



Integrating Spatiotemporal Vision Transformer into Digital Twins for High-Resolution Heat Stress Forecasting in Campus Environments

Wenjing Gong¹ , Xinyue Ye² , Keshu Wu¹ ,
 Suphanut Jamonnak¹, Wenyu Zhang¹ , Yifan Yang¹ ,
 and Xiao Huang³

Abstract

Extreme heat events, exacerbated by climate change, pose significant challenges to urban resilience and planning. This study introduces a climate-responsive digital twin framework integrating the Spatiotemporal Vision Transformer (ST-ViT) model to enhance heat stress forecasting and decision-making. Using a Texas campus as a testbed, we synthesized high-resolution physical model simulations with spatial and meteorological data to develop fine-scale human thermal predictions. The ST-ViT-powered digital twin enables efficient, data-driven insights for planners and stakeholders, supporting targeted heat mitigation strategies and advancing climate-adaptive urban design. This campus-scale demonstration offers a foundation for future applications across broader and more diverse urban contexts.

Keywords

digital twins, heat waves, outdoor heat exposure, human thermal comfort, Spatiotemporal Vision Transformer, campus

Introduction

Extreme heat events, intensified by climate change, pose growing health and socio-economic risks, particularly in urban and campus environments (Anderson and Bell 2011; Chu and Rotta Loria 2024; Georgescu, Broadbent, and Krayenhoff 2024). Unlike other hazards, heat is invisible, exacerbating its impact on vulnerable populations and contributing to rising emergency visits and health issues such as cardiovascular and mental health disorders (Campbell et al. 2018; Mallen et al. 2020). The complexity of urban heat dynamics necessitates proactive, data-driven planning to enhance resilience and mitigate risks.

Campuses pose distinct heat exposure challenges due to their spatial structure, activity patterns, and population dynamics. Diverse microclimates, shaped by open spaces, greenery, and outdoor activity areas, interact with frequent movement between buildings, exposing students and staff to fluctuating thermal conditions (Jiang et al. 2024; Mallen et al. 2020). Compared with urban cores, campus shading and cooling infrastructure is often insufficient for extreme heat events (Göçer et al. 2019). At the same time, modeling thermal comfort at the micro scale remains challenging due to the fine-grained spatial and temporal variability of key variables such as air temperature and wind speed. Prior studies have employed both physics-based models and data-driven

artificial intelligence (AI) approaches to assess urban heat stress at high resolution (Bruse and Fleer 1998; Jia et al. 2024). However, the complex and dynamic microclimatic conditions on campuses call for more adaptive, scalable, and context-specific solutions.

Digital twins offer a promising solution for mitigating extreme heat exposure on campuses by providing virtual replicas that simulate and manage the built environment (White et al. 2021; Ye et al. 2023). These data-driven platforms integrate real-time inputs, models, and simulations to assess environmental stressors and inform decision-making (Omran and Al-Obaidi 2024; Riaz, McAfee, and Gharbia 2023; Ye et al. 2024). By enabling dynamic analysis of heat mitigation strategies, digital twins allow stakeholders to predict thermal conditions, evaluate interventions, and enhance climate resilience without physical implementation.

Initial submission, February 2025; revised submissions, July and August 2025; final acceptance, October 2025

¹Texas A&M University, College Station, TX, USA

²The University of Alabama, Tuscaloosa, AL, USA

³Emory University, Atlanta, GA, USA

Corresponding Author:

Xinyue Ye, The University of Alabama, Box 870322, Tuscaloosa, AL 35487-0322, USA.
 Email: xyel0@ua.edu



Despite their potential, climate digital twins face two key challenges. First, traditional physical climate models are computationally intensive and costly, limiting their feasibility for real-time predictions and integration into digital twins (Hu et al. 2023; Xiaojiang Li et al. 2024). Their high resource demands and slow simulation speeds hinder rapid decision-making and reduce accessibility, particularly in resource-constrained settings. Second, many existing digital twins primarily serve as static visualization tools, lacking predictive capabilities to simulate heatwave progression or the impact of mitigation strategies (P. Liu et al. 2023; Ramani et al. 2023). Without forward-looking modeling, these systems fail to support proactive decision-making, leaving stakeholders reliant on representations of past and present conditions rather than anticipating future risks.

Artificial intelligence (AI) models offer a promising solution to the limitations of traditional climate modeling by reducing computational demands while maintaining accuracy, enabling real-time, data-driven heat predictions within digital twins (Bibri et al. 2024; Ketzler et al. 2020; Lehtola et al. 2022). However, two key challenges remain. First, AI models relying solely on meteorological data often fail to capture the nonlinear and nonstationary dynamics of atmospheric processes, leading to potential prediction errors (Jing Shi, Guo, and Zheng 2012; Zhu et al. 2024). Second, effective climate prediction requires integrating diverse climate and spatial datasets to model spatiotemporal dependencies critical for understanding localized heat variations (Zhu et al. 2024). Addressing these challenges is essential to enhancing AI-driven high-resolution climate forecasts and advancing their role in digital twin applications for climate resilience.

To overcome these challenges, this study presents a human-centric climate digital twin framework, using the Texas A&M University (TAMU) campus as a testbed. High-resolution three-dimensional (3D) campus models, derived from LiDAR point clouds and land cover maps, alongside meteorological data, were integrated to represent urban geometry and localized climatic contexts. The Spatiotemporal Vision Transformer (ST-ViT) model incorporated multimodal data, including Universal Thermal Climate Index (UTCI) (Bröde et al. 2012) mappings simulated via physics-based methods, to capture fine-scale spatiotemporal dependencies and enable real-time, high-accuracy heat predictions, with “real-time” denoting rapid processing and inference within seconds (Fujiwara et al. 2024; Jihao Shi et al. 2025). By combining physics-based simulations to retain atmospheric dynamics with a Transformer architecture that models long-range dependencies across space and time, the ST-ViT model enhances climate prediction precision and efficiency. Integrating the ST-ViT model into a digital twin platform, this study establishes a data-driven decision-support tool for heat resilience planning. The research makes three key contributions:

1. We proposed a human-centric climate digital twin framework that integrated the ST-ViT model with

real-time prediction capabilities and practical components to support informed decision-making and planning for TAMU campus heat resilience.

2. We developed an advanced ST-ViT model that coupled output from physics-based simulations to preserve atmospheric dynamical physical characteristics and leveraged self-attention mechanisms to provide a computationally efficient approach for precise predictions of human heat stress.
3. We leveraged multimodal data to identify high-resolution campus hot spots across different stages of a heatwave and diurnal cycles, providing detailed insights to guide resource allocation, planning, and strategies for mitigating extreme heat impacts.

Literature Review

Physical, AI, and Hybrid Models in Urban Climate Prediction

In the field of urban climate modeling, two main approaches are currently predominant: physics-based climate models and data-driven AI models. The former includes established models such as ENVI-met (Bruse and Fleer 1998), SOLWEIG (Lindberg, Holmer, and Thorsson 2008), WRF (Weather Research and Forecasting) (Skamarock et al. 2019), and Fluent (ANSYS Fluent 2011). These models rely on complex mathematical calculations and numerical simulations, offering high spatial resolution and detailed meteorological predictions, while often incurring high computational costs and time demands (Han et al. 2024; L. Zheng and Lu 2024). With the advancement of AI, an increasing number of researchers have turned to machine learning and deep learning models for urban climate predictions (Fujiwara et al. 2024; Jia et al. 2024). These AI models leverage their ability to capture nonlinear relationships between inputs and outputs while reducing computational resources and runtime.

Among AI architectures for spatiotemporal climate prediction, various approaches have been explored. Convolutional Neural Networks combined with Long Short-Term Memory networks (CNN-LSTM) are commonly used to extract spatial and temporal features but often struggle with capturing long-range temporal dependencies (Xinyu Li et al. 2023). U-Net variants with temporal extensions, such as ConvLSTM or 3D convolutions, can preserve fine spatial resolution, but their encoder-decoder structure and repeated processing of large feature maps often lead to high computational and memory costs (Briegel et al. 2023). Recently, Vision Transformers, particularly ST-ViT, have emerged as promising alternatives. Their self-attention mechanisms enable efficient parallel modeling of both local and global spatiotemporal dependencies, allowing flexible focus across time and space (Lin et al. 2023). Unlike convolution-based models with fixed receptive fields, ST-ViT can dynamically learn contextual relationships, making it well-suited for complex urban microclimate patterns that exhibit heterogeneous

spatial structures and temporal lags. However, relying solely on spatial or meteorological data may cause AI models to fail in retaining the physical characteristics of atmospheric dynamics embedded in nonlinear and nonstationary climate data, potentially resulting in significant prediction errors (Jing Shi, Guo, and Zheng 2012).

Hybrid models, which leverage external coupling by combining the outputs of one or more base models as features and integrating them with advanced models, have recently gained attention (Briegel et al. 2024; Wu, Wang, and Zeng 2022; Zhu et al. 2024). This approach allows for the integration of strengths from multiple individual models, creating complementary advantages and optimizing the overall model structure. This potential offers a new perspective and methodology for achieving accurate and efficient urban climate predictions, as well as for building robust digital twin systems. For instance, Briegel et al. used SOLWEIG (SOlar and LongWave Environmental Irradiance Geometry) model input and output data as the dataset for a U-Net model, achieving high-precision predictions of mean radiant temperature (T_{mrt}) at a one-meter scale (Briegel et al. 2023). However, this study did not account for the temporal dependencies in the data; temporal information was linearly transformed and then added to the compressed latent spatial representations. Another study introduced a framework that used building energy modeling and CFD (Computational Fluid Dynamics) simulation results as datasets for a graph attention network to predict outdoor thermal comfort at an urban micro-scale (L. Zheng and Lu 2024). While this model focused on capturing the spatial interrelationships of urban features, it lacked considerations for the cooling effects of greenery and did not incorporate temporal dependencies, limiting its ability to predict thermal comfort across different time steps. Recently, Zhu et al. (2024) proposed a hybrid model combining WRF and temporal fusion transformers (TFT), where WRF outputs served as the prediction dataset for the TFT model. This model achieved high predictive performance in forecasting urban air temperatures in central Guangzhou, China. However, the resolution was limited to 0.5 km, and air temperature alone does not fully represent human thermal comfort.

These studies collectively underscore the necessity and potential of developing computationally efficient and accurate frameworks to support urban climate modeling. Compared with traditional physics-based models, data-driven approaches provide a scalable and flexible foundation for building responsive, high-resolution digital twins in urban climate contexts. When coupled with outputs from physics-based models to form hybrid models, they gain additional advantages. However, data-driven methods in urban climate modeling are still in the early stages of development (Yang et al. 2023), and research on spatiotemporal predictions of human thermal comfort at fine scales, particularly their integration into digital twins, remains limited.

Digital Twins in Urban Climate and Heat Mitigation

The digital twin has emerged as a transformative tool, integrating digital innovations with urban operations to support planning and climate analysis (Peldon et al. 2024; Xia et al. 2022). Studies highlight their effectiveness in visualizing thermal environments, forecasting climate conditions, and assessing intervention strategies (Cárdenas-León et al. 2024; T. Liu and Fan 2023). However, many digital twins primarily focus on real-world data visualization, integrating 3D city models, thermal imaging, and meteorological data to map microclimates and temperature variations (Ramani et al. 2023). While valuable for monitoring, their full potential for proactive heat mitigation planning remains unrealized due to the lack of predictive and decision-support capabilities.

Prediction and simulation are fundamental to digital twins, enabling dynamic climate forecasting and mitigation planning (Deren, Wenbo, and Zhenfeng 2021). For instance, a smart city digital twin framework incorporating the SARIMA time series model and crowd simulation demonstrated how digital twins can help urban officials identify, predict, and mitigate extreme heat exposure (Pan et al. 2024). However, statistical models often fail to capture complex nonlinear relationships and struggle to leverage spatial characteristics inherent in climate data. In urban climate modeling, AI-driven spatiotemporal predictions offer a promising approach to dynamically evaluate and forecast human thermal comfort within digital twins (Čulić et al. 2021; Karyono et al. 2024). For example, P. Liu et al. (2023) applied a GraphSAGE model with street-view imagery to predict outdoor thermal comfort, suggesting its integration into urban digital twins. While this method effectively captures spatial dependencies, it overlooks temporal dynamics, which are essential for modeling seasonal, diurnal, and transient climate patterns necessary for accurate heat predictions.

Moreover, there is a notable lack of hybrid models that integrate physics-based simulations with AI-driven predictions within digital twin platforms. This gap limits the ability to preserve atmospheric dynamics and leverage the complementary strengths of both approaches, particularly in complex urban environments where high-resolution climate predictions are critical. University campuses, such as TAMU, represent thermal vulnerability hotspots, with dense populations, extensive outdoor activities, and diverse building configurations creating localized heat exposure patterns. Conventional models often fail to resolve these patterns at a fine scale, leading to gaps in accurately assessing campus heat risks. Additionally, many existing studies do not fully integrate their models into digital twin platforms or translate insights into actionable heat mitigation strategies. This is particularly consequential for campus environments, where real-time decision support for facility managers is essential

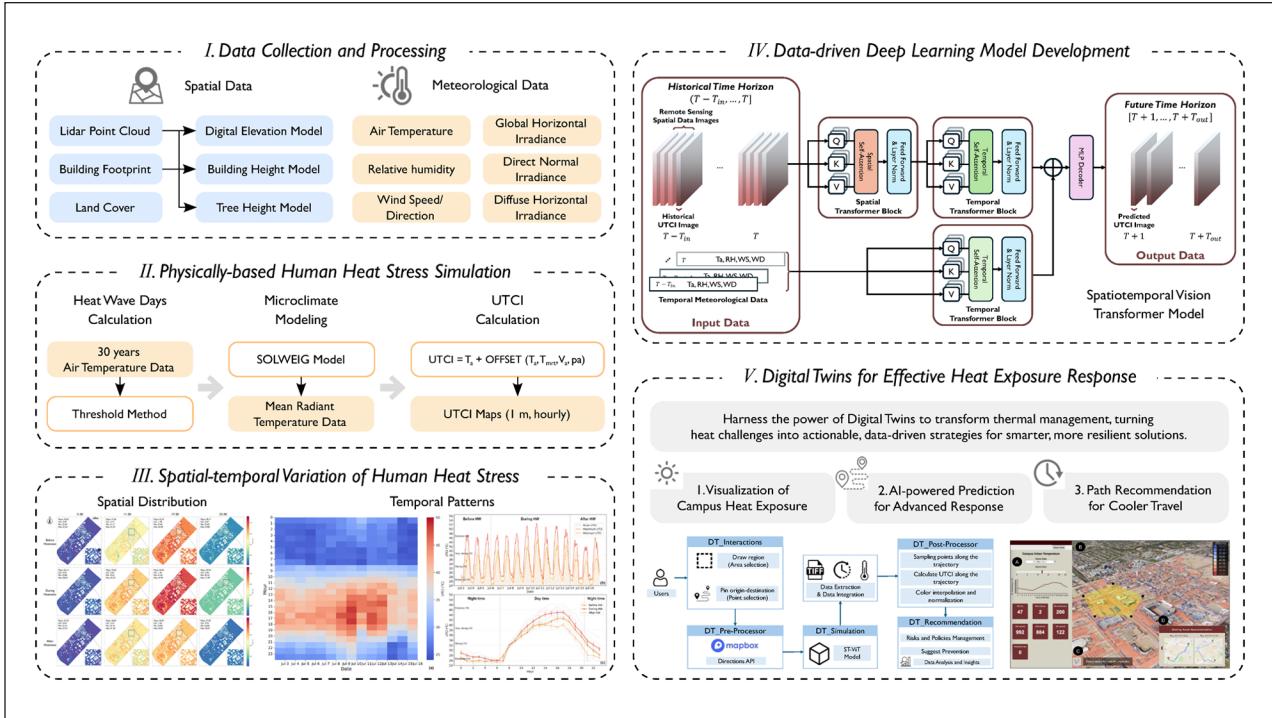


Figure 1. Analytical framework.

to mitigating heat stress impacts. These challenges highlight the urgent need for AI-powered campus digital twins to bridge the gap between advanced modeling techniques and practical applications in urban climate resilience.

Data and Methodology

Study Area and Analytical Framework

The TAMU main campus in College Station, Texas, USA, spans about 6.13 km² and serves as the study area (Supplementary Figure S1). With a subtropical climate of hot summers and mild winters, the region is projected to experience intensifying extreme heat, surpassing early twentieth-century levels by 2036 (Texas 2036 2023). The campus's varied land cover and building layouts provide an ideal setting for analyzing extreme heat impacts on human heat stress and advancing digital twin development for informed decision-making.

This study presents a novel human-centric climate digital twin framework for urban heat stress monitoring and prediction, integrating data collection, physical simulation, spatiotemporal analysis, deep learning models, and digital twin application into a cohesive system (Figure 1). It begins with the collection of spatial and meteorological data to establish the foundational data layer. Subsequently, through a physically-based human heat stress simulation module, we used the SOLWEIG model to calculate T_{mrt} and generated UTCI

maps. The third module analyzes the spatiotemporal variation of heat stress, visualizing the dynamic changes in the thermal environment. The system further employs an ST-ViT model, combined with UTCI output from the physics-based approach, efficiently capturing complex spatiotemporal dependencies in multimodal data to deliver rapid thermal predictions. Finally, the framework leads to a digital twin application that transforms theoretical analysis into practical heat exposure response strategies, facilitating campus heat exposure overview, route planning, and advanced response predictions for more resilient urban thermal management solutions.

Heat Wave Days Calculation

To explore human outdoor heat stress under extreme conditions, the heat wave days were calculated to define the simulation period. They were identified based on thirty years (1991–2020) of air temperature data from a campus meteorological station (ASOS 2024). The 98th percentile of daily maximum temperatures (38.33°C) served as the heat wave threshold (Z. Zheng, Zhao, and Oleson 2021). Periods in 2022 exceeding this threshold for at least three consecutive days were analyzed, with the longest event, July 6–13, lasting eight days, defined as the heat wave (Meehl and Tebaldi 2004). The simulation period included three days before and after the heat wave, spanning fourteen days to assess heat exposure impacts.

Physically-Based Human Heat Stress Simulation

This study used the UTCI, a widely recognized index for quantifying human heat stress and assessing urban microclimate impacts on outdoor thermal comfort (Yang et al. 2023). UTCI calculations mainly rely on T_{mrt} , air temperature, relative humidity, and wind speed (Bröde et al. 2012), with T_{mrt} being the most critical input (Xiaojiang Li et al. 2024). We employed the physics-based SOLWEIG model (Lindberg, Holmer, and Thorsson 2008) to simulate hourly T_{mrt} at a one-meter resolution across a fourteen-day heat wave period, covering day and night. The SOLWEIG model has been widely applied and validated globally for urban microclimate simulations (Chaowen and Fricker 2021; Hu et al. 2023). While SOLWEIG supports UTCI calculations only at specific points, a custom script was developed to generate UTCI maps for the entire study area. The SOLWEIG model configuration and justification for high-resolution UTCI simulation are detailed in Supplementary Note S1.

We used spatial and meteorological data to measure human heat stress. Spatial data included LiDAR point clouds, Digital Elevation Model (DEM), building footprints, and land cover data, all essential inputs for the SOLWEIG model. LiDAR and DEM data (one-meter resolution) were obtained from USGS (2024), with the LiDAR processed using PointCNN (“Point Cloud Classification Using PointCNN” 2024) to classify tree points and generate a tree height model. Building footprints from TAMU were combined with DEM and LiDAR data to create a calibrated building height model, while land cover data was also sourced from TAMU. Meteorological data included hourly air temperature, relative humidity, wind speed, and wind direction from an ASOS (2024) station on campus, as well as hourly radiation data from NREL’s NSRDB (2024) for the heat wave periods, including global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI).

ST-ViT Model for Human Heat Stress Forecast

This study developed the ST-ViT model to enable fine-scale spatiotemporal UTCI prediction within digital twin frameworks. A detailed summary of data characteristics and roles is provided in Supplementary Table S1. It includes four spatial datasets and seven temporal meteorological variables, which are also required inputs for the SOLWEIG model. To better capture the high-resolution spatiotemporal dynamics of human heat stress, historical UTCI map sequences were also included as input features. While SOLWEIG simulates each UTCI map independently, using the sequence allows the model to learn temporal patterns, such as diurnal heat retention, that enhance predictive accuracy beyond what static spatial features and spatially uniform meteorological inputs alone can offer. Notably, other variables required for SOLWEIG simulations, such as the Sky View Factor (SVF), derived from building height and tree height through

additional pre-processing, were excluded to reduce redundancy and computational cost.

Spatial and temporal data were normalized for training, with predicted UTCI maps denormalized for interpretability. Using 336 hourly time steps, the model predicted UTCI for the next twenty-four hours ($T_{out} = 24$) based on the preceding twenty-four hours ($T_{in} = 24$). Different sliding window steps were tested to evaluate the model’s sensitivity to input configurations. The dataset was split chronologically (first 70% for training), with data augmentation (random cropping of 64×64 subregions) applied to improve generalization and robustness (Dyk and Meng 2001; Maharana, Mondal, and Nemade 2022). Testing involved dividing the study area into overlapping 64×64 subsets with ten-pixel overlaps, and merging predictions to calculate final evaluation metrics.

The ST-ViT model is a transformer-based, decoder-only architecture designed for efficient and precise spatiotemporal dependency modeling (N. Liu et al. 2021; Shim et al. 2023). Inputs, including spatial images and temporal meteorological features, were embedded into a shared hidden space through linear projection layers. The model employed three parallel attention mechanisms: spatial attention for local and global spatial relationships, and two temporal attention mechanisms for capturing long-term dependencies of human heat stress and meteorological features. After parallel processing, spatial-temporal features and temporal meteorological features were fused through an additive mechanism. Temporal features were expanded to match the spatial resolution, and the unified representation was passed through a final linear projection layer, ensuring the preservation of spatial and temporal resolutions. Further details on the model architecture are provided in Supplementary Figure S2.

We compared the ST-ViT model with two benchmark models, CNN+LSTM and U-Net+LSTM, commonly used for spatiotemporal prediction (Xinyu Li et al. 2023; Yin et al. 2023). All models were implemented in Python with PyTorch and trained on two NVIDIA A800 Tensor Core GPUs (80GB memory each). Mean Squared Error Loss (MSELoss) was used as the loss function with the Adam optimizer (learning rate=0.0001) for optimization (Kingma and Ba 2017). Early stopping was applied to prevent overfitting by monitoring validation loss. Model performance was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The calculation formula is shown from equations (1)–(4). Additional details of the models can be found in Supplementary Table S2.

$$MSE = \frac{1}{m} \sum_{i=1}^m (\widehat{UTCI}_i - UTCI_i)^2 \quad (1)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (\widehat{UTCI}_i - UTCI_i)^2} \quad (2)$$

$$MAE = \frac{1}{m} \sum_{i=1}^m \left| \widehat{UTCI}_i - UTCI_i \right| \quad (3)$$

$$MAPE = \frac{100\%}{m} \sum_{i=1}^m \left| \frac{\widehat{UTCI}_i - UTCI_i}{UTCI_i} \right| \quad (4)$$

where m is the total number of pixels; \widehat{UTCI}_i denotes the ST-ViT predicted UTCI for pixel i ; and $UTCI_i$ denotes simulated UTCI based on SOLWEIG for pixel i .

Application: Interactive Campus Climate Digital Twin

The digital twin platform integrates high-resolution UTCI predictions from the ST-ViT model into an interactive web-based interface, transforming complex thermal data into actionable insights for campus heat mitigation. Through photorealistic 3D Tiles technology, the platform creates a detailed virtual representation of the campus environment, which is dynamically overlaid with UTCI predictions visualized as color heatmaps at a spatial resolution of one meter and hourly temporal granularity. This interactive visualization enables users to quickly identify critical heat exposure zones across the campus environment. Beyond basic visualization, the platform offers two advanced decision-support functions: Path Recommendation—generating optimized routes to minimize heat exposure, and Predictive Heatmap—using the ST-ViT model to forecast future thermal conditions. This seamless integration of thermal modeling with intuitive visualization and practical features creates an effective tool for making informed thermal comfort choices in the campus environment.

Built on advanced technologies and frameworks, the interactive web-based digital twin platform ensures scalability, usability, and accuracy across data processing, modeling, visualization, and user interaction. The backend architecture consists of two main components. First, a Python-based data preprocessing pipeline handles the computation of UTCI predictions using the ST-ViT model on tile-based imagery. This pipeline leverages specialized libraries like NumPy for numerical computations and GeoPandas for spatial data operations, enabling precise overlay of UTCI predictions onto campus geographical layers. Second, a Flask framework serves as a lightweight backend server, hosting the prediction endpoints and facilitating seamless integration between the ST-ViT model and the web interface. The front-end architecture was built on React JS for responsive user interactions across devices, using Deck.GL (a powerful JavaScript library) to render photorealistic 3D tiles and dynamically overlay UTCI predictions as interactive heat maps onto the campus model.

The digital twin platform is deployed on Amazon Web Services (AWS), leveraging scalable cloud infrastructure. AWS EC2 instances provide the core computational resources for prediction and visualization tasks. The backend services are containerized using Docker to ensure consistent performance across different environments and enable efficient resource management and automated scaling. This cloud-based deployment architecture not only ensures current platform reliability but also supports future scalability through the seamless integration of additional data sources and features.

Results

This section began with a spatiotemporal analysis of human heat stress simulated using physical climate models. A prediction model database was then constructed based on these results and input into the ST-ViT model for UTCI forecasting. The ST-ViT model's performance was validated against physical model results, achieving accurate and rapid predictions. Finally, the model was integrated into the climate digital twin platform to support risk warnings and planning for human heat stress.

Spatiotemporal Variation of Human Heat Stress Based on SOLWEIG

Based on the temporal and diurnal UTCI patterns observed in Supplementary Figures S3 and S4, which show that 17:00 is typically the peak UTCI hour across the heatwave, we selected four representative times (05:00, 11:00, 17:00, and 23:00) to show the spatial distributions of UTCI in Figure 2. During the heatwave stage, UTCI at 17:00 reached up to 47.99 °C, compared to 45.90 °C before the heatwave and 43.07°C after it, indicating intensified thermal stress during the peak period. Notably, spatial heterogeneity was highest at 11:00 across all heatwave stages, with the largest standard deviation observed during the heatwave period (2.22), compared to before (2.17) and after (2.15) the heatwave. Impervious surfaces exhibited consistently higher UTCI, while vegetated, water, and shaded zones remained relatively cooler across all stages and times.

The distribution of human heat stress categories across the campus based on pixel proportion is illustrated in Supplementary Figure S5 (see Supplementary Table S3 for category definitions). During all heatwave periods, no campus areas experienced “no thermal stress” in the daytime, indicating widespread discomfort. Before the heatwave, much of the campus faced strong or very strong heat stress. During the heatwave, extreme heat stress became prevalent, while nighttime stress levels also increased. Post-heat wave, thermal conditions improved, with no extreme heat stress in the daytime and a rise in moderate heat stress and no thermal stress at night, indicating partial recovery.

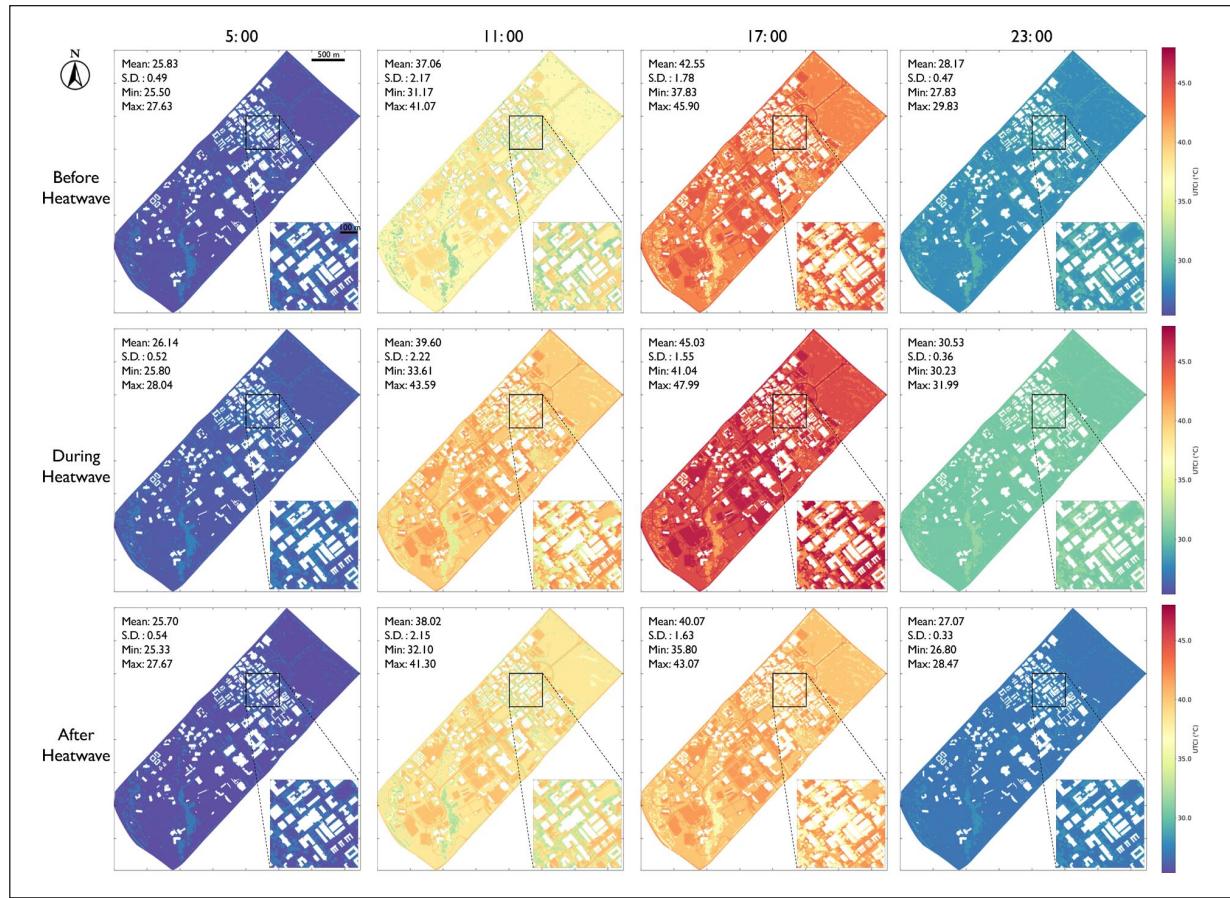


Figure 2. UTCI distributions at four representative hours during different heatwave stages, with each map showing the average UTCI across all days within the stage at the specified hour.

ST-ViT Model Performance

We evaluated the performance of deep learning models across various time step lengths and selected the one with the highest predictive accuracy for further analysis (results for the ST-ViT model are provided in Supplementary Table S4). Supplementary Table S5 compares the overall performance, where the ST-ViT model outperformed the baseline models (CNN+LSTM and U-Net+LSTM) across all metrics, achieving the lowest RMSE (2.163°C), MAE (1.770°C), and MAPE (5.811%). These results demonstrate the ST-ViT model's superiority in UTCI prediction. Additionally, the ST-ViT model exhibited real-time capability, predicting a 1200×1200 pixel image in about 7.2 seconds—sixteen times faster than the QGIS UMEP SOLWEIG-based approach, and 680 times faster when SOLWEIG preprocessing is considered (e.g., SVF calculation) on the same computer.

Figure 3A–P illustrates the diurnal distribution of UTCI across typical campus areas on July 15th at four intervals, comparing ST-ViT predictions (A–D) with SOLWEIG simulations (E–H) and their differences (I–P). Both methods effectively captured the influence of urban morphology on UTCI, including building and tree shadows, as well as thermal

comfort variations across different land cover types. Nighttime predictions (05:00 and 23:00) showed strong agreement between the two models. However, the ST-ViT model underestimated UTCI at 23:00, while it overestimated UTCI at 05:00, 11:00, and 17:00 in most areas, particularly in areas with dense vegetation and building shadows. Additionally, the ST-ViT model tended to smooth daytime UTCI predictions, particularly for extreme values, resulting in narrower distribution ranges and more concentrated peak values.

Figure 3Q–R quantifies ST-ViT prediction errors across land cover types, revealing significant temporal dependence and distinct diurnal cycles. The MAE was generally higher during the daytime, peaking around noon and lowest during the early morning (2:00–6:00). Among land cover types, tree-covered areas exhibited the largest errors and most significant fluctuations, with median MAE around 2.0 and peak values of approximately 3.8 during 12:00–14:00. In contrast, water surfaces consistently showed the smallest errors throughout the diurnal cycle, indicating better model performance in predicting water body thermal environments. Paved surfaces and grass followed the average MAE pattern throughout the day, aligning with the overall temporal trends.

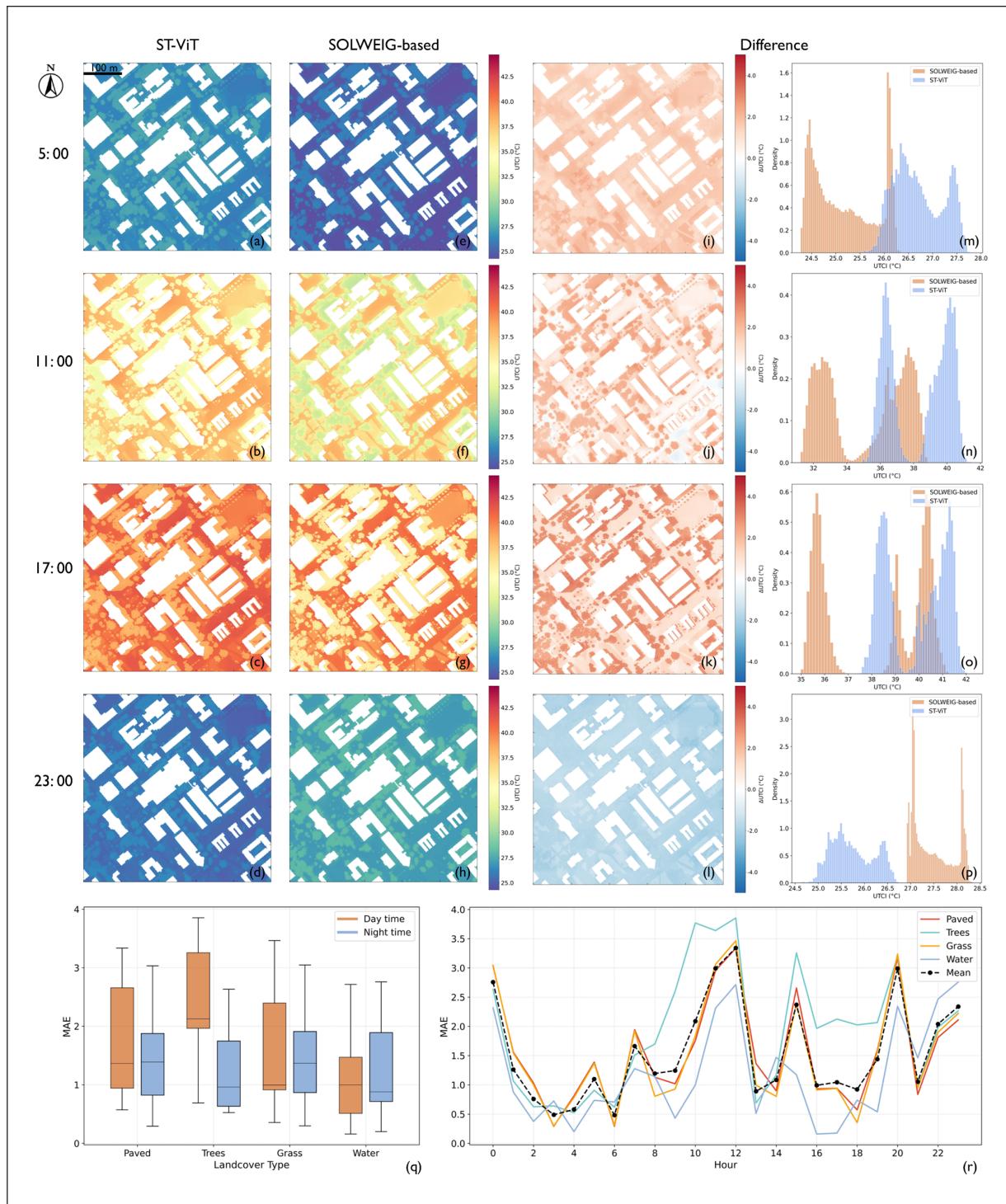


Figure 3. Diurnal distribution of UTCI predictions from ST-ViT (A-D), SOLWEIG-based simulations (E-H), and their differences (I-P) across typical campus areas (500×500 m) on July 15th, 2022; UTCI prediction errors (MAE) of ST-ViT model during daytime and nighttime (Q), and hourly variation (R) across land cover types for the entire study area.

See Supplementary Figure S6 for the landcover classifications.

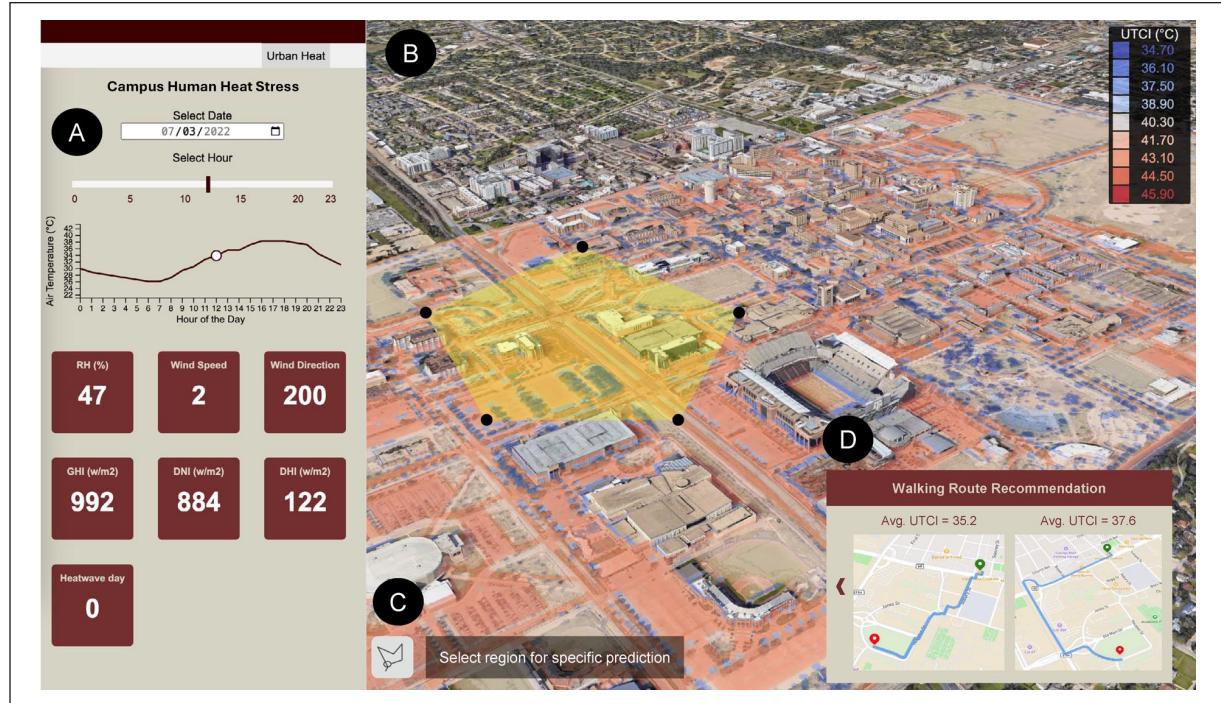


Figure 4. Digital twin platform interface: (A) Time-of-day controller and data summary view: allows users to adjust hour granularity along with detailed data summary; (B) Photorealistic 3D tiles map: visualize UTCI in color heatmap per one-meter; (C) Function: select the region for UTCI forecasting powered by ST-ViT model; and (D) Function: path recommendation based on user's OD selection.

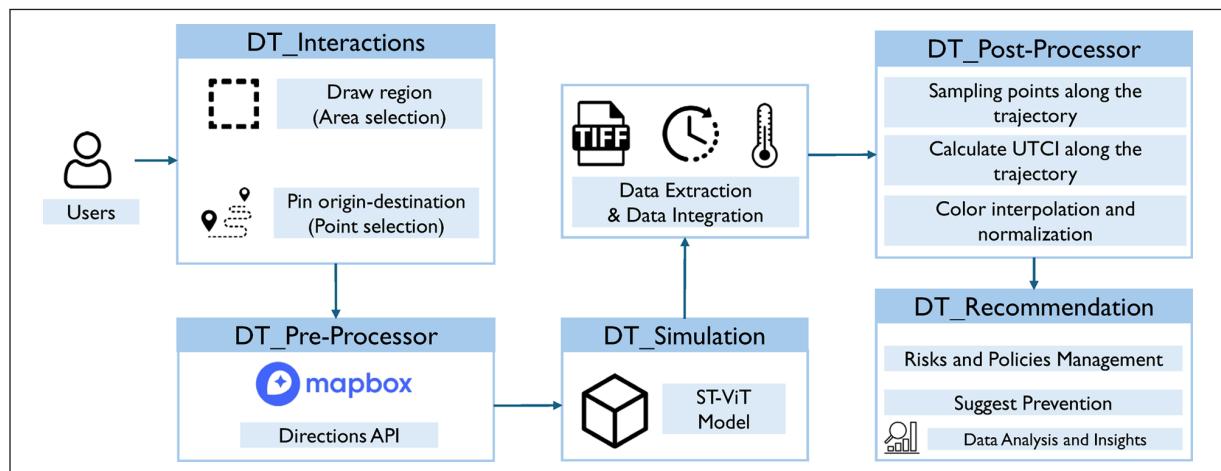


Figure 5. Back-end of the digital twin platform.

Dynamic Digital Twin Platform for Effective Response

The dynamic digital twin platform provides an interactive web-based interface that integrates comprehensive analytical and operational capabilities to address urban thermal comfort challenges. The user interface features four core modules. First, the Time-of-Day Controller and Data Summary View (Figure 4A) enables users to adjust the temporal

granularity (e.g., hourly or daily) for analyzing dynamic UTCI changes, helping understand thermal comfort patterns for resilient planning. Second, the Photorealistic 3D Tiles Map (Figure 4B) visualizes the spatial distribution of UTCI at a high resolution of one meter, using color-coded heatmaps to display thermal hotspots. Third, the Region Selection Function (Figure 4C) allows users to select specific areas for ST-ViT-powered UTCI simulations, enabling tailored analyses for specific user-defined scenarios. Lastly, the Path

Recommendation Module (Figure 4D) generates optimized walking routes with average UTCI values based on user-defined origin and destination. Together, these modules create a highly interactive interface that supports decision-making for urban thermal comfort management.

The back-end operation flow of the platform, illustrated in Figure 5, relies on three core technological components: high-resolution visualization, integrated prediction models, and path recommendation algorithms. For visualization, the platform leverages 3D tiled maps combined with multi-layer data integration to dynamically display thermal comfort distributions. Heatmaps, using color coding at a one-meter resolution, allow users to detect spatial disparities in thermal comfort, while the interface supports interactive functions such as panning, zooming, and overlaying additional layers. These features enable users to identify problem areas and assess the potential impact of various mitigation strategies effectively.

The second component is the integration of advanced prediction models. The platform incorporates pre-trained ST-ViT to predict and simulate dynamic UTCI conditions. The workflow for model integration includes multi-source data preprocessing, model inference, and post-processing for visual output. Input data, such as weather parameters, urban morphology, and vegetation indices, are harmonized in terms of spatial and temporal resolution before being fed into the ST-ViT model. The predictions are then processed into user-friendly visual outputs, such as heat maps and simulation results. This seamless integration of deep learning models enhances computational efficiency while enabling spatiotemporal forecasting of thermal conditions up to twenty-four hours in advance, equipping users with predictive insights to better plan for and mitigate heat stress impacts in dynamic campus environments.

Path recommendation is another key feature of the platform, leveraging grid-based algorithms to optimize walking routes for improved thermal comfort. The grid structure is defined as $G = (V, E)$, where V represents a set of nodes (grid cells) and E represents a set of edges (connections between adjacent grid cells). A walking trajectory P is defined as an ordered node sequence traversing the grid: $P = (v_1, v_2, \dots, v_n)$, where v_1 is the origin (V_i) and V_n is the destination (V_j). The average UTCI for path P is calculated by aggregating the node-specific thermal comfort weights along the trajectory. The formula is:

$$\text{Avg UTCI}(P) = \frac{1}{|P|} \sum_{k=1}^{|P|} W(v_k) \quad (5)$$

where $|P| = n$: Total number of nodes in path P ; $W(v_k)$: UTCI value of the node $v_k \in V$.

The platform leverages algorithms such as Dijkstra, A*, and multi-criteria optimization to generate alternative routes. These algorithms consider both the shortest distance and thermal comfort along the route, allowing users to adjust the weight of each criterion to suit their preferences. This feature

provides users with actionable insights for choosing convenient and comfortable walking paths.

Discussion

Campus Heat Stress Patterns and Mitigation Implications

The spatiotemporal variations of UTCI, calculated based on SOLWEIG simulations, reveal that most campus areas experience strong human heat stress or higher risk levels, underscoring the urgent need for mitigation under extreme heat conditions. The findings emphasize the critical role of land cover types in shaping campus thermal conditions. Paved areas, especially large impervious surfaces like parking lots, consistently exhibit the highest levels of heat stress during the day. This observation aligns with previous studies highlighting the thermal burden of such surfaces, supporting the use of mitigation strategies—such as reflective coatings, permeable pavements, or shading structures—that have been shown in the literature to reduce surface temperatures and enhance thermal comfort (Cheela et al. 2021; H. Li et al. 2013).

Tree-covered areas and water bodies provide substantial daytime cooling benefits by reducing UTCI. However, during nighttime, dense tree canopies may inhibit radiative cooling and airflow (Wujeska-Klause and Pfautsch 2020), resulting in slightly elevated UTCI values. The potential side effects of nighttime warming can be mitigated through thoughtful landscape design, including the selection of tree species with good canopy ventilation, careful control of planting density, and the integration of open spaces (Norton et al. 2015).

Evaluating ST-ViT Model Performance: Efficiency Gains and Systematic Biases

In this study, we developed the ST-ViT model, a novel deep-learning approach for rapidly forecasting spatiotemporal heat stress patterns—unlike static map provision, which offers only snapshots in time. To the best of our knowledge, this is the first application of a Transformer-based architecture for micro-scale modeling of human heat stress at a one-meter resolution on an hourly basis.

Integrated into a decision-support digital twin platform, the ST-ViT model enhances practical usability by effectively modeling UTCI and closely approximating the complexities of the SOLWEIG model, even with unseen data. The hybrid approach, leveraging physics-based SOLWEIG outputs as training data within the ST-ViT architecture, not only preserves essential atmospheric dynamics but also delivers substantial improvements in computational efficiency. This finding aligns with prior studies highlighting the advantages of hybrid models in balancing physical realism and operational feasibility (Jihao Shi et al. 2025; Zhu et al. 2024). By addressing key computational barriers that have traditionally

hindered the deployment of fine-scale climate models, this method makes advanced thermal comfort analysis more accessible to campus planners and facility managers who may lack high-performance computing infrastructure. Furthermore, the ST-ViT model outperformed conventional architectures such as CNN+LSTM and U-Net+LSTM across all evaluation metrics, demonstrating the superior capacity of Transformer-based models to capture complex spatiotemporal dependencies in urban microclimate data.

Despite its advantages, the ST-ViT model exhibited certain systematic biases. It slightly overestimated UTCI during daytime hours (11:00 and 17:00) and early morning (05:00) across most areas, while underestimating it at 23:00. This temporal pattern suggests a degree of diurnal bias, similar to patterns reported in previous studies showing greater UTCI prediction errors during daytime hours in Germany (Briegel et al. 2024). This bias may reflect the model's incomplete representation of nighttime radiative cooling and thermal inertia processes. In addition, the model tended to overestimate UTCI values in areas characterized by dense vegetation and building shadows during daytime hours. This may indicate limitations in its ability to capture localized cooling effects and an incomplete representation of shading and emissivity in the input features (Briegel et al. 2023).

Moreover, the ST-ViT model was prone to smoothing extreme UTCI values, particularly during periods of peak thermal stress. This behavior likely reflects the model's predisposition to prioritize global patterns over localized microclimatic extremes. Similar tendencies have been reported in other deep learning models, including U-Net, which has shown systematic biases when predicting thermal extremes (Briegel et al. 2023).

Human-Centered AI-Based Digital Twins for Planning and Education

The AI-powered digital twin platform developed in this study represents an emerging paradigm for advancing climate-informed curricula and fostering climate-adaptive, sustainable campus environments. This approach also aligns with broader objectives of equity and technological democratization. Universities, as microcosms of urban systems, provide ideal testbeds for applying and teaching intelligent resilience technologies (Rosak-Szyrocka and Wolniak 2025). The ST-ViT-powered digital twin enables deeper exploration of the interconnections between climate, infrastructure, and human behavior, while fostering interdisciplinary collaboration across urban planning, computer science, and climate science. Furthermore, by lowering computational barriers, such platforms hold the potential to empower resource-constrained institutions to adopt similar capabilities through technology transfer and collaborative frameworks. This opens pathways for participation in advanced, participatory resilience planning and contributes to greater equity and accessibility across the higher education sector.

The human-centric design of the digital twin framework further enhances its educational and operational value across multiple dimensions. First, the selection of UTCI as the thermal comfort indicator, rather than basic air temperature metrics, better captures the nuances of actual human thermal perception (Yang et al. 2023). Its high spatial resolution (one meter) enables person-scale insights into microclimatic variability. Second, the platform features an intuitive interface, spatially explicit thermal visualizations, and actionable functions that enable campus stakeholders to forecast dynamic thermal conditions, identify high-risk areas in advance, and explore alternative walking routes. These capabilities translate complex climate model outputs into accessible, practical insights, empowering users to proactively respond to extreme heat. The platform also lays the foundation for future extensions that could support scenario-based evaluations of heat mitigation strategies, reflecting a shift in climate adaptation from retrospective assessment to forward-looking, human-centered planning.

Finally, the deployment of AI-driven digital twins offers a transferable model for broader urban planning applications. Lessons learned from campus environments—whose spatial complexity, pedestrian dynamics, and diverse land uses closely mirror those of urban cores—can inform city-wide climate adaptation strategies, infrastructure resilience planning, data integration approaches, and the development of AI-driven digital twin systems at the city scale. In this way, campus digital twins contribute not only to localized climate risk management but also to the global knowledge base for sustainable urban futures amid escalating climate uncertainty.

Limitations and Outlook

There are some limitations to this study. First, this study focuses on evaluating the performance of the ST-ViT model as a potential alternative to physics-based models, rather than reassessing the SOLWEIG model, which has already been extensively validated in various urban environments, as mentioned in Section 3.3. Second, improvements in model performance can be achieved through hyperparameter optimizations and architectural enhancements, as well as the inclusion of more detailed spatial and meteorological predictors (Zhu et al. 2024). However, the overall results also depend on the performance of the underlying SOLWEIG model.

Third, while the ST-ViT model shows strong performance within the TAMU campus, its generalizability to larger or more heterogeneous regions remains uncertain. Although random cropping of spatial subregions was used during training to improve local spatial generalization, the model has not been explicitly evaluated on entirely unseen spatial areas. Applying the model to regions with different urban morphology, climatic conditions, or data characteristics will require additional validation and potential adaptation to ensure transferability.

Furthermore, the ST-ViT model follows a sequence-to-sequence paradigm, using the preceding twenty-four hours

of data to model spatiotemporal dependencies and forecast the subsequent twenty-four hours of UTCI. While the current implementation focuses on short-term forecasting due to data and computational constraints, the framework is designed to be scalable. It can be retrained on new datasets to accommodate different time periods and adapted to incorporate additional input variables through configuration adjustments. As more historical training data become available or stakeholder needs evolve, the model can also be extended for long-term forecasting, with reduced reliance on recent inputs by leveraging learned seasonal and meteorological patterns to lower computational demands.

Finally, our current framework falls primarily within the predictive stage of the digital twin continuum. Future development could incorporate stakeholder-driven scenario simulation modules, intelligent agents, and true bidirectional connectivity, gradually advancing the system toward the simulative and cognitive stages to support more adaptive, evidence-based decision-making.

Conclusion

Extreme heat poses significant risks to public health and urban environments, particularly in complex and vulnerable settings like university campuses. Addressing these challenges requires innovative approaches that support sustainable campus planning and urban climate adaptation. In this study, we proposed an innovative climate digital twin framework that coupled physics-based simulations to construct a comprehensive database and developed the ST-ViT model for fine-scale spatiotemporal prediction of human heat stress. By incorporating the ST-ViT model into the digital twin framework, this research bridged the gap between advanced AI modeling and practical urban climate management. It demonstrated the potential of combining high-resolution predictive models with decision-support systems to equip campus stakeholders with effective tools to address the growing challenges of extreme heat events in complex urban settings. This approach provided a pathway for enhancing urban resilience planning and sustainability in the face of intensifying climate extremes. The key findings of this study are as follows.

First, most areas on campus experienced very high levels of heat stress under extreme conditions, with paved surfaces exhibiting the highest thermal burden. In contrast, tree canopies and water bodies provided substantial cooling benefits, highlighting the need for sustainable design strategies that enhance thermal comfort while fostering heat resilience.

Second, the ST-ViT model showed its ability to integrate multimodal data and achieve fine-scale spatiotemporal predictions of the UTCI. It outperformed baseline models in accuracy, achieving an MAE of 1.77°C and a computational efficiency sixteen times greater than physics-based models (excluding preprocessing time). These capabilities enable real-time predictions and seamless integration into digital twin platforms, offering a scalable and efficient tool for climate-responsive planning.

Third, the digital twin platform further equips stakeholders with actionable insights for proactive heat stress planning and management. Its features, such as thermal comfort predictions and route recommendations, enhance safety and mobility under extreme heat conditions, providing practical solutions to mitigate risks and improve campus heat resilience.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We greatly appreciate the helpful comments and suggestions from the editors and anonymous reviewers. The research was supported by the National Science Foundation (NSF) under grant CMMI-2430700 and CNS-2401860, NASA under 80NSSC22KM0052, and Texas A&M University Internal Funding. The funders had no role in the study design, data collection, analysis, or preparation of this article.

ORCID iDs

- Wenjing Gong  <https://orcid.org/0009-0003-5463-9125>
- Xinyue Ye  <https://orcid.org/0000-0001-8838-9476>
- Keshu Wu  <https://orcid.org/0000-0002-8085-0990>
- Wenyu Zhang  <https://orcid.org/0000-0002-8931-4123>
- Yifan Yang  <https://orcid.org/0009-0009-9496-6925>
- Xiao Huang  <https://orcid.org/0000-0002-4323-382X>

Supplemental Material

Supplemental material for this article is available online.

References

- Anderson, G. Brooke, and Michelle L. Bell. 2011. "Heat Waves in the United States: Mortality Risk during Heat Waves and Effect Modification by Heat Wave Characteristics in 43 U.S. Communities." *Environmental Health Perspectives* 119 (2): 210–18.
- ANSYS Fluent. 2011. *Ansys Fluent Theory Guide*. Canonsburg, PA: Ansys Inc.
- ASOS. 2024. "AWOS Network." <https://mesonet.agron.iastate.edu/ASOS/>.
- Bibri, Simon Elias, Jeffrey Huang, Senthil Kumar Jagatheesaperumal, and John Krogstie. 2024. "The Synergistic Interplay of Artificial Intelligence and Digital Twin in Environmentally Planning Sustainable Smart Cities: A Comprehensive Systematic Review." *Environmental Science and Ecotechnology* 20 (July): 100433.
- Briegel, Ferdinand, Osama Makansi, Thomas Brox, Andreas Matzarakis, and Andreas Christen. 2023. "Modelling Long-Term Thermal Comfort Conditions in Urban Environments Using a Deep Convolutional Encoder-Decoder as a Computational Shortcut." *Urban Climate* 47 (January): 101359.
- Briegel, Ferdinand, Jonas Wehrle, Dirk Schindler, and Andreas Christen. 2024. "High-Resolution Multi-Scaling of Outdoor

- Human Thermal Comfort and Its Intra-Urban Variability Based on Machine Learning.” *Geoscientific Model Development* 17 (4): 1667–88.
- Bröde, Peter, Dusan Fiala, Krzysztof Błażejczyk, Ingvar Holmér, Gerd Jendritzky, Bernhard Kampmann, Birger Tinz, and George Havenith. 2012. “Deriving the Operational Procedure for the Universal Thermal Climate Index (UTCI).” *International Journal of Biometeorology* 56 (3): 481–94.
- Bruse, Michael, and Heribert Fleer. 1998. “Simulating Surface-Plant–Air Interactions inside Urban Environments with a Three Dimensional Numerical Model.” *Environmental Modelling & Software* 13 (3): 373–84.
- Campbell, Sharon, Tomas A. Remenyi, Christopher J. White, and Fay H. Johnston. 2018. “Heatwave and Health Impact Research: A Global Review.” *Health & Place* 53 (September): 210–18.
- Cárdenas-León, Iván, Rodrigoandrés Morales-Ortega, Mila Koeva, Funda Atún, and Karin Pfeffer. 2024. “Digital Twin-Based Framework for Heat Stress Calculation.” *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences X-4-2024* (October): 67–74.
- Chaowen, Yao, and Pia Fricker. 2021. “How to Cool Down Dense Urban Environments? A Discussion on Site-Specific Urban Mitigating Strategies.” *Journal of Digital Landscape Architecture* 2021 (6): 101–11.
- Cheela, V. R. Sankar, Michele John, Wahidul Biswas, and Prabir Sarker. 2021. “Combating Urban Heat Island Effect—A Review of Reflective Pavements and Tree Shading Strategies.” *Buildings* 11 (3): 93.
- Chu, Zhonghao, and Alessandro F. Rotta Loria. 2024. “Modeling Underground Climate Change across a City Based on Data about a Building Block.” *Sustainable Cities and Society* 114 (November): 105775.
- Čulić, Ana, Sandro Nižetić, Petar Šolić, Toni Perković, and Velimir Čongradac. 2021. “Smart Monitoring Technologies for Personal Thermal Comfort: A Review.” *Journal of Cleaner Production* 312 (August): 127685.
- Deren, Li, Yu Wenbo, and Shao Zhenfeng. 2021. “Smart City Based on Digital Twins.” *Computational Urban Science* 1 (1): 4.
- Dyk, David A. van, and Xiao-Li Meng. 2001. “The Art of Data Augmentation.” *Journal of Computational and Graphical Statistics* 10 (1): 1–50.
- Fujiwara, Kunihiko, Maxim Khomiakov, Winston Yap, Marcel Ignatius, and Filip Biljecki. 2024. “Microclimate Vision: Multimodal Prediction of Climatic Parameters Using Street-Level and Satellite Imagery.” *Sustainable Cities and Society* 114 (November): 105733.
- Georgescu, Matei, Ashley M. Broadbent, and E. Scott Krayenhoff. 2024. “Quantifying the Decrease in Heat Exposure through Adaptation and Mitigation in Twenty-First-Century US Cities.” *Nature Cities* 1 (1): 42–50.
- Göçer, Özgür, Kenan Göçer, Barış Özcan, Mujesira Bakovic, and M. Furkan Kıracı. 2019. “Pedestrian Tracking in Outdoor Spaces of a Suburban University Campus for the Investigation of Occupancy Patterns.” *Sustainable Cities and Society* 45 (February): 131–42.
- Han, Jintong, Adrian Chong, Joie Lim, Savitha Ramasamy, Nyuk Hien Wong, and Filip Biljecki. 2024. “Microclimate Spatio-Temporal Prediction Using Deep Learning and Land Use Data.” *Building and Environment* 253 (April): 111358.
- Hu, Jia, Yuyu Zhou, Yingbao Yang, Gang Chen, Wei Chen, and Mohamad Hejazi. 2023. “Multi-City Assessments of Human Exposure to Extreme Heat during Heat Waves in the United States.” *Remote Sensing of Environment* 295 (September): 113700.
- Jia, Siqi, Yuhong Wang, Nyuk Hien Wong, Chun Liang Tan, Shisheng Chen, Qihao Weng, and Cheuk Ming Mak. 2024. “Estimation of Mean Radiant Temperature across Diverse Outdoor Spaces: A Comparative Study of Different Modeling Approaches.” *Energy and Buildings* 310 (May): 114068.
- Jiang, Yuan, Congzhe Liu, Chenjie Wen, and Yuelin Long. 2024. “Study of Summer Microclimate and PM2.5 Concentration in Campus Plant Communities.” *Scientific Reports* 14 (1): 3311.
- Karyono, Kanisius, Badr M. Abdullah, Alison Cotgrave, Ana Bras, and Jeff Cullen. 2024. “Field Studies of the Artificial Intelligence Model for Defining Indoor Thermal Comfort to Acknowledge the Adaptive Aspect.” *Engineering Applications of Artificial Intelligence* 133 (July): 108381.
- Ketzler, Bernd, Vasilis Naserentin, Fabio Latino, Christopher Zangelidis, Liane Thuvander, and Anders Logg. 2020. “Digital Twins for Cities: A State of the Art Review.” *Built Environment* 46 (4): 547–73.
- Kingma, Diederik P., and Jimmy Ba. 2017. “Adam: A Method for Stochastic Optimization.” *arXiv*.
- Lehtola, Ville V., Mila Koeva, Sander Oude Elberink, Paulo Raposo, Juho-Pekka Virtanen, Faridaddin Vahdatkhaki, and Simone Borsci. 2022. “Digital Twin of a City: Review of Technology Serving City Needs.” *International Journal of Applied Earth Observation and Geoinformation* 114 (November): 102915.
- Li, H., J. T. Harvey, T. J. Holland, and M. Kayhanian. 2013. “The Use of Reflective and Permeable Pavements as a Potential Practice for Heat Island Mitigation and Stormwater Management.” *Environmental Research Letters* 8 (1): 015023.
- Li, Xiaojiang, Guoqing Wang, Benjamin Zaitchik, Angel Hsu, and T. C. Chakraborty. 2024. “Sensitivity and Vulnerability to Summer Heat Extremes in Major Cities of the United States.” *Environmental Research Letters* 19 (9): 094039.
- Li, Xinyu, Yang Xu, Xiaohu Zhang, Wenzhong Shi, Yang Yue, and Qingquan Li. 2023. “Improving Short-Term Bike Sharing Demand Forecast through an Irregular Convolutional Neural Network.” *Transportation Research Part C: Emerging Technologies* 147 (February): 103984.
- Lin, Fudong, Summer Crawford, Kaleb Guillot, Yihe Zhang, Yan Chen, Xu Yuan, Li Chen, et al. 2023. “MMST-ViT: Climate Change-Aware Crop Yield Prediction via Multi-Modal Spatial-Temporal Vision Transformer.” 2023 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, Paris, 5751–61.
- Lindberg, Fredrik, Björn Holmer, and Sofia Thorsson. 2008. “SOLWEIG 1.0—Modelling Spatial Variations of 3D Radiant Fluxes and Mean Radiant Temperature in Complex Urban Settings.” *International Journal of Biometeorology* 52 (7): 697–713.
- Liu, Nayu, Xian Sun, Hongfeng Yu, Wenkai Zhang, and Guangluan Xu. 2021. “D-MmT: A Concise Decoder-Only Multi-Modal Transformer for Abstractive Summarization in Videos.” *Neurocomputing* 456 (October): 179–89.
- Liu, Pengyuan, Tianhong Zhao, Junjie Luo, Binyu Lei, Mario Frei, Clayton Miller, and Filip Biljecki. 2023. “Towards Human-Centric Digital Twins: Leveraging Computer Vision and Graph Models to Predict Outdoor Comfort.” *Sustainable Cities and Society* 93 (June): 104480.

- Liu, Tong, and Chao Fan. 2023. "A Digital Twin Framework to Simulate Urban Microclimates." *ASCE Inspire* 2023, 617–22.
- Maharana, Kiran, Surajit Mondal, and Bhushankumar Nemade. 2022. "A Review: Data Pre-Processing and Data Augmentation Techniques." *Global Transitions Proceedings, International Conference on Intelligent Engineering Approach (ICIEA-2022)* 3 (1): 91–99.
- Mallen, Evan, Joshua Bakin, Brian Stone, Ramachandra Sivakumar, and Kevin Lanza. 2020. "Thermal Impacts of Built and Vegetated Environments on Local Microclimates in an Urban University Campus." *Urban Climate* 32 (June): 100640.
- Meehl, Gerald A., and Claudia Tebaldi. 2004. "More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century." *Science* 305 (5686): 994–97.
- Norton, Briony A., Andrew M. Coutts, Stephen J. Livesley, Richard J. Harris, Annie M. Hunter, and Nicholas S. G Williams. 2015. "Planning for Cooler Cities: A Framework to Prioritise Green Infrastructure to Mitigate High Temperatures in Urban Landscapes." *Landscape and Urban Planning* 134 (February): 127–38.
- NSRDB. 2024. <https://nsrdb.nrel.gov/>.
- Omrany, Hossein, and Karam M. Al-Obaidi. 2024. "Application of Digital Twin Technology for Urban Heat Island Mitigation: Review and Conceptual Framework." *Smart and Sustainable Built Environment*. Published electronically October 2. doi: 10.1108/SASBE-05-2024-0189.
- Pan, Xiyu, Dimitris Mavrokannidis, Hoang T. Ly, Neda Mohammadi, and John E. Taylor. 2024. "Assessing and Forecasting Collective Urban Heat Exposure with Smart City Digital Twins." *Scientific Reports* 14 (1): 9653.
- Peldon, Dechen, Saeed Banihashemi, Khuong LeNguyen, and Sybil Derrible. 2024. "Navigating Urban Complexity: The Transformative Role of Digital Twins in SmartCity Development." *Sustainable Cities and Society* 111 (September): 105583.
- "Point Cloud Classification Using PointCNN." 2024. ArcGIS API for Python. <https://developers.arcgis.com/python/guide/point-cloud-segmentation-using-pointcnn/>.
- Ramani, Vasantha, Marcel Ignatius, Joie Lim, Filip Biljecki, and Clayton Miller. 2023. "A Dynamic Urban Digital Twin Integrating Longitudinal Thermal Imagery for Microclimate Studies." *Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, BuildSys '23*, Association for Computing Machinery, New York, 421–28.
- Riaz, Khurram, Marion McAfee, and Salem S. Gharbia. 2023. "Management of Climate Resilience: Exploring the Potential of Digital Twin Technology, 3D City Modelling, and Early Warning Systems." *Sensors* 23 (5): 2659.
- Rosak-Szyrocka, Joanna, and Radosław Wolniak. 2025. *Smart Universities in Smart Cities: Shaping the Future of Education and Urban Innovation*. Abingdon: Taylor & Francis.
- Shi, Jihao, Junjie Li, Haoran Zhang, Bin Xie, Zonghao Xie, Qing Yu, and Jinyue Yan. 2025. "Real-Time Gas Explosion Prediction at Urban Scale by GIS and Graph Neural Network." *Applied Energy* 377 (January): 124614.
- Shi, Jing, Jinmei Guo, and Songtao Zheng. 2012. "Evaluation of Hybrid Forecasting Approaches for Wind Speed and Power Generation Time Series." *Renewable and Sustainable Energy Reviews* 16 (5): 3471–80.
- Shim, Jae-hun, Hyunwoo Yu, Kyeongbo Kong, and Suk-Ju Kang. 2023. "FeedFormer: Revisiting Transformer Decoder for Efficient Semantic Segmentation." *Proceedings of the AAAI Conference on Artificial Intelligence* 37 (2): 2263–71.
- Skamarock, William, Joseph Klemp, Jimy Dudhia, David O Gill, Zhiquan Liu, Judith Berner, Wei Wang, et al. 2019. "A Description of the Advanced Research WRF Model Version 4.1." *NCAR Technical Note* 475 (125). <https://opensky.ucar.edu/islandora/object/%3A3888>.
- Texas 2036. 2023. "Extreme Heat Remains on Track to Become Texas' New Normal." Texas 2036 (blog). December 12. <https://texas2036.org/posts/extreme-heat-remains-on-track-to-become-texas-new-normal/>.
- USGS. 2024. <https://apps.nationalmap.gov/downloader/#/>.
- White, Gary, Anna Zink, Lara Codecá, and Siobhán Clarke. 2021. "A Digital Twin Smart City for Citizen Feedback." *Cities* 110 (March): 103064.
- Wu, Binrong, Lin Wang, and Yu-Rong Zeng. 2022. "Interpretable Wind Speed Prediction with Multivariate Time Series and Temporal Fusion Transformers." *Energy* 252 (August): 123990.
- Wujeska-Klause, Agnieszka, and Sebastian Pfautsch. 2020. "The Best Urban Trees for Daytime Cooling Leave Nights Slightly Warmer." *Forests* 11 (9): 945.
- Xia, Haishan, Zishuo Liu, Maria Efremochkina, Xiaotong Liu, and Chunxiang Lin. 2022. "Study on City Digital Twin Technologies for Sustainable Smart City Design: A Review and Bibliometric Analysis of Geographic Information System and Building Information Modeling Integration." *Sustainable Cities and Society* 84 (September): 104009.
- Yang, Senwen, Liangzhu (Leon) Wang, Ted Stathopoulos, and Ahmed Moustafa Marey. 2023. "Urban Microclimate and Its Impact on Built Environment—a Review." *Building and Environment* 238 (June): 110334.
- Ye, Xinyue, Jiaxin Du, Yu Han, Galen Newman, David Retchless, Lei Zou, Youngjib Ham, and Zhenhang Cai. 2023. "Developing Human-Centered Urban Digital Twins for Community Infrastructure Resilience: A Research Agenda." *Journal of Planning Literature* 38 (2): 187–99.
- Ye, Xinyue, Suphanut Jamonnak, Shannon Van Zandt, Galen Newman, and Patrick Suermann. 2024. "Developing Campus Digital Twin Using Interactive Visual Analytics Approach." *Frontiers of Urban and Rural Planning* 2 (1): 9.
- Yin, Lirong, Lei Wang, Tingqiao Li, Siyu Lu, Jiawei Tian, Zhengtong Yin, Xiaolu Li, and Wenfeng Zheng. 2023. "U-Net-LSTM: Time Series-Enhanced Lake Boundary Prediction Model." *Land* 12 (10): 1859.
- Zheng, Lang, and Weisheng Lu. 2024. "Urban Micro-Scale Street Thermal Comfort Prediction Using a 'Graph Attention Network' Model." *Building and Environment* 262 (August): 111780.
- Zheng, Zhonghua, Lei Zhao, and Keith W. Oleson. 2021. "Large Model Structural Uncertainty in Global Projections of Urban Heat Waves." *Nature Communications* 12 (1): 3736.
- Zhu, Hao-Cheng, Chen Ren, Junqi Wang, Zhuangbo Feng, Fariborz Haghhighat, and Shi-Jie Cao. 2024. "Fast Prediction of Spatial Temperature Distributions in Urban Areas with WRF and Temporal Fusion Transformers." *Sustainable Cities and Society* 103 (April): 105249.

Author Biographies

Wenjing Gong is a PhD student at the Department of Landscape Architecture and Urban Planning & Center for Geospatial Sciences, Applications and Technology, Texas A&M University. Her research interests lie at the intersection of Urban AI, Urban Planning, and Climate Resilience.

Xinyue Ye is an Endowed Shelby Distinguished Professor of GeoAI (Geospatial AI) at the University of Alabama. His research integrates GeoAI with disciplines spanning the humanities to engineering, addressing challenges such as infrastructure resilience, extreme weather preparedness, urban revitalization, and community perceptions.

Keshu Wu is a postdoctoral research associate at Texas A&M University. His research aims to develop interactive, trustworthy intelligent systems comprised of autonomous agents, capable of understanding and making rational decisions in the physical environment.

Suphanut Jamonnak is a postdoctoral research associate at Texas A&M University. His research interests include geospatial analysis, visual analytics, Human-computer interaction (HCI), and eXplainable AI (XAI).

Wenyu Zhang is a PhD student at the Department of Landscape Architecture and Urban Planning & Center for Geospatial Sciences, Applications and Technology, Texas A&M University. Her research interests lie in GeoAI and Urban Informatics.

Yifan Yang is a PhD student at the Department of Geography, Texas A&M University. His research interests lie in Autonomous GeoAI Science, Spatial Data Science, GeoAI, AI4Science, and GIScience.

Xiao Huang is an Assistant Professor at the Department of Environmental Sciences, Emory University. His research primarily focuses on human-environment interaction, computational social sciences, vulnerability and resilience, urban informatics, disaster mapping and mitigation, GeoAI and deep learning, and disaster remote sensing.