GeoAI-Based High-Resolution Mapping of Thermal Stress and Surface Interventions in Bangladesh

Bijoy Mitra, Guiming Zhang

Department of Geography & the Environment, College of Natural Sciences and Mathematics, University of Denver, Denver, CO 80208, USA

1. Introduction:

Global warming is worsening heat loads, negatively impacting human health, labor productivity, biodiversity, the environment, and the economy [1,2]. Consequently, it is essential to comprehend the appropriate mitigation measures to prevent heat stress. The Universal Thermal Climate Index (UTCI) is a metric used to assess anthropogenic thermal comfort from a set of meteorological parameters. However, the only open-source UTCI is available from the ERA5-HEAT in a resolution of ~27.5 km, with a daily temporal coverage [3]. Such a dataset is too coarse and is not suitable for small-scale studies in a data-scarce region like Bangladesh. Bangladesh faces severe heat exposure due to its dense population and reliance on outdoor agricultural labor. No high-resolution UTCI product currently exists for Bangladesh, thereby, studies solely depend on interpolated images missing critical microclimatic thermal information.

This study bridges this gap by developing a UTCI downscaling strategy based on a CNN algorithm and biophysical and topographic features. Explainable AI analysis was utilized to examine the collective and localized impact of the features on UTCI. Furthermore, model-based counterfactual analysis was adopted to support heat stress management decision-making by quantifying the effects of anthropogenic actions on UTCI. Collectively, these approaches constitute a new combination of biophysical downscaling, explainability, and counterfactual analysis for fine-scale UTCI mapping in a data-scarce area.

2. Methodology:

2.1 Study Area:

Bangladesh is located at the south of Asia (20°34' - 26°38' N, 88°01' - 92°41' E). It has a tropical monsoon climate with three distinct seasons. The temperature often goes over comfort level in Bangladesh, particularly during summer. A spatial distribution of mean UTCI between 2001 and 2025 (based on ERA-5 HEAT UTCI dataset [4]) is presented in Fig. 1.

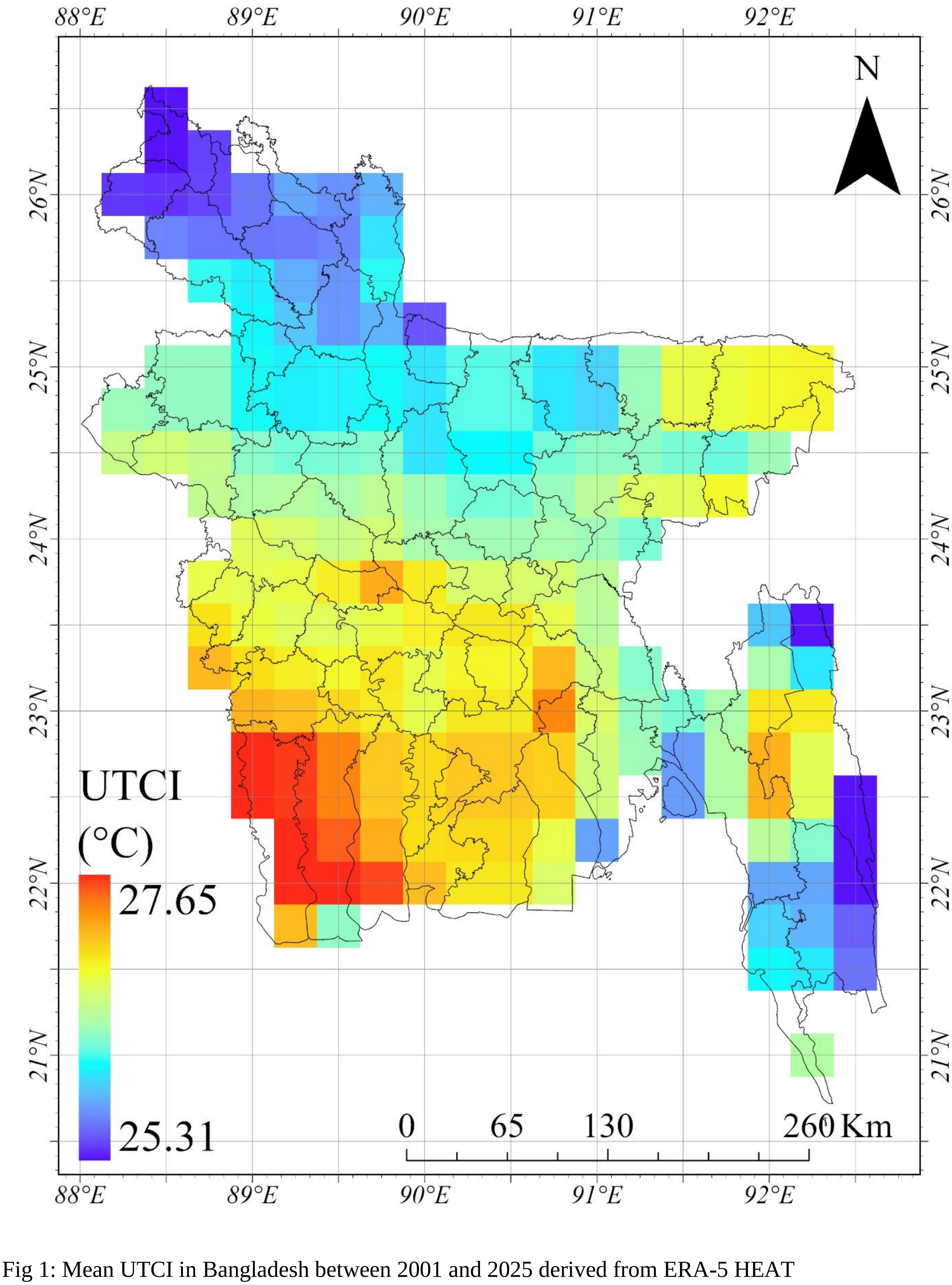

Fig 1: Mean UTCI in Bangladesh between 2001 and 2025 derived from ERA-5 HEAT

2.2 Data:

The monthly mean UTCI data for Bangladesh was collected from ERA5-HEAT for the period January 2001 and March 2026. For biophysical and topographic features, NDVI, NDMI, NDBI, MNDWI, Albedo, and LST were collected from the MODIS Terra instrument [5] and DEM was collected from Copernicus-DEM [6]. From the DEM raster, Slope and Aspect was calculated in ArcGIS Pro [7]. All feature raster layers, regardless of native resolution, were interpolated to 500 m using IDW for downscaling UTCI. The list of the selected features, their respective data source, unit, and spatial resolution is presented in Table 1.

Table 1: Dataset Description and Resolution

Variable	Unit	Data Source	Spatial Resolution	Temporal Resolution
Universal Thermal Climate Index (UTCI)	°C	ERA-5 HEAT	~27.5 km	Model prediction period: 2001-2025, Unseen prediction period January-March, 2026
Digital Elevation Model (DEM)	m	Copernicus DEM GLO-30: Global 30m Digital Elevation Model	30 m	
Slope	Degrees			
Aspect	Degrees			
Normalized Difference Vegetation Index (NDVI)	-	MODIS Terra Surface Reflectance (MOD09A1.061)	500 m	
Normalized Difference Moisture Index (NDMI)				
Normalized Difference Built-up Index (NDBI)				
Modified Normalized Difference Water Index (MNDWI)				
Albedo				
Land Surface Temperature (LST)	°C	MODIS Terra Land Surface Temperature and Emissivity Daily Global (MOD11A1.061)	1000 m	

2.3 Sampling Technique:

For the entire Bangladesh, the ERA-5 UTCI data set comprises only 205 pixels. Therefore, we took UTCI values at those 205 sample locations on a monthly resolution during the study period (2001-2025) and extracted feature values using zonal mean from the feature layers, in the corresponding monthly period for biophysical features, resulting in a total number of 58500 samples. Along with that, transformed month as sine and cosine features were incorporated as independent features to represent seasonal cycles of UTCI. Samples in 2001-2019, 2020-2022, and 2023-2025 were used for model training, validation, and testing, respectively. Apart from the training and the testing set, the validation set helped tune model parameters and prevent overfitting before final testing. The developed model was then used to predict the monthly mean UTCI in Bangladesh at 500 m resolution for January - March, 2026.

2.4 Model Development:

We adopted a CNN algorithm [8] to develop the UTCI downscaling strategy and offer effective insights for heat stress management policy making based on feature sensitivity analyses (Fig. 2). The CNN architecture employs four convolutional blocks with batch normalization and dropout to reduce overfitting risk. Global average pooling improves model stability, allowing for more robust and dependable UTCI predictions for Bangladesh. The model design is suitable for modeling our time-series data as it captures dependencies within the monthly mean values of the samples, stacked across months over the years.

2.5 Model Evaluation Statistics:

During model development, R^2, RMSE, and MAE was calculated on training, validation, and testing datasets.

2.6 Explainable AI methods:

To comprehend the global feature importance in the CNN model, we analyzed the SHAP (SHapley Additive exPlanations) value [9]. This offers critical insights regarding the magnitude and direction of each of the features' impacts on UTCI.

Further, to understand the localized feature importance on model prediction, Partial Dependence Plots (PDP) [10] and Accumulated Local Effects (ALE) [11] was employed to examine the mean marginal effects on original (PDP) and correlation-unbiased (ALE) anomaly scale of UTCI.

2.7 Counterfactual analysis:

Instead of relying solely on Pearson's correlation or Explainable AI techniques, which are highly reliant on model learning, we adopted the counterfactual analysis method, perturbing individual features and observing model-predicted UTCI changes, to aid effective heat stress management decision-making by analyzing what changes would be required on the features to achieve a 2°C reduction in UTCI. Counterfactual analysis forecasts what the model would predict if the input features changed. We also examined pairwise sensitivity using three-dimensional response surfaces, capturing two-way interactions only and excluding higher-order combinations.

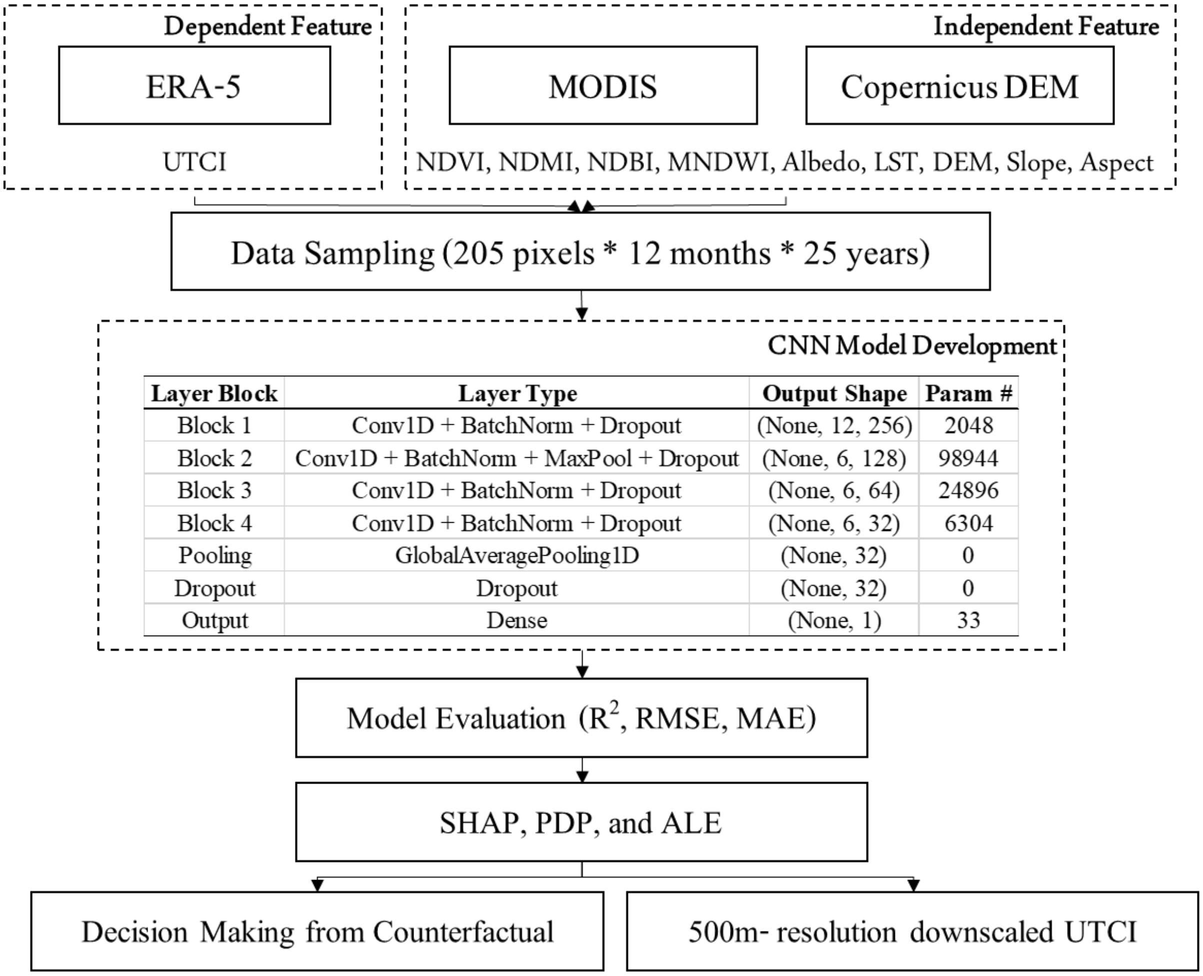

Fig 2: Methodological workflow

3. Result Analysis:

The UTCI has a significant but moderately low correlation with NDBI, Albedo, DEM, and Slope, while a positive association with NDMI and LST (Fig. 3).

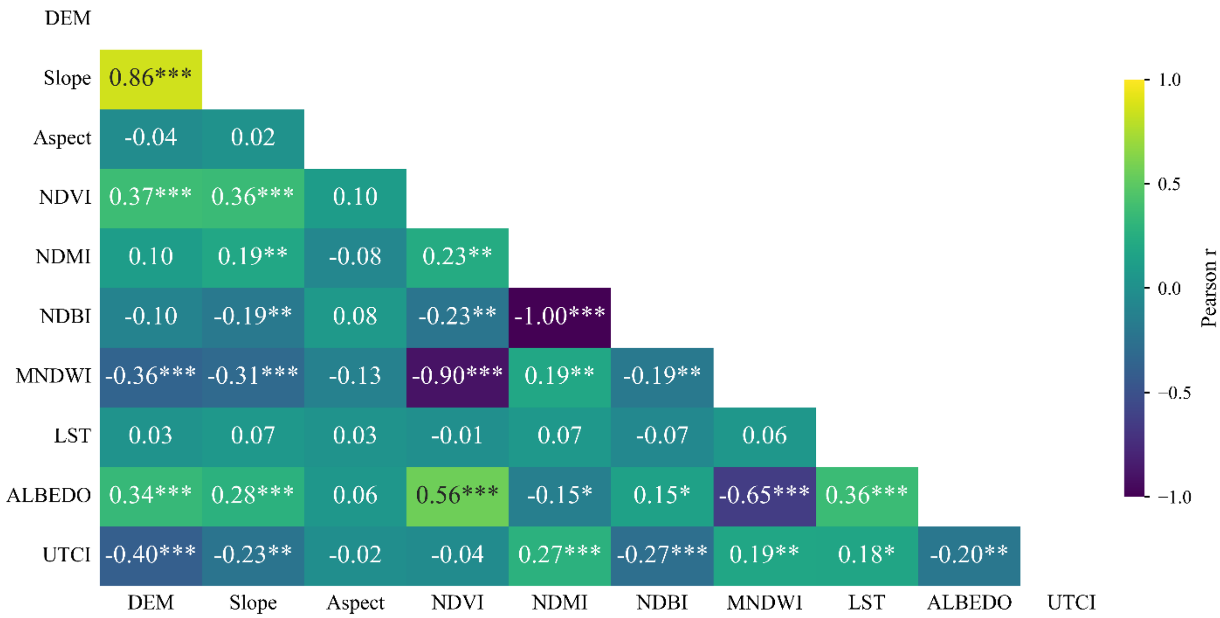

Fig 3: Pearson’s correlation between the selected features

The scatterplot (Fig. 4) shows the observed and predicted UTCI across the three datasets, with an R^2 of 0.934, RMSE of 1.27°C, and MAE of 1.04°C on the testing dataset.

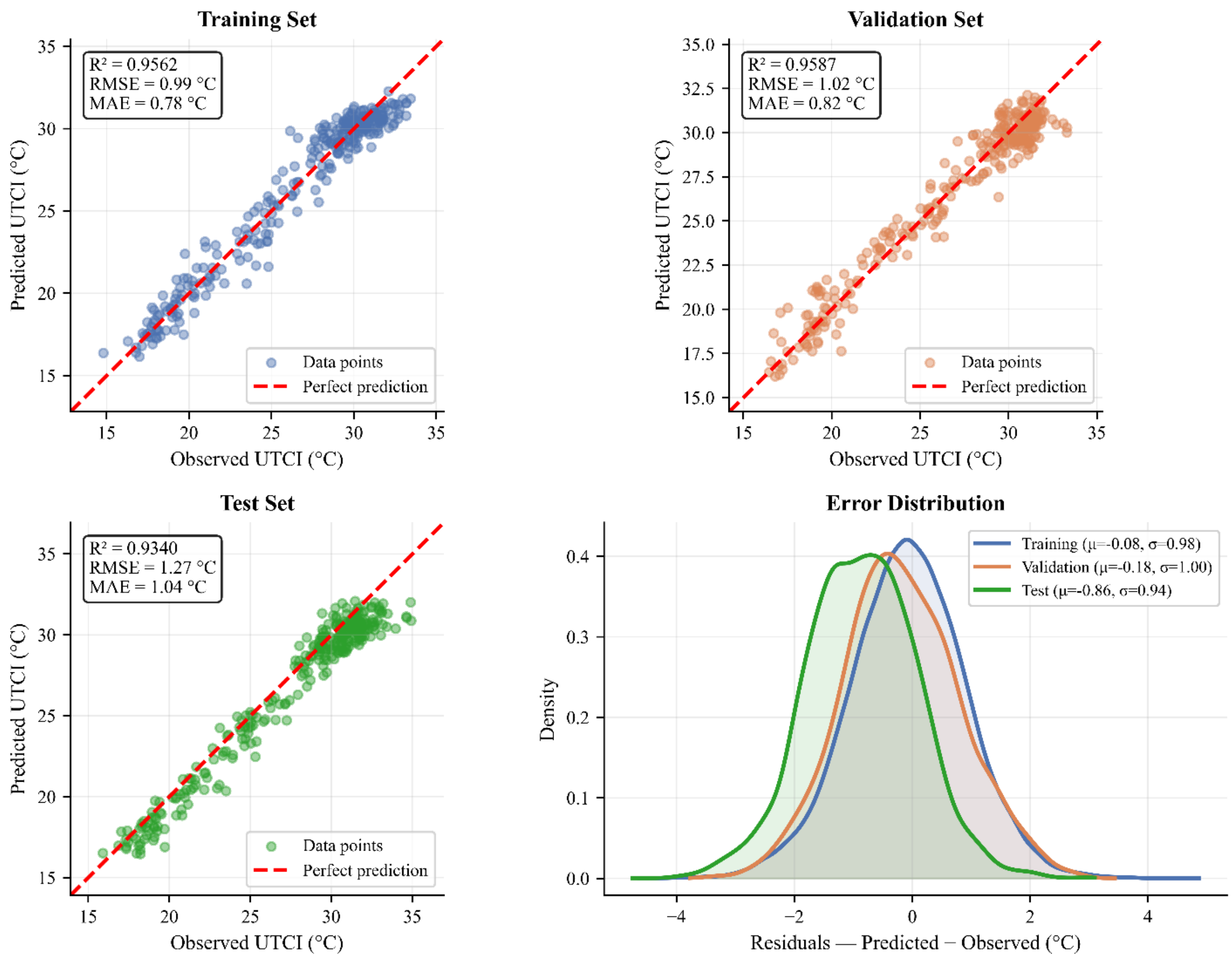

Fig 4: Model performance evaluation on the training, validation, and testing sets

The SHAP analysis reveals that LST has the highest mean SHAP value, and high LST values are associated with high UTCI values. LST was followed by DEM and NDVI, where the features had positive and negative impacts on UTCI, respectively.

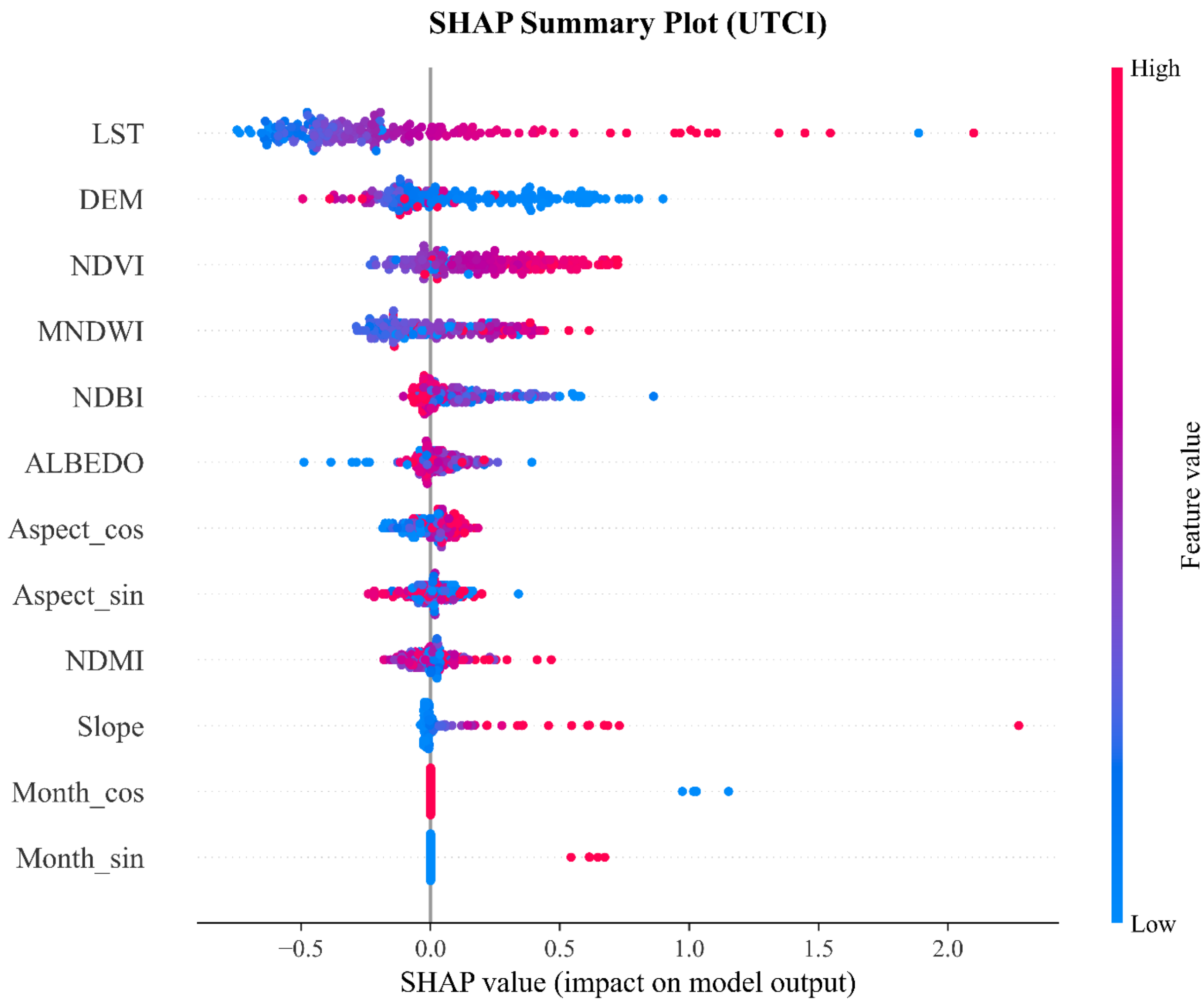

Fig 5: SHAP summary analysis to predict UTCI

From the PDP and ALE analysis, LST had the most marginal impact, where it reveals that a 20°C LST value is the triggering point to start UTCI increase (Fig. 6). The most non-linear effect was observed for NDVI, as UTCI increased till an NDVI value of -0.25, then plummeted till an NDVI value of 0.25, and escalated sharply after that.

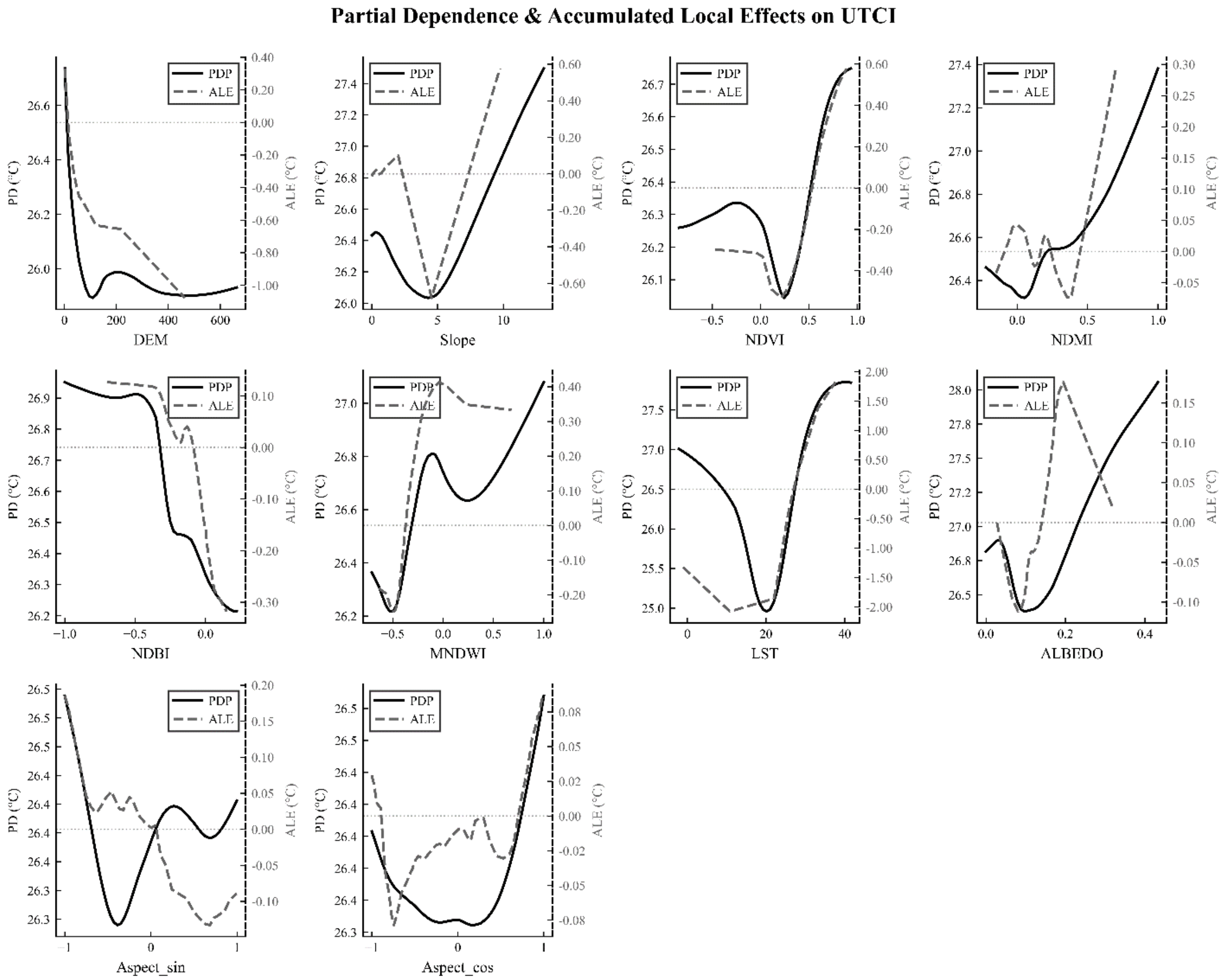

Fig 6: Partial dependence and accumulated local effects for key features

Fig. 7 shows observed and predicted UTCI in Bangladesh from January to March 2026. The CNN model demonstrates good potential to produce high-resolution (500 m) UTCI data to support heat stress management at fine spatial scales. The prediction MAE score was between 0.892 – 1.198°C.

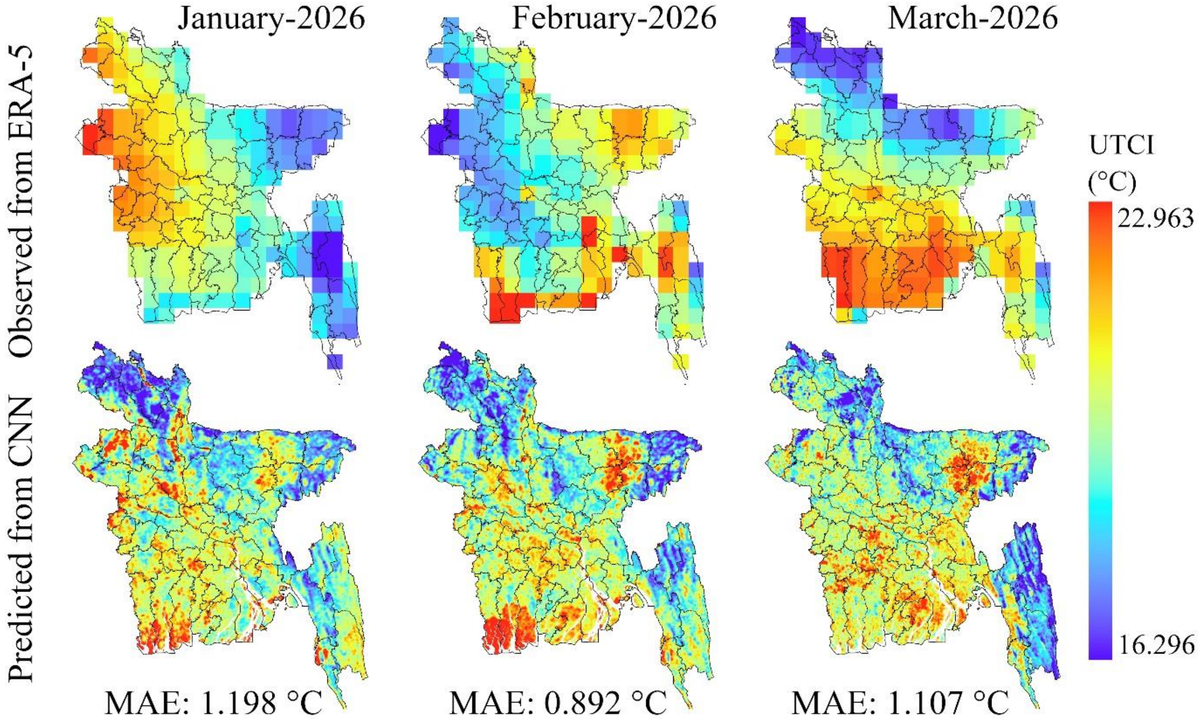

Fig 7: Observed (~27.5 km resolution) and model predicted (500 m resolution) UTCI in Bangladesh for January - March, 2026.

4. Discussion:

We achieved a high accuracy of $R^2 > 0.9$ in predicting UTCI from topographic and biophysical variables. Explainable AI analyses reveal that UTCI is highly nonlinear to the variables across locations. Counterfactual analysis was performed to determine how much each feature must change to reduce UTCI by a prescribed amount (e.g., 2°C). The analysis reveals that LST and NDMI have the most influence on UTCI reduction (Fig. 8).

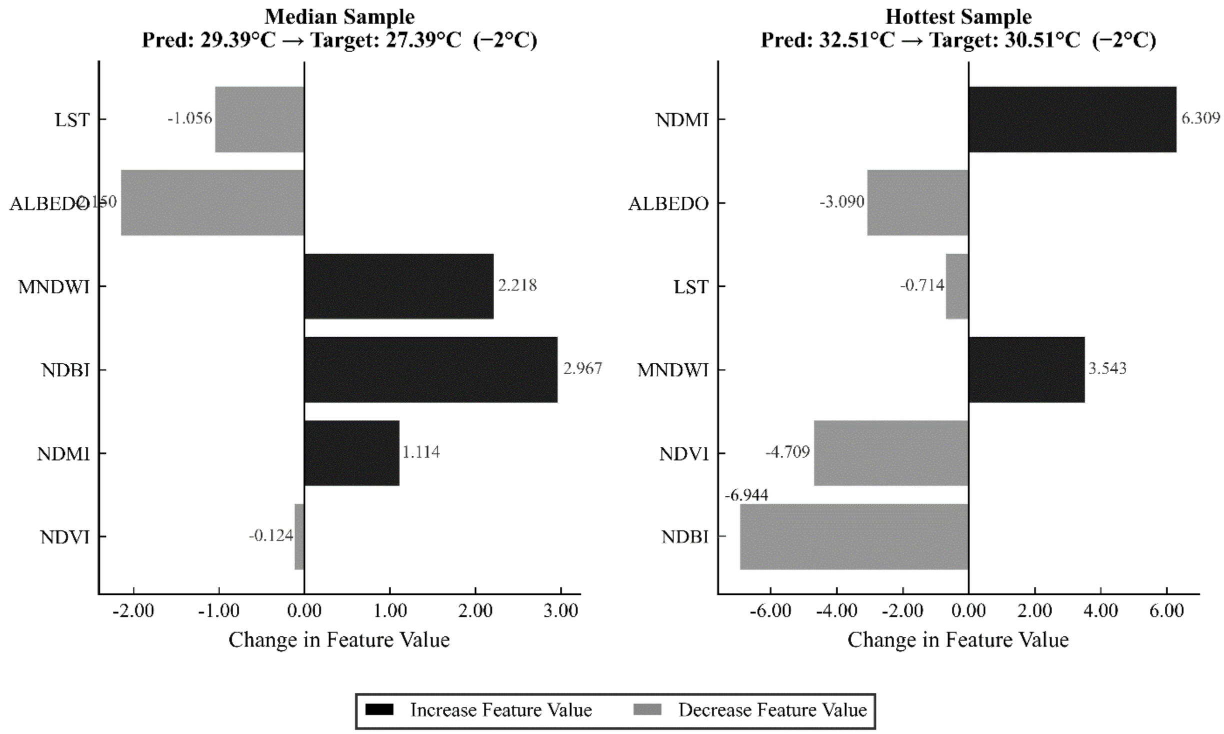

Fig 8: Counterfactual feature perturbations required to achieve a 2°C reduction at locations of the median and maximum UTCI, respectively.

The three-dimensional pairwise sensitivity visualization (Fig. 9) shows that LST, NDBI and MNDWI-driven cooling is more effective in reducing UTCI than albedo and vegetation-based interventions.

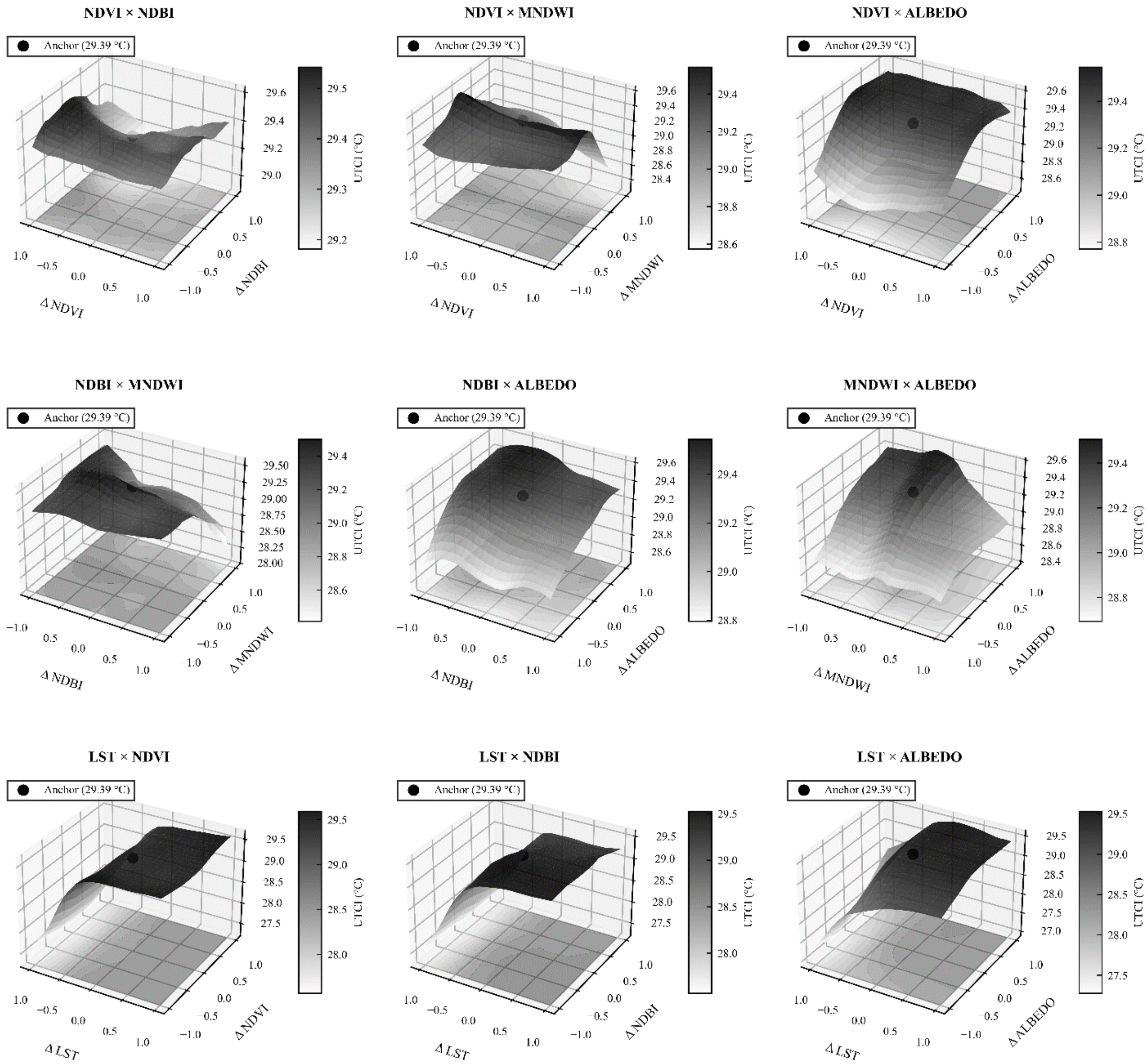

Fig 9: Three-dimensional response surfaces illustrating the sensitivity of predicted UTCI to pairwise variations in land surface biophysical parameters

Overall, counterfactual analysis indicates that effective measures should be adopted to control land surface temperatures and built-up area, and to increase vegetation and surface water to reduce UTCI in Bangladesh.

5. Conclusion

The study presents a novel GeoAI-based strategy (CNN model) to downscale UTCI to 500 m resolution using topographic and biophysical variables, achieving a mean absolute error of 1.04°C on the testing dataset. Model interpretability analysis reveals that LST has the greatest positive impact on UTCI, and NDVI shows the least consistent relationship with UTCI. Based that the model, a monthly UTCI prediction (forecast) was produced for January - March, 2026, demonstrating the potential of using the downscaled UTCI data to support heat stress management at fine spatial scales.

The counterfactual analysis revealed that changing LST, NDVI, and NDMI are the most effective in minimizing thermal stress in Bangladesh. While the CNN architecture is not new, this work introduces a unique application for fine-scale UTCI mapping. Since the model relies on interpolated, coarse-resolution ERA5-derived UTCI, future research should incorporate station-based meteorological data with short temporal intervals for cross-validating model output and 30-m resolution Landsat data products to downscale UTCI to even finer spatial and temporal scales.

References

1. Kyaw, A.K.; Hamed, M.M.; Shahid, S. Spatiotemporal Changes in Universal Thermal Climate Index over South Asia. *Atmos. Res.* **2023**, *292*, 106838, doi:10.1016/j.atmosres.2023.106838.

2. Rahman, F.; Hossain, M.D.; Akhter, Md.A.E. Study of Heat Wave Events by the Thermophysiological Comfort Indexes in Khulna City, Bangladesh, during Pre-Monsoon and Monsoon Seasons. *City and Environment Interactions* **2025**, *28*, 100275, doi:10.1016/j.cacint.2025.100275.

3. Urban, A.; Di Napoli, C.; Cloke, H.L.; Kyselý, J.; Pappenberger, F.; Sera, F.; Schneider, R.; Vicedo-Cabrera, A.M.; Acquaotta, F.; Ragettli, M.S.; et al. Evaluation of the ERA5 Reanalysis-Based Universal Thermal Climate Index on Mortality Data in Europe. *Environ. Res.* **2021**, *198*, 111227, doi:10.1016/j.envres.2021.111227.

4. Di Napoli, C.B.C.P.C.C.H.L.P.F. *Thermal Comfort Indices Derived from ERA5 Reanalysis*; 2020;

5. Vermote, E. *MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V061*; 2021;

6. CDSE *Copernicus DEM GLO-30*; 2021;

7. Esri Inc. *ArcGIS Pro (Version 3.6)*; Redlands, C A, 2025;

8. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE* **1998**, *86*, 2278–2324, doi:10.1109/5.726791.

9. Shapley, L.S. Stochastic Games. *Proceedings of the national academy of sciences* **1953**, *39*, 1095–1100.

10. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics* **2001**, *29*, doi:10.1214/aos/1013203451.

11. Apley, D.W.; Zhu, J. Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models. *J. R. Stat. Soc. Series B Stat. Methodol.* **2020**, *82*, 1059–1086, doi:10.1111/rssb.12377.