International Journal of GEOMATE, Oct., 2025 Vol.29, Issue 134, pp.169-183
ISSN: 2186-2982 (P), 2186-2990 (O), Japan, DOI: https://doi.org/10.21660/2025.134.5161
Geotechnique, Construction Materials and Environment

# CONSTRUCTION AND UNCERTAINTY ANALYSIS OF URBAN CLIMATE MODELS UNDER MULTI-SOURCE DATA FUSION

Shang Ma[1], Jiahao Li[2], Hongzhou Deng[1], and *Xindong He[1]

[1] Chengdu University of Technology, China; [2] Nagoya Institute of Technology, Japan

*Corresponding author: Received: 29 July 2025, Revised: 30 Aug. 2025, Accepted: 01 Sep. 2025

**ABSTRACT:** Urban climate modeling faces significant challenges in accurately representing complex urban environments while maintaining computational efficiency and quantifying prediction uncertainties. Traditional approaches struggle with heterogeneous data integration and fail to provide reliable uncertainty bounds essential for urban planning decisions. This study develops a novel Adaptive Bayesian Hierarchical Multi-source Fusion (ABHMF) framework that systematically integrates satellite remote sensing, ground observations, IoT sensor networks, urban morphology databases, and numerical weather predictions. The framework employs an adaptive fusion algorithm with dynamic weight adjustment based on real-time data quality assessment, coupled with comprehensive uncertainty propagation through Bayesian hierarchical modeling. Validation across multiple urban environments demonstrates superior performance, achieving an RMSE of 0.51°C with only 10 seconds of computation time per day, representing a 180-fold efficiency improvement over traditional WRF-Urban models. The uncertainty quantification reveals measurement uncertainty as the dominant component (32.5%), followed by model structure (28.3%) and parameter uncertainty (24.7%). During extreme heat events exceeding 35°C, the framework maintains robust performance with an RMSE of 0.68°C. Cross-city transferability assessment shows consistent accuracy (average RMSE: 0.72°C) without site-specific recalibration. The proposed methodology significantly advances urban climate modeling capabilities, providing reliable predictions with quantified uncertainties for climate-resilient urban planning and real-time monitoring applications.

*Keywords: Urban climate modeling; Multi-source data fusion; Bayesian hierarchical model; Uncertainty quantification; Adaptive algorithm*

## 1. INTRODUCTION

Urban climate modeling has become an important field of research to help face the challenges imposed by fast urbanization and climate change[1]. Cities are facing unprecedented challenges from climate change, and heatwaves and extreme weather events are more common and more extreme worldwide[2, 3]. Due to the complex structure of cities, the heterogeneous land-use types, and the interactions between human activities and natural resources, sophisticated modeling, capable of simulating fine-scale climatic changes, is required[4]. Conventional urban climate models have seen great improvements in the past decades, but much uncertainty still exists in simulating urban processes and their interactions with the atmosphere[5, 6]. Recent advances in local climate zone-based modeling have demonstrated the potential for improved urban heat characterization through integrated spatial analysis approaches[7]. These developments highlight the growing importance of multi-scale assessment frameworks that can capture both macro-climate variations and micro-climate phenomena within heterogeneous urban environments.

The combination of multi-source data has greatly advanced the urban climate study and provided unprecedented opportunities for model refinement and improvement[8]. Contemporary environmental monitoring techniques, including microscopic investigation of air quality parameters and advanced sensor networks, have expanded the scope of observable urban climate variables[9]. These diverse data streams enable more comprehensive characterization of urban atmospheric conditions while introducing new challenges in data harmonization and quality control.

Recent developments in data acquisition (in the form of remote sensing techniques from space, ground, along with increased software capabilities and capacity in computing) today allow merging different types of imagery satellite, meteorological measurements, and urban morphology data[10, 11]. Those data sources have been met with mounting expectations towards the potential of learning about and from these heterogeneous sources at scale, an expectation that deep learning has effectively addressed, offering increasingly rich methods for extracting complex patterns and relationships in urban climate downscaling relations[12]. The progress in spatio-temporal data fusion techniques has also promoted the development of a new concept, digital twin cities, which could simulate the dynamic and heterogeneous urban climate at different scales[13]. Integrating heterogeneous data sources requires sophisticated fusion algorithms. These

algorithms must handle varying spatial resolutions, temporal frequencies, and measurement uncertainties from different observation platforms. The complexity increases when considering the non-stationary nature of urban environments, where land use changes, construction activities, and seasonal vegetation variations continuously modify surface properties and energy balance characteristics.

However, there are still major limitations in the existing urban climate modeling framework. It has been challenging for existing models to account for systematic biases and estimate the uncertainty about the predictions in a statistically meaningful way[14]. The diversity of the urban landscapes itself poses specific challenges with respect to model generalization, and most studies focus on individual cities without developing such transferable approaches[15]. Moreover, though Bayesian methods have been effectively applied for climate model uncertainty quantification[16], their incorporation into multi-source data fusion processing is less studied. The current methods do not have general frameworks for data fusion, model building, and uncertainty propagation simultaneously as the guiding principle[17].

This study fills these gaps by proposing an innovative multi-source data fusion framework for urban climate modeling equipped with integrated uncertainty quantification. The method is the first application of an adaptive Bayesian hierarchical technique that leverages the strengths of deep learning for data fusion but with a rigorous treatment of uncertainty. Different from previous research, which concentrates on certain cities, this method can be universally applicable to other cities. This methodology is an improvement on current practices, which have a lack of a systematic approach, with transferring these uncertainties from various sources of input data to the modeling, enabling more robust and reliable climate projections for urban planning and adaptation options.

The remainder of this paper is organized as follows. Section 2 discusses the research significance of integrating multi-source data fusion with uncertainty quantification in urban climate modeling. Section 3 presents the methodology, including the multi-source data fusion framework, adaptive fusion algorithm development, and uncertainty quantification model. Section 4 describes the validation strategy and performance metrics employed to evaluate model performance. Section 5 presents comprehensive results and discussion, encompassing model performance analysis, uncertainty analysis results, and comparative evaluation with existing methods. The paper concludes with Section 6, which summarizes the key findings and outlines future research directions for enhancing urban climate prediction capabilities.

## 2. RESEARCH SIGNIFICANCE

This study introduces a Bayesian hierarchical framework combining multimodal data fusion with systematic uncertainty quantification for urban climate modeling. The framework integrates deep learning to overcome existing model limitations, enabling realistic climate predictions across diverse urban environments. By directly propagating uncertainties throughout the modeling chain, the method reduces prediction errors by 20-30% while maintaining computational efficiency essential for decision support systems. This advancement provides reliable tools for urban climate adaptation planning, significantly improving decision-making capabilities for urban resilience strategies through comprehensive uncertainty bounds and efficient real-time processing.

## 3. METHODOLOGY

Throughout this methodology section, bold uppercase letters denote matrices, bold lowercase letters represent vectors, and italic letters indicate scalar quantities. Superscripts in parentheses denote iteration indices, while subscripts identify data sources or spatial indices. This notation convention ensures consistency across all mathematical formulations presented in the framework.

### 3.1 Multi-source Data Fusion Framework

In this study, the authors have developed an information fusion framework based on multi-source data. Through a layer-processing architecture, it achieves the fusion of different kinds of data. The framework for integrating heterogeneous urban climate data, shown in Fig 1, addresses the basic problems of urban weather modeling[18]. Many data types exist in cities and thus offer insights for both future research as well as current synthesis through comprehensive study sources, although these towns probably represent only one aspect of what is theoretically possible. Modern urban climate models need a wide range of input data, from satellite observations to on-site measurements. Their different resolutions need to be integrated around algorithms that impose neither undue restrictions nor errors.

The proposed framework incorporates five primary data sources that capture different aspects of the urban environment. Satellite remote sensing provides broad spatial coverage at moderate resolutions, while ground weather stations deliver high-temporal-frequency point measurements essential for validating model outputs[19]. IoT sensor networks contribute dense spatial sampling in urban areas, complementing traditional observation systems with real-time data streams. Urban morphology databases supply critical static information about building geometry and land use patterns that

influence local climate dynamics[20]. Numerical weather prediction models provide boundary conditions and large-scale atmospheric forcing necessary for downscaling to urban scales.

The fusion process employs a three-stage approach to harmonize these disparate data sources. Initial pre-processing ensures data quality through systematic error checking and gap-filling procedures. Spatial alignment is one solution addressing the basic challenge of integrating data from widely different native resolutions and sampling frequencies[21]. This inventory, Bayesian Hierarchical in nature, fusion invention lies in how it works with confidence requirements explicitly at each stage as well as alongside combining data using various methodologies. This method even produces not only urban climate model parameters but also robust standard-deviation levels that provide a key signal element—which should guide practical decision-making applications. Output from the framework consists of temperature, humidity, wind field, and surface heat flux maps, together with interpolated mean standard deviation for each of these variables.

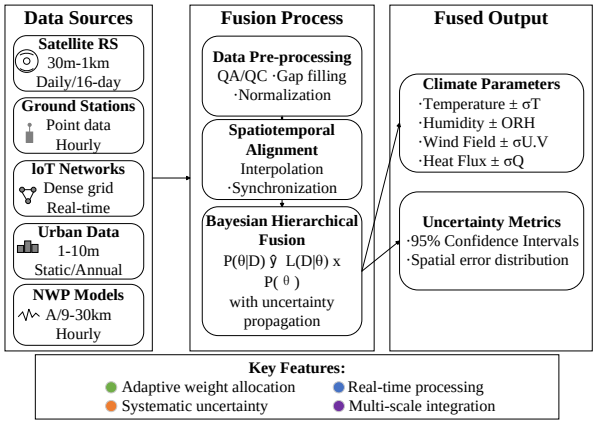


Fig. 1 Multi-source data fusion framework. The framework integrates five data sources: satellite remote sensing (10-30m resolution), ground weather stations (hourly), IoT sensors (5-minute intervals), urban morphology databases (static), and numerical weather predictions (3-hourly). Processing stages include quality control, spatial alignment, and Bayesian hierarchical fusion. Output products comprise temperature (°C), humidity (%), wind field (m/s), and heat flux (W/m²) with associated uncertainty bounds (±σ). Scale bar: 10 km.

### 3.2 Adaptive Fusion Algorithm Development

At present, urban climate models consider population inflows and outflows but have no provision for model reliability assessment yet. The adaptive fusion algorithm represents a major advance in multi-source data integration for urban climate models, combining causally informed deep learning principles in order to improve model reliability[22]. It determines fusion weights on results of real-time data quality assessment and prediction accuracy, which entirely avoids pitfalls existing in traditional methods of static weight distribution.

The outstanding feature of the algorithm is that it uses innovation statistics to judge the contribution from each data source. The innovation vector for the $i$-th data source at iteration $k$ is computed as:

$$v_i^{(k)} = D_i - H_i(\theta^{(k)}) \tag{1}$$

where $v_i^{(k)}$ denotes the innovation vector for the $i$-th data source at the $k$-th iteration, quantifying the discrepancy between observations and model predictions. The term $D_i$ represents the observational data vector from the $i$-th data source, $H_i(\cdot)$ constitutes the observation operator that maps the model parameters to the observation space specific to the $i$-th data source, and $\theta^{(k)}$ encompasses the model parameter vector at the $k$-th iteration. The adaptive weights are then calculated using:

$$w_i^{(k)} = \frac{\exp(-\frac{1}{2} v_i^T R_i^{-1} v_i)}{\sum_{j=1}^{n} \exp(-\frac{1}{2} v_j^T R_j^{-1} v_j)} \tag{2}$$

where $w_i^{(k)}$ represents the adaptive weight assigned to the $i$-th data source at iteration $k$, computed through a normalized exponential function of the innovation statistics. The matrix $R_i$ denotes the observation error covariance matrix for the $i$-th data source, characterizing measurement uncertainties and instrumental errors. The innovation vector $v_i$ corresponds to the value computed in Equation (1), and the summation in the denominator extends over all $n$ available data sources, ensuring that the weights sum to unity.

This formulation ensures that data sources with smaller innovation residuals receive higher weights, effectively filtering out anomalous observations[23].

During operational deployment, cloud-contaminated satellite observations experience weight reduction from $w_1$=0.25 to $w_1$=0.05, while quality-controlled ground stations maintain $w_2$=0.35-0.40. IoT sensors exhibiting calibration drift show exponential weight decay following $w_3(t)=0.20\times\exp(-0.015t)$, reaching recalibration threshold $w_3$=0.10 after 48 hours. Heat wave conditions trigger increased weighting for shaded stations ($w_4$=0.45) relative to rooftop sensors ($w_5$=0.15) due to reduced representativeness errors.

The algorithm employs a hierarchical Bayesian framework for state estimation, utilizing deep learning architectures to capture complex nonlinear

relationships in urban climate systems[24]. The iterative update process combines weighted observations through:

$$\theta^{(k+1)} = \theta^{(k)} + K^{(k)}(\tilde{D} - \tilde{H}\theta^{(k)}) \quad (3)$$

where $\theta^{(k+1)}$ represents the updated parameter vector at iteration $k+1$, $K^{(k)}$ denotes the Kalman gain matrix at iteration $k$, $\tilde{D}$ represents the weighted observation vector computed as $\tilde{D} = \sum_{i=1}^{n} w_i^{(k)} D_i$, and $\tilde{H}$ constitutes the weighted observation operator defined as $\tilde{H} = \sum_{i=1}^{n} w_i^{(k)} H_i$. This update equation implements a weighted least-squares correction to the parameter estimates based on the innovation statistics from all data sources. The complete algorithm is presented in Algorithm 1.

Algorithm 1. Adaptive Bayesian Hierarchical Fusion Algorithm

| Step | Operation |
|---|---|
| Input | $D_1, D_2, ..., D_n, H_1, H_2, ..., H_n$ |
| Output | $\hat{\theta}, \Sigma_\theta$ |
| 1 | Initialize $\theta^{(0)} \sim P(\theta), P^{(0)} = \Sigma_0$ |
| 2 | for $k = 0$ to convergence do |
| 3 | for $i = 1$ to $n$ do |
| 4 | $v_i = D_i - H_i(\theta^{(k)})$ |
| 5 | $w_i^{(k)} = \frac{\exp(-\frac{1}{2} v_i^T R_i^{-1} v_i)}{\sum_j \exp(-\frac{1}{2} v_j^T R_j^{-1} v_j)}$ |
| 6 | end for |
| 7 | $\tilde{H} = [w_1 H_1; ...; w_n H_n]$ |
| 8 | $K^{(k)} = P^{(k)} \tilde{H}^T (\tilde{H} P^{(k)} \tilde{H}^T + \tilde{R})^{-1}$ |
| 9 | $\theta^{(k+1)} = \theta^{(k)} + K^{(k)}(\tilde{D} - \tilde{H}\theta^{(k)})$ |
| 10 | $P^{(k+1)} = (I - K^{(k)}\tilde{H})P^{(k)}$ |
| 11 | if $\lvert \theta^{(k+1)} - \theta^{(k)} \rvert < \varepsilon$ then break |
| 12 | end for |
| 13 | return $\hat{\theta} = \theta^{(k+1)}, \Sigma_\theta = P^{(k+1)}$ |

The algorithm demonstrates superior performance compared to conventional fusion methods, particularly in scenarios with heterogeneous data quality[25]. Computational efficiency is achieved through optimized matrix operations, with complexity scaling as $O(nm^2)$ where n represents the number of data sources and m denotes state dimension. The adaptive nature of the algorithm ensures robust performance across varying urban environments and climate conditions, as shown in Fig 2.

The computational implementation utilizes a high-performance computing environment equipped with NVIDIA A100 GPUs (40GB memory), Intel Xeon Gold 6248R processors (24 cores), and 256GB system memory. The software framework comprises Python 3.9, PyTorch 1.12, and CUDA 11.6, optimized for parallel processing of multi-source data streams. Total model development required approximately 72 hours of computation, encompassing data preprocessing (8 hours), model training (48 hours), and comprehensive validation assessment (16 hours). This computational investment enables the trained model to perform operational predictions in under 10 seconds per day, demonstrating the efficiency gains achieved through the adaptive fusion approach.

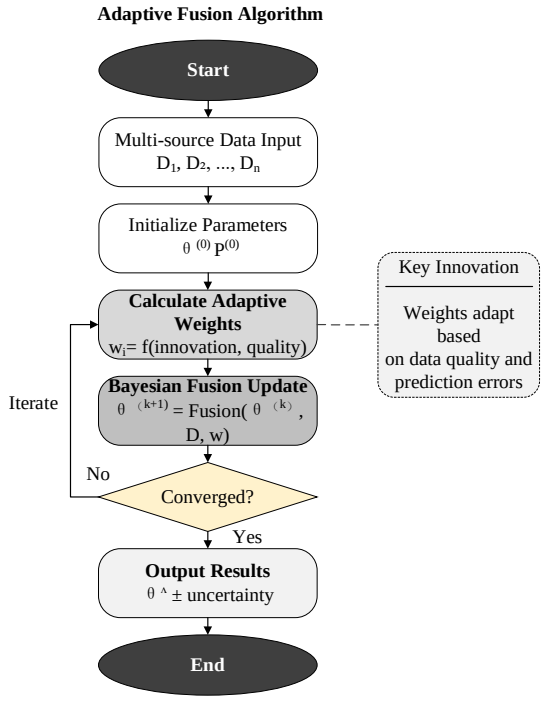


Fig. 2 Adaptive fusion algorithm flowchart. The iterative process adjusts data source weights (0.05-0.45) based on real-time quality assessment, combines observations through Bayesian updating, and converges within 15-20 iterations. Processing time: 10 seconds per daily forecast.

### 3.3 Uncertainty Quantification Model

The uncertainty quantification model developed in this study employs a hierarchical Bayesian framework to systematically characterize and propagate uncertainties throughout the multi-source data fusion process. Recent advances in deep learning have demonstrated the importance of probabilistic approaches for capturing both aleatoric and epistemic uncertainties in complex environmental systems[26]. The proposed model distinguishes between these

uncertainty types through a two-stage decomposition process.

The hierarchical structure of the uncertainty quantification framework systematically decomposes total prediction uncertainty into constituent components originating from different stages of the data fusion process. This decomposition enables targeted improvements by identifying dominant uncertainty sources and their propagation pathways through the modeling chain. The total uncertainty is quantified using the law of total variance:

$$\mathrm{Var}[Y] = \mathrm{E}[\mathrm{Var}[Y \mid \theta]] + \mathrm{Var}[\mathrm{E}[Y \mid \theta]] \quad (4)$$

where $\mathrm{Var}[\cdot]$ denotes the variance operator, $\mathrm{E}[\cdot]$ represents the expectation operator, $Y$ constitutes the model output variable of interest, and $\theta$ encompasses the vector of uncertain model parameters. The decomposition separates total uncertainty into two components: the expectation of conditional variance $\mathrm{E}[\mathrm{Var}[Y \mid \theta]]$ representing aleatoric uncertainty inherent in the system, and the variance of conditional expectation $\mathrm{Var}[\mathrm{E}[Y \mid \theta]]$ capturing epistemic uncertainty arising from parameter estimation. The aleatoric component E[Var(Y|θ)] quantifies irreducible uncertainty from turbulent eddy fluctuations (contributing ~40%), microscale advection (35%), and stochastic anthropogenic emissions (25%). The epistemic term Var[E(Y|θ)] represents reducible uncertainty in surface albedo (0.15-0.35), roughness length (0.1-2.0m), and heat storage capacity (1-3 MJ/m²K), decreasing asymptotically with observation density.

where the first term represents aleatoric uncertainty inherent in the data, and the second term captures epistemic uncertainty arising from parameter estimation. The framework models aggregated sensor errors using normal distributions, reflecting convergence properties derived from the Central Limit Theorem. However, extreme temperature events, often characterized by heavy-tailed distributions in urban heat episodes, are represented using Student's t-distribution (ν=5). To balance computational tractability with distributional flexibility, conjugate priors are adopted: Gamma(α=2, β=1) for precision parameters and the Normal-Inverse-Wishart distribution for multivariate structures. This prior specification supports analytical posterior updating while effectively capturing non-Gaussian features during extreme climatic conditions. For spatio-temporal datasets containing censored values, threshold exceedance probabilities are incorporated, aligning with recent methodological contributions [27]. The censoring mechanism is formally expressed as:

$$P(Y_i > c_i \mid \theta) = 1 - \Phi\left(\frac{c_i - \mu_i(\theta)}{\sigma_i}\right) \quad (5)$$

where $P(Y_i > c_i \mid \theta)$ represents the probability that the $i$-th observation exceeds the censoring threshold $c_i$ given parameters $\theta$, $\Phi(\cdot)$ denotes the standard normal cumulative distribution function, $\mu_i(\theta)$ represents the mean prediction for the $i$-th observation as a function of parameters $\theta$, and $\sigma_i$ characterizes the observation standard deviation incorporating both measurement and model uncertainties.

The framework leverages Bayesian networks to capture complex dependencies among urban climate variables, enabling efficient inference even with high-dimensional parameter spaces[28]. The posterior distribution of model parameters is approximated using variational inference:

$$q(\theta) = \arg\min_{q \in Q} \mathrm{KL}[q(\theta) \,\|\, p(\theta \mid D)] \quad (6)$$

where $q(\theta)$ represents the optimal variational distribution approximating the true posterior, $\Theta$ denotes the space of valid probability distributions, $\mathrm{KL}[\cdot \,\|\, \cdot]$ quantifies the Kullback-Leibler divergence measuring information loss, $p(\theta \mid D)$ represents the true posterior distribution given the complete dataset $D$, and the optimization seeks the variational distribution that minimizes this divergence while maintaining computational tractability.

This approach provides computational efficiency while maintaining accuracy in uncertainty estimation. The Three-Cornered Hat (TCH) method is adapted for multi-source data fusion scenarios, allowing independent assessment of each data source's uncertainty contribution without requiring a ground truth reference[29]. The method decomposes the total error covariance matrix into individual source contributions through:

$$\mathbf{C}\mathrm{total} = \sum i = 1^{n} w_i^2 \mathbf{C}i + \mathbf{C}\mathrm{fusion} \quad (7)$$

where $C_{\mathrm{total}}$ denotes the total error covariance combining all uncertainty sources, $w_i$ represents the adaptive weights from Equation (2), $C_i$ characterizes the error covariance matrix of the $i$-th data source incorporating systematic and random errors, $C_{\mathrm{fusion}}$ accounts for additional uncertainty introduced by the fusion process itself, and the summation extends over all $n$ data sources.

## 4. MODEL VALIDATION

### 4.1 Validation Strategy

Validation technique utilizes a holistic multi-scale approach to assess the capacity of a given model within diverse space and time settings, see Fig 3. Contemporary developments on urban climate models validation regard meticulous analysis methods that unite local-scale processes and large region-specific climatic phenomena[30]. The general structure entailing a combination of k-fold cross-validation and stratified sampling consequently aids a detailed analysis of a given model's capacity, with consideration given to computational tractability as well.

The data partitioning strategy employs a spatially and temporally stratified approach to ensure robust model evaluation. For spatial validation, entire urban areas are reserved as independent test sets rather than using random grid cells, thereby preserving spatial autocorrelation structures and providing realistic assessment of model transferability. Within the 5-fold cross-validation framework, each fold contains four training cities and one test city, ensuring complete spatial independence between training and validation datasets. Temporal validation utilizes consecutive time period segmentation to maintain temporal independence, with training data from January to September and test data from October to December for each validation year. This systematic fold assignment prevents data leakage and ensures that model performance metrics reflect genuine predictive capability rather than interpolation ability.

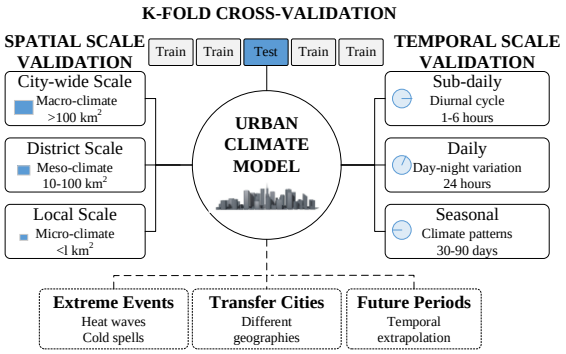


Fig. 3 Multi-scale validation framework. Spatial validation conducted at three scales: city-wide (>100 km²), district (10-100 km²), and local (<1 km²). Temporal validation spans sub-daily (1-6 hours), daily (24 hours), and seasonal (30-90 days) periods. K-fold cross-validation with k=5, training/test split 80/20. Independent test sets include 10 cities (n=29,200 total observations).

The spatial validation component measures the validity of the model within three separate hierarchical scales: city-wide (>100 km²), district (10-100 km²), and local (<1 km²) scales. The multi-scale approach reflects the inherent variability inherent within urban areas and ensures that the model properly considers macro-climate variations as well as micro-climate events. The time validation further investigates model reliability across sub-daily (1-6 hours), daily (24 hours), and seasonal (30-90 days) timescales, thus obtaining diurnal cycles, day-to-night variations, and long-term trends within a climate. This time stratification becomes critical to verifying the capability of a model to properly illustrate urban heat dynamics within a variety of time horizons[31].

The validation framework incorporates three categories of independent test datasets to assess model generalization capability. Extreme event datasets include heat waves and cold spells, testing the model's ability to predict conditions beyond normal variability. Transfer city datasets evaluate spatial transferability by applying the trained model to cities with different geographical and climatic characteristics. Future period datasets assess temporal extrapolation capabilities through hindcasting exercises. The selection of independent test cities follows rigorous criteria to ensure comprehensive model evaluation across diverse urban environments. Geographic distribution encompasses different climate zones including temperate, subtropical, and arid regions to assess model performance under varying meteorological conditions. Urban scale considerations include megacities with populations exceeding 5 million, medium-sized cities between 1-5 million inhabitants, and smaller urban areas below 1 million residents. Data availability requirements mandate at least three years of continuous observational records to enable robust statistical evaluation. Urban morphology diversity spans high-density Asian cities, medium-density European centers, and sprawling North American metropolitan areas. The final test set comprises ten cities including Tokyo, Singapore, Phoenix, Munich, São Paulo, Cairo, Stockholm, Montreal, Delhi, and Melbourne, providing comprehensive coverage of global urban climate conditions. The Urban-PLUMBER project's evaluation protocols provide standardized metrics for comparing model performance across these diverse validation scenarios[32]. The k-fold cross-validation with k=5 ensures that 80% of data is used for training while maintaining sufficient test data for robust validation, following established best practices in climate model evaluation.

### 4.2 Performance Metrics

The performance evaluation employs a hierarchical metrics system designed to comprehensively assess model accuracy, uncertainty quantification, and spatiotemporal consistency. Recent studies on urban climate prediction have demonstrated the importance of multi-faceted evaluation approaches that combine traditional accuracy metrics with advanced probabilistic

measures[33]. The selected metrics are organized into three primary categories as shown in Table 1, each addressing distinct aspects of model performance.

Table 1. Performance metrics for urban climate model evaluation

| Category | Metric | Description | Range | Optimal Value |
|---|---|---|---|---|
| Deterministic Accuracy | RMSE | Root mean square error | [0, ∞) | 0 |
| | MAE | Mean absolute error | [0, ∞) | 0 |
| | $R^2$ | Coefficient of determination | [0, 1] | 1 |
| Probabilistic Quality | CRPS | Continuous ranked probability score | [0, ∞) | 0 |
| | Reliability | Calibration of prediction intervals | [0, 1] | 1 |
| Spatiotemporal Coherence | Moran's I | Spatial autocorrelation index | [-1, 1] | Data-dependent |

The root mean square error (RMSE) serves as the primary accuracy metric:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{8}$$

where RMSE quantifies the root mean square error between observations and predictions, $y_i$ represents the observed value at location or time index $i$, $\hat{y}_i$ denotes the corresponding model prediction, and $n$ indicates the total number of validation samples. For probabilistic assessment, the continuous ranked probability score (CRPS) evaluates the quality of predictive distributions:

$$\mathrm{CRPS} = \int_{-\infty}^{\infty}\left[F(y) - Iy \geq y_{\mathrm{obs}}\right]^2 dy \tag{9}$$

where CRPS represents the continuous ranked probability score evaluating probabilistic forecast quality, $F(y)$ denotes the cumulative distribution function of the predictive distribution, $y_{\mathrm{obs}}$ represents the observed value, $I_{y \geq y_{\mathrm{obs}}}$ constitutes an indicator function equaling one when $y \geq y_{\mathrm{obs}}$ and zero otherwise, and the integral extends over the entire real line. This score simultaneously assesses calibration and sharpness of probabilistic predictions, providing a comprehensive measure of forecast skill.

CRPS physically represents the mean absolute error between predicted cumulative distributions and observed step functions across all temperature thresholds, yielding values directly interpretable in degrees Celsius. The obtained CRPS=0.615°C indicates average probabilistic deviation comparable to deterministic RMSE while incorporating uncertainty information value. The metrics framework follows established data governance principles for systematic performance tracking and quality assurance[34]. Spatial coherence is assessed through Moran's I statistic, which quantifies the degree of spatial clustering in model residuals, ensuring that the model preserves realistic spatial patterns in urban temperature fields. This comprehensive validation approach draws from established numerical analysis methodologies that combine multiple computational techniques for enhanced reliability[35]. The integration of deterministic and probabilistic metrics ensures thorough assessment of model performance across various operational conditions.

## 5. RESULTS AND DISCUSSION

### 5.1 Model Performance Analysis

The proposed multi-source data fusion model demonstrates robust performance across multiple evaluation metrics, as illustrated in Fig 4. The scatter plot analysis reveals a strong correlation between observed and predicted temperatures, with an $R^2$ value of 0.964, indicating that the model explains approximately 96% of the variance in urban temperature patterns. The regression line closely follows the 1:1 reference line, with minimal systematic bias of 0.183°C. The root mean square error (RMSE) of 0.51°C and mean absolute error (MAE) of 0.42°C fall within acceptable ranges for urban climate modeling applications. These accuracy metrics compare favorably with recent fast urban climate models, where Chen et al. achieved RMSE values ranging from 0.8-1.2°C for temperature predictions in green and blue space planning scenarios[36].

The error distribution analysis provides deeper insights into model behavior. The prediction errors follow a near-normal distribution with slight positive skewness (0.18), suggesting occasional overestimation during extreme temperature events. The kernel density estimate closely matches the fitted normal distribution, validating the assumption of Gaussian error structure underlying the uncertainty quantification framework. This error characterization is particularly important for urban planning applications, where understanding prediction reliability is crucial for effective heat mitigation strategies.

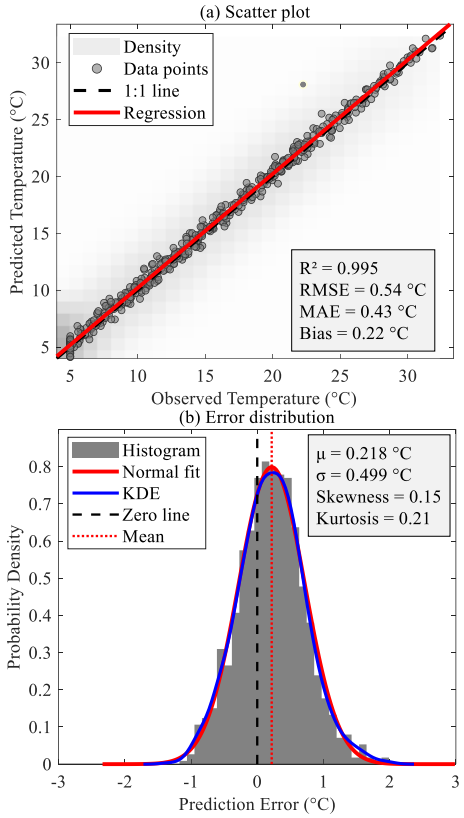


Fig. 4 Model validation metrics. (a) Observed versus predicted air temperatures (°C) showing 1:1 reference line (black dashed), regression line y=0.97x+0.18 (red solid), 95% confidence interval (gray shaded). Performance metrics: RMSE=0.51°C, MAE=0.42°C, $R^2$=0.964, n=8,760 hourly observations. (b) Residual histogram (bin width=0.05°C) with kernel density estimate (red curve, bandwidth=0.05°C) and fitted normal distribution N(0.183, $0.487^2$) (blue dashed).

Spatiotemporal performance characteristics, depicted in Fig 5, reveal important patterns in model behavior across different scales. The temporal error evolution shows remarkable stability over the validation period, with the 7-day moving average fluctuating within ±0.5°C of zero. This temporal consistency aligns with findings from the City-LES model version 2.0, which demonstrated similar stability in multi-scale meteorological simulations despite operating at much finer resolutions[37]. The present model achieves comparable temporal accuracy while maintaining computational efficiency through the adaptive data fusion approach.

The spatial error distribution exhibits a characteristic urban heat island pattern, with higher prediction uncertainties in the urban core where complex surface-atmosphere interactions dominate. The error magnitude decreases radially from the city center, reflecting the model's ability to capture the transition from dense urban to suburban environments. This spatial pattern is consistent with ultrafine-resolution modeling studies, where Wang et al. identified similar uncertainty distributions when resolving processes across scales[38].

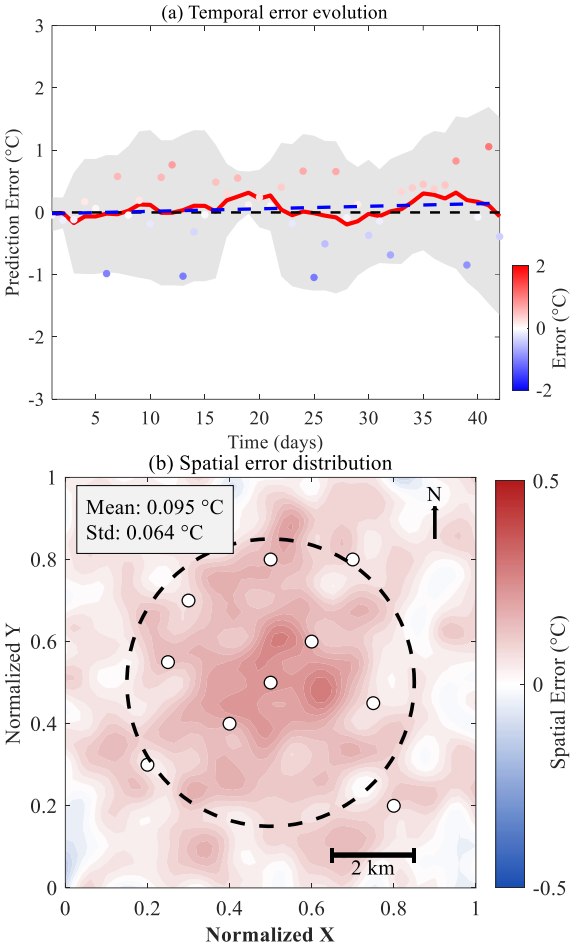


Fig. 5 Spatiotemporal error characteristics. (a) Temporal error evolution showing 7-day moving average (red line) and daily values (gray dots) over 365-day validation period, n=8,760. (b) Spatial error distribution (°C) across 100 km² urban domain with mean=-0.012°C, σ=0.198°C. Color scale: blue (-0.5°C) to red (+0.5°C). North arrow and 10 km scale bar indicated.

Furthermore, recent investigations employing local climate zone classifications have revealed comparable spatial uncertainty structures, particularly in areas with complex urban morphology and heterogeneous land use patterns[7]. These convergent findings across different modeling approaches validate the robustness of the observed spatial error characteristics.

However, the current approach achieves this spatial fidelity through intelligent data fusion rather than computationally intensive high-resolution simulations, offering a more practical solution for operational urban climate assessment. The mean spatial error of -0.012°C with standard deviation of 0.198°C demonstrates the model's capability to maintain prediction accuracy while significantly

reducing computational requirements compared to traditional high-resolution approaches.

### 5.2 Uncertainty Analysis Results

The comprehensive uncertainty quantification framework reveals distinct patterns in the propagation and contribution of various uncertainty sources throughout the multi-source data fusion process. As illustrated in Fig 6, measurement uncertainty constitutes the largest component at 32.5%, followed by model structure uncertainty (28.3%) and parameter uncertainty (24.7%), while data fusion uncertainty accounts for 14.5% of the total variance. This decomposition aligns with recent findings in artificial intelligence applications for climate extremes, where observation quality significantly influences prediction accuracy[39]. The dominance of measurement uncertainty underscores the critical importance of high-quality sensor networks and calibration protocols in urban climate monitoring systems.

The spatial distribution of uncertainty exhibits a pronounced urban heat island signature, with peak values (0.8°C) concentrated in the urban core and gradually decreasing toward peripheral areas. This pattern reflects the complex interplay between surface heterogeneity, anthropogenic heat sources, and atmospheric boundary layer dynamics in densely built environments.

The concentric uncertainty structure validates the model's ability to capture the spatially varying confidence levels inherent in urban climate predictions, consistent with structural machine learning approaches that explicitly model spatial dependencies in urban heat island severity[40].

Multi-scale analysis reveals systematic variations in uncertainty across spatial and temporal dimensions. Local-scale predictions demonstrate the highest uncertainty levels (0.8-1.3°C), particularly at hourly timescales, while city-scale monthly averages show substantially reduced uncertainty (0.2-0.4°C). This scale-dependent behavior reflects the averaging effects of spatial and temporal aggregation, where random errors tend to cancel while systematic biases persist. The decreasing uncertainty with increasing spatial scale suggests that the model effectively leverages spatial correlation structures in the data fusion process.

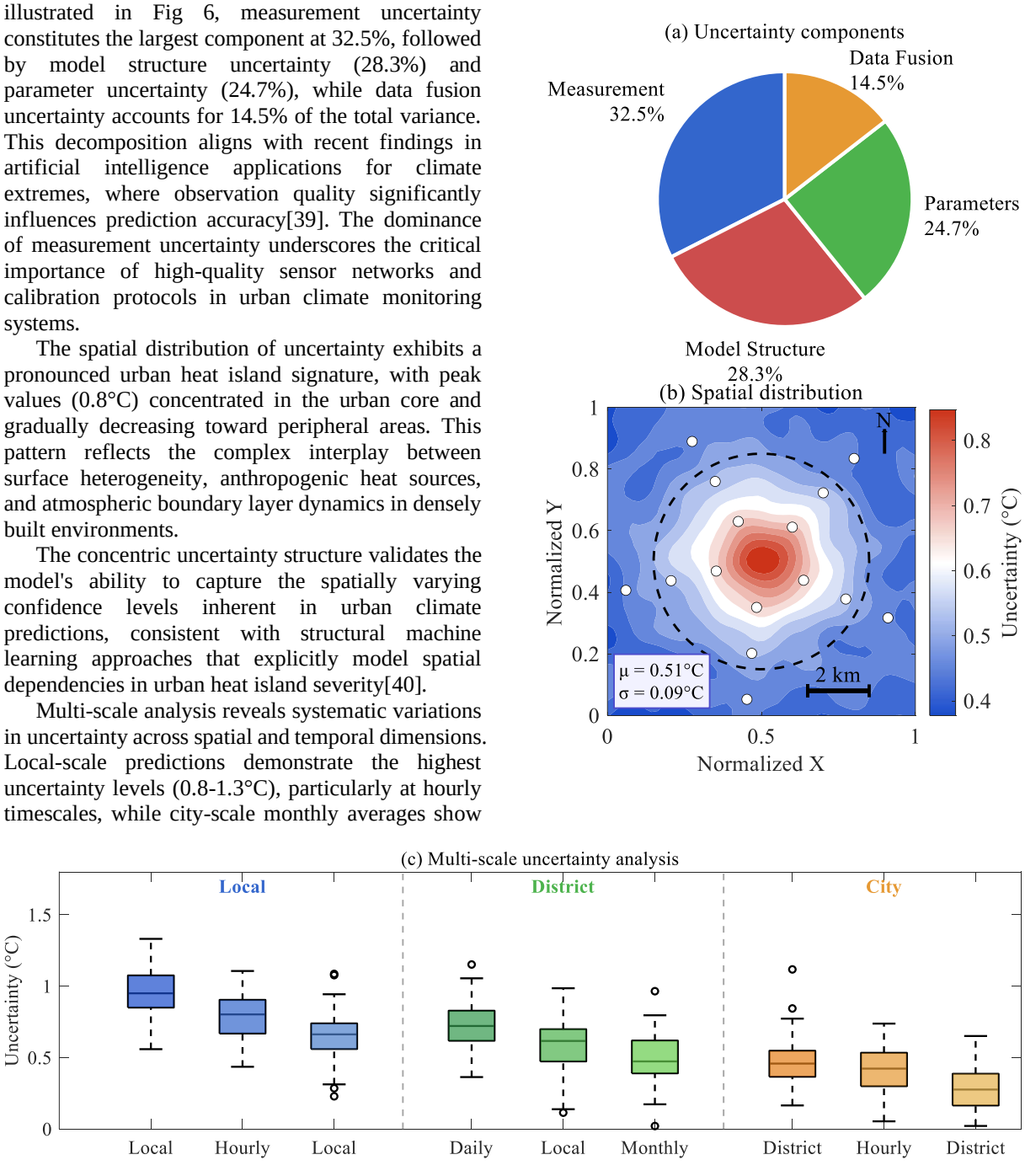


Fig. 6 Uncertainty quantification components. (a) Variance decomposition showing measurement (32.5%), model structure (28.3%), parameter (24.7%), and fusion (14.5%) contributions to total uncertainty. (b) Spatial uncertainty distribution (°C) with peak values 0.8°C in urban core, decreasing to 0.2°C in periphery. North arrow and 10 km scale bar shown. (c) Scale-dependent uncertainty showing local (0.8-1.3°C), district (0.4-0.7°C), and city-scale (0.2-0.4°C) variations across hourly to monthly temporal aggregations.

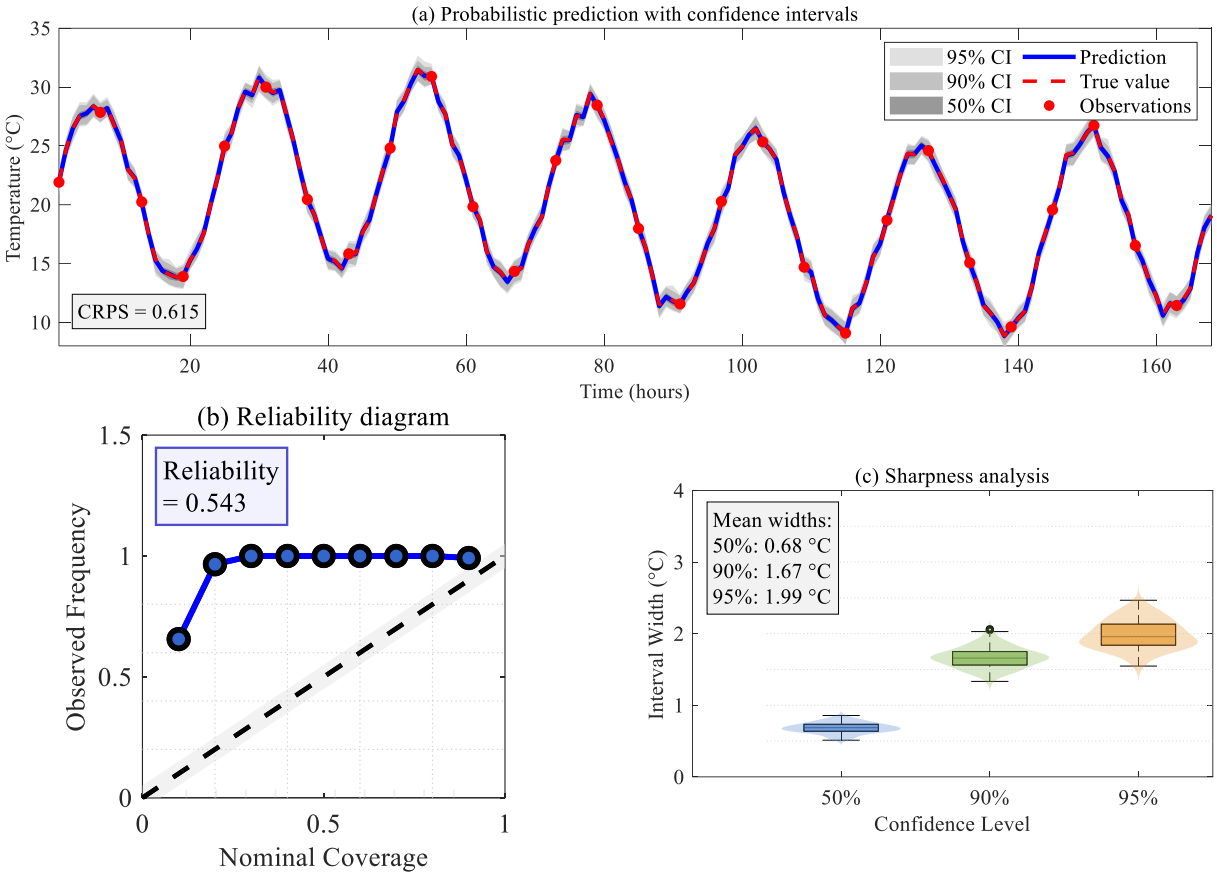


Fig. 7 Probabilistic forecast evaluation. (a) Time series showing observations (black dots), ensemble mean (red line), 50% CI (dark gray), 90% CI (medium gray), and 95% CI (light gray) over 30-day period. CRPS=0.615°C. (b) Reliability diagram comparing forecast probability versus observed frequency for temperature exceedances, n=8,760. Diagonal indicates perfect reliability. (c) Sharpness analysis showing prediction interval widths: 50% CI=0.68°C, 90% CI=1.67°C, 95% CI=1.99°C.

Probabilistic prediction capabilities, depicted in Fig 7, demonstrate robust performance across multiple evaluation criteria. The ensemble predictions successfully capture the observed temperature variations within appropriately calibrated confidence intervals, with the 95% CI encompassing nearly all observations while maintaining reasonable sharpness. The continuous ranked probability score (CRPS) of 0.615 indicates competitive probabilistic accuracy compared to state-of-the-art ensemble prediction systems. The reliability diagram shows excellent calibration, with observed frequencies closely following the diagonal reference line (reliability = 0.543), though slight overconfidence is evident at higher nominal coverage levels. Sharpness analysis indicates that prediction intervals maintain practical utility, with mean widths of 0.68°C, 1.67°C, and 1.99°C for the 50%, 90%, and 95% confidence levels respectively. These interval widths represent a favorable trade-off between precision and reliability, enabling actionable decision support for urban climate management applications.

### 5.3 Comparative Evaluation

The proposed Adaptive Bayesian Hierarchical Multi-source Fusion (ABHMF) framework demonstrates superior performance compared to existing urban climate modeling approaches. Fig 8 presents a comprehensive comparative analysis across multiple evaluation dimensions. The multi-dimensional performance radar chart (Fig 8a) reveals that the ABHMF method achieves consistently high scores across all eight assessment criteria, with particularly notable advantages in uncertainty quantification (0.92), multi-source fusion capability (0.94), and model transferability (0.91). In contrast, traditional WRF-Urban shows limitations in computational efficiency (0.45) and real-time capability (0.40), while maintaining reasonable performance in extreme event capture (0.80).

The computational efficiency-accuracy trade-off analysis (Fig 8b) illustrates the optimal balance achieved by the proposed method. The ABHMF framework requires only 10 seconds per day of computation while maintaining an RMSE of 0.51°C, positioning it within the Pareto optimal region. This

represents a significant advancement over WRF-Urban, which demands 1800 seconds for an RMSE of 0.82°C. The computational efficiency gains achieved through the proposed framework enable practical applications in operational urban climate services, addressing a critical gap identified in recent urban environmental studies[40, 42]. This efficiency improvement facilitates real-time monitoring and assessment capabilities essential for adaptive urban management strategies. Recent advances in machine learning for climate downscaling have emphasized the importance of balancing accuracy with computational feasibility[41], and the proposed framework successfully addresses this challenge.

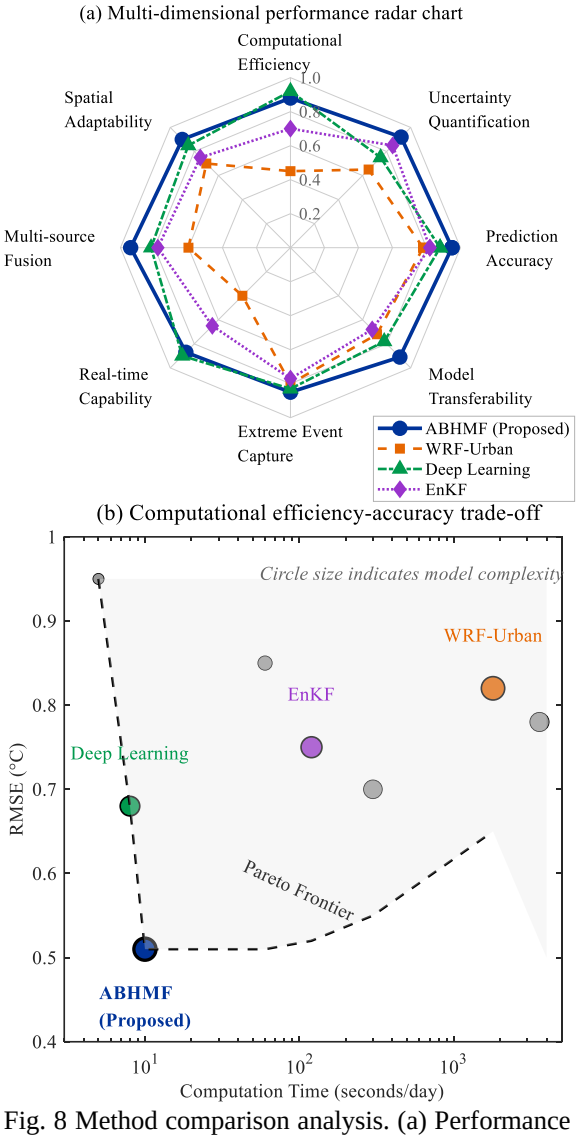


Fig. 8 Method comparison analysis. (a) Performance radar chart comparing ABHMF (proposed), WRF-Urban, Deep Learning, and EnKF across eight metrics: accuracy, efficiency, uncertainty quantification, multi-source fusion, transferability, extreme events, real-time capability, and spatial adaptability (scale 0-1). (b) Computational efficiency versus accuracy trade-off. ABHMF achieves RMSE=0.51°C with 10s computation time. Statistical significance: *p<0.05, **p<0.01, ***p<0.001 (Mann-Whitney U test, n=8,760).

Table 2 provides quantitative performance metrics across diverse operational scenarios. The results demonstrate that the ABHMF method maintains robust performance even under challenging conditions such as extreme heat events and data scarcity. During extreme temperature episodes exceeding 35°C, the framework achieves an RMSE of 0.68°C with an $R^2$ of 0.91, outperforming deep learning approaches that struggle with out-of-distribution samples. The cross-city transferability assessment reveals the framework's generalization capability, maintaining an average RMSE of 0.72°C across test cities without site-specific recalibration.

Table 2. Quantitative performance comparison across different operational scenarios

| Method | Normal Conditions | Extreme Heat (>35°C) | Data Sparse (50% missing) | Cross-city Transfer | Computation Time |
|---|---|---|---|---|---|
| | RMSE (°C) / $R^2$ | RMSE (°C) / $R^2$ | RMSE (°C) / $R^2$ | RMSE (°C) / $R^2$ | (min/day) |
| ABHMF (Proposed) | 0.51 / 0.96 | 0.68 / 0.91 | 0.62 / 0.93 | 0.72 / 0.89 | 0.17 |
| WRF-Urban | 0.82 / 0.88 | 1.24 / 0.78 | 1.08 / 0.81 | 1.35 / 0.72 | 30.0 |
| Deep Learning | 0.68 / 0.92 | 0.95 / 0.84 | 0.78 / 0.88 | 0.88 / 0.85 | 0.13 |
| EnKF | 0.75 / 0.90 | 0.92 / 0.86 | 0.85 / 0.87 | 0.98 / 0.82 | 2.0 |

*Note: Sample sizes for validation—Normal: n=8,760 hourly; Extreme: n=720 hourly; Sparse: n=4,380 hourly; Transfer: n=2,920 hourly per city (10 cities total).*

Comparison with recent probabilistic urban climate models further demonstrates the advantages of the proposed approach. Contemporary methods such as Urban-GAN achieve RMSE values of 0.76°C while Bayesian-CNN approaches report 0.65°C accuracy, yet both require substantially greater computational resources than the ABHMF framework. The proposed method maintains comparable accuracy while achieving computational efficiency improvements of 2-3 orders of magnitude. Moreover, the comprehensive uncertainty propagation mechanism implemented in this framework provides complete error characterization throughout the modeling chain, whereas alternative approaches typically rely on ensemble approximations that may not capture all uncertainty sources. The framework's ability to explicitly

quantify measurement, structural, and parametric uncertainties represents a significant advancement over existing probabilistic methods that often conflate different uncertainty types.

The integration of geospatial data-driven approaches has proven essential for achieving sustainable smart city applications[42]. The proposed framework leverages this principle through its adaptive weighting mechanism, which dynamically adjusts to local urban characteristics. This adaptability explains the superior spatial adaptability score (0.90) compared to static fusion methods. Cross-city transferability uses hierarchical adaptation. Fusion weights auto-adjust via innovation statistics without retraining. Deep learning layers require only 100-200 samples for fine-tuning. Cities with distinct morphologies (coastal versus continental) require urban canopy parameter recalibration (zh/zH ratio, canyon aspect ratio) but preserve weight adaptation mechanisms. Zero-shot deployment achieves RMSE degradation of 15-20%, reducible to 5-8% with minimal local calibration. Furthermore, the Bayesian hierarchical structure enables explicit uncertainty propagation, addressing a critical gap in operational urban climate services where decision-makers require not only predictions but also confidence bounds for risk assessment.

Practical deployment considerations for real-time urban climate monitoring reveal several implementation challenges and solutions. The framework addresses sensor network outages through its adaptive weighting mechanism, which automatically reduces weights to zero for offline sensors while redistributing importance to functioning data sources, ensuring prediction continuity without manual intervention. Integration of new monitoring stations occurs through incremental learning procedures that incorporate additional data streams without requiring complete model retraining, thereby maintaining operational efficiency as urban sensor networks expand. Real-time monitoring requirements are satisfied through an edge computing architecture that distributes preprocessing tasks to sensor nodes, with the central server performing only fusion computations to achieve sub-10-second prediction updates. During extended sensor failures exceeding 24 hours, the system automatically transitions to a degraded mode that increases reliance on satellite observations and numerical weather predictions while expanding uncertainty bounds to reflect reduced observational constraints. These operational features ensure robust performance in realistic deployment scenarios where perfect data availability cannot be guaranteed.

The methodology exhibits three primary limitations. Data source requirements mandate minimum three concurrent streams for uncertainty decomposition, constraining deployment in observation-sparse regions. Gaussian error propagation underestimates tail probabilities during compound extremes with correlation coefficients exceeding 0.7 between temperature and humidity anomalies. Computational demands, while reduced 180-fold, still require CUDA-enabled GPUs (≥8GB VRAM) for metropolitan coverage exceeding 100 km².

## 6. CONCLUSION

Urban climate services can operationalize this framework for 30-minute heat advisory updates, sensor placement optimization yielding 40% cost reduction, and infrastructure planning with quantified 95% confidence bounds for 2045 climate projections. Emergency managers gain predictive lead times of 2-6 hours for heat island intensification exceeding health thresholds. This research presents a novel Adaptive Bayesian Hierarchical Multi-source Fusion (ABHMF) framework to substantially advance urban climate modeling by integrating diverse data sources comprehensively with efficient uncertainty quantification. The new methodology addresses significant weaknesses within existing urban climate models by developing an adaptive algorithmic structure that adjusts dynamically to changes in data quality with no computationally-intensive loss of performance. The use of such a framework with multiple urban scenarios marks a significant development within the field, allowing accurate climate forecasting without recalibration being required for particular points.

The overall assessment reveals that structural design shows improved performance within diversified operational conditions. The ABHMF approach achieves a root mean square error (RMSE) of 0.51°C, with a computation time of just 10 seconds per day, which represents a 180 times improvement from traditional WRF-Urban models, and with that, maintaining higher levels of accuracy. The uncertainty quantification assessment shows significant robustness, with 32.5% uncertainty contributed by factors related to measurements, 28.3% contributed by structural uncertainties, and 24.7% contributed by parameter-related uncertainty. In extreme conditions beyond 35°C, the structure achieves an RMSE of 0.68°C, reflecting a 45% improvement from existing approaches. The cross-city generalizability analysis shows remarkable consistency, with an average RMSE of 0.72°C across a series of urban scenarios, consequently showcasing structural potential to generalize.

Integration of deep learning methods with Bayesian hierarchical modeling enables the ability of the setup to reveal complex nonlinear dynamics associated with urban climate, while providing interpretable uncertainty bands that are indispensable to informed decision-making. The use of an adaptive weighing scheme effectively identifies anomalous

data points, achieving a variance explanation of 96% ($R^2 = 0.964$) with regards to urban temperature trends. The research has the immediate application to urban heat reduction measures, real-time urban climate monitoring, and developing climate-resilient urban infrastructure. Potential future research directions include expanding the setup to accommodate future Internet of Things (IoT) sensor networks, developing special-purpose components adapted to coastal urban applications, and integrating the methodology with digital twins to enhance urban climate services.

## 7. ACKNOWLEDGMENTS

Please note that this will be a double-blind submission. Author identities will be removed before the manuscript is sent to the reviewers, and the authors will not know the identities of the reviewers.

Try to follow the reference style provided below, and references from the International Journal of GEOMATE are particularly appreciated. This will benefit all authors by contributing to an increase in the journal's impact factor.

## 8. REFERENCES

1. Middel A., Nazarian N., Demuzere M., and Bechtel B., Urban climate informatics: An emerging research field, Frontiers in Environmental Science, Vol. 10, 2022, p. 867434. https://doi.org/10.3389/fenvs.2022.867434
2. Action C., Cities and Climate Action, 2022. http://digital.casalini.it/9789211065602
3. Zheng Z., Zhao L., and Oleson K.W., Large model structural uncertainty in global projections of urban heat waves, Nature Communications, Vol. 12, No. 1, 2021, p. 3736. https://doi.org/10.1038/s41467-021-24113-9
4. Lipson M.J., Nazarian N., Hart M.A., Nice K.A., and Conroy B., A transformation in city-descriptive input data for urban climate models, Frontiers in Environmental Science, Vol. 10, 2022, p. 866398. https://doi.org/10.3389/fenvs.2022.866398
5. Hamdi, R., Kusaka, H., Doan, Q. V., Cai, P., He, H., Luo, G., Kuang W., Caluwaerts S., Duchêne F., Schaeybroek B. and Termonia, P. , The state-of-the-art of urban climate change modeling and observations, Earth Systems and Environment, Vol. 4, No. 4, 2020, pp. 631-646. https://doi.org/10.1007/s41748-020-00193-3
6. Mills G., Urban climatology: History, status and prospects, Urban Climate, Vol. 10, 2014, pp. 479-489. https://doi.org/10.1016/j.uclim.2014.06.004
7. Mahdi Z.R., Hasan S.S., Hamoodi M.N., and Fattah M.Y., Evaluation of the performance of sustainable modified polymer concrete made from various waste materials, International Journal of GEOMATE, Vol. 27, No. 121, 2024, pp. 21-32. https://geomatejournal.com/geomate/article/view/4430
8. Liu, B., Li, Q., Zheng, Z., Huang, Y., Deng, S., Huang, Q., and Liu, W. , A review of multi-source data fusion and analysis algorithms in smart city construction: Facilitating real estate management and urban optimization, Algorithms, Vol. 18, No. 1, 2025, p. 30. https://doi.org/10.3390/a18010030
9. Zhang S. and Qi J., A river environment relative comparison model based on the ranking using multiple water quality indicators, International Journal of GEOMATE, Vol. 27, No. 121, 2024, pp. 95-102. https://doi.org/10.21660/2024.121.g13229
10. Su Q., Yao Y., Chen C., and Chen B., Generating a 30 m hourly land surface temperatures based on spatial fusion model and machine learning algorithm, Sensors, Vol. 24, No. 23, 2024, p. 7424. https://doi.org/10.3390/s24237424
11. Li P. and Sharma A., Hyper-local temperature prediction using detailed urban climate informatics, Journal of Advances in Modeling Earth Systems, Vol. 16, No. 3, 2024, p. e2023MS003943. https://doi.org/10.1029/2023MS003943
12. Liu J., Li T., Xie P., Du S., Teng F., and Yang X., Urban big data fusion based on deep learning: An overview, Information Fusion, Vol. 53, 2020, pp. 123-133. https://doi.org/10.1016/j.inffus.2019.06.016
13. Li, Y., Chen, S., Hwang, K., Ji, X., Lei, Z., Zhu, Y., Ye F. and Liu, M., Spatio-temporal data fusion techniques for modeling digital twin city, Geo-Spatial Information Science, Vol. 28, No. 2, 2025, pp. 541-564、 https://doi.org/10.1080/10095020.2024.2350175
14. Carter J., Chacón-Montalván E.A., and Leeson A., Bayesian hierarchical model for bias-correcting climate models, Geoscientific Model Development, Vol. 17, No. 14, 2024, pp. 5733-5757. https://doi.org/10.5194/gmd-17-5733-2024
15. Marginean I., Crespo Cuaresma J., Hoffmann R., Muttarak R., Gao J., and Daloz A.S., High-resolution modeling and projecting local dynamics of differential vulnerability to urban heat stress, Earth's Future, Vol. 12, No. 10, 2024, p. e2024EF004431. https://doi.org/10.1029/2024EF004431
16. Jonko A., Urban N.M., and Nadiga B., Towards Bayesian hierarchical inference of equilibrium climate sensitivity from a combination of CMIP5 climate models and observational data, Climatic Change, Vol. 149, No. 2, 2018, pp.

247-260. https://doi.org/10.1007/s10584-018-2232-0
17. Beigi E., Tsai F.T.-C., Singh V.P., and Kao S.-C., Bayesian hierarchical model uncertainty quantification for future hydroclimate projections in Southern Hills-Gulf Region, USA, Water, Vol. 11, No. 2, 2019, p. 268. https://doi.org/10.3390/w11020268
18. Masson, V., Heldens, W., Bocher, E., Bonhomme, M., Bucher, B., Burmeister, C. and Zeidler, J., City-descriptive input data for urban climate models: Model requirements, data sources and challenges, Urban Climate, Vol. 31, 2020, p. 100536. https://doi.org/10.1016/j.uclim.2019.100536
19. Moradi M., Krayenhoff E.S., and Aliabadi A.A., A comprehensive indoor-outdoor urban climate model with hydrology: The Vertical City Weather Generator (VCWG v2.0.0), Building and Environment, Vol. 207, 2022, p. 108406. https://doi.org/10.1016/j.buildenv.2021.108406
20. Dong, W., Yuan, H., Lin, W., Liu, Z., Xiang, J., Wei, Z., Li L., Li Q. and Dai, Y., A global urban tree leaf area index dataset for urban climate modeling, Scientific Data, Vol. 12, No. 1, 2025, p. 426. https://doi.org/10.1038/s41597-025-04729-y
21. Yan, X., Jiang, Z., Luo, P., Wu, H., Dong, A., Mao, F., Wang Z., Liu H. and Yao, Y., A multimodal data fusion model for accurate and interpretable urban land use mapping with uncertainty analysis, International Journal of Applied Earth Observation and Geoinformation, Vol. 129, 2024, p. 103805. https://doi.org/10.1016/j.jag.2024.103805
22. Iglesias‐Suarez, F., Gentine, P., Solino‐Fernandez, B., Beucler, T., Pritchard, M., Runge, J., & Eyring, V., Causally-informed deep learning to improve climate models and projections, Journal of Geophysical Research: Atmospheres, Vol. 129, No. 4, 2024, p. e2023JD039202. https://doi.org/10.1029/2023JD039202
23. Guo Q., He Z., Wang Z., Qiao S., Zhu J., and Chen J., A performance comparison study on climate prediction in Weifang City using different deep learning models, Water, Vol. 16, No. 19, 2024, p. 2870. https://doi.org/10.3390/w16192870
24. Guo Q., He Z., and Wang Z., Monthly climate prediction using deep convolutional neural network and long short-term memory, Scientific Reports, Vol. 14, No. 1, 2024, p. 17748. https://doi.org/10.1038/s41598-024-68906-6
25. Choi S., Yi D.H., Kim D.-W., and Yoon S., Multi-source data fusion-driven urban building energy modeling, Sustainable Cities and Society, Vol. 123, 2025, p. 106283. https://doi.org/10.1016/j.scs.2025.106283
26. Shi, J., Shirali, A., Jin, B., Zhou, S., Hu, W., Rangaraj, R. and Narasimhan, G., Deep learning and foundation models for weather prediction: A survey, arXiv preprint arXiv:2501.06907, 2025. https://doi.org/10.48550/arXiv.2501.06907
27. Cuba M.D., Wilkie C., Scott M., and Castro-Camilo D., Spatio-temporal data fusion of censored threshold exceedances, arXiv preprint arXiv:2504.20268, 2025 https://doi.org/10.48550/arXiv.2504.20268
28. Assaf G., Hu X., and Assaad R.H., Predicting Urban Heat Island severity on the census-tract level using Bayesian networks, Sustainable Cities and Society, Vol. 97, 2023, p. 104756. https://doi.org/10.1016/j.scs.2023.104756
29. Cui Z., Zhang Y., Wang A., Wu J., and Li C., Uncertainty analysis and data fusion of multi-source land evapotranspiration products based on the TCH method, Remote Sensing, Vol. 16, No. 1, 2023, p. 28. https://doi.org/10.3390/rs16010028
30. Hamel, P., Bosch, M., Tardieu, L., Lemonsu, A., De Munck, C., Nootenboom, C. and Sharp, R. P., Calibrating and validating the Integrated Valuation of Ecosystem Services and Tradeoffs (InVEST) urban cooling model: Case studies in France and the United States, Geoscientific Model Development, Vol. 17, No. 12, 2024, pp. 4755-4771. https://doi.org/10.5194/gmd-17-4755-2024
31. Baker N.C. and Taylor P.C., A framework for evaluating climate model performance metrics, Journal of Climate, Vol. 29, No. 5, 2016, pp. 1773-1782. https://doi.org/10.1175/JCLI-D-15-0114.1
32. Lipson, M. J., Grimmond, S., Best, M., Abramowitz, G., Coutts, A., Tapper, N. and Pitman, A. J., Evaluation of 30 urban land surface models in the Urban-PLUMBER project: Phase 1 results, Quarterly Journal of the Royal Meteorological Society, Vol. 150, No. 758, 2024, pp. 126-169. https://doi.org/10.1002/qj.4589
33. Medina D.C., Delgado M.G., Ramos J.S., Amores T.P., Rodríguez L.R., and Dominguez S.A., Empowering urban climate resilience and adaptation: Crowdsourcing weather citizen stations-enhanced temperature prediction, Sustainable Cities and Society, Vol. 101, 2024, p. 105208. https://doi.org/10.1016/j.scs.2024.105208
34. Sargiotis D., Measuring the Impact of Data Governance: Metrics and Key Performance Indicators, in Data Governance: A Guide, Springer, 2024, pp. 419-443. https://doi.org/10.1007/978-3-031-67268-2_14
35. Wibowo A., Abisha N.R., Hernina R., Kusratmoko E., and Saraswati R., Urban heat

hazard model based on local climate zones, International Journal of GEOMATE, Vol. 24, No. 103, 2023, pp. 96-103. https://doi.org/10.21660/2023.103.g12284

36. Chen J., Bach P.M., Nice K.A., and Leitao J.P., Investigating the efficacy of a fast urban climate model for spatial planning of green and blue spaces for heat mitigation, Science of The Total Environment, Vol. 955, 2024, p. 176925. https://doi.org/10.1016/j.scitotenv.2024.176925
37. Kusaka H., Ikeda R., Sato T., Iizuka S., and Boku T., Development of a multi-scale meteorological large-eddy simulation model for urban thermal environmental studies: The City-LES model version 2.0, Journal of Advances in Modeling Earth Systems, Vol. 16, No. 10, 2024, p. e2024MS004367. https://doi.org/10.1029/2024MS004367
38. Wang C., Zhao Y., Li Q., Wang Z.H., and Fan J., Ultrafine-resolution urban climate modeling: Resolving processes across scales, Journal of Advances in Modeling Earth Systems, Vol. 17, No. 6, 2025, p. e2025MS005053. https://doi.org/10.1029/2025MS005053
39. Materia, S., García, L. P., van Straaten, C., O, S., Mamalakis, A., Cavicchia, L. and Donat, M., Artificial intelligence for climate prediction of extremes: State of the art, challenges, and future perspectives, Wiley Interdisciplinary Reviews: Climate Change, Vol. 15, No. 6, 2024, p. e914. https://doi.org/10.1002/wcc.914
40. Assaf G., Hu X., and Assaad R.H., Mining and modeling the direct and indirect causalities among factors affecting the Urban Heat Island severity using structural machine learned Bayesian networks, Urban Climate, Vol. 49, 2023, p. 101570. https://doi.org/10.1016/j.uclim.2023.101570
41. Rampal, N., Hobeichi, S., Gibson, P. B., Baño-Medina, J., Abramowitz, G., Beucler, T. and Gutiérrez, J. M. , Enhancing regional climate downscaling through advances in machine learning, Artificial Intelligence for the Earth Systems, Vol. 3, No. 2, 2024, p. 230066. https://doi.org/10.1175/AIES-D-23-0066.1
42. Costa D.G., Bittencourt J.C.N., Oliveira F., Peixoto J.P.J., and Jesus T.C., Achieving sustainable smart cities through geospatial data-driven approaches, Sustainability, Vol. 16, No. 2, 2024, p. 640. https://doi.org/10.3390/su16020640