85 % of the estimation variance. This explanation rate increases to 0.92 the MERRA-only model with just five predictors once a moisture product is added. Table 4 suggests that in the MODIS-only model the variables related to humidity were not selected in early selection stages, and the model needed a larger number of inputs to explain the HI variance. By contrast, for the MERRA-only configuration, specific humidity was among the first selected variables, suggesting that the availability of an explicit moisture representative in the model can improve the HI estimation accuracy.

The listed variables were then used as the input of XGBoost model and the HI error was estimated using the same spatial 10-fold cross-validation, as performed in full SHINE model. Fig. 4 shows that the MERRA-only had better alignment with the 1:1 line, whereas the MODIS-only deviated more from that line. Both models still struggle to model the HI at high values, thus they show no advantage over the fully integrated SHINE model in that aspect.

Across the 4.6 million evaluated samples, the two alternative configurations produce different HI estimation profiles with opposite limitations: MERRA-only shows more accurate HI estimations (RMSE = 1.9°C), however, they came at the cost of blurred kilometer-scale contrasts. On the other hand, the estimation performance of high resolution MODIS-only shows nearly 1°C decrease (RMSE = 2.9°C) compared to other alternative models. While MERRA-only to some extent outperforms both SHINE and MODIS-only, yet its coarse spatial resolution cannot address the complex heat dynamics in urban, coastal or irrigated hot-spots.

Considering this, integration of these two sources allows the estimation model to fuse large-scale surface humidity information with fine scale surface temperature variables, recovering the ~1°C error of MODIS-only model while preserving the 1 km texture lost in pure

Table 4
Results of forward feature selection and variable contribution to error statistics for MODIS-only and MERRA-only models.

MODIS-only	RMSE (°C)	MAE (°C)	R^2	VIF
Max Temp, summer months average from 1990 to 2021 (TC)	5.9	4.3	0.51	1.3
Julian Day	4.8	3.7	0.59	1.0
PWV of Terra at 10:30 a.m. (MODIS)	4.2	3.2	0.69	1.8
LST of Terra at 10:30p.m. (MODIS)	3.9	2.9	0.73	1.4
Average of Elevation (SRTM)	3.6	0.7	0.77	1.4
LST of Terra at 10:30 a.m. (MODIS)	3.4	2.5	0.80	1.7
PWV of Aqua at 01:30p.m. (MODIS)	3.2	2.3	0.82	1.8
LST of Aqua at 01:30 a.m. (MODIS)	3.1	2.3	0.83	1.4
LST of Aqua at 01:30p.m. (MODIS)	3.0	2.2	0.84	1.7
Precipitation, summer months average from 1990 to 2021 (TC)	3.0	2.2	0.84	1.6
% of Forest LULC within a 10 km buffer (NLCD)	2.9	2.1	0.85	1.2
% of Water LULC within a 10 km buffer (NLCD)	2.9	2.1	0.85	1.2
MERRA-only	RMSE (°C)	MAE (°C)	R^2	VIF
Surface layer height, Mean daily (MERRA-2)	2.8	2.2	0.85	3.3
Air temperature, Maximum daily (MERRA-2)	2.6	1.9	0.88	2.9
Precipitation, summer months average from 2000 to 2021 (TC)	2.5	1.8	0.89	1.7
% of Cropland LULC within a 10 km buffer (NLCD)	2.2	1.6	0.91	1.3
Specific humidity, Mean daily (MERRA-2)	2.2	1.5	0.92	4.2
% of Developed LULC within a 10 km buffer (NLCD)	2.1	1.5	0.92	1.3
Julian Day	2.0	1.4	0.92	1.0
Vapor pressure, summer months StDev from 2000 to 2021 (TC)	1.9	1.3	0.93	1.6
Wind speed, Maximum daily (MERRA-2)	1.8	1.3	0.93	1.1
Max Temp, summer months average from 2000 to 2021 (TC)	1.8	1.2	0.93	2.0

reanalysis-based estimation approach. The fully integrated SHINE model delivers both low error rate and kilometer-scale fidelity under different cloud contamination conditions, making it an operational approach for continental wide, wall-to-wall HI mapping.

3.4. Accuracy assessment across cloud conditions

The final SHINE approach included two variables that are susceptible to cloud interference, the daytime LSTs obtained by MODIS. Here, we present a sensitivity analysis to assess the impact of this cloud interference. Based on Table 5 only 40 percent (~1.8 million) of samples had valid values for both Aqua and Terra LST observations. The model took advantage of clear sky observations resulting in much lower errors compared to cloudy conditions. Unfortunately, it is common that when one MODIS LST is obscured the other daytime observation may also be missing resulting in 38 % of the total samples being affected. The errors there were the highest, however they still fell within the reasonable ranges of 2°C for the RMSE and below 2°C for the MAE. When only one daytime LST was missing the results were somewhere in between, with a slight edge in having the 10:30AM LST over the 1:30PM LST. For all conditions, the R^2 metric remained high showcasing that the predictor variables could explain the variation in the response variable in all cloud conditions.

3.5. Accuracy assessment across weather stations

Spatial analysis of error distribution across CONUS was carried out in Fig. 5. To create this figure, the SHINE RMSE was calculated for each of the stations across the entire time period. Measured RMSEs were then categorized under five main classes shown by distinct colors in the legend.

Overall, the results showed that nearly 70 percent of stations had RMSE below 2°C, while only 8 percent of stations exceeded 2.5°C. In general, stations with RMSE < 1.5°C were located in the Southwest-West regions and the Atlantic. The 1127 stations falling in the yellow class with RMSEs between 1.5 and 2°C had a stronger presence in the Midwest and Northeast areas. Moving to the worst performing stations in orange

Table 5
Influence of MODIS LST missing observations in model performance.

Scenario	Valid 10:30 am Terra LST	Valid 1:30 pm Aqua LST	Number of Samples (Percentages)	RMSE (°C)	MAE (°C)	R^2
1	Yes	Yes	1,811,605 (39 %)	1.6	1.2	0.93
2	Yes	No	568,615 (12 %)	1.9	1.3	0.93
3	No	Yes	485,682 (11 %)	2.0	1.4	0.92
4	No	No	1,748,338 (38 %)	2.2	1.6	0.92

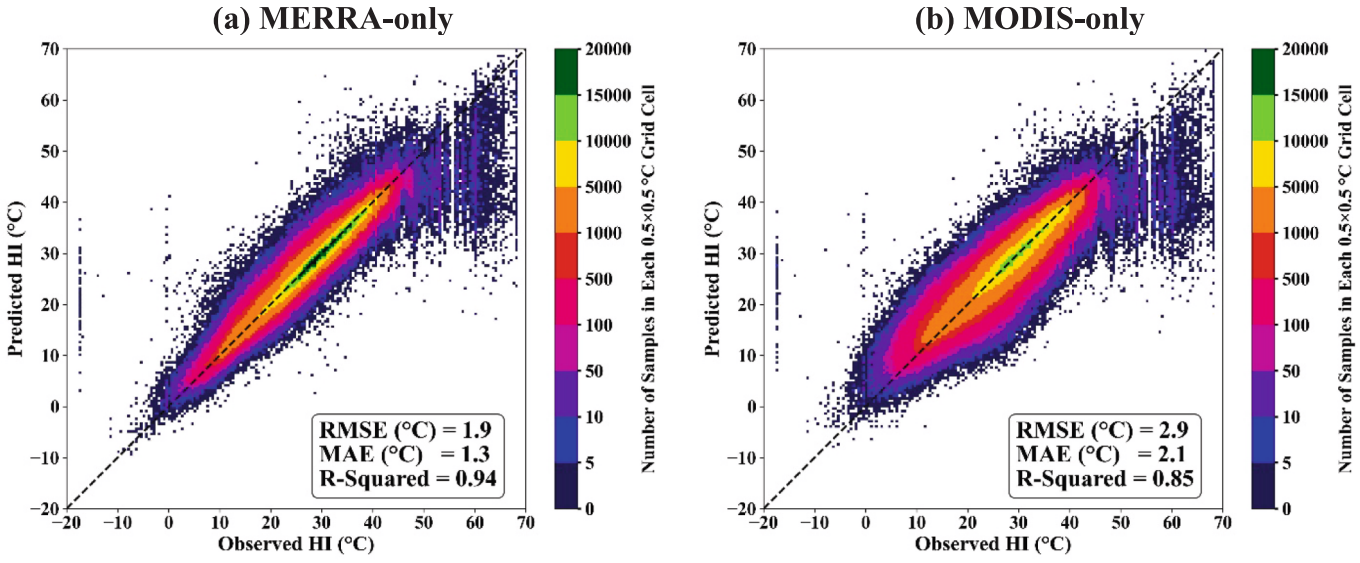

Fig. 4. Density–scatter plots of observed versus predicted daily-maximum HI for (a) MERRA-only and (b) MODIS-only.

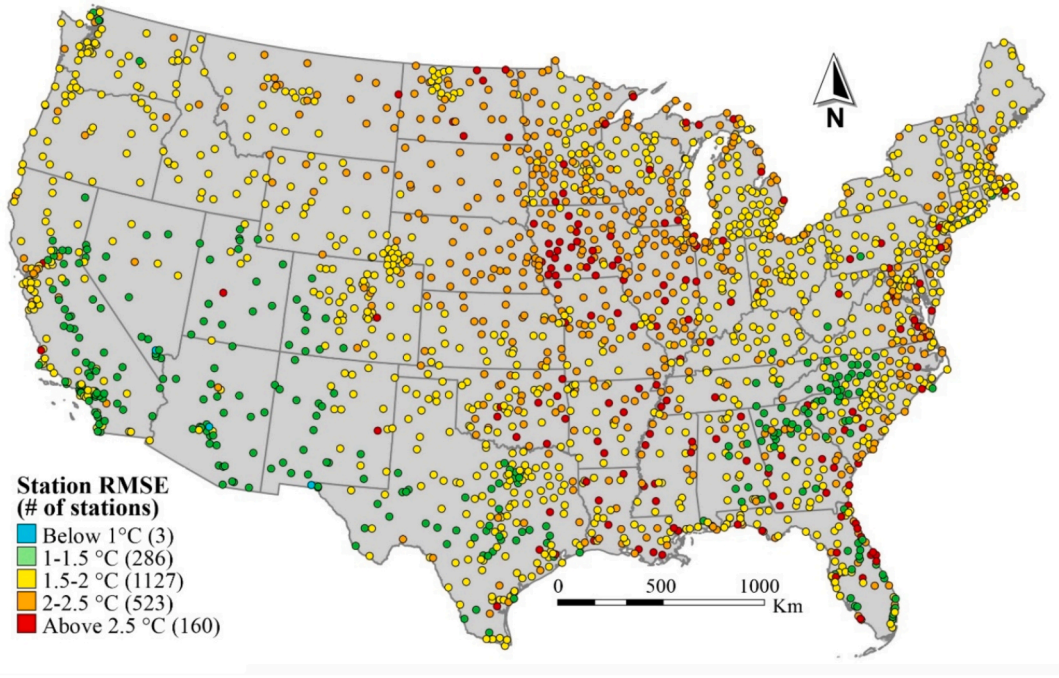

Fig. 5. Spatial distribution of per station RMSE (°C).

(RMSE 2°C–2.5°C) and red (RMSE over 2.5°C) classes, the number of stations dropped to 523 and 160, respectively. The majority of these stations were concentrated in the Dakotas, Iowa and Illinois. In terms of temporal residual analysis, temporal analysis revealed a consistent and stable residual structure across years. More specifically, evaluation of the annual RMSE showed no more than ~0.3°C variation from the 1.9°C baseline, therefore no evidence of performance degradation or overfitting to specific years.

3.6. Accuracy assessment across climatic zones

In this section error was assessed across different climatic zones. The intent was dual fold: identify climatic conditions of underperformance for future refinement and consider generalization capabilities to other regions outside our study area of the continental United States. From the calculation perspective, first the RMSE for every station was estimated, then the mean and standard deviation of the RMSEs for stations within each climate region were computed. Table 6 tabulated the mean and standard deviation of RMSEs for each climate region, corresponding number of stations and regions description based on this climate classifications. Based on the Köppen-Geiger classified map with resolution of approximately 90 m (0.00083 degrees), CONUS included a total of twenty different climate divisions (Beck et al., 2018). Here only divisions with 15 stations or higher were included to increase reliability of results.

cross the eleven Köppen-Geiger regions examined, mean RMSE stayed below 2.0°C in all arid environments, including the hot-desert (BWh, 1.3°C) and cold-desert (BWk, 1.5°C) classes, as well as the hot semi-arid steppe (BSh, 1.6°C). Mediterranean zones show slightly higher but still modest errors (~1.6°C) in the hot-summer (Csa) and warm-summer (Csb) types. Despite higher number of stations, with RMSEs of 2.0 and 2.1°C for Dfb and Dfa, respectively, humid-continental climates observed the highest mean RMSEs. The marine west-coast (Cfb) and subarctic dry-summer (Dsb) regions, each represented by fewer than 20 stations, fell in the mid-range with mean RMSEs of 1.8 to 1.9°C. The Cfb division showed the highest variability (StDev = 0.8°C) which may originate from the low number of associated samples. While summer humid continental (Dfb) and cold steppe (BSk) had higher number of included stations, the variability is higher (~0.6°C) compared to other climatic divisions. Overall, the SHINE approach delivered sub-1.5°C RMSE in dry and Mediterranean climates but increased to about 2.0°C in humid-subtropical, humid-continental and cold-steppe zones, underscoring those errors increased where atmospheric moisture and surface heterogeneity were high. Finally, the standard deviation highlighted model instability in marine west-coast regions, humid continental interiors, and cold semi-arid steppes.

Table 6
Error distribution of HI estimation across different climate regions defined by Köppen-Geiger classification over the CONUS.

Köppen-Geiger Climate divisions	Average ± StDev RMSE	Station counts per division	Climate condition description
BWh	1.3 ± 0.3°C	53	Hot desert
BWk	1.5 ± 02°C	49	Cold desert
BSh	1.6 ± 0.3°C	63	Hot semi-arid steppe
Csa	1.6 ± 0.4°C	43	Hot-summer Mediterranean
Csb	1.7 ± 0.2°C	71	Warm-summer Mediterranean
BSk	1.8 ± 0.6°C	239	Cold semi-arid steppe
Dsb	1.8 ± 0.4°C	16	Mediterranean-pattern sub-arctic
Cfb	1.9 ± 0.8°C	16	Marine west-coast (no dry season, warm summer)
Cfa	2.0 ± 0.3°C	750	Humid subtropical (no dry season, hot summer)
Dfb	2.0 ± 0.6°C	261	Humid continental warm-summer
Dfa	2.1 ± 0.3°C	466	Humid continental hot-summer

3.7. Accuracy assessment for extreme heat events detection

Since SHINE is using past observations, the intent is to create historical wall-to-wall maps that will assist policy makers in trend analysis. From that perspective, and as mentioned in the introduction, a critical

objective is the accurate and relevant detection of Extreme Heat Events (EHEs). Depending on the study objective, several HI thresholds have been applied to define and extract EHEs. Tuholske et al. (2021) identified EHEs with $HI_{max} \geq 40.6°C$ to gauge global urban exposure. Studies such as Madrigano et al. (2015) and Black (2010) identified HI equal to 35°C and above as the critical point where heat related illnesses and emergency visits start to spike. Other studies, for example from Metzger et al. (2010), showed that HI values between 35°C and ~37°C force a sharp increase in daily heat-related mortalities. Furthermore, comprehensive epidemiological studies conducted by Lewandowski and Shaman (2022) and Blum et al. (2024) identified the 105°F (~40.5°C) threshold as the upper limit of an extreme HI event. While based on NWS policies the ~40°C threshold could rise up to ~43°C for southern states (Hawkins et al., 2017), several studies including Vecellio et al. (2022) and Grundstein et al. (2012) demonstrated that HI values above ~40°C added no new physiological information that is not already included in previous hazard tiers.

Considering the above, here we present model performance for EHE binary classification using an HI threshold ranging from 35°C to 42°C. Table 7 presents the classification results. Recall indicates how many true EHEs the model correctly captures, while precision indicates how often an EHE detection is correct. Due to the fact that from the public health perspective it is better to err on unnecessary precautions rather than an unissued warning, the cost of a missed EHE (false negative) is greater than that of a false alarm (false positive). This makes recall a more important performance criterion. Up to 37°C, recall showed that 4/5ths of EHEs are correctly captured, with a similar performance for precision. Analyzing the number of EHE occurrences reveals that as the threshold increases, the EHE count rapidly drops from ~20 % to below 2 %. This sharp decline in the detected positives is the reason behind the recall considerable drop, while the decrease in precision remains more gradual.

Fig. 6 shows the tradeoff between Precision and Recall for an increase in the EHE definition threshold. As the trend lines show, near the 40°C threshold there is an accelerated non-linear drop in both recall and precision. At the highest examined value of 42°C approximately half of the actual EHEs were correctly identified. Looking beyond our modeling results, this analysis shows the sensitivity of the produced EHE accuracy to their definition, even for the same model. While the EHE thresholds are a convenient way to define them, it may be useful to look at continuous values and trends that SHINE can produce. From a public-health perspective, SHINE adopts the primary EHE threshold of 40°C (≈ 105°F), which highly aligns with practical heat warning and epidemiology studies. Several studies, such as Weinberger et al. (2018) and Le Roy et al. (2024), show that human health outcomes deteriorate rapidly beyond a HI of roughly 40°C as death rates show exponential rise. Similarly, Blum et al. (2024) also recognize the same HI threshold (~40.6°C) for an extreme humid-heat event as their results show significant spike in mortality risk among dialysis patients.

For better interpretability of the reported classification metrics in Table 7 and SHINEs ability to detect EHEs at the defined threshold, we considered a "no-skill" baseline model that classifies all samples as an EHE. Considering such baseline model is a standard practice across several domains with similar classification schemes (Ford et al., 2018; Lopez-Gomez et al., 2023; Saito and Rehmsmeier, 2015). Since this baseline model (classifier) catches all the actual EHEs, it has a perfect recall (=1.0). However, the precision would be as low as the actual event occurrences, since from all labeled EHEs only a certain low number are actually correct. At the defined optimal threshold (=40°C), precision of "no-skill" baseline is equal 0.04 as the prevalence rate is 4.0 %. Based on Table 7, at the same threshold, SHINE achieves precision of 0.75 and recall of 0.66, a ~19-fold precision lift over the baseline while retaining two-thirds of events, indicating substantial skill beyond class prevalence.

4. Discussion

4.1. Relevance of selected variables to HI

The variable selection process included both physical and statistical criteria. Based on the comprehensive candidate pool (Table 2) statistical assessment was conducted and at least one representative from each variable group was passed on to the FFS for further consideration. That way each physically meaningful variable had the statistical opportunity to be selected at the final pruning step. Sensitivity checks that reintroduced excluded, physical-motivated variables produced negligible accuracy gains, indicating that the final set preserved essential drivers while removing redundant features.

4.1.1. Statistical importance

To quantify the statistical relevance of the selected inputs, their relationship with observed station-level heat indicators was investigated. To do this, first-, second- and third-order polynomial functions were fitted between each predictor and station level Tmax, RH and HI. For each fit, the R^2 values were calculated and the highest R^2 among these three polynomial fits was selected as the best statistical relationship and tabulated in Table 8.

Surface layer height shows the strongest statistical alignment with observed HI and Tmax, underscoring its role in regulating near-surface heat entrapment. Specific humidity and MODIS LSTs exhibit moderate correlations with HI, Tmax and RH, consistent with their direct physical relevance to thermal stress (Yang et al., 2021). In contrast, variables such as Julian day and the forest and water fractions have very low R^2 values across all three heat indicators, suggesting that their contributions within SHINE primarily arise through indirect modulation of surface energy balance or humidity rather than through strong direct relationships with heat variables (Li et al., 2016; Wang et al., 2022b). Elevation shows nearly zero correlation with station-based HI and Tmax but a moderate association with RH, indicating its influence on background moisture climatology rather than on temperature directly (Feld et al., 2013; Gheyret et al., 2020). Finally, the frequent near-zero or negative correlations with RH (e.g., for historical maximum temperature or surface layer height) emphasize that RH alone cannot explain HI variability (Min et al., 2025; Parasin and Amnuaylojaroen, 2023) and that SHINE relies on multivariate interactions rather than simple univariate dependencies.

Figs. S1 and S2 in the appendix further investigate the relationships summarized in Table 8. These figures present scatter-density plots

Table 7
Classification performance of the HI model based on different extreme-heat thresholds.

Threshold (°C)	Precision "SHINE"	Precision "no-skill"	Recall	True Negative	False Positive	True Positive	False Negative	Total EHEs (% of EHEs)
35	0.88	0.20	0.85	3,572,123	108,647	135,638	797,832	933,470 (20.2 %)
36	0.86	0.16	0.83	3,778,160	99,087	121,929	615,064	736,993 (15.9 %)
37	0.84	0.12	0.81	3,965,296	85,980	108,888	454,076	562,964 (12.2 %)
38	0.83	0.09	0.76	4,122,715	67,497	103,781	320,247	424,028 (9.1 %)
39	0.79	0.06	0.72	4,272,183	53,597	81,592	206,868	288,460 (6.2 %)
40	0.75	0.04	0.66	4,388,556	41,116	62,628	121,940	184,568 (4.0 %)
41	0.71	0.02	0.60	4,472,133	27,695	45,634	68,778	114,412 (2.4 %)
42	0.70	0.01	0.52	4,522,963	17,010	35,498	38,769	74,267 (1.6 %)

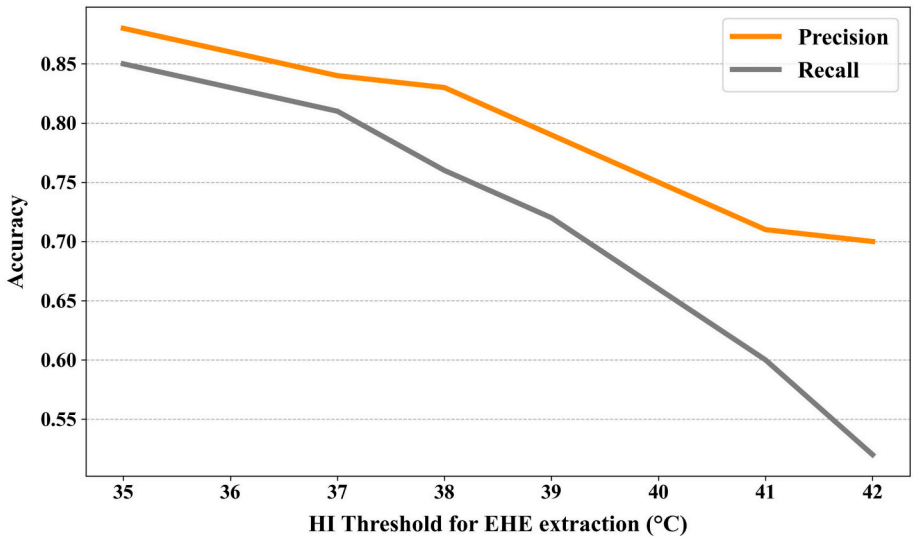

Fig. 6. Precision and Recall trend lines as a function of a changing EHE threshold.

Table 8
Statistical analysis between selected predictors and station-based HI, Tmax, and RH using Coefficient of determination (R^2).

Predictor variable	Observed HI	Tmax	RH
Surface layer height (MERRA-2)	0.86	0.77	0.01
LST of Terra (MODIS)	0.36	0.52	0.31
Average of Elevation (SRTM)	0.06	0.02	0.22
% Of Water LULC (NLCD)	0.01	0.03	0.08
Specific Humidity (MERRA-2)	0.42	0.20	0.30
LST of Aqua (MODIS)	0.32	0.49	0.34
Wind speed (TerraClimate)	0.05	0.04	0.01
Julian Day	0.21	0.18	0.01
Maximum Temp (TerraClimate)	0.34	0.31	0.02
% Of Forest LULC (NLCD)	0.02	0.03	0.07

showing how each environmental predictor relates to RH and Tmax. For RH, surface specific humidity and water land cover exhibit steep, saturating increases, consistent with their roles in supplying atmospheric moisture. In contrast, LST overpasses, surface layer height, and historical Tmax show decreasing RH trends, reflecting how intense surface heating and deeper boundary layers suppress near-surface humidity. Elevation also shows a gentle decline in RH, likely due to the cooler, drier air masses typical of higher altitudes. For Tmax, the dominant patterns shift toward thermal indicators as surface layer height, LSTs, and historical Tmax all show strong, monotonic increases, consistent with intensified heating under low mixing depths and higher skin temperatures. Interestingly, Tmax rises with increasing specific humidity up to a point, then slightly declines at the highest humidity levels, likely due to latent cooling. Forest and water land cover fractions tend to suppress Tmax but show limited or non-monotonic relationships with RH, reflecting their dual influence on energy balance and surface moisture.

4.1.2. Physical importance

The surface layer height, as one of two MERRA-based variables, was shown to be the most correlated variable with HI. This parameter directly governs the energy exchange, night time temperature trapping and evapotranspiration effect on humidity. Shallower height squeezes the heat and moisture into a smaller air volume, exacerbating their combined effects in the region (Budakoti and Singh, 2021; Miao et al., 2022). Specific Humidity (SH), the second selected MERRA-2 product provides information about the amount of latent heat in the air. Because relative humidity (RH) is temperature dependent, it can understate the moisture burden during hot events. In contrast, SH combines moisture content with temperature, providing a more direct measure of the body's ability to cool by evaporation and reduce heat stress (Bu et al., 2024; Coffel et al., 2017).

Evaluation of an alternative model that solely uses these MERRA-2 products (as the source of daily dynamic variables) provides a good initial HI estimation, however they are only offered at a coarse resolution of approximately 50 km. Studies such as Tuholske et al. (2021) also show that MERRA-2 generally underestimates temperature in case of high heat-humidity conditions but can be an influential input for HI calculation when used in combination with other input variables. Furthermore, a developed MODIS-only model (built on direct daily satellite observations) demonstrates acceptable estimation accuracy, but it struggles to capture the near surface moisture dynamics. MODIS PVW products do serve as moisture proxy, however, they represent water vapor based on atmospheric column integration which aggregates the moisture from earth surface up to the top of the atmosphere (Lin et al., 2013). This aggregation process reduces the applicability of PVW as a humidity proxy, since a dense cloud at higher altitudes can significantly inflate PVW while the surface humidity remains modest (Ruckstuhl et al., 2007; Zhao et al., 2022).

Based on this fact, integration of MERRA-2 reanalysis with MODIS products is a promising approach to provide an estimation model with high accuracy and spatial resolution. MODIS LST products were used for both their additional explanatory power (as their absence due to clouds shows their influence) and their ability to adjust locally the MERRA-2 estimation to a much finer 1 km cell size. Daily LST products from Aqua and Terra capture the surface temperature in the morning and at early afternoon, as well as the transition between these times for heat estimation (Albright et al., 2011; Benali et al., 2012). This daily LST pair has been widely used in heat studies as can be seen from Table 1. These MODIS- and MERRA-based variables were considered as dynamic variables, as they have several observations within a day capturing the within and intra-day variability.

In addition to the daily dynamic variables, several static variables were included. Among static variables, the percentage of a water and forest land cover in 5- and 10-kilometer radiuses, respectively, have the highest impact on HI estimation, compared to other tested land cover/use alternatives. Due to shading and albedo forest solely operates as a heat mitigation factor (Gage and Cooper, 2017; Gibbard et al., 2005). Proximity to water bodies can have a bidirectional HI effect. In case of high temperatures, the evaporation from water bodies could be a significant contributor to humidity increases, which exacerbates the heat

effect on human bodies (Jacobs et al., 2020; Yao et al., 2023). On the other hand, access to water bodies can serve as a heat mitigation factor since their high heat storage capacity keeps the surface cooler than adjacent land during the day. The created pressure gradient between land and water drives a breeze circulation that brings cooler marine air inland, often dropping the surface air temperature by several degrees (Jeong et al., 2015; Mutiibwa et al., 2015).

Furthermore, general climatic context can provide important insights. The summer average of two historical climatic variables, namely wind speed and maximum temperature, from TerraClimate were also found influential. These two variables act as static covariates that describe the background climate regime within which the day-to-day weather unfolds, thereby modulating the HI in complementary ways (Grotjahn and Huynh, 2018; Jackson et al., 2025). As another selected static variable, elevation was also shown to be a climate modifier factor by reducing the temperature-humidity pair as altitude increases (Krzysztof et al., 2021; Raymond et al., 2022). Finally, as the least computationally demanding variable, the Julian day was also shown to be an influential parameter for HI estimation. As an integer day-of-year counter, it encapsulates the seasonal phase of the Earth–sun geometry, which affects daily top-of-atmosphere insolation and drives the average pattern variability of near-surface temperature and atmospheric moisture, the two variables that are used to measure HI (Emamifar et al., 2013).

4.2. Evaluation of the model's comparative and practical value

In general, our literature review indicated a limited number of direct HI estimation efforts using satellite observations. A similar study conducted by Pede and Mountrakis (2022) led to an RMSE of 2.4°C and an R^2 of 0.83. While the geographic area was also CONUS, that work obtained lower accuracy while also limiting the evaluation to a single year, a much more favorable sampling that only included clear sky days, and one-tenth the number of observational sites. Due to the limited number of studies on HI estimation, the results of the present study can be compared with those based on alternative apparent temperature indices, among which the Universal Thermal Climate Index (UTCI) and Canada's Humidex are two prominent examples. Ho et al. (2016) estimated the Humidex heat detection criteria across Greater Vancouver, Canada and their accuracy assessment estimated an RMSE of 2.0°C and R^2 of 0.57. Wu et al. (2023) also employed a random forest based model to estimate Humidex over the Yangtze River Delta for four selected months (Jan, Apr, Jul and Oct) in 2020, reporting an RMSE ranging from 2.2 to 3.2°C (4.0–5.8°F) and R^2 between 0.86 and 0.93. While similar to Pede and Mountrakis (2022), these studies solely considered cloud free samples and their geographic coverage entailed only a limited number of land covers/uses. Furthermore, Hiroki et al. (2025) estimated the UTCI over 140 stations in Japan between 1980 and 2020, and their developed XGboost model returned an RMSE of 6.8°C and an R^2 of 0.92. While this study delivers the first 41-year, nation-wide record of UTCI change for Japan, its explanatory power and accuracy were constrained by the low number of observations, spatially coarse inputs and limited inclusion of land cover and climate-oscillation drivers.

4.2.1. Comparisons with other high spatial resolution models

Due to limitations in HI data availability, studies have been using land surface temperature (LST) and daily maximum air temperature (Tmax) as proxies. As our dataset can extract LSTs from MODIS and Tmax from the weather stations we have a unique opportunity to contrast here these commonly used approaches to both our estimated HI via SHINE and the reference HI data on these weather stations. Fig. 7 shows the error distribution of the three approaches (SHINE HI, Tmax and MODIS Aqua LST) within successive, fixed-width intervals of weather station HI as the reference variable. Each boxplot includes the median, inter-quartile range (IQR), and tails for its bin, allowing quick visual assessment of how both the magnitude and variability of the error change across the reference HI range. For each 5°C HI bins between 10°C and 60°C, grey bars (on the right y-axis) show the sample frequency in millions.

Due to the wall-to-wall global availability researchers have utilized MODIS LST for EHE identification through flagging multi-day exceedances above historical norms (Albright et al., 2011), tracking peak daily LSTs (Cotlier and Jimenez, 2022), detecting LST anomalies in the Mediterranean (Agathangelidis et al., 2022), and mapping continental-scale thermal extremes (Metz et al., 2017). Fig. 6 indicates that under extreme heat conditions (HI > 35°C), the LST-HI relationship becomes more consistent, which is possibly due to stronger surface–air coupling in hot and dry conditions (Dirmeyer et al., 2022; Ganeshi et al., 2022). However, LST use is not suggested as a good proxy for EHE detection due to high error variability. While the median error could be adjusted with appropriate normalizations across the HI range, the high variability cannot be addressed as a wide range of LST values could lead to very

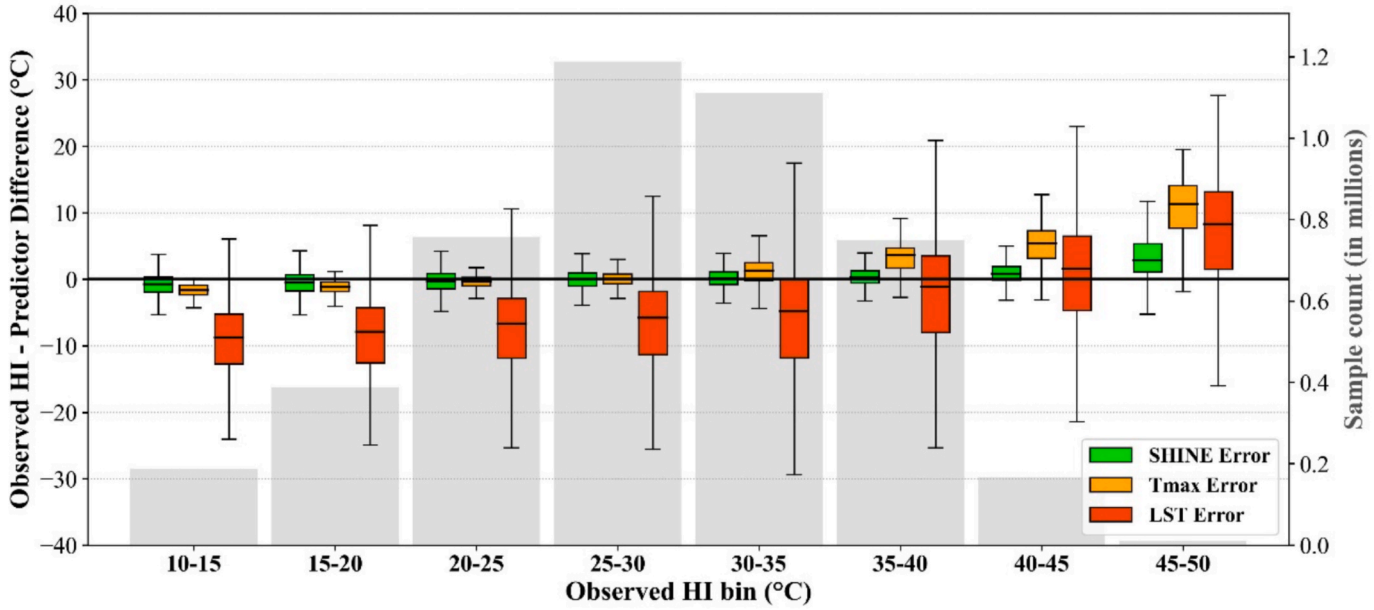

Fig. 7. Contrasting SHINE predictions with Air and Land temperature estimations for EHE detection.

similar HI estimations. This wide error variability is consistent across all HI ranges.

Although the HI estimation is conceptually distinct from estimating daily Tmax, the latter is often employed to identify EHEs in a wide range of interdisciplinary studies. This could be problematic as Tmax only captures the sensible-heat component of thermal stress and completely ignores the incremental role of moisture content in the air (Zhang et al., 2014). As a result, the disparities between Tmax-based and HI-based EHE detection are evident in the extreme heat-humidity tail (HI > 35°C), where public health interventions are most critical (Lee et al., 2021). Consequently, Tmax should not be used as the sole criterion for EHE identification, as it systematically underestimates and could miss dangerous hot-humid events. Another important consideration is that the presented Tmax analysis used weather station Tmax data, so from a practical perspective to create a wall-to-wall implementation a spatial distribution model would introduce additional variability.

In comparison the SHINE HI error analysis shows that the median error remains near zero up to the 40°C-45°C bin with a small IQR spread (<±2°C), indicating accurate performance over considered EHE thresholds. This further confirms that SHINE exhibits a low to moderate homoscedastic error across the public health warning range.

4.2.2. Comparisons with other high accuracy models

Tables 3 and 4 indicate that the full SHINE model exhibits very similar error to a MERRA-only model that does not incorporate MODIS products. While the latter model would reduce workload and dependencies on MODIS, the full SHINE model takes advantage of the MODIS 1 km products to capture fine-scale spatial detail. Fig. 8 visually contrasts the MERRA-only and the full SHINE HI outputs across the state of New Mexico (top row) and within the urban environment of Albuquerque (bottom row). These maps were generated using the trained models of each configuration with the selected input variables noted in Tables 3 and 4. To highlight HI variation, July 18th, 2012, one of the hottest summers in the U.S., was chosen as a case study (Anagnostopoulos et al., 2022; Romps and Lu, 2022). While MERRA-only provides accurate estimations (Fig. 4), its coarse resolution (~50 km) fails to capture important fine-scale thermal contrasts between the Albuquerque urban area and its surrounding desert and mountainous landscapes. Based on the reference spatial resolution of MERRA-2 products, the CONUS domain is represented by approximately 2700 pixels, whereas the full SHINE captures the same domain with nearly 12 million pixels, providing much finer spatial detail essential for local-scale exposure assessment. Together, these maps highlight the complementary strengths of the reanalysis-based and integrated modeling approaches and reinforce the need for a fused framework that combines accuracy with high spatial fidelity.

To further demonstrate the advantages of SHINE, its conceptual framework could be compared with similar heat-related spatiotemporal fusion models. For example, studies such as Meng et al. (2025), Guo et al. (2024) and Tang et al. (2023) produced 1-km gridded LST datasets with global coverage by fusing meteorological products from Landsat, MODIS and Moderate Resolution Spectral Imager (MERSI). These LST-fusion models focused on reconstructing LST and are considered valuable and accurate for near surface temperature mapping. SHINE on the other hand is designed for heat exposure estimation, where HI estimation is practically superior to LST estimations as Fig. 7 demonstrates.

SHINE is the first HI-based national scale dataset at 1 km scale, however, meteorological datasets such as PRISM and Daymet are also widely used to quantify the heat-exposure with similar spatial and temporal resolutions (Clark et al., 2024; Manware et al., 2022). PRISM and Daymet could theoretically provide HI estimations through RH estimations with the Daymet's vapor pressure or PRISM's dew-point variables. The issue though would be that PRISM and Daymet utilize all the available weather stations in their spatial interpolation models through the ASOS and AWOS networks (Thornton et al., 2021), therefore there would be no stations left to conduct an independent assessment between these models and SHINE, which are also utilized to define SHINE's HI

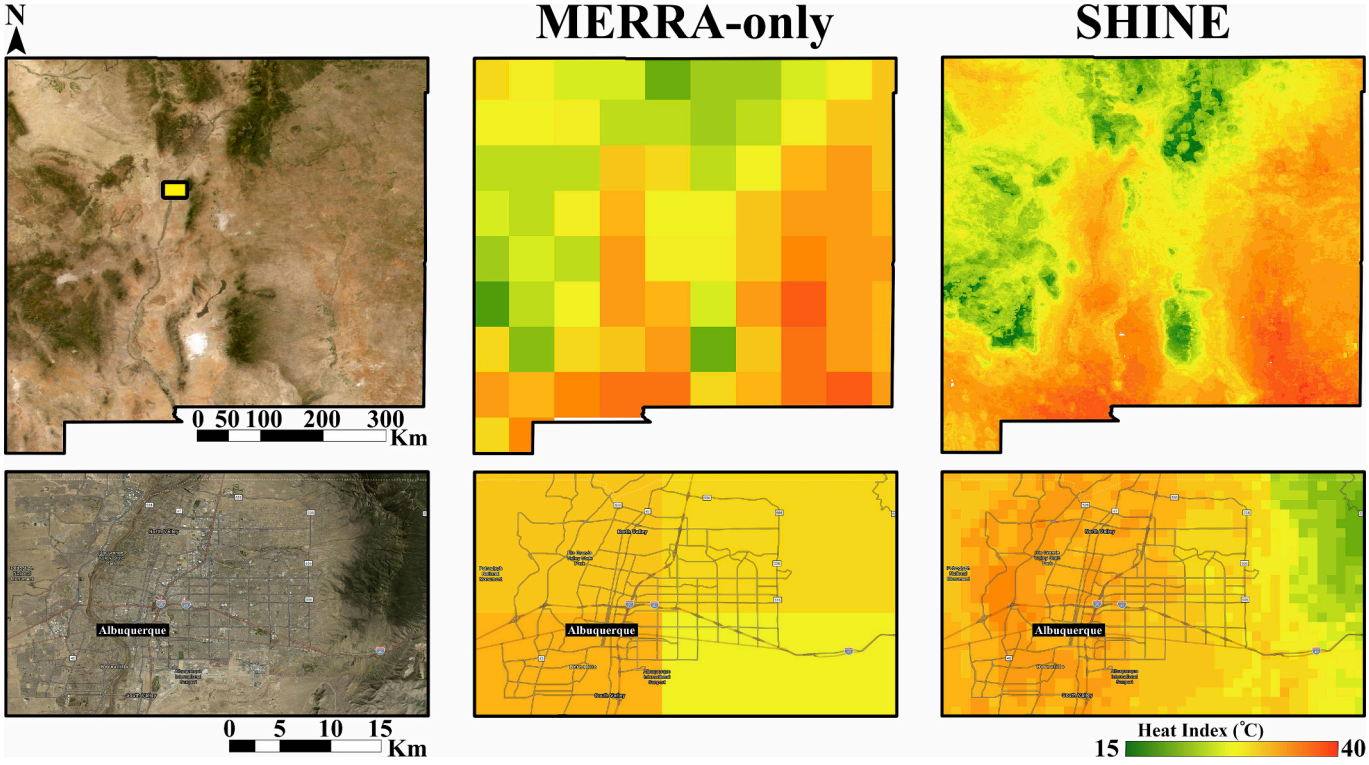

Fig. 8. Comparison between MERRA-only and SHINE spatial resolution in HI (°C) estimation on July 18, 2012, across New Mexico. The yellow box in the top-left map marks the Albuquerque metropolitan area depicted in the bottom row. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ground truth.

4.3. Assessment of spatial error patterns in HI estimations

Station-level error assessment showed a large portion of high error stations concentrated in the Midwest region, an area known as corn-belt (Fig. 5). While additional variables were tested to capture the heavily agricultural nature of these underperforming regions, they were not successful. A potential contributing factor is the high evapotranspiration rate in mid-summer caused from vast corn and soy agricultural fields (Ting et al., 2021). This factor forces a spike in surface dew-point temperature which is often smoothed by gridded meteorological products (Alter et al., 2018; Souri et al., 2020). Due to this smoothing effect, the model is fed with lower humidity values than those that actually occur, leading to HI underestimation during peak events, a behavior also observed by our model. Furthermore, this region is impacted on both sides by extreme Gulf moisture and Canadian air masses (Haigh et al., 2015; Rohli et al., 2024). Unlike deserts in southwest and plains of southeast, which have relatively stable climate governance, Midwest and Northeast regions can be exposed to hot-dry or hot-humid condition dynamics within a short time span (Agel et al., 2021; Archer and Predick, 2008). This rapid change in climate condition poses heteroscedasticity which is difficult for a regression model to capture (Gelfand, 2015; Paliwal and Kumar, 2011; Sall', 2015).

A climate-specific analysis of model residuals by Köppen–Geiger climate class reveals systematic, climate-dependent patterns in HI estimation error. Evaluation of model performance across different Köppen-Geiger regions showed low errors in desert and Mediterranean climates reflecting their relative thermodynamic simplicity. Low moisture content keeps HI close to air temperature, so uncertainty in humidity contributes little additional variance (Alfano et al., 2011; Anderson et al., 2013). By contrast, the Midwest and Southeast regions, which are represented in the Cfa, Dfa, and Dfb climatic divisions, couple high atmospheric humidity with sharp synoptic swings between Gulf and continental air masses, amplifying the HI variability and producing both higher means and broader error ranges (Obarein and Lee, 2025; Raymond et al., 2024). The exceptionally wide spreads in the Cfa and BSk divisions further point to sub-grid heterogeneity: "corn-sweat" moisture spikes in irrigated farmland and patchy irrigation across High-Plain's steppes can challenge a model calibrated on 1-km grids (Gibson et al., 2017; Qiao et al., 2022). Given that the three most populated classes (Cfa, Dfa, Dfb) contain nearly 1500 stations and dominate national statistics, future error reductions should prioritize improved boundary-layer humidity representation and land-cover heterogeneity in these humid and transition-zone climates rather than in already well-modelled arid regions.

Several underlying mechanisms could possibly lead to higher errors in these specific regions. At high temperature conditions, HI grows non-linearly with the humidity and dew point (Xu et al., 2025b). As a result, in moist climate, marginal error in RH will disproportionally propagate to estimated HIs and shift them by multiple °Cs (Simpson et al., 2023). Additionally, those humid regions have frequent and fast-evolving cloudy condition (Feng et al., 2019; Lynn et al., 2001; Romps et al., 2021), which reduces the MODIS clear sky availability and increases the model's reliance on coarser MERRA-2 products. While our error analysis showed that MERRA2-only has close performance with SHINE, its coarser resolution is unable to capture sharp moisture gradients which are caused by sea-lake breezes, frontal zones and convective outflow boundaries at small spatial scales (Apke et al., 2020; Justin et al., 2025; Wang et al., 2023a). To overcome this issue, the model needs to go through a smoothening stage, which increases HI estimation error. Finally, the dominant LULCs in Cfa and Dfa climatic regions are generally forests and crops lands with strong latent heat flux at midday (Hiestand and Carleton, 2020; Jansen et al., 2023). While this flux has little effect on air and dew point temperatures, it reduces the LST through evaporation and weakens its linkage to HI.

To assess whether local LULC composition contributed to the elevated errors in the Midwest, an additional diagnostic was employed which used station-level HI errors and LULC shares within 1 km buffer. For each station, the distribution of SHINE errors was calculated and then linked to binned fractions of cropland, developed, forest, and barren LULCs. Results show very similar error distributions across LULC fraction bins, with no systematic shift in bias or variance as the share of any class increases. This suggests that local LULC composition is not a primary driver of the higher errors in these regions, and that atmospheric and radiative factors are more likely responsible.

The modest underestimation of extreme HI in those regions indicates that SHINE may yield conservative estimates in threshold-based EHE detection, and errors in regions with sharp moisture gradients can restrict its reliability for very fine-scale operational applications. Understanding these behaviors is important both for guiding future methodological improvements and for ensuring that appropriate caution would be considered for interpreting SHINE in such environments.

4.4. Limitations and future suggestions

The targeted 1 km^2 spatial resolution, while an advancement for the current state-of-the-art, may be inadequate to capture the micro-climate fluctuation in small heterogenous regions. Leveraging higher-resolution LST from Landsat (~100 m) and ECOSTRESS (~70 m) can refine the temperature component of HI at sub-kilometer scale. This will help SHINE resolve sharp land–water, urban–rural, and terrain-driven gradients in hot-humid regions that are under-represented currently at 1 km. Another potential issue is the decommission of MODIS and potential replacement by VIIRS products. After 2022, the MODIS overpass times have drifted (10:30am to ~9:30am for Terra and from 1:30 pm to ~2:30 pm for Aqua), therefore the potential of future MODIS/VIIRS use should be investigated.

Another approach to reduce the cloud contamination is the use of advanced machine learning imputation models. These models consider the entire time-series and predict missing values. While they have been used by the remote sensing community to fill the gaps in LST products of Landsat (Tang, 2024; Wang et al., 2024) and MODIS (Mokari et al., 2021; Xiao et al., 2021), their ability to improve the estimation accuracy is still under-studied.

Although SHINE demonstrates superior accuracy relative to comparable studies, error rates are higher in humid climates, notably across the agriculture-dominated Midwest. While several scientific factors likely contribute to this underperformance, a more detailed, region-targeted investigation is required to identify and mitigate the sources of error and uncertainty and to further improve model's performance. One approach is the application of dataset fusion, which could further reduce the pixel-based estimation bias of SHINE by including fields from gridded datasets such as PRISM and Daymet. Produced HIs from these sources can be incorporated as auxiliary covariates within SHINE to exploit their spatial coherence in data-sparse regions while preserving independence in evaluation, possibly through a CoKriging statistical estimation.

Finally, the imbalance of HI samples across the entire operating range poses modeling constraints, most notably the low frequency of extreme events. While SHINE used calculated HI from ASOS and AWOS stations as the target variable, addition of samples from national scale networks such as Remote Automated Weather Stations (RAWS) and more localized and state-specific Mesoscale Meteorological Network (Mesonet) can increase the sample size of those under-sampled high HI occurrences. Application of synthetic sample generators and neural network based resampling methods, such as Generative Adversarial Networks, could offer improvements. These methods have been successful in generating new samples in remote sensing tasks (Han et al., 2020; Ma et al., 2019; Reyes et al., 2022); however, they are complex and computational demanding.

5. Conclusions

This study presented a model for CONUS-wide, wall-to-wall daily HI estimation with 1 km resolution spanning 21 seasons from 2003 to 2023. The obtained overall RMSE of 1.9°C (MAE = 1.4°C, $R^2 = 0.93$) suggests that SHINE could provide accurate HI estimations. A comparative analysis was conducted to demonstrate the trade-offs between different input alternatives. The MERRA-only model demonstrated slightly more accurate HI estimation but at a much coarser spatial resolution, while the MODIS-only model retained kilometer-scale detail but with a higher error. The proposed integrated SHINE framework reconciles these limitations, recovering fine-scale spatial scale without sacrificing accuracy, and thus represents a balanced operational solution. Furthermore, SHINE retains high accuracy under different cloud contamination scenarios enabling dependable estimates where satellite products may be missing. Analysis on the spatial and climatic distribution of the error indicated good performance across the board with the highest potential for improvement in the humid Midwest and Northeast, where "corn-sweat" evapotranspiration and sharp air-mass transitions may amplify humidity uncertainty. From the policy perspective and considering the HI operational range for public-health heat effects, the proposed SHINE approach outperforms typically used proxies such as land and air temperature. Moreover, SHINE reliably detected Extreme Heat Events, maintaining strong EHE correct identification up to the 37°C threshold and operational accuracy at the critical 40°C level used in health warning systems.

The study demonstrated that integrating multi-source Earth-observation data with machine learning algorithms can deliver operational, continent scale HI estimations of sufficient accuracy for public health decision-making, climate risk attribution and adaptation research. Expanding and refining this framework through global deployment, real-time updating and coupled health-impact modelling represents a promising pathway toward more proactive and precise heat-risk management in a warming world.

CRediT authorship contribution statement

Seyed Babak Haji Seyed Asadollah: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation. **Giorgos Mountrakis:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Stephen B. Shaw:** Writing – review & editing, Supervision.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2026.01.018.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5, 1–12.

Adab, H., Kanniah, K.D., Solaimani, K., Tan, K.P., 2013. Estimating atmospheric humidity using MODIS cloud-free data in a temperate humid region. In: 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS. IEEE, pp. 1827–1830.

Adinolfi, M., Raffa, M., Reder, A., Mercogliano, P., 2023. Investigation on potential and limitations of ERA5 Reanalysis downscaled on Italy by a convection-permitting model. Clim. Dyn. 61, 4319–4342.

Agathangelidis, I., Cartalis, C., Polydoros, A., Mavrakou, T., Philippopoulos, K., 2022. Can satellite-based thermal anomalies be indicative of heatwaves? an investigation for MODIS land surface temperatures in the Mediterranean region. Remote Sens. 14, 3139.

Agel, L., Barlow, M., Skinner, C., Colby, F., Cohen, J., 2021. Four distinct Northeast US heat wave circulation patterns and associated mechanisms, trends, and electric usage. npj Clim. Atmos. Sci. 4, 31.

Aghazadeh, F., Bageri, S., Garajeh, M.K., Ghasemi, M., Mahmodi, S., Khodadadi, E., Feizizadeh, B., 2023. Spatial-temporal analysis of day-night time SUHI and its relationship between urban land use, NDVI, and air pollutants in Tehran metropolis. Appl. Geomat. 15, 697–718.

Akter, T., Gazi, M.Y., Mia, M.B., 2021. Assessment of land cover dynamics, land surface temperature, and heat island growth in northwestern Bangladesh using satellite imagery. Environ. Process. 8, 661–690.

Albert, P., Bennartz, R., Preusker, R., Leinweber, R., Fischer, J., 2005. Remote sensing of atmospheric water vapor using the moderate resolution imaging spectroradiometer. J. Atmos. Ocean. Technol. 22, 309–314.

Albright, T.P., Pidgeon, A.M., Rittenhouse, C.D., Clayton, M.K., Flather, C.H., Culbert, P. D., Radeloff, V.C., 2011. Heat waves measured with MODIS land surface temperature data predict changes in avian community structure. Remote Sens. Environ. 115, 245–254.

Alfano, d'Ambrosio, F.R., Palella, B.I., Riccio, G., 2011. Thermal environment assessment reliability using temperature—humidity indices. Ind. Health 49, 95–106.

Alqasemi, A.S., Hereher, M.E., Al-Quraishi, A.M.F., Saibi, H., Aldahan, A., Abuelgasim, A., 2022. Retrieval of monthly maximum and minimum air temperature using MODIS aqua land surface temperature data over the United Arab Emirates. Geocarto Int. 37, 2996–3013.

Alter, R.E., Douglas, H.C., Winter, J.M., Eltahir, E.A.B., 2018. Twentieth century regional climate change during the summer in the Central United States attributed to agricultural intensification. Geophys. Res. Lett. 45, 1586–1594.

Anagnostopoulos, G.C., Menesidou, S.A.I., Efthymiadis, D.A., 2022. The March 2012 heat wave in Northeast America as a possible effect of strong solar activity and unusual space plasma interactions. Atmosphere 13 (6), 926.

Anderson, G.B., Bell, M.L., Peng, R.D., 2013. Methods to calculate the heat index as an exposure metric in environmental health research. Environ. Health Perspect. 121, 1111–1119.

Angelini, L.P., Biudes, M.S., Machado, N.G., Geli, H.M.E., Vourlitis, G.L., Ruhoff, A., Nogueira, J. de S., 2021. Surface Albedo and temperature models for surface energy balance fluxes and evapotranspiration using SEBAL and Landsat 8 over Cerrado-Pantanal, Brazil. Sensors 21, 7196.

Apke, J.M., Hilburn, K.A., Miller, S.D., Peterson, D.A., 2020. Towards objective identification and tracking of convective outflow boundaries in next-generation geostationary satellite imagery. Atmos. Meas. Tech. 13, 1593–1608.

Archer, S.R., Predick, K.I., 2008. Climate change and ecosystems of the southwestern United States. Rangelands 30, 23–28.

Asadollah, S.B.H.S., Khan, N., Sharafati, A., Shahid, S., Chung, E.-S., Wang, X.-J., 2021. Prediction of heat waves using meteorological variables in diverse regions of Iran with advanced machine learning models. Stoch. Environ. Res. Risk Assess. 1–16.

Ballester, J., Quijal-Zamorano, M., Méndez Turrubiates, R.F., Pegenaute, F., Herrmann, F.R., Robine, J.M., Basagaña, X., Tonne, C., Antó, J.M., Achebak, H., 2023. Heat-related mortality in Europe during the summer of 2022. Nat. Med. 29, 1857–1866.

Balram, D., Lian, K.-Y., Sebastian, N., 2019. Air quality warning system based on a localized PM2. 5 soft sensor using a novel approach of Bayesian regularized neural network via forward feature selection. Ecotoxicol. Environ. Saf. 182, 109386.

Barbosa, N., Andreani, L., Gloaguen, R., Ratschbacher, L., 2021. Window-based morphometric indices as predictive variables for landslide susceptibility models. Remote Sens. 13, 451.

Baronian, I., Borna, R., Jafarpour Ghalehteimouri, K., Zohoorian, M., Morshedi, J., Khaliji, M.A., 2024. Unveiling the thermal impact of land cover transformations in Khuzestan province through MODIS satellite remote sensing products. Paddy Water Environ, 22, 503–520.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 1–12.

Bell, M.L., O'neill, M.S., Ranjit, N., Borja-Aburto, V.H., Cifuentes, L.A., Gouveia, N.C., 2008. Vulnerability to heat-related mortality in Latin America: a case-crossover study in Sao Paulo, Brazil, Santiago, Chile and Mexico City. Mexico. Int. J. Epidemiol. 37, 796–804.

Benali, A., Carvalho, A.C., Nunes, J.P., Carvalhais, N., Santos, A., 2012. Estimating air surface temperature in Portugal using MODIS LST data. Remote Sens. Environ. 124, 108–121.

Beniston, M., Stephenson, D.B., 2004. Extreme climatic events and their evolution under changing climatic conditions. Glob. Planet. Change 44, 1–9.

Benmarhnia, T., Schwarz, L., Nori-Sarma, A., Bell, M.L., 2019. Quantifying the impact of changing the threshold of New York City heat emergency plan in reducing heat-related illnesses. Environ. Res. Lett. 14, 114006.

Black, H., 2010. When to warn? Comparing heat indices to evaluate public health risks.

Blum, M.F., Feng, Y., Tuholske, C.P., Kim, B., DeMarco, M.A.M., Astor, B.C., Grams, M.E., 2024. Extreme humid-heat exposure and mortality among patients receiving dialysis. Am. J. Kidney Dis. 84, 582–592.

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.

Bu, Y., Sun, Z., Tao, Y., Zhao, X., Zhao, Y., Liang, Y., Hang, X., Han, L., 2024. The synergistic effect of high temperature and relative humidity on non-accidental deaths at different urbanization levels. Sci. Total Environ. 940, 173612.

Budakoti, S., Singh, C., 2021. Examining the characteristics of planetary boundary layer height and its relationship with atmospheric parameters over Indian sub-continent. Atmos. Res. 264, 105854.
Burkart, K., Schneider, A., Breitner, S., Khan, M.H., Krämer, A., Endlicher, W., 2011. The effect of atmospheric thermal conditions and urban thermal pollution on all-cause and cardiovascular mortality in Bangladesh. Environ. Pollut. 159, 2035–2043.
Cai, Y., Chen, G., Wang, Y., Yang, L., 2017. Impacts of land cover and seasonal variation on maximum air temperature estimation using MODIS imagery. Remote Sens. 9, 233.
Cai, Z., Guldmann, J.-M., Tang, Y., Han, G., 2022. Does city-water layout matter? Comparing the cooling effects of water bodies across 34 chinese megacities. J. Environ. Manage. 324, 116263.
Cebrián, A.C., Asín, J., Gelfand, A.E., Schliep, E.M., Castillo-Mateo, J., Beamonte, M.A., Abaurrea, J., 2022. Spatio-temporal analysis of the extent of an extreme heat event. Stoch. Environ. Res. Risk Assess. 36, 2737–2751.
Center, S.W.P., 2012. National oceanic and atmospheric administration.
Chapman, S.C., Watkins, N.W., Stainforth, D.A., 2019. Warming trends in summer heatwaves. Geophys. Res. Lett. 46, 1634–1640.
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, pp. 785–794.
Chen, C., He, Q., Li, Y., 2024. Downscaling and merging multiple satellite precipitation products and gauge observations using random forest with the incorporation of spatial autocorrelation. J. Hydrol. 632, 130919.
Chen, K., Horton, R.M., Bader, D.A., Lesk, C., Jiang, L., Jones, B., Zhou, L., Chen, X., Bi, J., Kinney, P.L., 2017. Impact of climate change on heat-related mortality in Jiangsu Province, China. Environ. Pollut. 224, 317–325.
Chen, A., Huang, L., Liu, Q., Piao, S., 2021. Optimal temperature of vegetation productivity and its linkage with climate and elevation on the Tibetan Plateau. Glob. Chang. Biol. 27, 1942–1951.
Chen, Y., Sun, H., Li, J., 2016. Estimating daily maximum air temperature with MODIS data and a daytime temperature variation model in Beijing urban area. Remote Sens. Lett. 7, 865–874.
Cheng, J., Sun, J., Yao, K., Xu, M., Cao, Y., 2022. A variable selection method based on mutual information and variance inflation factor. Spectrochim. Acta A Mol. Biomol. Spectrosc. 268, 120652.
Chinchanikar, S., Shaikh, A.A., 2022. A review on machine learning, big data analytics, and design for additive manufacturing for aerospace applications. J. Mater. Eng. Perform. 31, 6112–6130.
Chowdhury, M.Z.I., Leung, A.A., Walker, R.L., Sikdar, K.C., O'Beirne, M., Quan, H., Turin, T.C., 2023. A comparison of machine learning algorithms and traditional regression-based statistical modeling for predicting hypertension incidence in a Canadian population. Sci. Rep. 13, 13.
Chung, J., Lee, Y., Jang, W., Lee, S., Kim, S., 2020. Correlation analysis between air temperature and MODIS land surface temperature and prediction of air temperature using TensorFlow long short-term memory for the period of occurrence of cold and heat waves. Remote Sens. 12, 3231.
Clark, A., Grineski, S., Curtis, D.S., Cheung, E.S.L., 2024. Identifying groups at-risk to extreme heat: intersections of age, race/ethnicity, and socioeconomic status. Environ. Int. 191, 108988.
Coffel, E.D., Horton, R.M., De Sherbinin, A., 2017. Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century. Environ. Res. Lett. 13, 14001.
Cotlier, G.I., Jimenez, J.C., 2022. The extreme heat wave over western North America in 2021: an assessment by means of land surface temperature. Remote Sens. 14, 561.
Coughlan de Perez, E., Arrighi, J., Marunye, J., 2023. Challenging the universality of heatwave definitions: gridded temperature discrepancies across climate regions. Clim. Change 176, 167.
Cristo, R.D., Mazzarella, A., Viola, R., 2007. An analysis of heat index over Naples (Southern Italy) in the context of European heat wave of 2003. Nat. Hazards 40, 373–379.
De Meij, A., Vinuesa, J.F., 2014. Impact of SRTM and Corine Land Cover data on meteorological parameters using WRF. Atmos. Res. 143, 351–370.
Deppner, J., Cajias, M., 2024. Accounting for spatial autocorrelation in algorithm-driven hedonic models: a spatial cross-validation approach. J. Real Estate Financ. Econ. 68, 235–273.
Dirmeyer, P.A., Sridhar Mantripragada, R.S., Gay, B.A., Klein, D.K.D., 2022. Evolution of land surface feedbacks on extreme heat: Adapting existing coupling metrics to a changing climate. Front. Environ. Sci. 10, 949250.
Dobrowski, S.Z., 2011. A climatic basis for microrefugia: the influence of terrain on climate. Glob. Chang. Biol. 17, 1022–1035.
Dorji, N., Awange, J.L., Zerihun, A., 2025. Reliability of satellite, reanalysis and observation-based gridded temperature datasets for climate change impact studies in Bhutan. Sci. Rem. Sens., 100275
Emamifar, S., Rahimikhoob, A., Noroozi, A.A., 2013. Daily mean air temperature estimation from MODIS land surface temperature products based on M5 model tree. Int. J. Clim. 33, 3174–3181.
Engdaw, M.M., Ballinger, A.P., Hegerl, G.C., Steiner, A.K., 2022. Changes in temperature and heat waves over Africa using observational and reanalysis data sets. Int. J. Climatol. 42, 1165–1180.
Ermida, S.L., Soares, P., Mantas, V., Göttsche, F.-M., Trigo, I.F., 2020. Google earth engine open-source code for land surface temperature estimation from the landsat series. Remote Sens. 12, 1471.
Espinosa, L.A., Portela, M.M., 2025. Red-hot Portugal: mapping the increasing severity of exceptional maximum temperature events (1980–2024). Atmosphere (Basel). 16, 514.
Feld, S.I., Cristea, N.C., Lundquist, J.D., 2013. Representing atmospheric moisture content along mountain slopes: examination using distributed sensors in the Sierra Nevada, California. Water Resour. Res. 49, 4424–4441.
Feng, Z., Houze Jr, R.A., Leung, L.R., Song, F., Hardin, J.C., Wang, J., Gustafson Jr, W.I., Homeyer, C.R., 2019. Spatiotemporal characteristics and large-scale environments of mesoscale convective systems east of the Rocky Mountains. J. Clim. 32, 7303–7328.
Fernández-Duque, B., Vicente-Serrano, S.M., Maillard, O., Domínguez-Castro, F., Peña-Angulo, D., Noguera, I., Azorin-Molina, C., El Kenawy, A., 2023. Long-term observed changes of air temperature, relative humidity and vapour pressure deficit in Bolivia, 1950–2019. Int. J. Climatol. 43, 6484–6504.
Ford, T.W., Dirmeyer, P.A., Benson, D.O., 2018. Evaluation of heat wave forecasts seamlessly across subseasonal timescales. NPJ Clim. Atmos. Sci. 1, 20.
Gage, E.A., Cooper, D.J., 2017. Urban forest structure and land cover composition effects on land surface temperature in a semi-arid suburban area. Urban For. Urban Green. 28, 28–35.
Ganeshi, N.G., Mujumdar, M., Yuhei, T., Goswami, M.M., Singh, B.B., Krishnan, R., Terao, T., 2022. Assessing the impact of soil moisture-temperature coupling on temperature extremes over the Indian region. arXiv Prepr. arXiv2204.04079.
Gao, B., Kaufman, Y.J., 2003. Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels. J. Geophys. Res. Atmos. 108.
Gao, B.-C., Yang, P., Guo, G., Park, S.K., Wiscombe, W.J., Chen, B., 2003. Measurements of water vapor and high clouds over the Tibetan Plateau with the Terra MODIS instrument. IEEE Trans. Geosci. Remote Sens. 41, 895–900.
Gelfand, S.J., 2015. Understanding the impact of heteroscedasticity on the predictive ability of modern regression methods. Finally, "After 2022" (Page 30 line 940) is not a reference, it means after year 2022.
Gheyret, G., Mohammat, A., Tang, Z., 2020. Elevational patterns of temperature and humidity in the middle Tianshan Mountain area in Central Asia. J. Mt. Sci. 17, 397–409.
Gibbard, S., Caldeira, K., Bala, G., Phillips, T.J., Wickett, M., 2005. Climate effects of global land cover change. Geophys. Res. Lett. 32.
Gibson, J., Franz, T.E., Wang, T., Gates, J., Grassini, P., Yang, H., Eisenhauer, D., 2017. A case study of field-scale maize irrigation patterns in western Nebraska: implications for water managers and recommendations for hyper-resolution land surface modeling. Hydrol. Earth Syst. Sci. 21, 1051–1062.
Gobbi, A., Alikadic, A., Ylinen, K., Angaramo, F., Furlanello, C., 2017. A heat wave forecast system for Europe. In: 2017 IEEE International Conference on Big Data (Big Data). IEEE, pp. 3734–3738.
Grotjahn, R., Black, R., Leung, R., Wehner, M.F., Barlow, M., Bosilovich, M., Gershunov, A., Gutowski, W.J., Gyakum, J.R., Katz, R.W., 2016. North American extreme temperature events and related large scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends. Clim. Dyn. 46, 1151–1184.
Grotjahn, R., Huynh, J., 2018. Contiguous US summer maximum temperature and heat stress trends in CRU and NOAA climate Division data plus comparisons to reanalyses. Sci. Rep. 8, 11146.
Grundstein, A.J., Ramseyer, C., Zhao, F., Pesses, J.L., Akers, P., Qureshi, A., Becker, L., Knox, J.A., Petro, M., 2012. A retrospective analysis of American football hyperthermia deaths in the United States. Int. J. Biometeorol. 56, 11–20.
Guo, S., Li, M., Li, Y., Chen, J., Zhang, H.K., Sun, L., Wang, J., Wang, R., Yang, Y., 2024. The improved U-STFM: a deep learning-based nonlinear spatial-temporal fusion model for land surface temperature downscaling. Remote Sens. 16, 322.
Gupta, P., Verma, S., Bhatla, R., Chandel, A.S., Singh, J., Payra, S., 2020. Validation of surface temperature derived from MERRA-2 reanalysis against IMD gridded data set over India. Earth Sp. Sci. 7 e2019EA000910.
Haigh, T., Takle, E., Andresen, J., Widhalm, M., Carlton, J.S., Angel, J., 2015. Mapping the decision points and climate information use of agricultural producers across the US Corn Belt. Clim. Risk Manag. 7, 20–30.
Han, W., Wang, L., Feng, R., Gao, L., Chen, X., Deng, Z., Chen, J., Liu, P., 2020. Sample generation based on a supervised Wasserstein Generative Adversarial Network for high-resolution remote-sensing scene classification. Inf. Sci. (NY) 539, 177–194.
Hawkins, M.D., Brown, V., Ferrell, J., 2017. Assessment of NOAA National Weather Service methods to warn for extreme heat events. Weather Clim. Soc. 9, 5–13.
He, T., Liu, F., Wang, A., Fei, Z., 2023. Estimating Monthly Surface Air Temperature using MODIS LST Data and an Artificial Neural Network in the Loess Plateau, China. Chines. Geogr. Sci. 33, 751–763.
Hiestand, M.P., Carleton, A.M., 2020. Growing-season synoptic and phenological controls on heat fluxes over forest and cropland sites in the Midwest US Corn Belt. J. Appl. Meteorol. Climatol. 59, 381–400.
Hiroki, R., Varquez, A.C.G., Khanh, D.N., Rynazal, R., Renard, F., Alonso, L., Kanda, M., 2025. Long-term changes of Universal Thermal Climate Index (UTCI) estimated from weather stations and gradient-boosted decision trees throughout Japan. Int. J. Climatol. e8843.
Ho, H.C., Knudby, A., Xu, Y., Hodul, M., Aminipouri, M., 2016. A comparison of urban heat islands mapped using skin temperature, air temperature, and apparent temperature (Humidex), for the greater Vancouver area. Sci. Total Environ. 544, 929–938.
Hough, I., Just, A.C., Zhou, B., Dorman, M., Lepeule, J., Kloog, I., 2020. A multi-resolution air temperature model for France from MODIS and Landsat thermal data. Environ. Res. 183, 109244.
Ingleby, B., Moore, D., Sloan, C., Dunn, R., 2013. Evolution and accuracy of surface humidity reports. J. Atmos. Ocean. Technol. 30, 2025–2043.
Jackson, L.S., Birch, C.E., Chagnaud, G., Marsham, J.H., Taylor, C.M., 2025. Daily rainfall variability controls humid heatwaves in the global tropics and subtropics. Nat. Commun. 16, 3461.
Jacobs, C., Klok, L., Bruse, M., Cortesão, J., Lenzholzer, S., Kluck, J., 2020. Are urban water bodies really cooling? Urban Clim. 32, 100607.

Janatian, N., Sadeghi, M., Sanaeinejad, S.H., Bakhshian, E., Farid, A., Hasheminia, S.M., Ghazanfari, S., 2017. A statistical framework for estimating air temperature using MODIS land surface temperature data. Int. J. Climatol. 37, 1181–1194.
Jansen, F.A., Jongen, H.J., Jacobs, C.M.J., Bosveld, F.C., Buzacott, A.J.V., Heusinkveld, B.G., Kruijt, B., van der Molen, M., Moors, E., Steeneveld, G., 2023. Land cover control on the drivers of evaporation and sensible heat fluxes: an observation-based synthesis for the Netherlands. Water Resour. Res. 59 e2022WR034361.
Jeong, Y., Lee, G., Kim, S., 2015. Analysis of the relation of local temperature to the natural environment, land use and land coverage of neighborhoods. J. Asian Archit. Build. Eng. 14, 33–40.
Jiang, L., Zhang, J., Liu, Q., Meng, X., Shi, L., Zhang, D., Xing, M., 2023. Spatiotemporal variations of the global compound heat wave and the drivers of its spatial heterogeneity. J. Clean. Prod. 408, 137201.
Jin, S., Dewitz, J., Danielson, P., Granneman, B., Costello, C., Smith, K., Zhu, Z., 2023a. National Land Cover Database 2019: a new strategy for creating clean leaf-on and leaf-off Landsat composite images. J. Remote Sens. 3, 22.
Jin, S., Dewitz, J., Li, C., Sorenson, D., Zhu, Z., Shogib, M.R.I., Danielson, P., Granneman, B., Costello, C., Case, A., 2023b. National land cover database 2019: a comprehensive strategy for creating the 1986–2019 forest disturbance product. J. Remote Sens. 3, 21.
Jordan, N.S., Hoff, R.M., Bacmeister, J.T., 2010. Validation of goddard earth observing system-version 5 MERRA planetary boundary layer heights using CALIPSO. J. Geophys. Res. Atmos. 115.
Joy, A., Satheesan, K., Paul, A., 2025. High-resolution maximum air temperature estimation over India from MODIS data using machine learning. Remote Sens. Appl. Soc Environ., 101463
Jucker, T., Hardwick, S.R., Both, S., Elias, D.M.O., Ewers, R.M., Milodowski, D.T., Swinfield, T., Coomes, D.A., 2018. Canopy structure and topography jointly constrain the microclimate of human-modified tropical landscapes. Glob. Chang. Biol. 24, 5243–5258.
Justin, A.D., McGovern, A., Allen, J.T., 2025. FrontFinder AI: Efficient identification of frontal boundaries over the continental United States and NOAA's unified surface analysis domain using the UNET3+ model architecture. Artif. Intell. Earth Syst. 4, e240043.
Khatana, S.A.M., Werner, R.M., Groeneveld, P.W., 2022. Association of extreme heat and cardiovascular mortality in the United States: a county-level longitudinal analysis from 2008 to 2017. Circulation 146, 249–261.
Kheyruri, Y., Teymourkhani, A., Neshat, A., Sharafati, A., Hameed, A.S., 2025. An intelligent deep learning-based approach for downscaling atmospheric general circulation model outputs. Acta Geophys. 73, 2081–2095.
Kitsara, G., Papaioannou, G., Retalis, A., Paronis, D., Kerkides, P., 2018. Estimation of air temperature and reference evapotranspiration using MODIS land surface temperature over Greece. Int. J. Remote Sens. 39, 924–948.
Krehbiel, C., Henebry, G.M., 2016. A comparison of multiple datasets for monitoring thermal time in urban areas over the US Upper Midwest. Remote Sens. 8, 297.
Krzysztof, B., Pavol, N., Oleh, S., Agnieszka, H., Olesya, S., Anna, B., Katarina, M., 2021. Influence of geographical factors on thermal stress in northern Carpathians. Int. J. Biometeorol. 65, 1553–1566.
Le Bideau, D., Mandin, P., Benbouzid, M., Kim, M., Sellier, M., 2019. Review of necessary thermophysical properties and their sensivities with temperature and electrolyte mass fractions for alkaline water electrolysis multiphysics modelling. Int. J. Hydrogen Energy 44, 4553–4569.
Le Roy, B., Dixon, K.W., Adams-Smith, D., 2024. High-resolution urban climate simulations for heat and health applications in Philadelphia. Urban Clim. 57, 102114.
Lee, J.Y., Röösli, M., Ragettli, M.S., 2021. Estimation of heat-attributable mortality using the cross-validated best temperature metric in Switzerland and South Korea. Int. J. Environ. Res. Public Health 18, 6413.
Lewandowski, S.A., Shaman, J.L., 2022. Heat stress morbidity among US military personnel: Daily exposure and lagged response (1998–2019). Int. J. Biometeorol. 66, 1199–1208.
Li, J., An, X., Li, Q., Wang, C., Yu, H., Zhou, X., Geng, Y., 2022. Application of XGBoost algorithm in the optimization of pollutant concentration. Atmos. Res. 276, 106238.
Li, X., Chen, Y., Zhu, Y., Shi, Y., An, N., Liao, Z., 2023a. Underestimated increase and intensification of humid-heat extremes across southeast China due to humidity data inhomogeneity. Front. Environ. Sci. 10, 1104039.
Li, Z., Wu, H., Duan, S., Zhao, W., Ren, H., Liu, X., Leng, P., Tang, R., Ye, X., Zhu, J., 2023b. Satellite remote sensing of global land surface temperature: definition, methods, products, and applications. Rev. Geophys. 61.
Li, L., Zha, Y., 2018. Mapping relative humidity, average and extreme temperature in hot summer over China. Sci. Total Environ. 615, 875–881.
Li, Y., Zhao, M., Mildrexler, D.J., Motesharrei, S., Mu, Q., Kalnay, E., Zhao, F., Li, S., Wang, K., 2016. Potential and actual impacts of deforestation and afforestation on land surface temperature. J. Geophys. Res. Atmos. 121, 14–372.
Liang, S., 2001. Narrowband to broadband conversions of land surface albedo I: Algorithms. Remote Sens. Environ. 76, 213–238.
Liao, Q.-Y., Leng, P., Li, Z.-L., Ren, C., Sun, Y.-Y., Gao, M.-F., Duan, S.-B., Shang, G.-F., 2020. A method for deriving relative humidity from modis data under all-sky conditions. IEEE Trans. Geosci. Remote Sens. 59, 8992–9006.
Lin, X., Hubbard, K.G., 2004. Comparison of ASOS dew point temperatures: HO-1088 and DTS1, in: Preprints, Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface, Seattle, WA, Amer. Meteor. Soc.
Lin, S., Moore, N.J., Messina, J.P., Wu, J., 2013. Evaluation of MODIS surrogates for meteorological humidity data in east Africa. Int. J. Remote Sens. 34, 4669–4679.
Lin, X., Zhang, W., Huang, Y., Sun, W., Han, P., Yu, L., Sun, F., 2016. Empirical estimation of near-surface air temperature in China from MODIS LST data by considering physiographic features. Remote Sens. 8, 629.
Liu, M., Li, X., Chai, Z., Chen, A., Zhang, Y., Zhang, Q., 2023. Dense temperature mapping and heat wave risk analysis based on multisource remote sensing data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16, 3148–3157.
Lopez-Gomez, I., McGovern, A., Agrawal, S., Hickey, J., 2023. Global extreme heat forecasting using neural weather models. Artif. Intell. Earth Syst. 2, e220035.
Lu, Y.-C., Romps, D.M., 2022. Extending the heat index. J. Appl. Meteorol. Climatol. 61, 1367–1383.
Lynn, B.H., Stauffer, D.R., Wetzel, P.J., Tao, W.-K., Alpert, P., Perlin, N., Baker, R.D., Muñoz, R., Boone, A., Jia, Y., 2001. Improved simulation of Florida summer convection using the PLACE land model and a 1.5-order turbulence parameterization coupled to the Penn State–NCAR Mesoscale Model. Mon. Weather Rev. 129, 1441–1461.
Ma, D., Tang, P., Zhao, L., 2019. SiftingGAN: Generating and sifting labeled samples to improve the remote sensing image scene classification baseline in vitro. IEEE Geosci. Remote Sens. Lett. 16, 1046–1050.
MacDonald, H., McKenney, D.W., Papadopol, P., Lawrence, K., Pedlar, J., Hutchinson, M. F., 2020. North American historical monthly spatial climate dataset, 1901–2016. Sci. Data 7, 411.
Madrigano, J., Ito, K., Johnson, S., Kinney, P.L., Matte, T., 2015. A case-only study of vulnerability to heat wave–related mortality in New York City (2000–2011). Environ. Health Perspect. 123, 672–678.
Mann, M.E., Steinman, B.A., Brouillette, D.J., Miller, S.K., 2021. Multidecadal climate oscillations during the past millennium driven by volcanic forcing. Science (80-.) 371, 1014–1019.
Manware, M., Dubrow, R., Carrión, D., Ma, Y., Chen, K., 2022. Residential and race/ethnicity disparities in heat vulnerability in the United States. GeoHealth 6 e2022GH000695.
McElroy, S., Schwarz, L., Green, H., Corcos, I., Guirguis, K., Gershunov, A., Benmarhnia, T., 2020. Defining heat waves and extreme heat events using sub-regional meteorological data to maximize benefits of early warning systems to population health. Sci. Total Environ. 721, 137678.
Meili, N., Manoli, G., Burlando, P., Carmeliet, J., Chow, W.T.L., Coutts, A.M., Roth, M., Velasco, E., Vivoni, E.R., Fatichi, S., 2021. Tree effects on urban microclimate: Diurnal, seasonal, and climatic temperature differences explained by separating radiation, evapotranspiration, and roughness effects. Urban For. Urban Green. 58, 126970.
Meng, Q., Chen, S., Zhang, L., Zhu, X., Zhang, Y., Atkinson, P.M., 2025. GLOSTFM: a global spatiotemporal fusion model integrating multi-source satellite observations to enhance land surface temperature resolution. Remote Sens. Environ. 319, 114640.
Menne, M.J., Williams Jr, C.N., 2009. Homogenization of temperature series via pairwise comparisons. J. Clim. 22, 1700–1717.
Metz, M., Andreo, V., Neteler, M., 2017. A new fully gap-free time series of land surface temperature from MODIS LST data. Remote Sens. 9, 1333.
Metzger, K.B., Ito, K., Matte, T.D., 2010. Summer heat and mortality in New York City: how hot is too hot? Environ. Health Perspect. 118, 80–86.
Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Nauss, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environ Model Softw. 101, 1–9.
Miao, Y., Guo, J., Liu, S., Liu, H., Li, Z., Zhang, W., Zhai, P., 2017. Classification of summertime synoptic patterns in Beijing and their associations with boundary layer structure affecting aerosol pollution. Atmos. Chem. Phys. 17, 3097–3110.
Miao, Y., Che, H., Liu, S., Zhang, X., 2022. Heat stress in Beijing and its relationship with boundary layer structure and air pollution. Atmos. Environ. 282, 119159.
Michelozzi, P., Accetta, G., De Sario, M., D'Ippoliti, D., Marino, C., Baccini, M., Biggeri, A., Anderson, H.R., Katsouyanni, K., Ballester, F., 2009. High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities. Am. J. Respir. Crit. Care Med. 179, 383–389.
Min, J., Lee, W., Oh, J., Kang, C., Ha, E., 2025. Comparing the role of relative and absolute humidity in heat-related mortality: a case time series study in South Korea. Environ. Health Perspect.
Mokari, E., Mohebzadeh, H., Samani, Z., DuBois, D., Daggupati, P., 2021. Spatiotemporal imputation of MODIS land surface temperature using machine learning techniques (Case study: New Mexico's lower Rio Grande Valley). Remote Sens. Appl.: Soc. Environ. 24, 100651.
Muro, J., Strauch, A., Heinemann, S., Steinbach, S., Thonfeld, F., Waske, B., Diekkrüger, B., 2018. Land surface temperature trends as indicator of land use changes in wetlands. Int. J. Appl. Earth Obs. Geoinf. 70, 62–71.
Mutiibwa, D., Strachan, S., Albright, T., 2015. Land surface temperature and surface air temperature in complex terrain. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 4762–4774.
Na, Q., Li, H., Cao, B., Bian, Z., Du, Y., Liu, Q., 2024. Estimation of daily mean, maximum, and minimum land surface temperatures from modis data using machine learning. In: IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 1850–1853.
Navarro-Serrano, F., López-Moreno, J.I., Azorin-Molina, C., Alonso-González, E., Aznarez-Balta, M., Buisán, S.T., Revuelto, J., 2020. Elevation effects on air temperature in a topographically complex mountain valley in the Spanish Pyrenees. Atmosphere (Basel). 11, 656.
Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., Righetti, M., 2024. Applications of XGBoost in water resources engineering: a systematic literature review (Dec 2018–May 2023). Environ Model Softw. 174, 105971.

Noi, P.T., Kappas, M., Degener, J., 2016. Estimating daily maximum and minimum land air surface temperature using MODIS land surface temperature data and ground truth data in Northern Vietnam. Remote Sens. 8, 1002.
Nouri, A.S., Çalışkan, O., Charalampopoulos, I., Cheval, S., Matzarakis, A., 2022. Defining local extreme heat thresholds and Indoor Cooling Degree Necessity for vulnerable residential dwellings during the 2020 summer in Ankara–Part I: Air temperature. Sol. Energy 242, 435–453.
Nukala, S., 2023. A systematic review with meta-analysis on the impact of heatwaves/extreme heat on cardiovascular disease outcomes. Hear. Lung Circ. 32, S381.
O'brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673–690.
Obarein, O.A., Lee, C.C., 2025. Synoptic air mass controls on global and regional precipitation. J. Geophys. Res. Atmos. 130 e2025JD043370.
Østby, T.I., Schuler, T.V., Westermann, S., 2014. Severe cloud contamination of MODIS land surface temperatures over an arctic ice cap, Svalbard. Remote Sens. Environ. 142, 95–102.
Paliwal, M., Kumar, U.A., 2011. The predictive accuracy of feed forward neural networks and multiple regression in the case of heteroscedastic data. Appl. Soft Comput. 11, 3859–3869.
Parasin, N., Amnuaylojaroen, T., 2023. Development of a heat index related to air quality and meteorology for an assessment of work performance in Thailand's Urban areas. Urban Sci. 7, 124.
Parastatidis, D., Mitraka, Z., Chrysoulakis, N., Abrams, M., 2017. Online global land surface temperature estimation from Landsat. Remote Sens. 9, 1208.
Pede, T., Mountrakis, G., 2022. Towards daily maximum heat index estimation across the conterminous United States using satellite-derived products. Int. J. Remote Sens. 43, 2861–2884.
Peng, G., Li, J., Chen, Y., Norizan, A.P., Tay, L., 2006. High-resolution surface relative humidity computation using MODIS image in Peninsular Malaysia. Chinese Geogr. Sci. 16, 260–264.
Perera, T., Nayanajith, T.M.D., Jayasinghe, G.Y., Premasiri, H.D.S., 2022. Identification of thermal hotspots through heat index determination and urban heat island mitigation using ENVImet numerical micro climate model. Model. Earth Syst. Environ. 1–18.
Petitti, D.B., Hondula, D.M., Yang, S., Harlan, S.L., Chowell, G., 2016. Multiple trigger points for quantifying heat-health impacts: new evidence from a hot climate. Environ. Health Perspect. 124, 176–183.
Phan, T.N., Kappas, M., Nguyen, K.T., Tran, T.P., Tran, Q.V., Emam, A.R., 2019. Evaluation of MODIS land surface temperature products for daily air surface temperature estimation in northwest Vietnam. Int. J. Remote Sens. 40, 5544–5562.
Qiao, L., Will, R., Wagner, K., Zhang, T., Zou, C., 2022. Improvement of evapotranspiration estimates for grasslands in the southern Great Plains: Comparing a biophysical model (SWAT) and remote sensing (MODIS). J. Hydrol.: Reg. Stud. 44, 101275.
Ramírez-Beltrán, N.D., Salazar, C.M., Castro Sánchez, J.M., González, J.E., 2019. A satellite algorithm for estimating relative humidity, based on GOES and MODIS satellite data. Int. J. Remote Sens. 40, 9237–9259.
Raymond, C., Waliser, D., Guan, B., Lee, H., Loikith, P., Massoud, E., Sengupta, A., Singh, D., Wootten, A., 2022. Regional and elevational patterns of extreme heat stress change in the US. Environ. Res. Lett. 17, 64046.
Raymond, C., Matthews, T., Tuholske, C., 2024. Evening humid-heat maxima near the southern Persian/Arabian Gulf. Commun. Earth Environ. 5, 591.
Recondo, C., Peón, J.J., Zapico, E., Pendás, E., 2013. Empirical models for estimating daily surface water vapour pressure, air temperature, and humidity using MODIS and spatiotemporal variables. applications to peninsular Spain. Int. J. Remote Sens. 34, 8051–8080.
Reyes, M.F., d'Angelo, P., Fraundorfer, F., 2022. SyntCities: a large synthetic remote sensing dataset for disparity estimation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 10087–10098.
Rodrigues, J.F., Florea, L., De Oliveira, M.C.F., Diamond, D., Oliveira, O.N., 2021. Big data and machine learning for materials science. Discov. Mater. 1, 1–27.
Rohli, R. V, Vega, A.J., Henderson, K.G., 2024. United States, Canada, and Adjacent Oceans (Theory), in: Atmospheric and Oceanic Circulation: An Explanation of Earth's Climate Patterns. Springer, pp. 213–240.
Romps, D.M., Lu, Y.-C., 2022. Chronically underestimated: a reassessment of US heat waves using the extended heat index. Environ. Res. Lett. 17, 94017.
Romps, D.M., Öktem, R., Endo, S., Vogelmann, A.M., 2021. On the life cycle of a shallow cumulus cloud: is it a bubble or plume, active or forced? J. Atmos. Sci. 78, 2823–2833.
Rothfusz, L.P., Headquarters, N.W.S.S.R., 1990. The heat index equation (or, more than you ever wanted to know about heat index). Fort Worth, Texas Natl. Ocean. Atmos. Adm. Natl. Weather Serv. Off. Meteorol. 9023, 640.
Ruckstuhl, C., Philipona, R., Morland, J., Ohmura, A., 2007. Observed relationship between surface specific humidity, integrated water vapor, and longwave downward radiation at different altitudes. J. Geophys. Res Atmos. 112.
Sadiq Khan, M., Ullah, S., Sun, T., Rehman, A.U.R., Chen, L., 2020. Land-use/land-cover changes and its contribution to urban heat Island: a case study of Islamabad, Pakistan. Sustainability 12, 3861.
Safieddine, S., Clerbaux, C., Muñoz-Sabater, J., Thépaut, J.-N., 2025. Local hourly trends in near-surface and land surface temperatures. Sci. Rep. 15, 29915.
Saito, T., Rehmsmeier, M., 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 10 e0118432.
Sall', M.A., 2015. Climate risks: temporal trends and heteroscedasticity. Russ. Meteorol. Hydrol. 40, 489–494.
Sangelantoni, L., Sobolowski, S., Lorenz, T., Hodnebrog, Ø., Cardoso, R.M., Soares, P.M. M., Ferretti, R., Lavín-Gullón, A., Fernandez, J., Goergen, K., 2023. Investigating the representation of heatwaves from an ensemble of km-scale regional climate simulations within CORDEX-FPS convection. Clim. Dyn. 1–37.
Schaefer, M., Salari, H.E., Köckler, H., Thinh, N.X., 2021. Assessing local heat stress and air quality with the use of remote sensing and pedestrian perception in urban microclimate simulations. Sci. Total Environ. 794, 148709.
Serra, C., Lana, X., Martínez, M.D., Roca, J., Arellano, B., Biere, R., Moix, M., Burgueño, A., 2020. Air temperature in Barcelona metropolitan region from MODIS satellite and GIS data. Theor. Appl. Climatol. 139, 473–492.
Shaik, N.B., Jongkittinarukorn, K., Bingi, K., 2024. XGBoost based enhanced predictive model for handling missing input parameters: a case study on gas turbine. Case Stud. Chem. Environ. Eng. 10, 100775.
Sharifnezhadazizi, Z., Norouzi, H., Prakash, S., Beale, C., Khanbilvardi, R., 2019. A global analysis of land surface temperature diurnal cycle using MODIS observations. J. Appl. Meteorol. Climatol. 58, 1279–1291.
Shen, X., Liu, Y., Liu, B., Zhang, J., Wang, L., Lu, X., Jiang, M., 2022. Effect of shrub encroachment on land surface temperature in semi-arid areas of temperate regions of the Northern Hemisphere. Agric. For. Meteorol. 320, 108943.
Sheridan, S.C., Dixon, P.G., Kalkstein, A.J., Allen, M.J., 2021. Recent trends in heat-related mortality in the United States: an update through 2018. Weather Clim. Soc. 13, 95–106.
Shi, N., Li, Y., Wen, L., Zhang, Y., 2022. Rapid prediction of landslide dam stability considering the missing data using XGBoost algorithm. Landslides 19, 2951–2963.
Simpson, C.H., Brousse, O., Ebi, K.L., Heaviside, C., 2023. Commonly used indices disagree about the effect of moisture on heat stress. NPJ Clim. Atmos. Sci. 6, 78.
Singh, K.K., Vogler, J.B., Shoemaker, D.A., Meentemeyer, R.K., 2012. LiDAR-Landsat data fusion for large-area assessment of urban land cover: balancing spatial resolution, data volume and mapping accuracy. ISPRS J. Photogramm. Remote Sens. 74, 110–121.
Skandalos, N., Wang, M., Kapsalis, V., D'Agostino, D., Parker, D., Bhuvad, S.S., Peng, J., Karamanis, D., 2022. Building PV integration according to regional climate conditions: BIPV regional adaptability extending Köppen-Geiger climate classification against urban and climate-related temperature increases. Renew. Sustain. Energy Rev. 169, 112950.
Sobolewski, A., Młynarczyk, M., Konarska, M., Bugajska, J., 2021. The influence of air humidity on human heat stress in a hot environment. Int. J. Occup. Saf. Ergon. 27, 226–236.
Souri, A.H., Wang, H., González Abad, G., Liu, X., Chance, K., 2020. Quantifying the impact of excess moisture from transpiration from crops on an extreme heat wave event in the midwestern US: a top-down constraint from Moderate Resolution Imaging Spectroradiometer water vapor retrieval. J. Geophys. Res. Atmos. 125 e2019JD031941.
Steadman, R.G., 1979. The assessment of sultriness. Part I: a temperature-humidity index based on human physiology and clothing science. J. Appl. Meteorol. Climatol. 18, 861–873.
Steadman, R.G., 1984. A universal scale of apparent temperature. J. Appl. Meteorol. Climatol. 23, 1674–1687.
Sulikowska, A., Wypych, A., 2021. Seasonal variability of trends in regional hot and warm temperature extremes in Europe. Atmosphere (Basel). 12, 612.
Sun, B., Baker, C.B., Karl, T.R., Gifford, M.D., 2005. A comparative study of ASOS and USCRN temperature measurements. J. Atmos. Ocean. Technol. 22, 679–686.
Suthar, G., Singh, S., Kaul, N., Khandelwal, S., Singhal, R.P., 2023. Prediction of maximum air temperature for defining heat wave in Rajasthan and Karnataka states of India using machine learning approach. Remote Sens. Appl.: Soc. Environ. 32, 101048.
Tang, Y., 2024. Cloud removal in landsat images based on spatial-spectral-temporal random forests. In: Journal of Physics: Conference Series. IOP Publishing, p. 12001.
Tang, Y., Wang, Q., Atkinson, P.M., 2023. Filling then spatio-temporal fusion for all-sky MODIS land surface temperature generation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16, 1350–1364.
Thompson, C.G., Kim, R.S., Aloe, A.M., Becker, B.J., 2017. Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. Basic Appl. Soc. Psych. 39, 81–90.
Thornton, P.E., Shrestha, R., Thornton, M., Kao, S.-C., Wei, Y., Wilson, B.E., 2021. Gridded daily weather data for North America with comprehensive uncertainty quantification. Sci. Data 8, 190.
Ting, M., Seager, R., Li, C., Liu, H., Henderson, N., 2021. Future summer drying in the US Corn Belt and the role of midlatitude storm tracks. J. Clim. 34, 9043–9056.
Trancoso, R., Syktus, J., Toombs, N., Ahrens, D., Wong, K.-K.-H., Dalla Pozza, R., 2020. Heatwaves intensification in Australia: a consistent trajectory across past, present and future. Sci. Total Environ. 742, 140521.
Tsagris, M., Pandis, N., 2021. Multicollinearity. Am. J. Orthod. Dentofac. Orthop. 159, 695–696.
Tuholske, C., Caylor, K., Funk, C., Verdin, A., Sweeney, S., Grace, K., Peterson, P., Evans, T., 2021. Global urban population exposure to extreme heat. Proc. Natl. Acad. Sci. 118 e2024792118.
Vaidyanathan, A., Saha, S., Vicedo-Cabrera, A.M., Gasparrini, A., Abdurehman, N., Jordan, R., Hawkins, M., Hess, J., Elixhauser, A., 2019. Assessment of extreme heat and hospitalizations to inform early warning systems. Proc. Natl. Acad. Sci. 116, 5420–5427.
Van Zyl, J.J., 2001. The Shuttle Radar Topography Mission (SRTM): a breakthrough in remote sensing of topography. Acta Astronaut. 48, 559–565.
Vandal, T.J., Nemani, R.R., 2021. Temporal interpolation of geostationary satellite imagery with optical flow. IEEE Trans. Neural Networks Learn. Syst. 34, 3245–3254.

Vargas Zeppetello, L.R., Raftery, A.E., Battisti, D.S., 2022. Probabilistic projections of increased heat stress driven by climate change. Commun. Earth Environ. 3, 183.
Vecellio, D.J., Wolf, S.T., Cottle, R.M., Kenney, W.L., 2022. Utility of the Heat Index in defining the upper limits of thermal balance during light physical activity (PSU HEAT Project). Int. J. Biometeorol. 66, 1759–1769.
Wan, H., Shao, Y., Campbell, J.B., Deng, X., 2019. Mapping annual urban change using time series Landsat and NLCD. Photogramm. Eng. Remote Sens. 85, 715–724.
Wang, Z., Franke, J.A., Luo, Z., Moyer, E.J., 2021. Reanalyses and a high-resolution model fail to capture the "high tail" of CAPE distributions. J. Clim. 34, 8699–8715.
Wang, B., Geddes, J.A., Adams, T.J., Lind, E.S., McDonald, B.C., He, J., Harkins, C., Li, D., Pfister, G.G., 2023a. Implications of sea breezes on air quality monitoring in a coastal urban environment: evidence from high resolution modeling of NO2 and O3. J. Geophys. Res. Atmos. 128 e2022JD037860.
Wang, X., He, L., Ma, X., Bie, Q., Luo, L., Xiong, Y., Ye, J., 2022a. The emergence of prolonged deadly humid heatwaves. Int. J. Climatol. 42, 8607–8618.
Wang, Y., Ouyang, W., Zhan, Q., Zhang, L., 2022b. The cooling effect of an urban river and its interaction with the littoral built environment in mitigating heat stress: a mobile measurement study. Sustainability 14, 11700.
Wang, Q., Tang, Y., Tong, X., Atkinson, P.M., 2024. Filling gaps in cloudy Landsat LST product by spatial-temporal fusion of multi-scale data. Remote Sens. Environ. 306, 114142.
Wang, N., Tian, J., Su, S., Tian, Q., 2023b. A downscaling method based on MODIS product for hourly ERA5 reanalysis of land surface temperature. Remote Sens. 15, 4441.
Wang, K., Wan, Z., Wang, P., Sparrow, M., Liu, J., Haginoya, S., 2007. Evaluation and improvement of the MODIS land surface temperature/emissivity products using ground-based measurements at a semi-desert site on the western Tibetan Plateau. Int. J. Remote Sens. 28, 2549–2565.
Wasif Ali, N. ul A.B., Amir, S., Iqbal, K.M.J., Shah, A.A., Saqib, Z., Akhtar, N., Ullah, W., Tariq, M.A.U.R., 2022. Analysis of land surface temperature dynamics in Islamabad by using MODIS Remote Sensing Data. Sustainability 14, 9894.
Weinberger, K.R., Zanobetti, A., Schwartz, J., Wellenius, G.A., 2018. Effectiveness of National Weather Service heat alerts in preventing mortality in 20 US cities. Environ. Int. 116, 30–38.
Werner, M., 2001. Shuttle radar topography mission (SRTM) mission overview. Frequenz 55, 75–79.
Wu, X., Liu, Q., Huang, C., Li, H., 2022. Mapping heat-health vulnerability based on remote sensing: a case study in Karachi. Remote Sens. 14, 1590.
Wu, X., Ge, Y., Gong, D., Zhang, X., Hu, S., Liu, Q., 2023. Reconstruction of the hourly fine-resolution apparent temperature (Humidex) with the aerodynamic parameters. Sci. Total Environ. 866, 161253.
Wu, Q., Tan, J., Guo, F., Li, H., Chen, S., 2019. Multi-scale relationship between land surface temperature and landscape pattern based on wavelet coherence: the case of metropolitan Beijing, China. Remote Sens. 11, 3021.
Xiao, Y., Zhao, W., Ma, M., He, K., 2021. Gap-free LST generation for MODIS/Terra LST product using a random forest-based reconstruction method. Remote Sens. 13, 2828.
Xu, W.R., Dixon, K.W., Zenes, N., Adams-Smith, D., 2025b. Sometimes missing the heat: the risk of underestimating extreme heat days with daily maximum heat index approximation. Int. J. Biometeorol. 1–15.
Xu, J., Zhang, H., Qiao, Y., Yuan, H., Xu, W., Xia, X., 2025a. Effects of microtopography on neighborhood diversity and competition in subtropical forests. Plants 14, 870.
Yang, Y.Z., Cai, W.H., Yang, J., 2017. Evaluation of MODIS land surface temperature data to estimate near-surface air temperature in Northeast China. Remote Sens. 9, 410.
Yang, Z., Dominguez, F., Zeng, X., 2019. Large and local-scale features associated with heat waves in the United States in reanalysis products and the NARCCAP model ensemble. Clim. Dyn. 52, 1883–1901.
Yang, S., Li, S., Chen, B., Xie, Z., Peng, J., 2021. Responses of heat stress to temperature and humidity changes due to anthropogenic heating and urban expansion in South and North China. Front. Earth Sci. 9, 673943.
Yao, L., Sailor, D.J., Yang, X., Xu, G., Zhao, L., 2023. Are water bodies effective for urban heat mitigation? evidence from field studies of urban lakes in two humid subtropical cities. Build. Environ. 245, 110860.
Yiou, P., Jézéquel, A., 2020. Simulation of extreme heat waves with empirical importance sampling. Geosci. Model Dev. 13, 763–781.
Yoo, C., Im, J., Park, S., Quackenbush, L.J., 2018. Estimation of daily maximum and minimum air temperatures in urban landscapes using MODIS time series satellite data. ISPRS J. Photogramm. Remote Sens. 137, 149–162.
Zeng, L., Hu, Y., Wang, R., Zhang, X., Peng, G., Huang, Z., Zhou, G., Xiang, D., Meng, R., Wu, W., 2021. 8-Day and daily maximum and minimum air temperature estimation via machine learning method on a climate zone to global scale. Remote Sens. 13, 2355.
Zeng, C., Long, D., Shen, H., Wu, P., Cui, Y., Hong, Y., 2018. A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud. ISPRS J. Photogramm. Remote Sens. 141, 30–45.
Zhang, X., Alexander, L., Hegerl, G.C., Jones, P., Tank, A.K., Peterson, T.C., Trewin, B., Zwiers, F.W., 2011. Indices for monitoring changes in extremes based on daily temperature and precipitation data. Wiley Interdiscip. Rev. Clim. Chang. 2, 851–870.
Zhang, M., Gao, Y., 2023. Time of emergence in climate extremes corresponding to Köppen-Geiger classification. Weather Clim. Extrem. 41, 100593. https://doi.org/10.1016/J.WACE.2023.100593.
Zhang, K., Li, Y., Schwartz, J.D., 2014. What weather variables are important in predicting heat-related mortality? a new application of statistical learning methods. Environ. Res. 132, 350–359.
Zhang, Z., Li, P., Zheng, X., Zhang, H., 2024. Remotely sensed estimation of daily near-surface air temperature: a comparison of metop and MODIS. Remote Sens. 16, 3754.
Zhang, W., Zhao, J., Zhu, W., Kong, Y., Wan, B., Liao, Y., 2025. Comprehensive validation of MODIS-derived instantaneous air temperature and daily minimum temperature at nighttime. Remote Sens. 17, 1732.
Zhang, C., Zhou, J., Li, C., Fu, W., Peng, T., 2017. A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. Energy Convers. Manag. 143, 360–376.
Zhao, Y., Zhao, H., Li, J., Xiao, G., 2022. Comprehensive validation and calibration of MODIS PWV over mainland China. Atmosphere (Basel). 13, 1763.
Zheng, M., Zhang, J., Wang, J., Yang, S., Han, J., Hassan, T., 2022. Reconstruction of 0.05° all-sky daily maximum air temperature across Eurasia for 2003–2018 with multi-source satellite data and machine learning models. Atmos. Res. 279, 106398.
Zhong, Y., Hong, S., Wei, Z., Walker, J.P., Wang, Y., Huang, C., 2024. Spatial downscaling of SMAP soil moisture estimation using multiscale geographically weighted regression during SMAPVEX16. J. Hydrol. 637, 131348.
Zhou, W., Qian, Y., Li, X., Li, W., Han, L., 2014. Relationships between land cover and the surface urban heat island: seasonal variability and effects of spatial and thematic resolution of land cover data on predicting land surface temperatures. Landsc. Ecol. 29, 153–167.