The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-9-2025
30th CIPA Symposium “Heritage Conservation from Bits:
From Digital Documentation to Data-driven Heritage Conservation”, 25–29 August 2025, Seoul, Republic of Korea

Towards Sustainable Heritage Conservation: Transformer-Based temperature forecasting in the city of Florence

Fabiana Di Ciaccio[1], Paolo Russo[2], Erica Isabella Parisi [1], Riccardo Angelini[1], Grazia Tucci[1]

[1] Deparment of Civil and Environmental Engineering, University of Florence, Italy - (fabiana.diciaccio, ericaisabella.parisi, riccardo.angelini, grazia.tucci)@unifi.it

[2] Department of Computer, Control and Management Engineering, Sapienza University of Rome, Italy - paolo.russo@diag.uniroma1.it

Keywords: Climate Change Impact, Cultural Heritage Preservation, Deep Learning, Geospatial Analysis, Satellite Remote Sensing, Surface Land Temperature.

Abstract

Urban Heat Islands (UHIs) are intensifying under climate change, posing serious threats to the conservation of cultural heritage in historic cities. This study presents a deep learning framework for forecasting land surface temperature (LST) at high temporal resolution, aimed at supporting thermal risk mitigation in heritage-sensitive urban environments. Using hourly LST data from the Copernicus Land Monitoring Service (CLMS), we developed a custom autoregressive Transformer model capable of predicting 72 future hourly temperature values based on the previous 168 hours of satellite observations. The model follows an encoder-decoder architecture: the encoder processes the full input sequence with spatial and temporal embeddings, while the decoder generates the forecast step-by-step, leveraging past predictions. We applied our approach to the city of Florence, Italy, using available data from the years 2021, 2022, 2023, and 2024. The model was evaluated on the four summer months (June 2021, July 2022, August 2023 and September 2024) achieving a minimum MAE of 0.81° C and confirming its applicability for extreme heat forecasting.

1. Introduction

Urban Heat Islands (UHIs) represent one of the most critical climate-related challenges for contemporary and historic cities. The phenomenon, driven by the concentration of anthropogenic heat sources, impermeable surfaces, and limited vegetation, leads to elevated temperatures in urban cores compared to surrounding rural areas, often around 7°C (Heisler and Brazel, 2010). This thermal anomaly becomes particularly severe during heatwaves, amplifying environmental discomfort, increasing energy consumption for cooling, and, most significantly, accelerating the degradation processes of built heritage (Camuffo, 2019).

These risks are especially acute in cities with a high density of culturally and architecturally significant sites (Quesada-Ganuza et al., 2023). Florence, Italy, stands as a paradigmatic case: its historic center, a UNESCO World Heritage Site, hosts masterpieces of Renaissance architecture and art whose conservation is tightly coupled to environmental stability. For example, during the 2017 heatwave, peak temperatures forced the temporary closure of the Uffizi Gallery, underlining the vulnerability of such spaces to climate extremes (Morabito et al., 2017).

The study and mitigation of UHIs has traditionally relied on remote sensing, Geographic Information Systems (GIS), and statistical modeling (Shi et al., 2021). More recently, the availability of large-scale satellite data, such as those from the Copernicus Earth observation programme, has enabled more granular and temporally rich analyses of urban microclimates (Bühler et al., 2021). In particular, the Copernicus Land Monitoring Service (CLMS) provides hourly Land Surface Temperature (LST) products at approximately 5 km resolution, based on geostationary satellite observations. These data are highly relevant for monitoring surface-level heat dynamics, particularly in urban contexts where surface temperature often diverges significantly from ambient air temperature (Copernicus Land Monitoring Service, 2025).

Despite the growing availability of such datasets, relatively few studies have focused on the specific impact of UHIs on cultural heritage assets (Pioppi et al., 2020). Furthermore, there is a growing need for tools that not only monitor but also predict short-term thermal dynamics at city scale, providing actionable insights for heritage protection, urban adaptation strategies, and early warning systems (Li et al., 2024).

Recent advancements in Artificial Intelligence, and in particular Deep Learning, have demonstrated strong capabilities in learning from spatiotemporal environmental data (Han et al., 2023). Architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and more recently, Transformers, have been successfully applied to a range of heterogeneous tasks, from temperature forecasting (Alerskans et al., 2022) and high-resolution mapping of urban thermal conditions (Tehrani et al., 2024) to anomaly detection (Russo and Schaerf, 2023) and sequence to sequence mapping (Yun et al., 2020). Among them, Transformer-based models offer unique advantages in capturing long-range dependencies in time-series data, making them particularly suitable for medium-term forecasting applications (Zhou et al., 2021).

In this work, we propose a custom autoregressive Transformer model designed to forecast urban land surface temperatures at hourly resolution over a three-day horizon, using as input the previous seven days of CLMS hourly LST data. The model follows an encoder-decoder structure: the encoder receives the entire past signal, encoding its spatiotemporal characteristics, while the decoder generates the output signal step-by-step, conditioning each new prediction on the previously generated values. This approach mimics the natural temporal causality of prediction tasks and allows the network to refine its estimates as the sequence unfolds.

The objective is twofold: (1) to demonstrate the feasibility of Transformer-based approaches for short-term forecasting of LST fields in urban environments, and (2) to assess their potential in supporting risk-informed conservation planning for historic city centers. The city of Florence is adopted as a case study due to its climatic sensitivity and cultural value. This research aims to support the integration of predictive climate modeling into decision-support tools for urban heritage management. Our results show in fact promising performance, with reasonable forecast accuracy even under challenging seasonal generalization settings.

2. Related Works

The phenomenon of Urban Heat Islands has been widely documented in the literature, with numerous studies analyzing its causes, spatial dynamics, and implications for urban resilience (O’Malley et al., 2015). Originating from a combination of factors, including the concentration of artificial surfaces, reduced vegetation, and anthropogenic heat sources, UHIs result in significantly higher temperatures in densely built urban cores compared to surrounding rural areas (Shahmohamadi et al., 2011). This effect, first measured in the mid-20th century, has become increasingly pronounced due to the intensification of climate change and urban expansion. In recent years, UHI studies have extended beyond purely meteorological analyses, increasingly intersecting with concerns related to public health, energy sustainability, and environmental justice (Das et al., 2024).

However, a more specific and less explored dimension of UHI research pertains to its impact on cultural heritage assets. While much attention has been given to the structural and chemical degradation of heritage materials under conditions of humidity, pollution, and biological colonization, the role of heat stress has only recently started to gain recognition (He et al., 2023). Cultural landmarks, especially those located in open-air settings or lacking modern climate control systems, are particularly vulnerable to continuous thermal fluctuations and extreme temperature events. As pointed out in works such as (Camuffo, 2019), thermal anomalies can accelerate deterioration processes, especially in materials like marble, sandstone, and frescoes. However, the integration of UHI monitoring into conservation planning is still not widely spread, and predictive tools to support preventive measures remain scarce.

Remote sensing has played a pivotal role in advancing the study of urban heat patterns. Earth observation missions such as MODIS, Landsat, and Sentinel-3 have made it possible to systematically retrieve Land Surface Temperature (LST) data at various spatial and temporal resolutions (Onačillová et al., 2022). These datasets are widely used to characterize the spatial footprint of UHIs, quantify surface temperature anomalies, and track seasonal dynamics. Among recent contributions, the Copernicus Land Monitoring Service (CLMS) provides hourly LST data derived from geostationary platforms, offering an unprecedented temporal granularity suitable for short-term climate monitoring in European cities. This rich stream of satellite data has been exploited for urban heat vulnerability assessments, land cover correlation studies, and thermal mapping (Reis et al., 2022). Nonetheless, a considerable portion of the literature remains descriptive, focusing on retrospective analysis rather than forward-looking prediction.

Parallel to the growth of remote sensing applications, recent years have witnessed the rapid expansion of Deep Learning techniques for environmental monitoring. Neural networks, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated significant capabilities in extracting patterns from complex geospatial time-series data. More recently, Transformer-based architectures, initially developed for natural language processing tasks (Vaswani et al., 2017), have shown exceptional performance in learning long-range dependencies in sequential data, making them highly suitable for climate forecasting applications. Their self-attention mechanism enables the model to capture temporal and spatial correlations simultaneously, an essential feature when dealing with non-linear and multi-scale climate phenomena.

Several studies have begun to apply Transformers and hybrid models (e.g., CNN+Transformer) to temperature forecasting, air pollution estimation, and weather pattern classification (Tehrani et al., 2024). These approaches often combine multiple data sources, including satellite observations, reanalysis fields, and ground station measurements (Balsamo et al., 2018). Despite these promising advancements, the vast majority of such works focus on regional or global scales, and rarely address urban-scale, high-resolution forecasting tasks. Even less common are applications aimed at supporting the conservation of historic urban environments, where the granularity and specificity of forecasts can have a tangible impact on risk mitigation. An example is the use of a Continuous Wavelet Transform (CWT) for improving the predictive accuracy of sea surface temperature models based on Copernicus data (Conforti et al., 2024). These methods demonstrate the effectiveness of combining remote sensing with Deep Learning and open promising perspectives for applications in environmental forecasting, ranging from marine to urban contexts.

In this context, our work contributes to bridging a critical gap between predictive environmental modeling and cultural heritage protection. We propose a deep learning framework based on a custom autoregressive Transformer model trained on CLMS hourly LST data, applied to the city of Florence, a historic urban environment particularly exposed to UHI effects. The model forecasts hourly surface temperatures over a three-day horizon, leveraging the temporal richness of satellite observations and the predictive capacity of the Transformer architecture. Unlike traditional time-series models, our approach generates the future temperature sequence step-by-step, conditioning each prediction on both the encoded historical features and the sequence of previously generated outputs. This autoregressive structure is particularly well-suited to capturing the compounding effects of thermal inertia and urban microclimate dynamics.

To our knowledge, this is one of the first studies to apply such a methodology specifically in the context of heritage risk forecasting. By focusing on short-term, high-resolution temperature prediction using purely satellite-derived data, our approach complements existing monitoring strategies and opens new avenues for anticipatory conservation. The results demonstrate that deep learning models, when carefully adapted and trained, can play a vital role in supporting the adaptive management of historic cities under climate pressure.

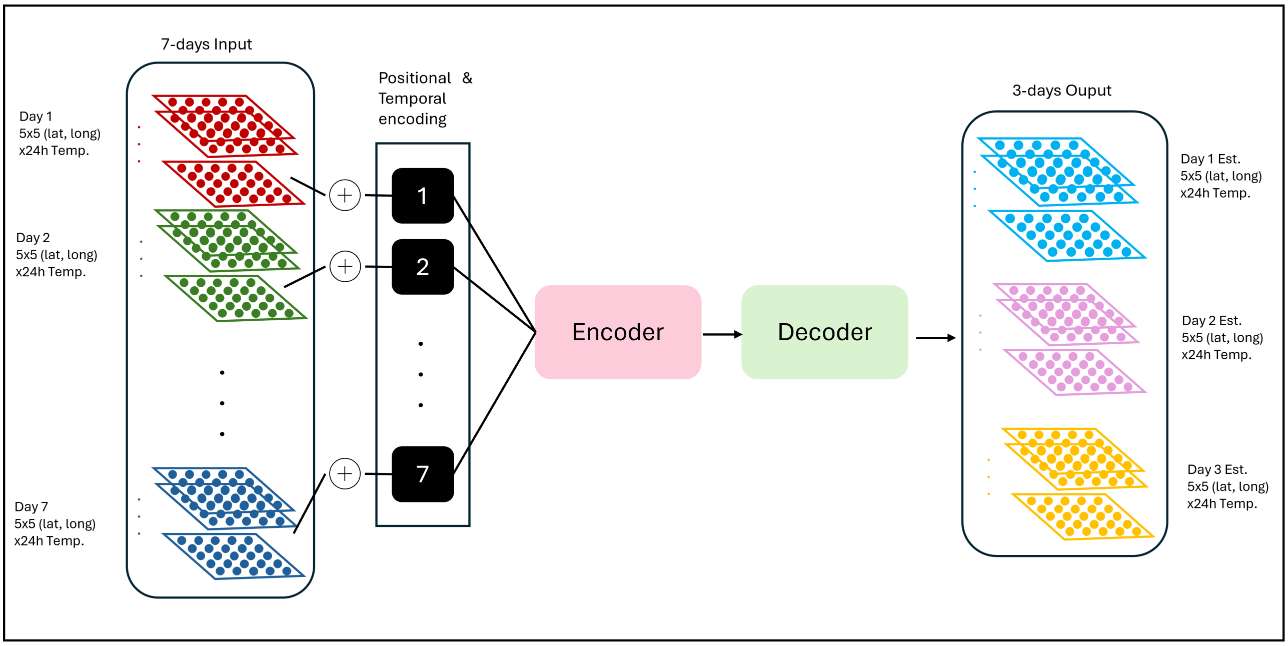

Figure 1. Overview of the model pipeline. The 7-day input grid (5x5 spatial) is embedded with positional and temporal encodings before entering the encoder. The decoder then generates 3 days of hourly temperature forecasts step-by-step.

3. Methodology

3.1 Dataset

The dataset employed in this study is the *Hourly Land Surface Temperature, Global (Version 2.0)*, provided by the Copernicus Land Monitoring Service(1). This product provides hourly Land Surface Temperature observations representing the skin temperature of the Earth’s surface, making it particularly suitable for assessing surface heat accumulation in densely built-up environments.

Regarding the spatial reference, the data are organized on a regular latitude/longitude grid based on the WGS 1984 ellipsoid (terrestrial radius = 6378 km). The grid resolution is defined as $5°/112$, and the coordinates refer to the center of each pixel.

From a temporal perspective, the LST product consists of instantaneous fields estimated hourly (00, 01, 02, ..., 23 UTC) using geostationary satellite data representative of each time step. The dataset is distributed globally in multi-band `NetCDF4` format, with metadata compliant with the Climate and Forecast (CF) conventions. Additionally, an INSPIRE-compliant metadata file in XML format, a corresponding XSLT stylesheet for visualization, and a subsampled color quicklook in `GeoTIFF` format are provided separately. Physical values (PV) of LST are derived from the digital numbers (DN) using the linear relation: $\text{PV} = \text{Scaling} \times \text{DN} + \text{Offset}$.

Validation of the LST product follows the protocol recommended by the CEOS Cal/Val LPV subgroup for Land Surface Temperature(2). The dataset is compared with in-situ observations from a set of ground stations spanning diverse land cover types and climate zones. To ensure representativeness, each station’s surroundings are analyzed using high-resolution information on orography, land cover, and Landsat-derived LST (processed via Google Earth Engine), in order to identify, within a 5×5 pixel window, the pixel that best corresponds to the ground truth measurements.

We extracted a 25×25 km window centered on Florence, Italy, structured as a 5×5 point grid. The dataset includes hourly data from 2021 to 2024 (except made from the first 18 days of January 2021). To evaluate the model under peak thermal conditions, we selected as test set the entire summer months of **June 2021, July 2022, August 2023 and September 2024**, totaling **82 overlapping windows of 10 days** (20 for June and September, 21 each for July and August). Each window includes 168 hours (7 days) of input and 72 hours (3 days) of output.

Unfortunately, we found out that in many regions there was a high number of missing data, for approximately the 48% of the total: to account for this issue, we implemented interpolation techniques. After testing cubic and spline interpolation (both of which led to non-realistic outputs) we opted for a linear interpolation, trading some temperature realism for stability and smoother results.

The data, originally in Kelvin, were converted to **Celsius** for improved interpretability and consistency in visualization. During the training procedure, all values were normalized by substracting the mean and dividing by the standard deviation. To prevent possible overfitting caused by the limited spatial resolution and repetitive sequences, we applied **early stopping** during training. We found that the use of added gaussian noise did not change the model performances.

3.2 Method

We propose an autoregressive Transformer-based model for the short-term prediction of land surface temperatures in urban environments. The model is trained to generate an hourly temperature forecast over a 72-hour horizon, given the preceding 168 hours of data. The overall architecture follows an encoder-decoder design, adapted to handle spatiotemporal data extracted from satellite-based temperature grids.

(1) `https://land.copernicus.eu/en/products/temperature-and-reflectance/hourly-land-surface-temperature-global-v2-0-5km`
(2) `https://lpvs.gsfc.nasa.gov/LSTE/LSTE_home.html`

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-399-2025

Model Formulation Let $X = \{x_{t-167}, \ldots, x_t\}$ denote a sequence of 168 hourly LST frames, each represented as a 5×5 grid. The goal is to predict the sequence $\hat{Y} = \{\hat{x}_{t+1}, \ldots, \hat{x}_{t+72}\}$, where each $\hat{x}_{t+i}$ is also a 5×5 grid representing the LST field at hour $t + i$.

The model consists of:

- **Encoder:** A stack of Transformer layers that processes the full input sequence. Each frame is flattened and embedded with spatiotemporal positional encodings to preserve both time order and grid location. Multi-head self-attention captures dependencies across the entire input window.
- **Decoder:** An autoregressive decoder that generates the output sequence one frame at a time. At each timestep $t + i$, the decoder uses the encoder output and the previously predicted frames $\hat{x}_{t+1}, \ldots, \hat{x}_{t+i-1}$ to compute $\hat{x}_{t+i}$. Masked self-attention ensures causality in the generation process.

Training Strategy The model is trained using the Mean Absolute Error (MAE) loss:

$$\mathcal{L}_{\text{MAE}} = \frac{1}{72} \sum_{k=1}^{72} |\hat{x}_{t+k} - x_{t+k}|$$

We experimented also with Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), noticing a degradation of the performances. In fact, as shown in the Results section, despite an increase of overfitting the MAE loss was able to produce the lowest estimation error. We used *Adam* optimizer, with early stopping based on validation MAE. A small Dropout value, equal to 0.1, helped prevent overfitting.

The decoder constructs the output autoregressively: after the first prediction $\hat{x}_{t+1}$ is generated, it is used to compute $\hat{x}_{t+2}$, and so on, building the 72-hour sequence recursively. In this way, it is possible to exploit the model to produce longer or shorter predictions while keeping good performances.

4. Results

We evaluate the performance of the model on 82 test windows drawn from three distinct summer months across 2022–2024. These months were held out during training to assess the model's ability to generalize across years and climatic conditions.

Evaluation Metrics

We report both *Root Mean Square Error* (RMSE), *Mean Squared Error* (MSE), and *Mean Absolute Error* (MAE), computed as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2},$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2,$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i|$$

Results and Temporal Forecast Behavior

Using MSE, the model achieves an average error of **1.48° C^2**, which is largely influenced by the smoothing and information loss due to linear interpolation. The use of L1 loss, however, results in a significantly improved MAE of **0.79° C**, demonstrating the effectiveness of using absolute error minimization in this context. Finally, the model trained and tested with RMSE scored an error value equal to **1.18° C**.

As relates to the forecast accuracy, it remains consistent over the first 48 hours and slightly degrades in the final third of the prediction window, reflecting the expected uncertainty accumulation typical of autoregressive decoders.

The following figures show the evolution of loss values over the 72-hour prediction window, averaged across all 82 test sequences. Specifically, Figure 2 presents the RMSE values over time, while Figure 3 highlights the corresponding MSE trend; Figure 4 demonstrates the improved accuracy of the MAE used both as training loss and as metric for assessing performances.

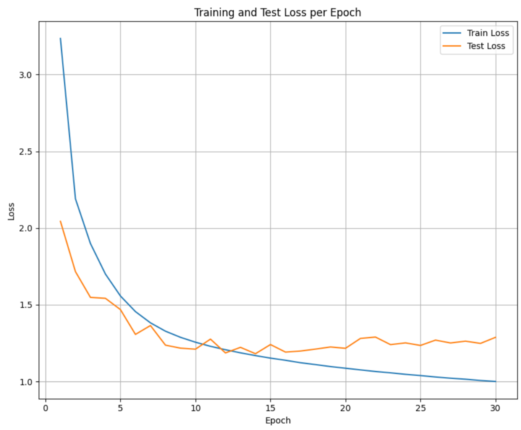

Figure 2. RMSE values over the 72-hour forecast horizon, averaged across all test windows.

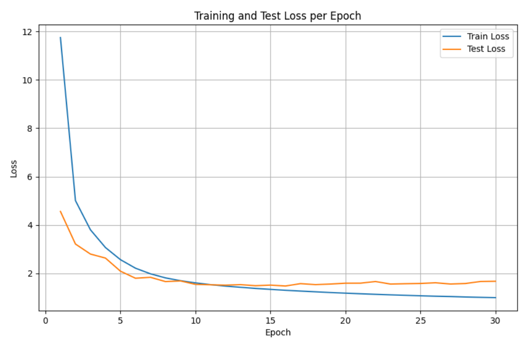

Figure 3. MSE values over the 72-hour forecast horizon, averaged across all test windows.

These plots illustrate that while RMSE and MSE highlight the model's sensitivity to peak errors, especially relevant in critical climate scenarios, MAE provides a more stable assessment of forecast reliability. Moreover, the MAE training-testing curves exhibit some overfitting: for this reason, we applied early stopping after 15-20 training epochs and thus obtained an accurate model capable of consistent previsions throughout the 3-day

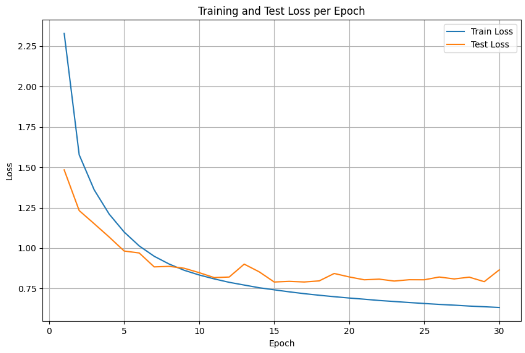

Figure 4. MAE values over the 72-hour forecast horizon, showing improved performance with L1 loss.

forecast window, which is especially valuable in urban climate risk contexts.

Case Study: Selected Grid Points in Florence

To further evaluate the quality of the model predictions, we analyzed the temperature forecasts for three specific spatial grid points selected within the city of Florence. These points were chosen along the same longitude (11°15'00"E) but at different latitudes according to the spatial structure of the CLMS dataset as follows:

- **Point A:** 43°45'00"N, 11°15'00"E
- **Point B:** 43°47'40"N, 11°15'00"E
- **Point C:** 43°50'21"N, 11°15'00"E

Figure 5 illustrates the selected locations, which can be assumed as representative of the thermal behavior across a latitudinal gradient spanning the urban and peri-urban zones of Florence. For each point, we plotted the 72-hour predicted temperature time series against the ground truth. Figures 6, 7 and 8 display the model's prediction vs. ground truth for each location.

The results confirm that the model captures the fundamental shape and timing of temperature variations across different latitudes within the grid. Despite the smoothing introduced by linear interpolation and the slight phase lag in some peaks, the forecast curves closely match the measured values.

An interesting observation emerges when comparing Point A (Fig. 6) and Point B (Fig. 7). Although Point A is located in a less densely built-up area than Point B, which lies in a residential urban zone, it consistently exhibits higher peak temperatures. This unexpected thermal behavior may be linked to the limitations of satellite-based LST data, particularly the relatively coarse resolution and the presence of missing values that required interpolation. While linear interpolation ensured continuity, it may not preserve local thermodynamic variations, especially in mixed land cover regions. These anomalies highlight the importance of integrating satellite data with additional information, such as **in-situ ground station measurements**, detailed land use classifications, and emissivity maps. Further analysis is needed to confirm whether the elevated values in Point A reflect microclimatic conditions, land surface characteristics or artifacts introduced during preprocessing. These plots support the overall conclusion that the model's prediction is accurate not only in terms of global metrics (RMSE, MSE, MAE), but also in localized point-based performance—a key requirement for effective application in microclimatic monitoring and heritage risk assessment.

5. Conclusions

In this work, we presented a deep learning framework for short-term forecasting of urban land surface temperature, with a specific focus on heritage-sensitive areas affected by the Urban Heat Island phenomenon. Leveraging hourly satellite observations from the Copernicus Land Monitoring Service, we trained a custom autoregressive Transformer model to predict temperature distributions over a 72-hour horizon at 5 km spatial resolution.

The results on the Florence case study demonstrate the model's ability to capture both temporal dynamics and spatial patterns of urban heat. Despite being trained on a limited set of three summer months across different years, the model achieved promising accuracy, with MAE values as low as 0.79°C. These findings highlight the feasibility of applying Transformer-based architectures to climate forecasting tasks using remote sensing data alone.

Beyond quantitative performance, the method provides a valuable tool for anticipatory temperature-related risk assessment in historic urban environments. By generating high-resolution forecasts of surface temperature, the framework supports informed decision-making for heritage conservation, such as planning mitigation measures, identifying thermal hotspots and deploying Nature-Based Solutions during heatwave episodes.

The analysis of specific locations within the Florence grid further revealed discrepancies that challenge the interpretations of LST patterns. For instance, a less urbanized grid cell (Point A) recorded consistently higher predicted temperatures than a more built-up counterpart (Point B), suggesting the need for more nuanced and spatially contextualized validation strategies. Such anomalies highlight the limitations of relying solely on satellite-derived LST data and emphasize the value of integrating auxiliary information from ground stations, land use datasets, and emissivity maps.

Looking ahead, enhancing the spatial resolution of forecasts by incorporating multi-source observations will be essential for making precise predictions at the microclimate level. This integration will enable targeted conservation strategies and improve the interpretability and applicability of UHI forecasts in heritage-rich urban areas. Future work will also explore temporal expansion using multi-year datasets and the application of attention-based interpretability to identify the driving factors behind the extreme thermal patterns.

In conclusion, our research aims to contribute to the growing intersection between artificial intelligence, remote sensing, and cultural heritage management, offering a scalable and transferable methodology to anticipate thermal stress in historic cities under increasing climate pressure.

Acknowledgment

This project is supported by the Spoke 7 within the Italian National Research Programme (NRP) - PE05 (CHANGES), CUP:

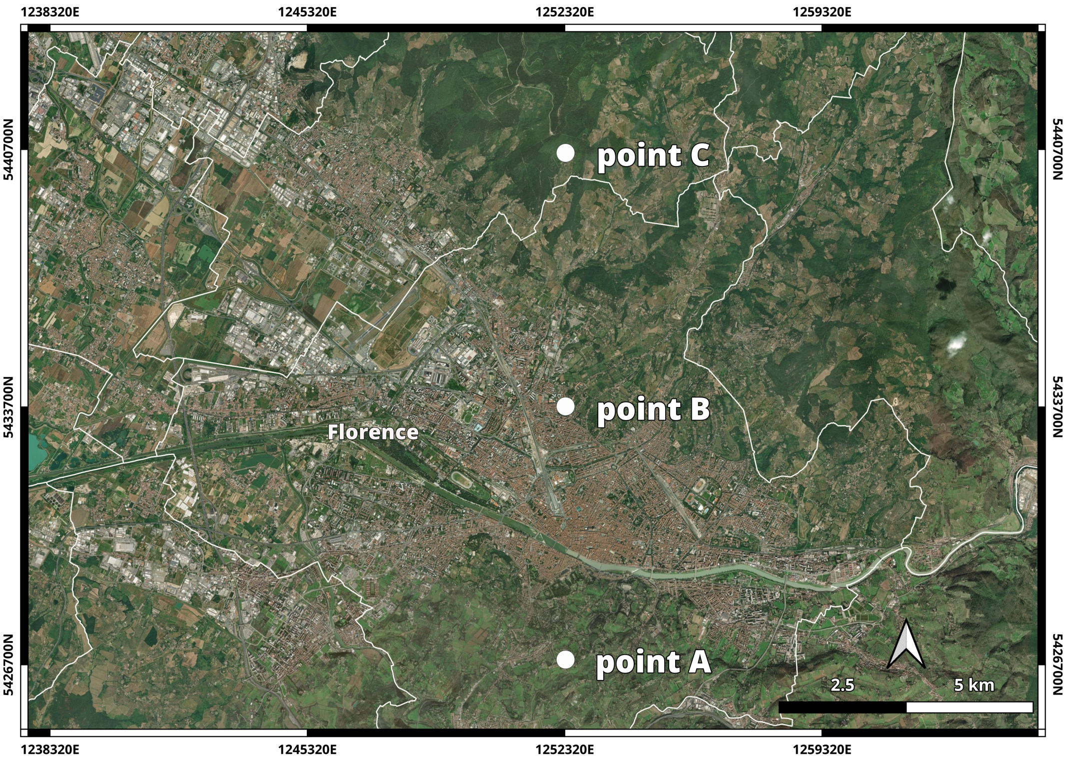

Figure 5. Satellite image of Florence and its surrounding area showing the three selected grid points used for model validation, labeled as Point A (43°45’00”N), Point B (43°47’40”N), and Point C (43°50’21”N), all aligned along longitude 11°15’00”E. The selected points exhibit different degrees of urbanization despite their geographic proximity, underscoring the importance of capturing microclimatic variability and motivating the integration of finer-resolution and ground-based data for future analyses.

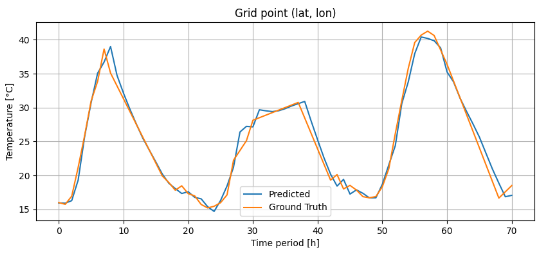

Figure 6. Predicted vs. actual temperature for Point A (43°45’00”N). The model successfully tracks both peak and trough dynamics.

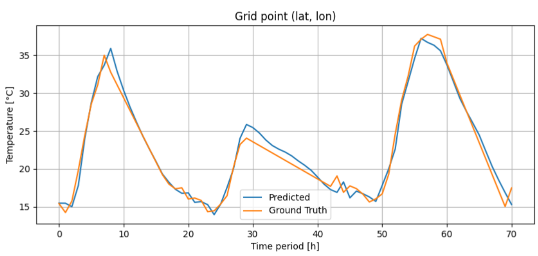

Figure 7. Predicted vs. actual temperature for Point B (43°47’40”N). Forecasts closely align with the diurnal cycle.

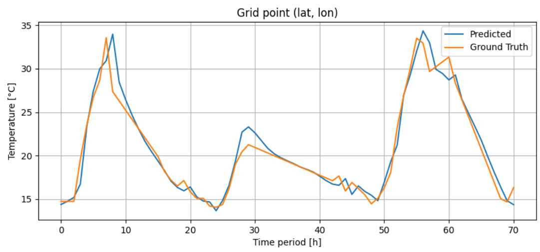

Figure 8. Predicted vs. actual temperature for Point C (43°50’21”N). Peak daytime temperatures are well captured.

B53C22004010006. The authors gratefully acknowledge the use of the Copernicus Land Monitoring Service (CLMS) dataset ’Hourly Land Surface Temperature – Global, v2 (0.05°)’, provided under the European Union’s Copernicus Programme.

References

Alerskans, E., Nyborg, J., Birk, M., Kaas, E., 2022. A transformer neural network for predicting near-surface temperature. *Meteorological Applications*, 29(5), e2098.

Balsamo, G., Agustì-Parareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Blyth, E., Bousserez, N., Boussetta, S., Brown, A. et al., 2018. Satellite and in situ observations for advancing global Earth surface modelling: A review. *Remote Sensing*, 10(12), 2038.

Bühler, M. M., Sebald, C., Rechid, D., Baier, E., Michalski, A., Rothstein, B., Nübel, K., Metzner, M., Schwieger, V., Harrs, J.-A. et al., 2021. Application of copernicus data for climate-relevant urban planning using the example of water, heat, and vegetation. *Remote sensing*, 13(18), 3634.

Camuffo, D., 2019. *Microclimate for cultural heritage: Measurement, risk assessment, conservation, restoration, and maintenance of indoor and outdoor monuments*. Elsevier.

Conforti, P. M., Russo, P., Di Ciaccio, F., 2024. Advancing sea surface temperature forecasting with deep learning techniques on copernicus data: An application in the gulf of trieste. *2024 IEEE International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters (MetroSea)*, IEEE, 83–88.

Copernicus Land Monitoring Service, 2025. Land surface temperature 2021-present (raster 5 km), global, hourly – version 2. Accessed: 2025-03-07.

Das, S., Choudhury, M. R., Chatterjee, B., Das, P., Bagri, S., Paul, D., Bera, M., Dutta, S., 2024. Unraveling the urban climate crisis: Exploring the nexus of urbanization, climate change, and their impacts on the environment and human well-being–A global perspective. *AIMS Public Health*, 11(3), 963.

Han, W., Zhang, X., Wang, Y., Wang, L., Huang, X., Li, J., Wang, S., Chen, W., Li, X., Feng, R. et al., 2023. A survey of machine learning and deep learning in remote sensing of geological environment: Challenges, advances, and opportunities. *ISPRS Journal of Photogrammetry and Remote Sensing*, 202, 87–113.

He, B.-J., Wang, W., Sharifi, A., Liu, X., 2023. Progress, knowledge gap and future directions of urban heat mitigation and adaptation research through a bibliometric review of history and evolution. *Energy and Buildings*, 287, 112976.

Heisler, G. M., Brazel, A. J., 2010. The urban physical environment: Temperature and urban heat islands. *Urban ecosystem ecology*, 55, 29–56.

Li, F., Yigitcanlar, T., Nepal, M., Thanh, K. N., Dur, F., 2024. A novel urban heat vulnerability analysis: Integrating machine learning and remote sensing for enhanced insights. *Remote Sensing*, 16(16), 3032.

Morabito, M., Crisci, A., Messeri, A., Messeri, G., Betti, G., Orlandini, S., Raschi, A., Maracchi, G., 2017. Increasing heatwave hazards in the southeastern European Union capitals. *Atmosphere*, 8(7), 115.

Onačillová, K., Gallay, M., Paluba, D., Péliová, A., Tokarčík, O., Laubertová, D., 2022. Combining Landsat 8 and Sentinel-2 data in Google Earth Engine to derive higher resolution land surface temperature maps in urban environment. *Remote Sensing*, 14(16), 4076.

O'Malley, C., Piroozfar, P., Farr, E. R., Pomponi, F., 2015. Urban Heat Island (UHI) mitigating strategies: A case-based comparative analysis. *Sustainable cities and society*, 19, 222–235.

Pioppi, B., Pigliautile, I., Piselli, C., Pisello, A. L., 2020. Cultural heritage microclimate change: Human-centric approach to experimentally investigate intra-urban overheating and numerically assess foreseen future scenarios impact. *Science of the Total Environment*, 703, 134448.

Quesada-Ganuza, L., Garmendia, L., Alvarez, I., Roji, E., 2023. Vulnerability assessment and categorization against heat waves for the Bilbao historic area. *Sustainable Cities and Society*, 98, 104805.

Reis, C., Lopes, A., Nouri, A. S., 2022. Assessing urban heat island effects through local weather types in Lisbon's Metropolitan Area using big data from the Copernicus service. *Urban Climate*, 43, 101168.

Russo, P., Schaerf, M., 2023. Anomaly detection in railway bridges using imaging techniques. *Scientific Reports*, 13(1), 3916.

Shahmohamadi, P., Che-Ani, A., Maulud, K., Tawil, N. M., Abdullah, N., 2011. The impact of anthropogenic heat on formation of urban heat island and energy consumption balance. *Urban Studies Research*, 2011(1), 497524.

Shi, H., Xian, G., Auch, R., Gallo, K., Zhou, Q., 2021. Urban heat island and its regional impacts using remotely sensed thermal data—a review of recent developments and methodology. *Land*, 10(8), 867.

Tehrani, A. A., Veisi, O., Delavar, Y., Bahrami, S., Sobhaninia, S., Mehan, A. et al., 2024. Predicting urban Heat Island in European cities: A comparative study of GRU, DNN, and ANN models using urban morphological variables. *Urban Climate*, 56, 102061.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Yun, C., Bhojanapalli, S., Rawat, A. S., Reddi, S. J., Kumar, S., 2020. Are transformers universal approximators of sequence-to-sequence functions? *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*, OpenReview.net.

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., Zhang, W., 2021. Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI conference on artificial intelligence*, 35number 12, 11106–11115.