

Fine-grained Urban Heat Island Effect Forecasting: A Context-aware Thermodynamic Modeling Framework

Anonymous Author(s)

Abstract

Climate change and rapid urbanization have led to the Urban Heat Island (UHI) effect, resulting in higher temperatures in metropolitan areas and negatively impacting urban communities. Accurate UHI forecasting is crucial for identifying high-risk periods and locations, especially in cities with vulnerable populations. Current methods are limited by data granularity and inadequate modeling of regional thermodynamics, which affects both accuracy and spatio-temporal granularity. In this paper, we propose DeepUHI, a data-driven context-aware framework for modeling local thermodynamics based on the heat equation, alongside the *SeoulTemp* dataset, the first multi-modal dataset for UHI effect predictions at the street level. Our framework utilizes a heat decomposition method to represent urban thermodynamics through thermodynamic cycles and thermal flows, effectively integrating urban environmental data. Extensive experiments show that our framework improves accuracy in UHI effect prediction and warning tasks, outperforming leading models. We have integrated DeepUHI into our SeoUHI platform to provide hourly street-level UHI forecasting for Seoul. The code, platform, and dataset are accessible at <https://anonymous.4open.science/r/DeepUHI-2F67/>.

Keywords

Urban Heat Island Effect, Thermodynamic Modeling, Time Series Forecasting, Spatio-Temporal Data Mining, Urban Computing

1 Introduction

Rapid urbanization has caused a host of environmental and sustainability challenges in cities and beyond [17, 19, 67, 71, 72]. Within cities, the dynamic interaction between natural elements and human activities leads to localized heat accumulation, thereby giving rise to the sophisticated **Urban Heat Island (UHI)** effect [23, 31, 64]. Recognizing the critical importance of this issue, the United Nations has identified "*cities and human settlements climate resilient and sustainable*" as a key sustainable development goal [30]. This mandate, coupled with the challenges posed by global warming [2, 47, 48] and the increasing vulnerability of elderly and at-risk populations [14, 38, 69], underscores the urgent need to develop robust thermodynamic models for accurate UHI forecasting.

Forecasting the UHI effect, usually quantified through urban field temperatures, presents distinct challenges compared to conventional climate forecasting. Most existing studies have predominantly utilized macro-scale and qualitative approaches, either by contrasting temperatures between urban cores and their surrounding suburban areas [45, 53, 60] or by interpolating land temperatures using remote sensing observations [7, 34, 40, 68]. However, both the UHI effect and the associated field temperature demonstrate

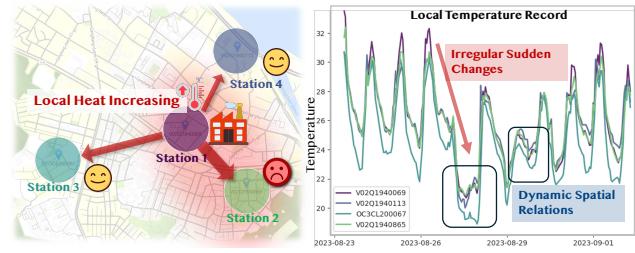


Figure 1: Introduction of Urban Heat Island effect. Left: Regional UHI effects are closely linked to the regional environment. Right: Urban thermodynamics is temporal irregular and has highly dynamic spatial relations.

remarkable variability within urban areas, exhibiting gradients that reflect differing levels of urban development [31, 64]. Neighborhood-specific factors, such as the presence of green spaces, water bodies, and land function, further contribute to this spatial heterogeneity. Consequently, the inherent complexity of these urban thermodynamics renders coarse-grained remote sensing data and traditional numerical inference methods [1, 62] insufficient for this application.

Recently, data-driven methods such as GenCast [42] and Panguweather [5, 59] have shown impressive accuracy in climate forecasting by modeling global atmospheric dynamics. However, when addressing regional climate phenomena at low altitudes, they tend to underperform compared to traditional numerical models [16, 37]. This discrepancy reveals the significant challenge of modeling urban thermodynamics for fine-grained UHI effect forecasting. As shown in Figure 1, firstly, UHI patterns are closely intertwined with various environmental factors, leading to complex spatial heterogeneity that further complicates modeling effects. Secondly, fine-grained regional temperature variations are more sensitive and irregular than those found in general atmospheric circulation [35, 41] or in other spatio-temporal phenomena such as regional traffic flow [22, 25, 63]. As a result, key challenges remain in effectively capturing spatio-temporal thermodynamics of local thermal systems and integrating essential environmental information without introducing noise.

In theory, the problem can be addressed using the heat equation by treating spatio-temporal thermodynamics as a flow physics problem. Accurate modeling of local thermodynamics becomes promising by incorporating urban environmental information as boundary conditions for the nonlinear partial differential heat equation [9, 21, 55], although an analytical solution for such an equation is not achievable. In this paper, guided by this equation, we introduce DeepUHI, a data-driven framework for modeling local thermodynamics with contextual information. To enable data-driven instantiation of the partial differential equation, local thermodynamics are decomposed into intra-region **thermodynamic cycles** [6, 43] and inter-region **thermal flows** [21, 70] (see Fig. 2).

Firstly, according to thermal equation, without disturbance, the daily thermal cycle of a region is stable and distinct, referred to as the intra-region **thermodynamic cycle**. This cycle contributes to the daily temperature variations and is determined by the region’s Specific Heat Capacity (SHC), which is influenced by the components of the surrounding environment. Inspired by these characteristics, we cluster stations into groups based on their similar environmental features and design a temporal periodicity learning block to explicitly instantiate the different temporal cycles. This approach efficiently models the regional thermodynamic cycle while avoiding complex over-fitting.

Secondly, inter-region **thermal flow** refers to the thermal transfer process of local heat variation between adjacent regions caused by temperature differences, contributing to irregular temporal variations in temperature. Although these mechanics can be naturally represented through spatial graph convolution [3, 29, 46, 52, 71], heat transfer in regional thermal equation is anisotropic in different directions and times due to various environment situations. This behavior goes beyond the capacity of a vanilla distance-based adjacency matrix. Therefore, we introduce a dynamic adaptive convolution block to instantiate the anisotropic thermal flow. Specifically, we introduce a context-aware adaptive graph convolution to learn the spatial anisotropy from environment features. For the temporal anisotropy, we design lightweight temporal dynamic learning mechanics to represent the temporal variation of heat transfer without introducing redundant model complexity.

Modeling the two thermal mechanics separately provides clear solutions. However, orderly recomposing them for UHI effect forecasting presents another challenge in terms of an imbalanced value range. The temperature temporal fluctuations caused by local thermal flow are minor compared to the daily thermal cycle, which may lead to these fluctuations being misinterpreted as noise by the cycle learning block, and vice versa. While explicitly modeling the daily cycle in our temporal learning block can help mitigate this issue, the training process remains unstable. To address this challenge, we introduce a fluctuation fine-tuning training strategy and a variation-aware loss function within our framework to effectively support our decomposition strategy.

Furthermore, we conduct an in-depth study to demonstrate the necessity and effectiveness of our framework in the context of fine-grain temperature predictions and warnings for extreme heat events. In addition, we deploy it on our SeoUHI platform to deliver real-world UHI effect forecasting and notifications to the citizens in the Seoul area. Our contributions are summarized as follows:

(1) We propose *DeepUHI*, the first data-driven context-aware framework for modeling local thermodynamics based on the heat equation. In our methodology, we introduce an explainable heat decomposition framework to represent thermodynamics from the perspectives of *thermodynamic cycles* and *thermal flows*, which systematically and concisely incorporates urban environmental data into the forecasting process.

(2) We conduct extensive experiments to demonstrate the necessity and effectiveness of our proposed framework. Comprehensive results indicate that it outperforms existing time-series and spatio-temporal forecasting baselines in UHI effect forecasting and warning tasks while maintaining high computational efficiency.

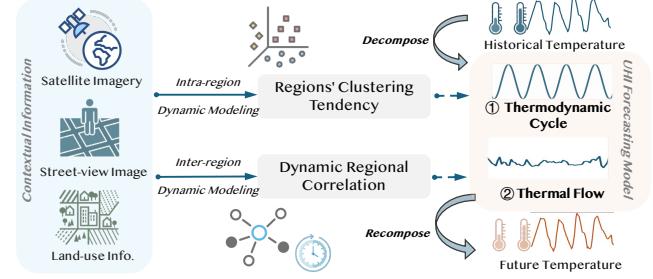


Figure 2: Insight of the framework.

(3) We collect and introduce *SeoulTemp*, the first fine-grained urban temperature dataset that includes field environment data across multiple modalities. This dataset encompasses a total of 947 temperature stations covering 605 km^2 of land in Seoul from 2021 to 2024, specifically targeting spatio-temporal UHI effect forecasting at the street level in urban areas.

(4) Our framework and model are deployed on the SeoUHI platform, providing precise short-term (day-ahead) UHI effect and extreme temperature notification services for over 10 million citizens in the Seoul area.

2 Preliminaries

DEFINITION 1. Urban Heat Island Effect. *Urban heat island effect refers to the quantification of the local thermal behavior of city within a complex field environment. A higher level of UHI effect indicates that people or urban infrastructures are likely to experience greater heat effects at that time. In practice, the UHI effect is quantified using field temperature series $\mathbf{X} \in \mathbb{R}^{N \times T}$, where $\mathbf{X}_{i,j}$ denotes the i_{th} attribute at time j .*

DEFINITION 2. Thermal Station Network. *Thermal station network consists of a group of local temperature monitoring stations within a city, denoted as $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, where \mathbf{V} represents the set of stations and $N = |\mathbf{V}|$ is the number of stations. The set \mathbf{E} denotes the edges that represent the relationships among the stations. Here, we use $\mathbf{A} \in \mathbb{R}^{N \times N}$ to denote the basic distance-based adjacency matrix of the thermal station network.*

DEFINITION 3. Contextual Feature. *Contextual features represent the environmental features surrounding each station. In practice, we use satellite imagery, street-view imagery, and land-use information to represent the surrounding environment. These diverse types of features can be flattened and concatenated to form a comprehensive representation. We denote the contextual features associated with each station as $\mathbf{C} \in \mathbb{R}^{N \times C}$, where C represents the feature dimension.*

PROBLEM STATEMENT. Urban Heat Island Effect Forecasting. *For graph \mathcal{G} , we compute pairwise distances between stations to derive the adjacency matrix based on a predefined distance threshold. Let $\mathbf{X} \in \mathbb{R}^{N \times T}$ denote the observed field temperatures from all stations, where T is the number of historical time steps. The historical observation at time t is defined as $\mathcal{H}^t = (\mathcal{G}, \mathbf{C}, \mathbf{X}^t)$. The forecasting problem is framed probabilistically as:*

$$p(\mathbf{Y}^{t+1:t+K} | \mathcal{H}^{t-P+1:t}) = \mathcal{F}_\theta(\mathcal{H}^{t-P+1}, \dots, \mathcal{H}^t) \quad (1)$$

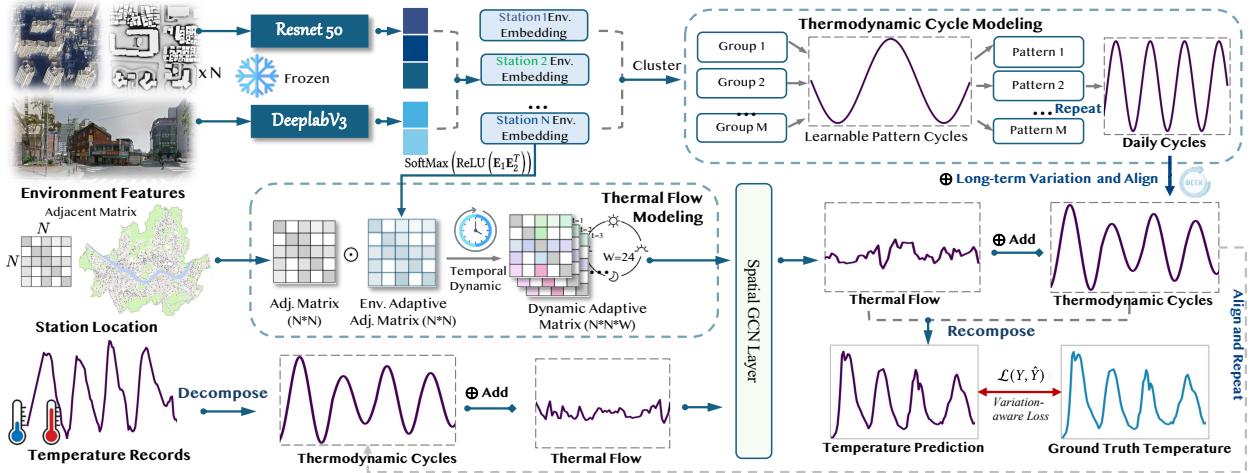


Figure 3: An illustration of our proposed DeepUHI framework.

where $p(\mathbf{y}^{t+1:t+K} | \mathcal{H}^{t-P+1:t})$ is the predictive probability distribution over future observations. To learn model parameters θ , we maximize the log-likelihood of the observed future sequences:

$$\theta^* = \arg \max_{\theta} \sum_{t=P}^{T-K} \log p(\mathbf{y}^{t+1:t+K} | \mathcal{H}^{t-P+1:t}) \quad (2)$$

3 Methodology

3.1 Overall Framework

Theoretical Statement. The study of the Urban Heat Island (UHI) effect involves understanding thermodynamics as a flow physics problem. In the context of flow physics, where diffusion occurs through a medium, it is common to establish a Cartesian coordinate system to analyze the classical heat function $u(x, y, z, t)$ with three spatial variables (x, y, z) and one time variable t . In this setting, u is considered a solution to the heat equation under isotropic conditions if it satisfies the following equation:

$$\frac{\partial u}{\partial t} = \alpha \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \quad (3)$$

Here, α is a positive coefficient known as the *thermal diffusivity* of the medium. However, in the context of fine-grain heat dynamics within a city, the assumption of isotropic conditions for air is inaccurate. The heat flow across different regions with varying environments is anisotropic, meaning that α is not a constant coefficient but rather a function that depends on the spatial coordinates (x, y, z) . Consequently, the equation assumes a nonlinear partial differential form that does not possess a precise analytical solution [44]:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(\alpha(x, y, z) \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(\alpha(x, y, z) \frac{\partial u}{\partial y} \right) + \frac{\partial}{\partial z} \left(\alpha(x, y, z) \frac{\partial u}{\partial z} \right) \quad (4)$$

where $\alpha(x, y, z)$ represents the value's variation of *thermal diffusivity* at different spatial locations.

Our thermodynamic modeling approach is founded on this equation. First, for a specified region, under steady-state conditions where we exclude influences from other regions, the temperature field will remain constant over time, which implies that $\frac{\partial u}{\partial t} = 0$. In

this context, the temperature in that region will converge to a stable daily cycle, referred to as the intra-region **thermodynamic cycle**. Second, by removing the inner thermodynamic cycle for each station, the remaining variations are attributed to spatial heat flow, as described by Equation 4 where $\frac{\partial u}{\partial t} \neq 0$. This heat flow characterizes the transfer of heat between urban regions, such as industrial heat flow from urban areas to suburban areas during peak times. We refer to this anisotropic flow as inter-region **thermal flow**.

Framework Design. The overall architecture of our proposed model is depicted in Figure 3. Initially, we extract contextual features from multi-modal data to create station environment embeddings to support the modeling of Equation 4. Based on the theoretical statement, we decompose the temperature records into *thermodynamic cycles* and *thermal flows*. To model these components in our framework, we introduce a temporal periodicity modeling block for the **thermodynamic cycles** ($\frac{\partial u}{\partial t} = 0$) and a dynamic adaptive convolution block for the **thermal flows** ($\frac{\partial u}{\partial t} \neq 0$). Additionally, we incorporate a spatial fine-tuning strategy with a variation-aware loss to facilitate the training of our framework.

3.2 Contextual Feature Representation

Contextual features are utilized to reveal the similarities or differences between regions for this research. The primary focus is to identify separability in geospatial space rather than intricate regional details. Therefore, we design a straightforward pipeline to extract contextual feature embeddings $C \in \mathbb{R}^{N \times C}$ with minimal computational cost.

We employ ResNet50 and DeepLabV3 to unsupervisedly extract visual features from satellite and street-view imagery, respectively. For land use semantic information, inspired by [71], we apply one-hot encoding to differentiate land types and multiply them by their area proportions. We normalize the extracted embeddings from different data sources to align their scales before concatenating them into C_{origin} . Subsequently, we use t-SNE to reduce the dimension of $C_{\text{origin}} \rightarrow C \in \mathbb{R}^{N \times C}$, with the aim of enhancing separability, where C is the compressed dimension of the contextual feature.

3.3 Thermodynamic Cycle Modeling

3.3.1 Context-aware Region Grouping. Conventional multi-channel models typically model each channel individually. This approach can introduce greater complexity and noise as the number of channels increases. This issue is particularly significant when modeling the thermodynamic cycle of temperature across various regions within a city, as numerous environmental studies have shown that temperature cycles in most regions of a city are synchronous [8, 12, 61]. Therefore, similar regions should share the same thermodynamic cycle modeling; otherwise, respective modeling may capture more noise from the transient inter-region thermal flow. In practice, this similarity is determined by their contextual environment features [12, 15].

To learn the thermodynamic cycle within a city, we first conduct context-aware region grouping to divide stations and their corresponding regions into M groups based on the context embedding C using K-Means clustering. The subsequent cycle modeling is then performed for each group rather than for each region. This reduces the computational complexity from $S \times N$ to $S \times M$, where S represents the complexity associated with modeling a single channel. More importantly, by explicitly setting the cycle modeling of similar regions to be the same, the framework becomes more robust to local fluctuations caused by noise or inter-region thermal flow.

3.3.2 Temporal Periodicity Modeling. Our goal is to identify the underlying stable thermal periodicity for each regional group. Although recent works such as Autoformer and PatchTST have emphasized the importance of data periodicity for long-term time series forecasting, their reliance on heavy Transformer backbones is computationally expensive and sensitive to noise. Additionally, these models lack interpretability, which is crucial for real-world deployment. Inspired by [26], we directly employ learnable daily cycles to explicitly model the temporal periodic pattern of the thermodynamic cycle.

Daily Patterns Modeling. Given M groups with a daily cycle length $W = 24$, we first generate learnable daily cycles $Q \in \mathbb{R}^{M \times D}$, all initialized to zeros. This can be mapped to $Q \in \mathbb{R}^{N \times D}$ according to the group mapping $M \Rightarrow N$. These daily cycles are globally shared among groups, which allows us to perform cyclic replications to obtain cyclic components C of the sequence X of the same length. The recurrent cycles Q undergo gradient backpropagation training alongside the backbone module for prediction, yielding learned explicit representations that reveal the internal cyclic patterns within the sequence.

When deriving subsequences C from the cyclic replications of sequence Q , we cannot directly ensure that they maintain the same periodic phase. To address this, we need to align Q by determining an alignment offset $t \bmod W$, where t is the current time point. After alignment, we extract the required subsequences from the aligned sequence $Q^{(t)}$. Mathematically, this process can be represented as:

$$Q^{(t)} = \text{Roll}(Q, t \bmod W), C_{t-L+1:t} = \overbrace{[Q^{(t)} \dots Q^{(t)}, Q^{(t)}_{0:L \bmod W}]}^{\lfloor L/W \rfloor \text{ times}} \quad (5)$$

Long-term Variation Modeling. In addition to daily patterns, regional temperature exhibits weekly variations due to human

activity and significant seasonal changes resulting from climate change. While these variations may not be significant for short-term predictions over several days, they become important when learning a stable daily cycle from data collected over several years. Otherwise, the regression of the daily cycle Q may become unstable as data transitions from summer to winter.

Seasonal variation manifests as a shift in the overall data distribution. To address this, we utilize Reversible Instance Normalization (ReVIN) [18] for each input data sequence to mitigate the effects of seasonal variation, defined by the following equation:

$$x = (x - \frac{1}{T} \sum_{t=1}^T x_t) / \sqrt{\frac{1}{T} \sum_{t=1}^T (x_t - \frac{1}{T} \sum_{t=1}^T x_t)^2 + \psi} \quad (6)$$

Here, ψ is the normalization constant that ensures stabilization. To address the issue of weekly variation, we introduce a weekly variation filtration layer that models and filters out weekly fluctuations, thereby facilitating a more accurate learning of the daily pattern. We compute the time embedding for the day of the week using $E_{\text{day}} = \text{Emb}(t_{\text{day}} \bmod 7, pos)$ and learn the weights and biases for the weekly daily cycle using the following equations:

$$\mathcal{W}_{\text{cycle}} = \text{ReLU}(\text{Linear}_w(E_{\text{day}})), \quad \mathcal{B}_{\text{cycle}} = \text{ReLU}(\text{Linear}_b(E_{\text{day}})) \quad (7)$$

Next, for each daily cycle, we update the cycle values in sequence Q^t to account for the weekly variation using the equation:

$$Q^t = Q^t \cdot \mathcal{W}_{\text{cycle}} + \mathcal{B}_{\text{cycle}}. \quad (8)$$

3.4 Thermal Flow Modeling

3.4.1 Residual Fluctuation Learning. Thermal flow refers to the process of heat transfer that occurs due to temperature differences between adjacent regions. This process is similar to local fluctuations in the stable climate system. Unlike the thermodynamic cycles observed within a single region, the global value variation resulting from thermal flow is typically minor, often limited to single-digit changes. As a result, traditional time series forecasting models for weather or climate prediction frequently overlook the individual modeling of thermal flow.

We introduce residual wave learning to address this issue as shown in Equation 9. Based on thermodynamic cycle modeling, we remove the cycle components Q^t from the observed field temperature $\mathbf{X} \in \mathbb{R}^{N \times P}$ to enhance the \mathcal{F} model's attention on the spatio-temporal wave caused by thermal flow. The output results of the backbone will be combined with the inferred cycle component Q^{t+P} to obtain the final prediction $\hat{\mathbf{X}} \in \mathbb{R}^{N \times K}$.

$$\hat{\mathbf{X}} = \mathcal{F}(\mathbf{X} - Q^t) + Q^{t+P} \quad (9)$$

3.4.2 Periodic Dynamic Graph Learning. To model thermal flow, it is necessary to introduce graph learning in \mathcal{F} . However, static and distance-based graph learning structures, such as conventional STGCNs, cannot effectively represent the anisotropy of thermal flow mechanics in both spatial and temporal dimensions. This anisotropy arises from the spatio-temporal variation of thermal diffusivity, as described in Equation 4. Although advanced works, such as GraphWaveNet [58], introduce adaptive adjacency matrices to automatically learn the spatial relationships between nodes, they are limited in representing the spatial variance of $\alpha(x, y, z)$ and do not account

for the time variation in $u(x, y, z, t)$ of the heat equation. Directly incorporating temporal dynamic graph learning to represent t can lead to a model that is overly complex for multiple time steps and difficult to train. We design a periodic dynamic graph learning framework to address the modeling of thermal flow.

Context-aware Adaptive Graph. We adapt the basic self-adaptive adjacency matrix learning method from [58] as the base matrix to represent spatial variance $\alpha(x, y, z)$ at static time t and enhance it with context information awareness. The adaptive adjacency matrix is computed as:

$$\tilde{A}_{adp} = \text{SoftMax} \left(\text{ReLU} \left(E_1 \cdot E_2^T \right) \right) \odot A \quad (10)$$

where E_1 and E_2 are node embeddings obtained through linear transformations of the contextual feature $C \in \mathbb{R}^{N \times C}$. It is important to note that we multiply the learned matrix by the original distance-based matrix A to maintain sparsity.

Temporal Periodic Dynamic Learning. To account for the temporal variation in thermal flow mechanics, we need to introduce variable t to adjust the learned spatial relations in \tilde{A}_{adp} at different time step t . However, the task becomes overly complex for increasing time steps. Therefore, we design a periodic dynamic learning method to manage complexity while accommodating an increasing number of time steps t .

In practice, regional thermal flow continuously varies over time while also exhibiting regularity similar to intra-region thermodynamic cycles. This regularity results from the consistent warming and cooling periods that occur daily, leading to periodic thermal flow variation. Consequently, we expand $\tilde{A}_{adp} \rightarrow \tilde{A}_{adp}^I \in \mathbb{R}^{N \times N \times W}$, where W is consistently set to 24, akin to thermodynamic cycle modeling, to represent the regular daily dynamics of thermal flow. For each $i \in W$, which represents the time of day, we individually construct \tilde{A}_{adp}^i for this time period and perform spatial convolution based on it when each time step t satisfies $t \bmod W = i$. Thus, the number of \tilde{A}_{adp}^i is limited to 24, offering a clear and efficient solution for the temporal dynamic representation.

Additionally, to extract fluctuation properties while maintaining efficiency, we employ a simple 2-layer MLP as the temporal backbone for the residual thermal flow series $\mathbf{X} - \mathbf{Q}^t$, denoted as \mathbf{X}_{res} below. The overall model's temporal periodic dynamic learning process can be formulated as:

$$\hat{\mathbf{X}}_{res}^t = \text{Conv}(\mathbf{X}_{res}^t, \tilde{A}_{adp}^{t \bmod W}), \quad \hat{\mathbf{Y}}_{res} = \text{MLP}(\hat{\mathbf{X}}_{res}) \quad (11)$$

3.5 Fine-tuning and Variation-aware Loss

3.5.1 Fluctuation Fine-tuning. Our residual learning framework enables the separate learning of thermodynamic cycles and thermal flow, which are decomposed and recomposed to predict the urban heat island (UHI) effect. However, the imbalance between these components can impede training. Local thermal flow fluctuations are minor compared to the daily thermal cycle, leading to potential misinterpretation as noise by the cycle learning block, and vice versa.

Additionally, the temporal cycle learning block may overly adapt to temporal variations, hindering the thermal flow learning block's ability to model spatial relationships. This issue is partly addressed through our explicit daily pattern modeling strategy, which restricts

the daily cycle modeling capacity of \mathbf{Q} . To enhance performance, it is crucial to address this during training. Therefore, we introduce fluctuation fine-tuning. We freeze the gradient updates of the temporal cycle learning block for 20 epochs during the fine-tuning stage to reduce the impact of temporal overfitting on thermal flow modeling. The residual framework remains unchanged, allowing the model to focus on capturing meaningful relationships within the static residual components.

3.5.2 Variation-aware Loss. Time series (TS) or spatial-temporal (ST) models typically minimize L_p norm distances as loss functions, such as Mean Absolute Error (MAE) or Mean Squared Error (MSE) for training. However, point-wise comparisons are insufficient for identifying continuous variations, such as trends and cycles, which can lead to uninformative predictions that do not account for temporal dynamics like peaks and troughs. For urban heat island (UHI) effect forecasting, we focus more on variation dynamics rather than single-value precision. Therefore, we need to design a variation-aware loss that considers both point-wise distances and temporal variations.

Inspired by [20], we construct our loss using Amplitude Shifting Invariance \mathcal{L}_{amp} , Phase Shifting Invariance \mathcal{L}_{phase} , and Point-wise Distance \mathcal{L}_{point} . Our variation-aware loss is formulated as follows:

$$\mathcal{L}(Y, \hat{Y}) = \alpha \mathcal{L}_{amp}(Y, \hat{Y}) + (1 - \alpha) \mathcal{L}_{phase}(Y, \hat{Y}) + \gamma \mathcal{L}_{point}(Y, \hat{Y}) \quad (12)$$

$$\mathcal{L}(Y, \hat{Y}) = \alpha \mathcal{L}_{amp} + \mathcal{L}_{phase} + \mathcal{L}_{point} \quad (13)$$

where $\alpha, \gamma \in [0, 1]$ are hyperparameters. We provide the introduction and equations for each loss component in Appendix A.1.

4 Experiments

In this section, we evaluate our proposed DeepUHI to address the following research questions:

- **RQ1:** Can DeepUHI outperform the latest time-series (TS) or spatio-temporal (ST) forecasting models on the regional temperature prediction task using real-world datasets?
- **RQ2:** How does the performance of DeepUHI in UHI warning tasks, from both spatial and temporal dimensions, compare to state-of-the-art TS or ST forecasting models?
- **RQ3:** Can DeepUHI clearly learn the thermodynamic cycle and spatial thermal flow relations using the heat decomposing framework to model the fine-grained thermal dynamics?
- **RQ4:** How effective are the designs of DeepUHI under various ablation settings?

4.1 Experimental Settings

4.1.1 Dataset. *SeoulTemp* dataset consists of three data modalities, covering 947 temperature stations across 605 km^2 of Seoul from 2021 to 2024. The satellite and street-view imagery provide detailed environmental context for each station. The Seoul Land-use Dataset, obtained from Korea EGIS, offers fine-grained land-use information for each station. The SDot Temperature Dataset, sourced from the Seoul SDot project, includes hourly average temperatures for the stations. With an average distance of approximately 500 meters between deployed stations, satellite imagery, street-view data, and land-use information are assigned to each station at a $500 \text{ m} \times 500 \text{ m}$ grid scale. The temperature dataset spans from January 1,

Table 1: Regional Temperature Prediction. The best results are in bold and second-best results are underlined.

Horizons	Metric	Traditional Methods			Time-series Methods			Spatio-temporal Methods			DeepUHI
		HA	LR	ARIMA	DLinear	PatchTST	GWN	MTGNN	STID	STAformer	
12	MAE	3.535	1.599	3.453	1.653	1.774	1.642	1.639	1.431	1.456	1.433
	sMAPE	0.485	0.408	0.486	0.256	0.309	0.257	0.261	<u>0.242</u>	0.244	0.241
24	MAE	3.672	2.031	3.563	2.015	2.111	2.014	1.987	<u>1.875</u>	1.944	1.811
	sMAPE	0.497	0.332	0.499	0.321	0.344	0.322	0.309	<u>0.305</u>	0.312	0.303
48	MAE	3.884	2.693	3.991	2.735	2.748	2.601	2.552	<u>2.495</u>	2.499	2.367
	sMAPE	0.514	0.403	0.524	0.412	0.434	0.405	0.392	<u>0.387</u>	0.382	0.375
96	MAE	4.210	3.431	4.325	3.428	3.670	3.419	3.335	3.331	<u>3.264</u>	3.231
	sMAPE	0.543	0.484	0.555	<u>0.477</u>	0.513	0.493	<u>0.477</u>	0.483	<u>0.478</u>	0.464

2021, to April 8, 2024, with values ranging from $-19.9 \sim +43.2$ °C. The maximum temperature difference between stations exceeds 10 °C during both winter and summer.

4.1.2 Implementation Details. Experiments are conducted on the SeoulTemp Dataset. The dataset is divided into training, validation, and test sets with a ratio of 8 : 1 : 1. We fix the look-back length of data to 96 steps to fairly compare the models. DeepUHI is implemented using PyTorch, with the Adam optimizer selected for optimization. The learning rate and weight decay are set to 0.01 and 0.0001, respectively. The batch size is fixed at 256, and the embedding dimension of the environmental features is set to 8. For the number of node groups, we follow the taxonomy studies [24, 39] of urban region landscapes to divide the nodes into 10 to 25 groups, and we standardize this to 10 for simplicity.

4.1.3 Baseline Methods. We include diverse baseline models from various research lines to ensure an extensive comparison. For traditional statistic methods, we include Historical Average (HA), Linear Regression (LR), and Auto-Regressive Integrated Moving Average (ARIMA). For TS methods, we include DLinear [66] and PatchTST [36] as the representative TS models. For ST methods that are devised for spatio-temporal forecasting tasks, we adopt Graph WaveNet [58] and MTGNN [57] as the state-of-the-art STGCNs model. We also include STID [51] and STAformer [27] as state-of-the-art spatio-temporal embedding-based ST methods.

4.1.4 Evaluation Protocols. For temperature prediction tasks, we adopt the commonly used Mean Absolute Error (MAE) and symmetric Mean Absolute Percentage Error (sMAPE) to evaluate the models. Lower values in MAE and sMAPE indicate higher prediction accuracy. For UHI warning tasks, we quantify the accuracy in identifying the hottest location and time in both spatial and temporal dimensions using Spatio-temporal Top-K Accuracy (STTop-K):

$$\mathcal{A} = \frac{1}{B} \sum_{b=1}^B \mathbb{I} \left(\bigcup_{i=1}^k (\mathcal{D}[Y_{b,i}, \hat{Y}_{b,0}] \leq \mathcal{D}\theta) \wedge (|\mathcal{T}[Y_{b,i}, \hat{Y}_{b,0}]| \leq \mathcal{T}\theta) \right) \quad (14)$$

Here, $Y_{b,0}$ represents the hottest point in batch b , and $\hat{Y}_{b,i}$ represents the top i hottest points in the prediction result. The terms \mathcal{D} and \mathcal{T} denote their spatial and temporal distances, while $\mathcal{D}\theta$ and $\mathcal{T}\theta$ are the corresponding thresholds. The function \mathbb{I} is the indicator function. We use STTop-K to evaluate the matching condition between the top K predicted values and the maximum true value. This method provides a more accurate reflection of the model's UHI warning quality and enhances robustness.

4.2 Results on Temperature Prediction (RQ1)

We first examine the performance of DeepUHI on the regional temperature value prediction task, which is comparable to conventional end-to-end TS or ST forecasting tasks. For a fair comparison, we adhere to the same deployment of baseline models as used in BasicTS [50]. The results summarized in Table 1 confirm that our model outperforms both the TS and ST baselines. Specifically, when compared to the previous state-of-the-art STID, the variants of our model demonstrate improvements of over 5% across most metrics. This indicates the effectiveness of our model in integrating environmental features and heat equation theory into the framework. We attribute the improved performance to the robust thermal decomposing framework, which offers a clearer optimization path for modeling thermal dynamics. Additionally, the variation-aware loss enables our model to effectively capture thermal variation information within the temperature time series, ultimately enhancing model training and downstream performance.

4.3 Results on UHI Warning (RQ2)

We further investigate the topic in RQ2, which examines whether our model can achieve more stable performance on future UHI warnings compared to other statistical models from both spatial and temporal perspectives. Unlike urban traffic prediction tasks, the UHI effect and regional temperature prediction encounter more unpredictable variations due to meteorological climate change. Therefore, focusing solely on the accuracy of temperature values is insufficient and unreliable for real-world deployment. Users are more interested in the future trends of the UHI effect, specifically regarding when and where conditions may become excessively hot and hazardous. To address this, we introduce the Spatio-temporal Top-K Accuracy

Table 2: UHI Warning Spatio-temporal Precision Evaluation.

Method	Complexity (multi-adds)	24 Hours		48 Hours	
		k=3	k=9	k=3	k=9
LR	1.69 M	0.395	0.513	0.294	0.331
DLinear	1.69 M	0.414	0.532	0.322	0.359
GWN	297,160 M	0.459	0.593	0.355	0.386
STID	7,560 M	0.421	0.556	0.331	0.358
SATEformer	236,330 M	0.472	0.604	0.376	0.423
Contextual Features Augmented Variants					
DLinear+Context	13.01M	0.423	0.545	0.329	0.365
SATEformer+Context	236,352M	<u>0.481</u>	0.644	0.389	<u>0.452</u>
DeepUHI	20.59 M	0.653	0.802	0.495	0.569

(STTop-K) metrics to quantify the model's performance in identifying the hottest points in both spatial and temporal dimensions. The results are presented in Table 2. Since existing conventional task-oriented models are usually incompatible to use environment context information of these tasks that may be highly related to the environment, we include context feature augmented with strong TS and ST models for fair and comprehensive evaluation. From the results, we can summarize two significant merits of DeepUHI:

- **Superior spatial temporal warning accuracy compared to existing models.** We can observe that the proposed DeepUHI can obtain superior results when comparing with state-of-the-art transformer-based ST models that are larger in size and trained for a longer time. Specifically, DeepUHI outperforms STAEfomer by a significant margin, which is 35.7% over 24-hour-prediction accuracy ($k=3$). This proves our thermal decomposing strategy and training framework effectively boost the model's understanding of regional thermal dynamics. Our model can quickly size the key variant component of regional temperature and provide more stable forecasting.
- **Light-weight but most efficient.** It is also noticeable that DeepUHI with only 20.95 million of multi-adds for computation achieves even stronger performances than baseline models with huge complexity. DeepUHI is reported to be significantly stronger on accuracy with TS models with huge computation cost like GWN. Although RLinear is light in size as DeepUHI and widely reported to achieve good performance in conventional time-series forecasting tasks, it shows unsatisfactory performance on this task. It is noticeable that our model also outperforms the contextual feature augmented variants. This proves the advantage of our context-aware framework to efficiently integrate environment information for this task.

4.4 Interpretability of Decomposing (RQ3)

To prove our theoretical decomposing assumption based on the heat equation and clarify the effectiveness of thermal decomposing mechanisms in enhancing model performance, we provide a comprehensive study of the thermal cycle modeling block and the dynamic heat flow learning block in this section.

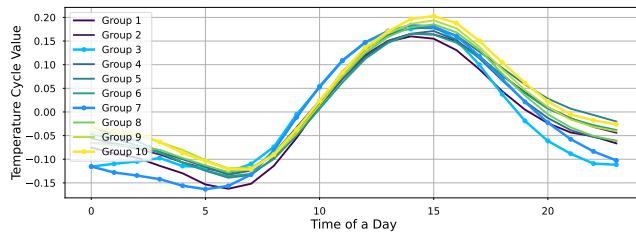


Figure 4: Daily thermal cycle for different station groups.

Thermal cycle modeling. We first visualize the learned daily thermal cycle for different station groups in our cycle learning block. In Figure 4, we observe that the learned daily temperature patterns show that stations warm up before 15:00 PM, cool down at night, and reach their lowest temperatures just before sunrise. This aligns well with the natural temperature patterns typical of Seoul, which has a temperate monsoon climate. Additionally, the differences observed between various station groups accurately

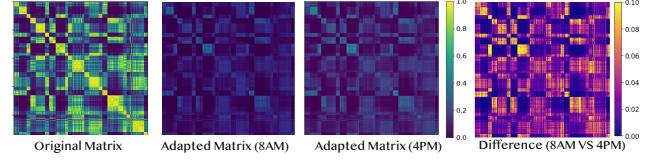


Figure 5: Adapted adjacent matrix for thermal flow modeling at different time steps.

reflect the thermal cycle variations in regions with different environments. For example, stations situated in dense urban areas (Group 10) consistently show higher regional temperatures, indicating an increased heat risk, while regions with water bodies (Group 7) and greenbelt (Group 3) display distinctive temperature cycles. These results and observations demonstrate that the thermal cycle learning block, along with our decomposition-based framework, effectively captures clear and precise thermal cycles from noisy historical temperature data.

Thermal flow modeling. In our decomposition-based framework, we utilize spatial graph convolution to model the mechanics of heat flow and introduce a dynamic adaptive adjacency matrix to capture its anisotropy at different time periods. We visualized both the original distance-based adjacency matrix and the learned adaptive matrix for thermal heat flow during the morning, noon, and evening. The differences between the original and learned matrices indicate the anisotropy of spatial heat flow learned by our model, as the vanilla distance-based adjacency matrix is isotropic. Additionally, it is evident that our learned matrix exhibits variations across different time periods, which corresponds well with the dynamic mechanics of thermal flow.

4.5 Ablation Study (RQ4)

To investigate the impact of each proposed component of our framework on its performance, we developed the following model variants for an ablation study:

- **(-) Cycle.** This variant removes the thermal cycle learning block and includes no daily cycle decomposing in the framework.
- **(-) Flow.** This variant removes the heat flow learning block and uses a 3-layer MLP for regression.
- **(-) Loss & Fine-tune.** This variant replaces the variant-aware loss with MSE loss and drops the fine-tuning training strategy.
- **(-) Context.** This variant drops the contextual features for region clustering and regional relations adapting.

We report the ablation results of DeepUHI and its variants on both temperature prediction and urban heat island (UHI) warning

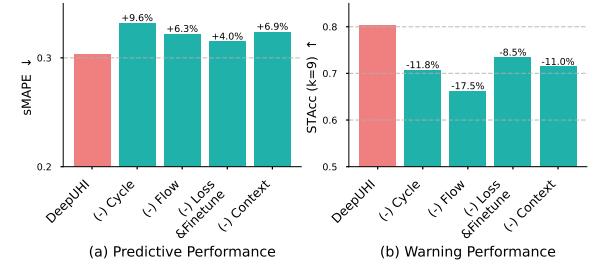


Figure 6: Ablation study on prediction and warning tasks.

performance in Figure 6. In summary, our findings are as follows: 1) For the predictive task, thermodynamic modeling plays the most significant role. This indicates that clear daily pattern information can greatly enhance the model’s performance in prediction. However, for the warning task, thermal flow modeling is crucial for maintaining high performance. Warning tasks require a deeper understanding of spatio-temporal dynamics, making the dynamic relationship modeling process essential for comprehending complex thermal flows. 2) The variation-aware loss and fine-tuning strategy can improve performance in both tasks to a similar extent. 3) Contextual features can enhance model performance in both tasks; however, if they are excluded, the self-adaptive version of DeepUHI still yields satisfactory results.

5 Real-world deployment

We deployed the proposed model into the SeoUHI system on the cloud to forecast the Urban Heat Island (UHI) effect. The system comprises four main modules. The data collector module continuously gathers real-time regional temperature data from the SDot project API. To address limited coverage and missing indoor temperature data, the data preprocessing module fills in gaps based on spatio-temporal neighbors. The prediction module forecasts regional temperatures for the next 24 hours at an hourly frequency, while the decision module identifies future hot times and locations, issuing UHI warnings. Additionally, the system automatically sends extreme heat warning emails to registered users during summer to protect vulnerable groups from heat exposure.

The system features a user-friendly web interface, as shown in Figure 7, displaying a dynamic heat map of Seoul. Users can select a target region to view predicted temperatures in a line graph format, allowing for comparisons between predicted and actual temperatures over different time spans. The predictive system executes hourly updates for the 947 temperature stations in Seoul. To maintain accuracy, the core model is retrained monthly with the latest regional temperature data, benefiting from its lightweight architecture. Our data-driven model does not rely on physical rule-based simulation, eliminating the need for detailed physical information or extensive computational resources, which facilitates practical deployment. However, it requires hourly average regional temperature data, necessitating the installation of temperature sensors within the city, currently limiting deployment to the Seoul area.

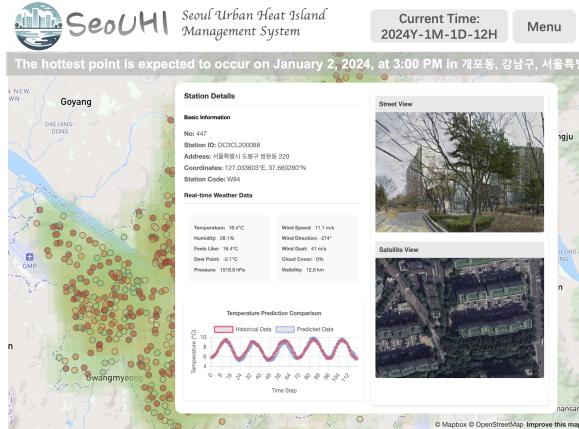


Figure 7: Web platform of SeoUHI system

6 Related Works

Climate Forecasting. There are two primary approaches for climate forecasting: the physical and statistical approaches. The physical approach, represented by Numerical Weather Prediction (NWP) methods [4], typically requires substantial computational resources to solve simulation equations with boundary conditions. In contrast, statistical approaches, such as GenCast [41], utilize historical observations to develop large-scale global weather forecasting models that demonstrate superior performance compared to NWP methods. However, both physical and statistical approaches encounter challenges in regional urban climate forecasting, such as UHI effect prediction. At fine-grained scales, physical simulations can become overly complex and time-consuming [33], while statistical methods are limited to general atmospheric circulation forecasting [49].

Multivariate Series Forecasting. Fine-grain UHI effect forecasting can be approached through multivariate time-series (TS) forecasting or spatio-temporal (ST) forecasting [50], the latter of which incorporates spatial information. Deep neural networks, particularly Transformers [54], have gained attention in TS forecasting, with models like iTransformer [28], Autoformer [56], and PatchTST [36] demonstrating high performance. Fully connected models, such as Dlinear [66], also serve as competitive alternatives. In spatiotemporal forecasting, STGCN utilizes graph convolution for spatial data processing. STGCN-like models form a broad category of methods capable of addressing both temporal and spatial dependencies [50, 65], often applied to traffic prediction tasks. Additionally, techniques like STID [51] offer a simple spatial-temporal identity-attaching approach to capture spatial dependencies without relying on graphs.

Data-driven Urban Thermal Study. Contextual environment data like satellite, and street-view imagery are commonly utilized as key data sources to support regional thermal condition studies. Existed context data-enhanced techniques often employ regression-based approaches to compute the relation between the local thermal effect and environmental situations. However, the results of these methods are limited to qualitative analysis. [10] analysis of the relation between local satellite image and UHI effect and compute the coefficient of determination factor R^2 of 0.68 between them in San Francisco. [11] emphasize the importance of terrain features in UHI spatial inference. [13] incorporate land use data to support the prediction of serval climate situations.

7 Conclusion

In this paper, we present DeepUHI, a context-aware thermodynamic modeling framework for fine-grained Urban Heat Island (UHI) effect forecasting. We introduce a data-driven heat decomposition method that captures urban thermodynamics from both thermodynamic cycle and thermal flow perspectives, systematically integrating urban environment data. Empirical results from regional temperature prediction and thermal warning tasks, compared with state-of-the-art baselines, demonstrate the effectiveness of DeepUHI. Additionally, we release the first fine-grained multimodal field temperature dataset, *SeoulTemp*, and deploy our model on the SeoUHI web platform to support further data-driven studies on micro-climates and solutions for urban sustainability.

References

- [1] I. Adilkhanova, J. Ngarambe, and G. Y. Yun. Recent advances in black box and white-box models for urban heat island prediction: Implications of fusing the two methods. *Renewable and Sustainable Energy Reviews*, 165:112520, 2022.
- [2] M. J. Alcoforado and H. Andrade. Global warming and the urban heat island. *Urban ecology: An international perspective on the interaction between humans and nature*, pages 249–262, 2008.
- [3] F. Alet, A. K. Jeewajee, M. B. Villalonga, A. Rodriguez, T. Lozano-Perez, and L. Kaelbling. Graph element networks: adaptive, structured computation and memory. In *International Conference on Machine Learning*, pages 212–222. PMLR, 2019.
- [4] P. Bauer, A. Thorpe, and G. Brunet. The quiet revolution of numerical weather prediction. *Nature*, 525(7567):47–55, 2015.
- [5] K. Bi, L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian. Pangu-weather: A 3d high-resolution model for fast and accurate global weather forecast. *arXiv preprint arXiv:2211.02556*, 2022.
- [6] H. Chen, D. Y. Goswami, and E. K. Stefanakos. A review of thermodynamic cycles and working fluids for the conversion of low-grade heat. *Renewable and sustainable energy reviews*, 14(9):3059–3067, 2010.
- [7] P. K. Diem, C. T. Nguyen, N. K. Diem, N. T. H. Diep, P. T. B. Thao, T. G. Hong, and T. N. Phan. Remote sensing for urban heat island research: Progress, current issues, and perspectives. *Remote Sensing Applications: Society and Environment*, 33:101081, 2024.
- [8] A. Dimoudi, A. Kantzioura, S. Zoras, C. Pallas, and P. Kosmopoulos. Investigation of urban microclimate parameters in an urban center. *Energy and Buildings*, 64:1–9, 2013.
- [9] J. L. Doob. A probability approach to the heat equation. *Transactions of the American Mathematical Society*, 80(1):216–280, 1955.
- [10] V. Equare, P. A. Mirzaei, and S. Riffat. Definition of a new morphological parameter to improve prediction of urban heat island. *Sustainable Cities and Society*, 56:102021, 2020.
- [11] V. Equare, P. A. Mirzaei, S. Riffat, and Y. Wang. Integration of topological aspect of city terrains to predict the spatial distribution of urban heat island using gis and ann. *Sustainable Cities and Society*, 69:102825, 2021.
- [12] F. Gómez, L. Gil, and J. Jabaloyes. Experimental investigation on the thermal comfort in the city: relationship with the green areas, interaction with the urban microclimate. *Building and environment*, 39(9):1077–1086, 2004.
- [13] J. Han, A. Chong, J. Lim, S. Ramasamy, N. H. Wong, and F. Biljecki. Microclimate spatio-temporal prediction using deep learning and land use data. *Building and Environment*, 253:111358, 2024.
- [14] C. Heaviside, H. Macintyre, and S. Vardoulakis. The urban heat island: implications for health in a changing environment. *Current environmental health reports*, 4:296–305, 2017.
- [15] G. M. Heisler and A. J. Brazel. The urban physical environment: Temperature and urban heat islands. *Urban ecosystem ecology*, 55:29–56, 2010.
- [16] L. Huang, Y. Wang, H. Bi, G. Zhu, L. Liu, and W. Jiang. Initial results of atmospheric weighted mean temperature estimation with pangu-weather in real-time gnss ppv retrieval for china. *GPS Solutions*, 29(1):48, 2025.
- [17] R. Keivani. A review of the main challenges to urban sustainability. *International Journal of Urban Sustainable Development*, 1(1-2):5–16, 2009.
- [18] T. Kim, J. Kim, Y. Tae, C. Park, J.-H. Choi, and J. Choo. Reversible instance normalization for accurate time-series forecasting against distribution shift. In *International Conference on Learning Representations*, 2021.
- [19] E. Krupat. *People in cities: The urban environment and its effects*. Number 6. Cambridge University Press, 1985.
- [20] H. Lee, C. Lee, H. Lim, and S. Ko. Tilde-q: a transformation invariant loss function for time-series forecasting. *arXiv preprint arXiv:2210.15050*, 2022.
- [21] R. W. Lewis, P. Nithiarasu, and K. N. Seetharamu. *Fundamentals of the finite element method for heat and fluid flow*. John Wiley & Sons, 2004.
- [22] F. Li, J. Feng, H. Yan, G. Jin, F. Yang, F. Sun, D. Jin, and Y. Li. Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution. *ACM Transactions on Knowledge Discovery from Data*, 17(1):1–21, 2023.
- [23] X. Li, Y. Zhou, S. Yu, G. Jia, H. Li, and W. Li. Urban heat island impacts on building energy consumption: A review of approaches and findings. *Energy*, 174:407–419, 2019.
- [24] Z. Li, B. Chen, S. Wu, M. Su, J. M. Chen, and B. Xu. Deep learning for urban land use category classification: A review and experimental assessment. *Remote Sensing of Environment*, 311:114290, 2024.
- [25] Y. Liang, Y. Xia, S. Ke, Y. Wang, Q. Wen, J. Zhang, Y. Zheng, and R. Zimmermann. Airformer: Predicting nationwide air quality in china with transformers. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 14329–14337, 2023.
- [26] S. Lin, W. Lin, H. Xinyi, W. Wu, R. Mo, and H. Zhong. Cyclenet: Enhancing time series forecasting through modeling periodic patterns. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.
- [27] H. Liu, Z. Dong, R. Jiang, J. Deng, J. Deng, Q. Chen, and X. Song. Spatio-temporal adaptive embedding makes vanilla transformer sota for traffic forecasting. In *Proceedings of the 32nd ACM international conference on information and knowledge management*, pages 4125–4129, 2023.
- [28] Y. Liu, T. Hu, H. Zhang, H. Wu, S. Wang, L. Ma, and M. Long. itransformer: Inverted transformers are effective for time series forecasting. *arXiv preprint arXiv:2310.06625*, 2023.
- [29] B. Lu, X. Gan, H. Jin, L. Fu, and H. Zhang. Spatiotemporal adaptive gated graph convolution network for urban traffic flow forecasting. In *Proceedings of the 29th ACM International conference on information & knowledge management*, pages 1025–1034, 2020.
- [30] Y. Lu, N. Nakicenovic, M. Visbeck, and A.-S. Stevance. Policy: Five priorities for the un sustainable development goals. *Nature*, 520(7548):432–433, 2015.
- [31] F. Lyu, S. Wang, S. Y. Han, C. Catlett, and S. Wang. An integrated cyberGIS and machine learning framework for fine-scale prediction of urban heat island using satellite remote sensing and urban sensor network data. *Urban Informatics*, 1(1):6, 2022.
- [32] H. Madsen. *Time series analysis*. Chapman and Hall/CRC, 2007.
- [33] C. F. Mass and Y.-H. Kuo. Regional real-time numerical weather prediction: Current status and future potential. *Bulletin of the American Meteorological Society*, 79(2):253–264, 1998.
- [34] P. A. Mirzaei and F. Haghhighat. Approaches to study urban heat island–abilities and limitations. *Building and environment*, 45(10):2192–2201, 2010.
- [35] T. Nguyen, R. Shah, H. Bansal, T. Arcomano, R. Maulik, V. Kotamarthi, I. Foster, S. Madireddy, and A. Grover. Scaling transformer neural networks for skillful and reliable medium-range weather forecasting. *arXiv preprint arXiv:2312.03876*, 2023.
- [36] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. *arXiv preprint arXiv:2211.14730*, 2022.
- [37] L. Olivetti and G. Messori. Do data-driven models beat numerical models in forecasting weather extremes? a comparison of ifs hres, pangu-weather, and graphcast. *Geoscientific Model Development*, 17(21):7915–7962, 2024.
- [38] C. Y. Park, J. H. Thorne, S. Hashimoto, D. K. Lee, and K. Takahashi. Differing spatial patterns of the urban heat exposure of elderly populations in two megacities identifies alternate adaptation strategies. *Science of the Total Environment*, 781:146455, 2021.
- [39] S. Pauleit and F. Duhme. Assessing the environmental performance of land cover types for urban planning. *Landscape and urban planning*, 52(1):1–20, 2000.
- [40] S. Peng, S. Piao, P. Ciais, P. Friedlingstein, C. Ottle, F.-M. Bréon, H. Nan, L. Zhou, and R. B. Myneni. Surface urban heat island across 419 global big cities. *Environmental science & technology*, 46(2):696–703, 2012.
- [41] I. Price, A. Sanchez-Gonzalez, F. Alet, T. R. Andersson, A. El-Kadi, D. Masters, T. Ewalds, J. Stott, S. Mohamed, P. Battaglia, et al. Gencast: Diffusion-based ensemble forecasting for medium-range weather. *arXiv preprint arXiv:2312.15796*, 2023.
- [42] I. Price, A. Sanchez-Gonzalez, F. Alet, T. R. Andersson, A. El-Kadi, D. Masters, T. Ewalds, J. Stott, S. Mohamed, P. Battaglia, et al. Probabilistic weather forecasting with machine learning. *Nature*, 637(8044):84–90, 2025.
- [43] S. Qian, J. Ling, Y. Hwang, R. Radermacher, and I. Takeuchi. Thermodynamics cycle analysis and numerical modeling of thermoelastic cooling systems. *International Journal of Refrigeration*, 56:65–80, 2015.
- [44] D. Reitzle, S. Geiger, and A. Kienle. Semi-analytical solution of the time-dependent heat equation for three-dimensional anisotropic multi-layered media. *International Journal of Heat and Mass Transfer*, 134:984–992, 2019.
- [45] A. M. Rizwan, L. Y. Dennis, and L. Chunho. A review on the generation, determination and mitigation of urban heat island. *Journal of environmental sciences*, 20(1):120–128, 2008.
- [46] A. Sanchez-Gonzalez, J. Godwin, T. Pfaff, R. Ying, J. Leskovec, and P. Battaglia. Learning to simulate complex physics with graph networks. In *International conference on machine learning*, pages 8459–8468. PMLR, 2020.
- [47] M. Santamouris. On the energy impact of urban heat island and global warming on buildings. *Energy and buildings*, 82:100–113, 2014.
- [48] M. Santamouris, C. Cartalis, A. Synnefa, and D. Kolokotsa. On the impact of urban heat island and global warming on the power demand and electricity consumption of buildingsâa review. *Energy and buildings*, 98:119–124, 2015.
- [49] M. G. Schultz, C. Betancourt, B. Gong, F. Kleinert, M. Langguth, L. H. Leufen, A. Mozaffari, and S. Stadtler. Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A*, 379(2194):20200097, 2021.
- [50] Z. Shao, F. Wang, Y. Xu, W. Wei, C. Yu, Z. Zhang, D. Yao, T. Sun, G. Jin, X. Cao, et al. Exploring progress in multivariate time series forecasting: Comprehensive benchmarking and heterogeneity analysis. *IEEE Transactions on Knowledge and Data Engineering*, 2024.
- [51] Z. Shao, Z. Zhang, F. Wang, W. Wei, and Y. Xu. Spatial-temporal identity: A simple yet effective baseline for multivariate time series forecasting. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 4454–4458, 2022.
- [52] C. Song, Y. Lin, S. Guo, and H. Wan. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages

- 914–921, 2020.
- [53] A. A. Tehrani, O. Veisi, Y. Delavar, S. Bahrami, S. Sobhaninia, A. Mehan, et al. Predicting urban heat island in european cities: A comparative study of gru, dnn, and ann models using urban morphological variables. *Urban Climate*, 56:102061, 2024.
 - [54] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun. Transformers in time series: A survey. *arXiv preprint arXiv:2202.07125*, 2022.
 - [55] D. V. Widder. *The heat equation*, volume 67. Academic Press, 1976.
 - [56] H. Wu, J. Xu, J. Wang, and M. Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in neural information processing systems*, 34:22419–22430, 2021.
 - [57] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang. Connecting the dots: Multivariate time series forecasting with graph neural networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 753–763, 2020.
 - [58] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang. Graph wavenet for deep spatial-temporal graph modeling. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 1907–1913, 2019.
 - [59] H. Xu, Y. Zhao, D. Zhao, Y. Duan, and X. Xu. Improvement of disastrous extreme precipitation forecasting in north china by pangu-weather ai-driven regional wrf model. *Environmental Research Letters*, 19(5):054051, 2024.
 - [60] L. Yang, F. Qian, D.-X. Song, and K.-J. Zheng. Research on urban heat-island effect. *Procedia engineering*, 169:11–18, 2016.
 - [61] S. Yang, L. L. Wang, T. Stathopoulos, and A. M. Marey. Urban microclimate and its impact on built environment—a review. *Building and Environment*, 238:110334, 2023.
 - [62] Y. Yang, S. Guangrong, Z. Chen, S. Hao, Z. Zhouyiling, and Y. Shan. Quantitative analysis and prediction of urban heat island intensity on urban-rural gradient: A case study of shanghai. *Science of the Total Environment*, 829:154264, 2022.
 - [63] X. Yin, G. Wu, J. Wei, Y. Shen, H. Qi, and B. Yin. Deep learning on traffic prediction: Methods, analysis, and future directions. *IEEE Transactions on Intelligent Transportation Systems*, 23(6):4927–4943, 2021.
 - [64] S. Yoo. Investigating important urban characteristics in the formation of urban heat islands: A machine learning approach. *Journal of Big Data*, 5(1):2, 2018.
 - [65] B. Yu, H. Yin, and Z. Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*, 2017.
 - [66] A. Zeng, M. Chen, L. Zhang, and Q. Xu. Are transformers effective for time series forecasting? In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pages 11121–11128, 2023.
 - [67] Y. Zheng, L. Capra, O. Wolfson, and H. Yang. Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(3):1–55, 2014.
 - [68] D. Zhou, J. Xiao, S. Bonafoni, C. Berger, K. Deilami, Y. Zhou, S. Frolking, R. Yao, Z. Qiao, and J. A. Sobrino. Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives. *Remote Sensing*, 11(1):48, 2018.
 - [69] W. Zhu and C. Yuan. Urban heat health risk assessment in singapore to support resilient urban design by integrating urban heat and the distribution of the elderly population. *Cities*, 132:104103, 2023.
 - [70] O. Zienkiewicz, E. Oñate, and J. Heinrich. A general formulation for coupled thermal flow of metals using finite elements. *International Journal for Numerical Methods in Engineering*, 17(10):1497–1514, 1981.
 - [71] X. Zou, J. Huang, X. Hao, Y. Yang, H. Wen, Y. Yan, C. Huang, and Y. Liang. Learning geospatial region embedding with heterogeneous graph. *arXiv preprint arXiv:2405.14135*, 2024.
 - [72] X. Zou, Y. Yan, X. Hao, Y. Hu, H. Wen, E. Liu, J. Zhang, Y. Li, T. Li, Y. Zheng, et al. Deep learning for cross-domain data fusion in urban computing: Taxonomy, advances, and outlook. *Information Fusion*, 113:102606, 2025.

A Appendix

A.1 Components of Variation-aware Loss.

1. **Amplitude Shifting Invariance:** Implemented using the softmax function to ensure that the gap between the prediction and the ground truth is equal at each step. The formula is as follows:

$$\mathcal{L}_{\text{a.shift}}(Y, \hat{Y}) = \frac{1}{T'} \sum_{i=1}^{T'} |1 - \text{Softmax}(d(y_i, \hat{y}_i))|$$

where $d(y_i, \hat{y}_i)$ is the signed distance between the true value and the predicted value, and T' is the length of the sequence.

2. **Phase Shifting Invariance:** Achieved by utilizing the gap between the Fourier coefficients to retain the dominant frequency components of the original time series while reducing noise in non-dominant frequencies. The formula is as follows:

$$\mathcal{L}_{\text{phase}}(Y, \hat{Y}) = \begin{cases} \|F(Y) - F(\hat{Y})\|_p & \text{for dominant freq.} \\ \|F(\hat{Y})\|_p & \text{otherwise} \end{cases}$$

where $F(Y)$ and $F(\hat{Y})$ are the Fourier transforms of the true and predicted values, respectively, and $\|\cdot\|_p$ is the L_p norm.

3. **Point-wise Distance:** Implemented through the mean average error to measure the average magnitude of the errors in a set of predictions, without considering their direction. The formula is as follows:

$$\mathcal{L}_{\text{distance}}(Y, \hat{Y}) = \frac{1}{T'} \sum_{i=1}^{T'} |y_i - \hat{y}_i|$$

where y_i and \hat{y}_i are the true and predicted values at time step i , respectively, and T' is the length of the sequence.

A.2 Details of Baselines

We describe the baselines as follows:

- **Historical Average (HA):** A simple forecasting method that uses the average value of the historical data as the predicted value for future time points.
- **Linear Regression (LR):** A basic model that establishes a linear relationship between the input variables and the output variable. It uses historical data to estimate the coefficients of the linear equation and predicts future values based on this equation.
- **Auto-Regressive Integrated Moving Average (ARIMA):** A popular time series forecasting model that combines autoregression, differencing, and moving average components to capture the temporal dependencies and trends in the data.
- **DLinear [66]:** A deep learning-based time series model that employs fully connected neural networks to capture the temporal patterns and dependencies in the data. It has shown strong performance in various time series forecasting tasks.
- **PatchTST [36]:** A time series model that utilizes the Transformer architecture and patching technique to effectively capture the long-range dependencies and local patterns in time series data. It divides the time series into patches and uses self-attention mechanisms to model the relationships between these patches.
- **Graph WaveNet [58]:** A spatio-temporal graph convolutional network that uses dilated convolutions and adaptive graph convolutional networks to capture the spatial and temporal dependencies in spatio-temporal data. It is designed for spatio-temporal

forecasting tasks and has achieved robust performance in many datasets.

- **MTGNN [57]:** A deep learning framework consists of three main components: a graph learning layer to automatically extract the relationships between variables, a graph convolution module to capture spatial dependencies, and a temporal convolution module to capture temporal patterns. The model uses a mix-hop propagation layer and dilated inception layers to effectively model the complex dependencies in the data.
- **STID [51]:** A spatio-temporal embedding-based model that captures the spatio-temporal dependencies by learning spatio-temporal embeddings and using them as input to a neural network for prediction. It is designed to handle the complex interactions and dependencies in spatio-temporal data.
- **STAEformer [27]:** A spatio-temporal embedding-based model that utilizes self-attention mechanisms and embeddings to capture the long-range spatio-temporal dependencies and enhance the model’s predictive performance. It combines the strengths of Transformer and embedding techniques to effectively model spatio-temporal data.

A.3 Spatio-temporal Periodicity.

In this work, we develop a thermodynamic model of the urban heat island effect primarily using a historical spatio-temporal field temperature dataset. To manage the complex variations in thermal issues, we integrate thermal equations to guide our model and simplify the modeling process through an equation-based decomposition framework. We introduce the design of each component from a thermal modeling perspective and demonstrate their effectiveness in the paper.

However, from a data-driven spatio-temporal forecasting perspective, some parameter designs based on thermal principles may still require further experimental validation. Specifically, our model is designed to leverage the spatio-temporal periodicity of field temperature data. Below, we will present additional experimental evidence regarding the spatio-temporal periodicity of temperature data.

Spatio-relation Periodicity. In Section 3.4.2, we efficiently model the temporal periodic dynamics of **thermal flow** by setting a period $W = 24$ (daily). This approach simplifies the modeling of the temporal dynamics of inter-region spatial correlation, allowing us to focus on daily variations rather than the entire time period, given the daily periodicity of thermal flow dynamics. To validate this periodicity within the data, we present the daily variation of the learned adjacency matrix in DeepUHI over the course of a day in Figure 8. The visualization results further demonstrate the necessity of modeling the temporal dynamics of spatial relations in thermal flow. Notably, the daily variation exhibits a self-regressive trend; the difference between matrix at time period 0 and other periods vary during warning and cooling time within a day but return to the initial state by the end of the day. This observed spatio-relation periodicity in the data-driven training reinforces the rationale behind our design.

Intra-region Temporal Periodicity. In Section 3.3.2, we model the intra-region thermodynamics by proposing learnable daily cycles with a cycle length of $W=24$ for Q . This approach allows us to concisely model the basic thermodynamic variations. We utilize the

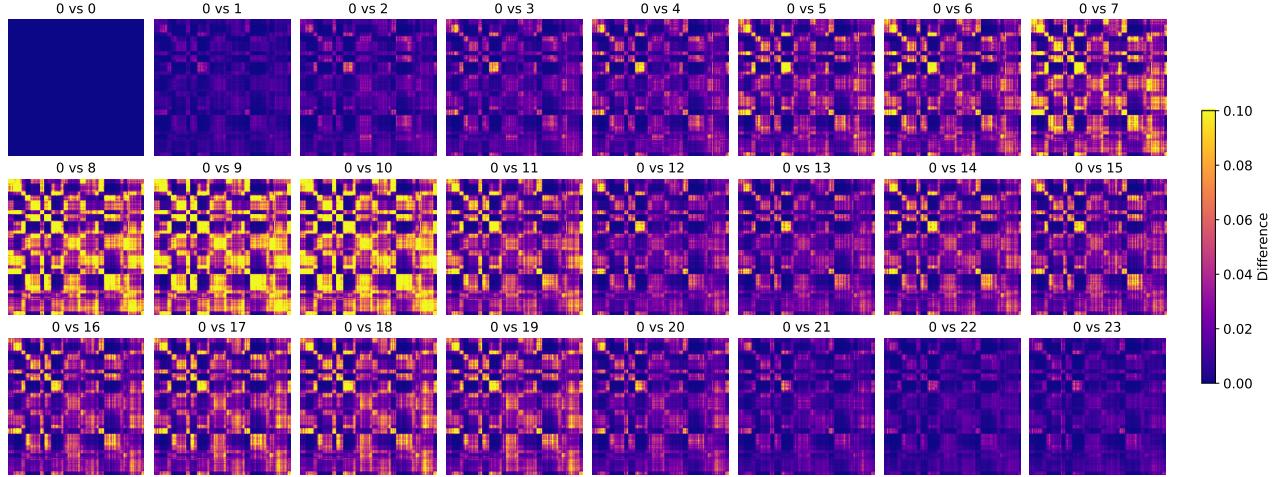


Figure 8: Spatio-relation Periodicity: daily variation of the learned adjacency matrix

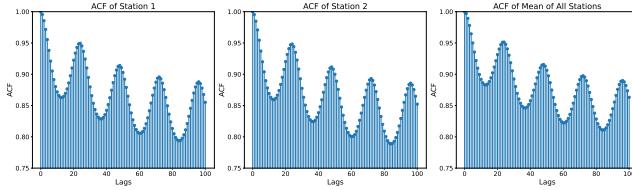


Figure 9: Intra-region Temporal Periodicity: visualization of ACF results on SeoulTemp Dataset

autocorrelation function (ACF) [32] to provide statistical support for our model design from a data perspective.

The autocorrelation function (ACF) is a powerful mathematical tool that helps detect periodicity within the data. It measures the correlation between a time series and its lagged values, indicating the presence of autocorrelation within the dataset. Mathematically, this can be expressed as:

$$ACF = \frac{\sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2}, \quad (15)$$

where N represents the total number of observations, x_t denotes the value of the time series at time t , k is the lag time, and \bar{x} is the mean of the time series values.

When the lag time k aligns with the cycle of the data, the ACF value exhibits a significant peak. Specifically, the largest peak corresponds to the lag that matches the length of the maximum cycle present in the dataset. Conversely, if the data lacks periodicity, no significant peaks or troughs will be observed.

We present the ACF results for the SeoulTemp dataset in Figure 9. It can be observed that intra-region temperature variations display evident periodicity, indicated by prominent peaks and troughs in the plots. More importantly, the maximum cycles shown in the plots align with our pre-inferred cycle length of $W=24$. This finding further supports the validity of our parameter design.