AI-Based Framework for Urban Climate Downscaling: A Case Study of Sofia

Koki Nakamura [1,2], Lidia Lazarova Vitanova [1]*, Quang-Van Doan [3], Dessislava Petrova-Antonova [1]

[1] GATE Institute, Sofia University "St. Kliment Ohridski", Sofia, Bulgaria
{koki.nakamura, lidia.vitanova, dessislava.petrova}@gate-ai.eu
[2] Graduate School of Science and Technology, University of Tsukuba, Japan
{s2521133}@u.tsukuba.ac.jp
[3] Center for Computational Sciences, University of Tsukuba, Japan
{doan.van.gb}@u.tsukuba.ac.jp

Keywords: Downscaling, Temperature, Convolutional Neural Networks, Weather Research and Forecasting Model

Abstract

With the increasing availability of powerful computational resources and artificial intelligence (AI), a data-driven approach has been gaining attention in atmospheric science, particularly in downscaling/emulating research. For example, Convolutional Neural Networks (CNNs) are used to emulate fine-resolution climate data fields by physics-based models. Although CNN-based downscaling has shown good performance for air temperature across various climate scales, studies at the urban scale remain limited due to the scarcity of long-term high-resolution climate data that can represent local urban effects. This study explores whether CNNs can accurately downscale/emulate urban-scale temperature distributions, with a focus on the urban heat island (UHI), using 250 m resolution Weather Research and Forecasting (WRF) simulation data over Sofia, Bulgaria, as training data. Input features, including air temperature, wind components, surface solar radiation, and elevation, are used, and their impacts on targeted 250 m resolution air temperature are evaluated. The results show that, while the CNN-based approach shows promise in generating urban-scale temperature distributions, the selection/combination of input features influences downscaling performance. Warm bias exceeding +2.00 °C was almost eliminated by including all input features. The downscaled result emulated the broad spatial pattern. However, detailed spatial and temporal variability, such as UHI, was still not captured. These findings revealed that the input feature selection can influence CNN-based temperature downscaling, providing useful insight into future regional-to-microscale downscaling toward urban digital twins (UDTs).

1. Introduction

Atmospheric datasets from global climate models (GCMs) and reanalyses are typically provided at coarse spatial resolutions, limiting their ability to capture local-scale processes and urban climate variability. Downscaling addresses this limitation by translating coarse atmospheric fields into finer-scale representations suitable for detailed climatological and meteorological analyses.

Downscaling approaches are generally divided into dynamical downscaling and statistical downscaling. Dynamical downscaling commonly employs regional climate models (RCMs), which solve the governing equations of the atmosphere over a limited domain using boundary and/or initial conditions derived from GCM outputs (Visconti, 2021). While physically consistent, dynamical downscaling is computationally expensive, sensitive to boundary conditions, and prone to systematic bias in long integrations. In contrast, statistical downscaling derives fine-scale information from statistical relationships between large -scale predictors and observations (Baño-Medina et al., 2020).

Recent advances in artificial intelligence (AI) and accessible GPU-onboard computational resources have accelerated the development of data-driven downscaling methods. Among these, convolutional neural networks (CNNs) can effectively learn complex spatial patterns from data, both global and regional cases, while operating at substantially lower computational cost than traditional dynamical approaches (Baño-Medina et al., 2020). CNN-based statistical downscaling has shown strong performance in temperature estimation. However, most studies focus on global-to-regional scales (Damiani et al., 2024), and the applicability of these findings to microscale environments remains uncertain (Yasuda et al., 2022). This limitation became critical as the demand for high-resolution climate information continues to grow, providing a valuable input for the development of urban digital twins (UDTs).

Sofia is no exception to this growing demand. As in many rapidly developing urban areas, detailed spatial information on near-surface air temperature is essential for understanding urban heat patterns and supporting climate-resilient planning. UDTs developed for Sofia rely on high-resolution meteorological datasets to represent local environmental conditions (Vitanova et al., 2025). Previous studies have generated these datasets using the Weather Research and Forecasting (WRF) model at a spatial resolution of 250 m to provide city-scale meteorological fields (L. L. Vitanova et al., 2026), reconstruct urban temperature distributions, and optimise the spatial deployment of meteorological sensor networks (Lidia L. Vitanova et al., 2026). However, this resolution remains insufficient and comprehensive citywide temperature fields that resolve fine-scale urban thermal variability are still lacking.

To directly address these gaps, this study proposes an AI-based framework based on Super-Resolution Convolutional Neural Network (SRCNN) (Dong et al., 2014; Onishi et al., 2019) to downscale/emulate air temperature fields from 0.1-degree resolution ERA5-Land to 250-m resolution data, which is generated by WRF for a case study of Sofia, Bulgaria.

Hence, the following research questions (RQ) are tackled:
RQ1) How accurately can the proposed SRCNN-based downscaling framework reproduce city-wide high-resolution air temperature patterns compared with WRF simulations to assist UDT?

* Corresponding author

https://doi.org/10.5194/isprs-archives-L-4-W2-2026-139-2026

RQ2) Which auxiliary input feature improves the performance of CNN-based downscaling in data-scarce urban environments?

The paper is structured as follows. Section 2 outlines the methodology used. The experimental results are described in Section 3. A discussion and conclusion are given in Sections 4 and 5, respectively

2. Methodology

The study workflow is illustrated in Figure 1 and includes data collection and preparation, data preprocessing, CNN training phase, and downscaling prediction.

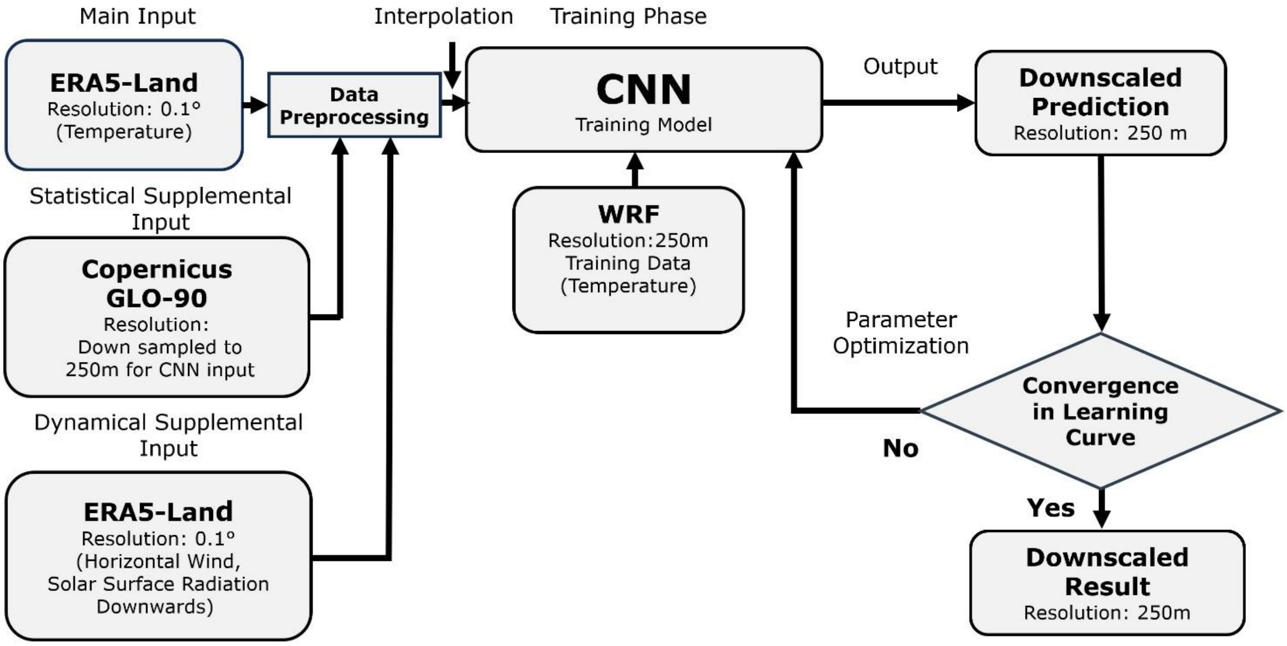

Figure 1. Study workflow.

2.1 Development of the SRCNN-based urban climate downscaling framework

The downscaling model used in this study is based on the SRCNN-type architecture originally developed by (Dong et al., 2014) for image super-resolution and later adapted to urban climate downscaling by (Onishi et al., 2019). While the previous studies demonstrated the capability of CNNs to generate high-resolution air temperature distributions from coarse-resolution inputs, the present study investigates whether an SRCNN-based approach can emulate high-resolution air temperature fields via ERA5-Land-to-WRF downscaling under limited training data.

2.2 Study Area

Sofia is the capital of Bulgaria and the country's largest city. The city is located in the Sofia Valley with an altitude of around 500 meters (Sofia Municipality, 2024). The study area is shown in Figure 2.

2.3 Data Collection

The study utilises historical reanalysis data for temperature, wind velocity, and downward surface solar radiation as dynamic input features, and elevation data as static input features for the CNN model (Figure 3). High-resolution training data for the Sofia region is produced by the WRF model. More information for each type of data is provided below.

Reanalysis Data: ERA5-Land is a reanalysis dataset provided by European Centre for Medium Range Weather Forecasts (ECMWF) (Muñoz-Sabater et al., 2021). The data supports past atmospheric conditions over the entire world domain since 1950 at hourly time steps across all land areas, with a spatial resolution of 0.1 degrees by 0.1 degrees for a variety of climate variables.

As described and shown in Figure 3, this study uses four ERA5-Land dynamic input features: 2m air temperature (t2m), 10m horizontal wind components (u10 and v10) and downward surface solar radiation (ssrd).

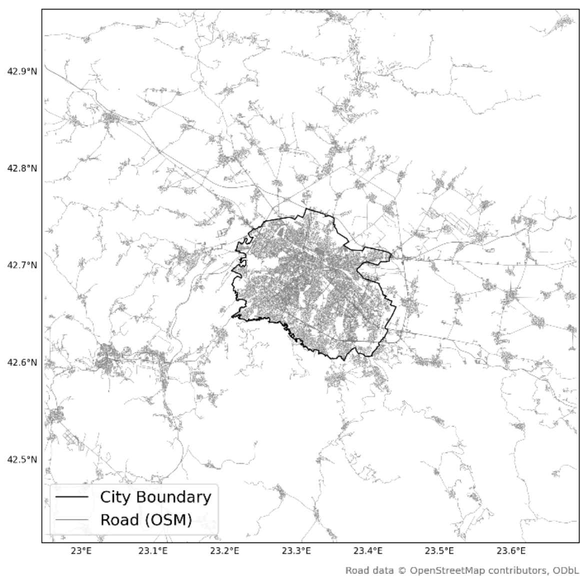

Figure 2. Study area. The City Boundary data is obtained by (Sofia Municipality, 2009), and Road data is obtained by ("OpenStreetMap," 2026)

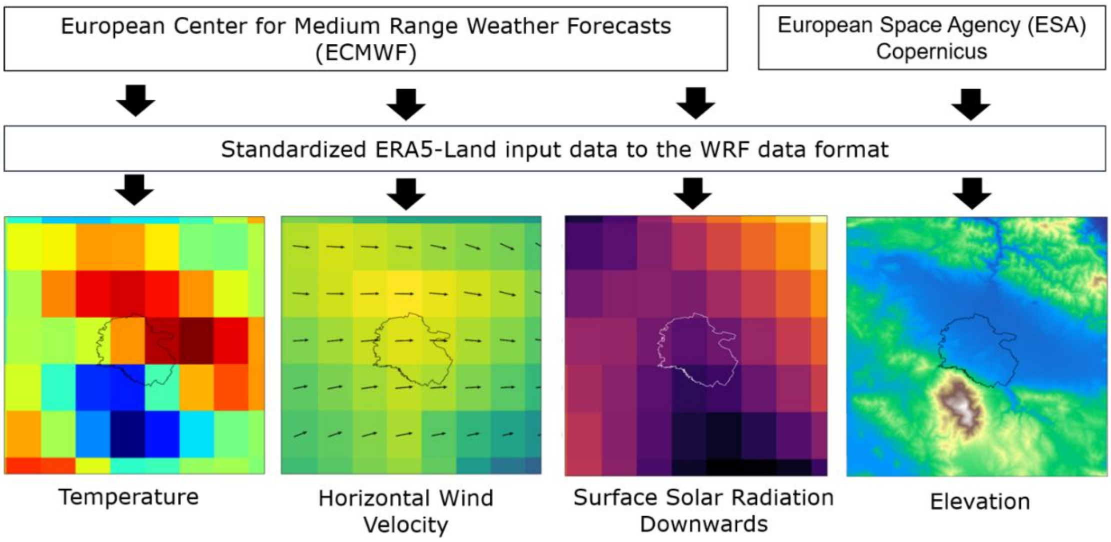

Figure 3. Data collection and data preprocessing.

Digital Elevation Model (DEM) data: The data is provided by the European Space Agency (ESA) and Copernicus provides digital elevation (DEM) data (European Space Agency and Airbus Defence and Space GmbH, 2022). DEM data is used as a static model input feature in this study. Since terrain largely contributes to temperature variation across climate scales, the terrain effect is not negligible. In this study, pre-processed Copernicus GLO-90 data with 90m spatial resolution is used. Original data has a coordinate system in WGS84 (EPSG:4326).

Output data from the WRF model: The study utilises WRF version 4.6 (Skamarock et al., 2019) outputs as training data. The NCEP FNL (Final) Operational Global Analysis dataset drove the model (National Centers for Environmental Prediction, 2000). For the purpose of this study, the fourth domain (0.25 x 0.25 km spatial resolution) is used as the ground truth of the mean 2m temperature distribution during the model evaluation period (Figure 4). The data availability is from July 14, 2024, to July 22, 2024, in EEST, and contains 175 hourly data points. The first 4 hours of the time step are omitted in this study because they fall within the WRF spin-up period. More information is provided in Section 2.7.

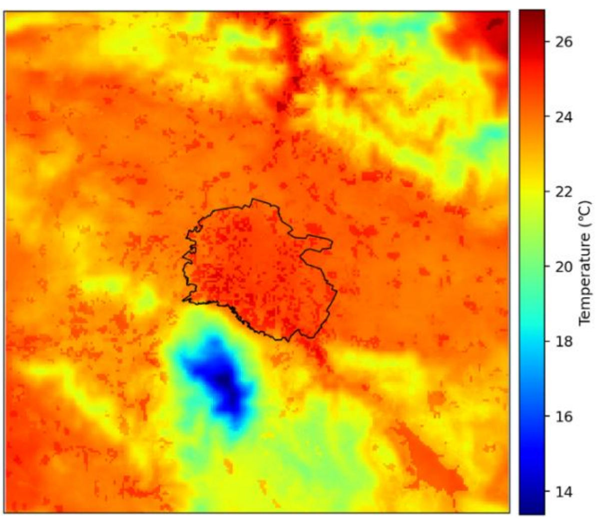

Figure 4. Simulated mean 2m temperature (°C) of the WRF model used as ground truth data.

2.4 Data Preprocessing

ERA5-Land data and elevation data are pre-processed to standardise the data format for training. In this preprocessing, the coordinate system and projection for both ERA5-Land and Copernicus GLO-90 are unified with WRF data (Lambert Conformal Conic Projection, WRF sphere; 6370km) and set to the same central coordinates as the WRF data (42.69 N, 23.32 E).

In addition, for CNN-based downscaling, static input features are preferred to match the training data spatial resolution (250m), following related studies (Yasuda et al., 2022). The DEM resolution is down sampled from the original 90m to 250m using the box-mean method to form an input for the CNN. The ERA5-Land and DEM data domain is clipped to 42.415 N-42.964 N and 22.945 E-23.695 E to match the WRF data.

Before being fed into the CNN, the dynamic input features were interpolated from the original ERA5-Land grid (0.1 degrees by 0.1degrees) to a 250 m WRF grid using either bilinear or bicubic interpolation. This preprocessing required by the SRCNN-type architecture adopted in this study. The network contains no up-sampling layers, and its convolutional layers preserve the input's spatial dimensions. The low-resolution predictors must be supplied on the same grid as high-resolution target field (Dong et al., 2014). The interpolated fields thus serve only as a smooth first guess from which the CNN learns to reconstruct the fine-scale spatial structure.

2.5 Model Architecture

Table 1 shows the CNN architecture used. Based on (Onishi et al., 2019) and (Yasuda et al., 2022), the original SRCNN architecture was modified to incorporate not only air temperature but also the additional predictors, including low-resolution horizontal wind field and surface solar radiation downwards, and terrain.

Layer	Operation	Kernel Size	Filters	Activation Function
Conv 1	Conv2D (pad4)	9×9	64	ReLU
Conv 2	Conv2D (pad0)	1×1	32	ReLU
Conv 3	Conv2D (pad2)	5×5	1	-

Table 1. CNN structure.

Hyperparameters: Since the aim of this study is to compare input feature combinations rather than to optimise the network itself, no exhaustive hyperparameter search was conducted. Instead, the hyperparameters were chosen to values widely used in CNN-based downscaling studies, following the standard SRCNN setting (Dong et al., 2014), which has also been adopted in urban scale downscaling applications (Onishi et al., 2019). Only the number of training epochs was determined empirically by monitoring the validation loss. The individual settings are described below.

Convolutional Layers: Convolutional layers extract spatial features from the input field using learnable filters. The number of convolutional layers determines whether the model is shallow or deep. The three-layer convolutional model is used in this study, as shown in Table 1. Padding is used to preserve the spatial size of feature maps after convolution.

Kernel Size: Local features are extracted by sliding small filters, called kernels, over the input data. In the CNN used in this study, the first layer uses a 9 by 9 kernel to extract broad spatial information, while the second and third layers use 1 by 1 and 5 by 5 kernels, respectively (Table 1).

Filters: The number of filters indicates the number of feature maps produced by each convolutional layer (64 for the first layer and 32 for the second). In the SRCNN-based model, a large number of filters in the early layer enables the extraction of spatial patterns from the input temperature field. These extracted features are then integrated into subsequent layers to reconstruct the high-resolution temperature field. The number of filters in the final layer was set to 1 because the output consists of a single temperature channel (Table 1).

Activation Function: The activation function provides nonlinearity to the CNN, enabling it to learn complex mappings between the input and output fields. Rectified Linear Unit (ReLU) was used after the convolutional layers to extract nonlinear spatial features from the temperature and auxiliary variables (Table 1).

Batch Size: The number of data samples processed at a time during model training. CNN learns by repeatedly making predictions on input data and calculating the error between the predictions and the ground truth to update its parameters. Batch size is selected based on computational cost and memory capacity and varies by research.

In this study, full-batch training was adopted due to the limited amount of training data. In full-batch training, all training samples are processed at once with each parameter update, rather than dividing the training data into smaller mini-batches. Because the training set contains only 103 hourly samples, the entire data set fits into GPU memory at once, and full-batch training provides deterministic gradient updates unaffected by mini-batch sampling noise.

Epoch Number: An epoch is one complete training cycle over the entire training dataset. In this study, the model is trained for up to 30 epochs. The number of epochs was empirically set to 30 because the training process tended to converge around epoch 30, after which signs of overfitting were observed.

Optimiser: In this study, Adam (Kingma and Ba, 2014) is used as an optimiser. The Adam optimiser is widely used in CNN downscaling research to adaptively control parameter updates based on past gradients, enabling fast and stable convergence. Adam was also employed in (Onishi et al., 2019).

Learning Rate: The learning rate controls the size of parameter updates during the model optimisation phase. The learning rate was set to 0.001 in this study, which is the default value recommended for the Adam optimiser (Kingma and Ba, 2014).

Loss Function: Loss function represents the difference between the predicted value and the actual value, and a smaller value indicates better model performance. A CNN model optimises its parameters to minimise loss. Since time-series air temperature data are continuous variables, mean squared error (MSE) is commonly used as a loss function for regression problems such as temperature downscaling. The study uses MSE as the loss function, following other downscaling studies (Onishi et al., 2019).

2.6 Interpolation Details and the Model Inputs

This section describes the interpolation details and the model inputs.

Interpolation: The dynamical variables from ERA5-Land were resized to the WRF domain using either bilinear or bicubic interpolation, and the effects of different interpolation methods on downscaling performance were investigated (Table 2). Bilinear interpolation was included as a baseline method due to its simplicity and widespread use for resampling gridded meteorological data (Dennis Shea, 2014). Bicubic interpolation follows the original CNN framework proposed by (Dong et al., 2014).

No	Input Features	Interpolation
1	t2m+dem+u10+v10+ssrd	Bilinear
2	t2m+dem+u10+v10+ssrd	Bicubic
3	t2m+u10+v10+ssrd	Bicubic
4	t2m+u10+v10+ssrd	Bilinear
5	t2m+dem+ssrd	Bicubic
6	t2m+dem+u10+v10	Bicubic
7	t2m+u10+v10	Bicubic
8	t2m+u10+v10	Bilinear
9	t2m+dem	Bilinear
10	t2m+dem	Bicubic
11	t2m+dem+ssrd	Bilinear
12	t2m+ssrd	Bilinear
13	t2m+ssrd	Bicubic
14	t2m	Bilinear
15	t2m	Bicubic

Table 2. Experimental conditions used in the study.

Input Features: Variables supplied to the CNN to estimate the high-resolution target field. In addition to 2m air temperature (t2m) to be downscaled, ERA5-Land dynamic variables (10 m horizontal wind components, u10 and v10), downward surface solar radiation (ssrd), and static (elevation) input features were supplied to the CNN as supplemental predictors. All input features were pre-processed onto the 250 m WRF data format configurations as described in Section 2.4, so that every input channel shares the spatial dimensions of the target field. To assess the contributions of each supplemental predictor, the model was trained with different combinations of input features as shown in Table 2.

2.7 Data Operation and Performance Evaluation

WRF data is split chronologically into training, validation, and test periods (Figure 5).

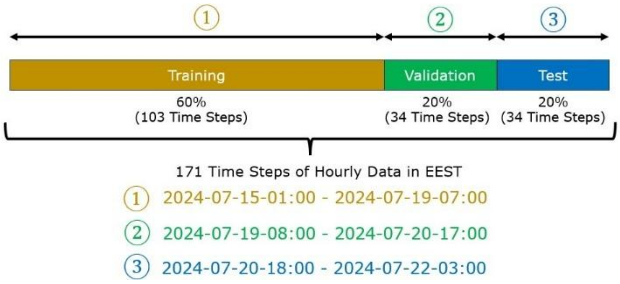

Figure 5. Data operation.

The first 60% is used for training, the next 20% for validation, and the last 20% for testing. While the original dataset contains 175 hourly data points, the first 4 hours are excluded as a spin-up period of the WRF model, and the remaining 171 hourly data points are used. The CNN downscaling results are evaluated using only the last 20% of the 171 hourly data.

Downscaling performance is evaluated by comparing the downscaled temperature fields with the WRF ground truth over the test period, which corresponds to the last 20% of the dataset (34 hourly time steps from 2024-07-20 18:00 EEST to 2024-07-22 03:00 EEST, Figure 5). Neither the training nor the validation period is included in the evaluation to assess the model's generalisation on unseen data.

This study was conducted using a computer equipped with an NVIDIA GeForce RTX 4080.

3. Results

This section presents results from the ERA5-Land-to-WRF downscaling and evaluates the downscaled outputs by comparing them with the WRF ground truth data.

3.1 Comparison of Experimental Conditions

In this study, root mean square error (RMSE), bias, Pearson correlation coefficient (r), and the coefficient of determination (R^2) are used as evaluation metrics. Downscaling performance is evaluated using all grids within the target domain and all test time steps. These scores are used to rank the experimental conditions, which consist of different combinations of input features and interpolation methods, to identify the condition that best emulates the WRF temperature field.

Table 3 summarises the scores for all 15 experimental conditions, ranked by RMSE. RMSE was adopted as the primary ranking metric because it quantifies the overall magnitude of prediction error in physical units of temperature. RMSE is directly consistent with the mean squared error loss function minimised during training. The condition of all input features (t2m, DEM, u10, v10 and ssrd) with bilinear interpolation achieved the best performance, with an RMSE of 2.31°C, a bias of -0.09°C, r of 0.86, and R^2 of 0.69. The same input combination with bicubic interpolation yielded nearly identical RMSE. The model trained with temperature alone yielded the largest errors and a bias exceeding 2.00 °C. The RMSE generally decreased as features were added, while the warm bias was largely removed only when all features were combined. The difference in scores between bilinear and bicubic interpolation was small under the input feature condition.

Rank	Input Features	RMSE (°C)	Bias (°C)	r	R^2
1	t2m+dem+u10+v10+ssrd	2.31	-0.09	0.86	0.69
2	t2m+dem+u10+v10+ssrd	2.33	-0.09	0.85	0.69
3	t2m+u10+v10+ssrd	2.38	1.36	0.89	0.67
4	t2m+u10+v10+ssrd	2.42	1.39	0.88	0.66
5	t2m+dem+ssrd	2.69	1.80	0.88	0.58
6	t2m+dem+u10+v10	2.74	1.82	0.87	0.57
7	t2m+u10+v10	2.74	0.73	0.77	0.57
8	t2m+u10+v10	2.79	1.00	0.78	0.55
9	t2m+dem	2.80	1.87	0.86	0.54
10	t2m+dem	2.87	1.93	0.86	0.52
11	t2m+dem+ssrd	2.90	2.03	0.87	0.51
12	t2m+ssrd	3.07	2.16	0.85	0.45
13	t2m+ssrd	3.07	2.19	0.86	0.45
14	t2m	3.07	2.15	0.86	0.45
15	t2m	3.13	2.22	0.86	0.43

Table 3. Downscaling performance of the 15 experimental conditions over the test period, ranked by RMSE. The interpolation method corresponding to each experimental setting is shown in Table 2.

3.2 Spatial Distribution of the Best-Performing Condition

Among the 15 experimental conditions, the best performance was obtained when all input features were combined with bilinear interpolation (Rank 1 in Table 3). This subsection examines the spatial characteristics of the condition over the test period. Figure 6 shows the test period of the downscaled mean simulated temperature under the best experimental condition. While the WRF model results clearly simulate a warm Sofia basin and a cooler area over the high-elevation terrain of Mountain Vitosha, the best downscaled result can roughly emulate the overall temperature distribution.

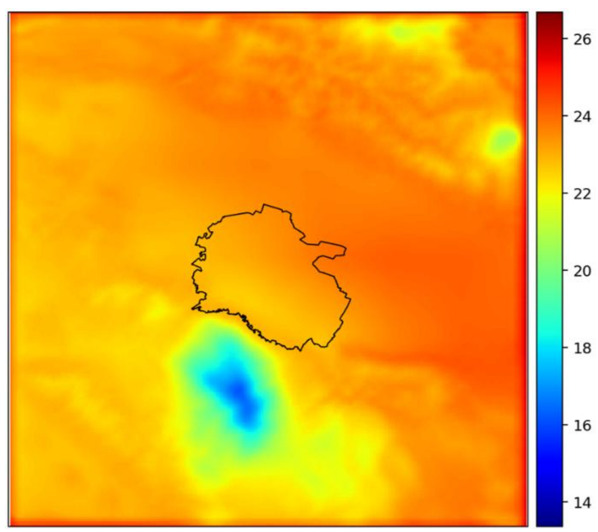

Figure 6. Mean 2m temperature (°C) of the downscaled result.

Figure 7 presents the spatial distribution of the mean bias, computed as the difference between the downscaled result and the ground truth. The results show that the simulated temperatures were underestimated in the urban area of Sofia, whereas they were overestimated in the surrounding mountainous area.

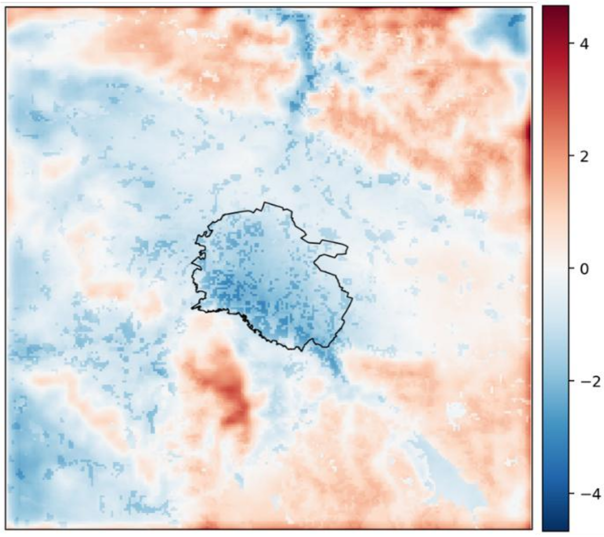

Figure 7. Differences between the downscaled result and ground truth (°C).

Note that the downscaled temperature field showed a general tendency toward underestimation across the spatial domain.

3.3 Taylor Diagram and Bias-RMSE

Figure 8 shows the Taylor diagram and the BIAS-RMSE plot from the test period at each grid in the study domain. In the Taylor diagram, most downscaled results had correlation coefficients between 0.8 and 0.95, and their normalised standard deviations were below 1.0. In the bias-RMSE scatter plot, RMSE mostly ranged from 1.00°C to 3.00°C, and the bias spread from approximately -3.00 to +3.00°C on both sides of zero.

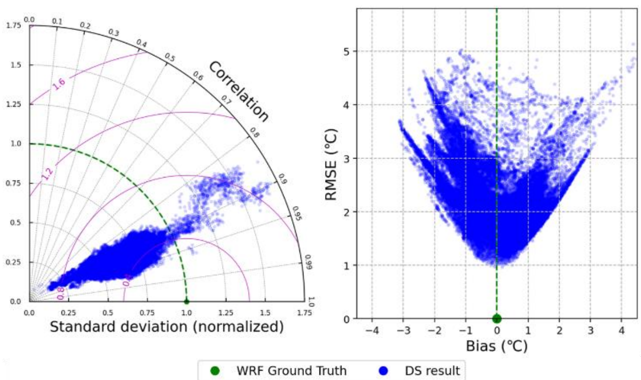

Figure 8. Taylor diagram and bias-RMSE for the best case (see Table 3, rank 1).

3.4 Training Curve

Figure 9 shows the training curve of the MSE loss over epochs for the best experimental condition in terms of RMSE (t2m+dem+u10+v10+ssrd with bilinear interpolation, ranked number 1 in Table 3). Both the training and validation losses decreased rapidly during the first several epochs and converged to similar values by epoch 30. The validation loss remained slightly lower than the training loss throughout most of the training.

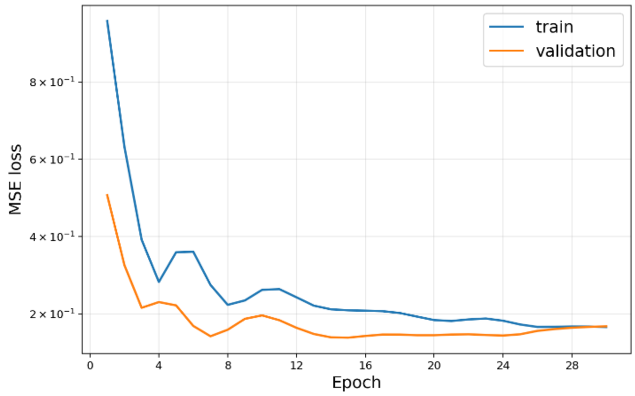

Figure 9. Training curve.

4. Discussion

This section primarily discusses the effects of input features, interpolation and the underestimation of temperature variability.

4.1 Effect of Input Features and Interpolation

The ranking in Table 3 indicates that the choice of input features had a substantially larger impact on performance than the choice of interpolation methods. Adding wind components, solar radiation, and elevation largely reduced both the RMSE and the warm bias compared to the temperature-only experiment. Elevation constrains the spatial pattern of near-surface temperature through the lapse rate, solar radiation modulates the diurnal heating cycle, and wind components convey information about advection and ventilation.

In particular, the warm bias exceeding +2.00 °C in the temperature-only case was almost eliminated when all supplemental features were included. It suggests that the supplemental features allowed the model to correct the discrepancy between the coarse ERA5-Land temperature and the WRF temperature field rather than merely sharpening the spatial pattern. The negligible difference between bilinear and bicubic interpolation is consistent with the original SRCNN, in which the CNN refines an interpolated input. The network's refinement capacity appears sufficient to absorb the small difference between the two interpolation methods.

4.2 Underestimation of Variability

The Taylor diagram in Figure 8 shows high correlation across most grids in the best-performance case for Rank 1 (Table 3). However, a normalised standard deviation below 1.0 indicates that the overall temperature variation was underestimated. The limited amount of training data for only 103 hourly data points within a single summer week likely reinforced this tendency, as the model had little opportunity to learn the single summer week data trend.

5. Main Findings and Future Work

This section discusses the main findings, limitations and future work.

5.1 Main Findings

This study investigated whether an SRCNN-based statistical downscaling model can emulate a high-resolution WRF air temperature field over Sofia, Bulgaria, from coarse regional-scale climate data under severely limited training data. The results show that the model can reproduce the broad spatial patterns of the WRF temperature field. However, the quantitative emulation of the fine-scale urban features requires further improvement.

(1) A combination of all input features that meteorologically affects temperature variation from coarse data (2m air temperature, wind components, surface solar radiation, and elevation) with bilinear interpolation performed the best, achieving an RMSE of 2.31°C, bias of -0.09°C, r of 0.86 and R^2 of 0.69 against the WRF ground truth (RQ1);

(2) More supplemental input features than the interpolated methods dominated downscaling performance, and their inclusion removed the warm bias of the temperature (RQ2);

(3) The downscaled fields emulate broad spatial patterns but underestimate their spatial and temporal variability. In particular, the Urban Heat Island signal for Sofia was insufficiently reproduced, resulting in underestimation in urban areas, offset by rural overestimation (RQ1);

5.2 Limitations and Future Work

This study served as a proof of concept for using a CNN to simulate/emulate the urban climate in Sofia. The results shown are based on short-term training data and may lead to overfitting. In detail, the study period covers only 9 days in a single summer, and the chronological split leaves only 34 hourly steps for testing. Therefore, results may not generalise to other seasons or synoptic conditions. Given the limitations above, future work is planned as follows:

(1) Extending the training data to longer periods and multiple seasons is expected to mitigate data scarcity that limits the downscaling performance;
(2) Incorporating urban surface descriptors such as high-resolution land use, building features, or satellite-derived indices such as NDVI as additional static input features may improve the downscaling performance of urban heat islands;

(3) Downscaling performance improved when meteorologically relevant supplemental input features were added to the target variable. The best result was obtained when all features were combined. The optimal number and combinations of inputs remain an open question. Future work will systematically evaluate feature combinations through sensitive experiments to identify the optimal set.

(4) The framework could be extended to urban temperature downscaling, using street-scale, super-high-resolution simulation data as training data to serve the development of the UDT of the Sofia use case.

Acknowledgements

We acknowledge the provided access to the Discoverer at Sofia Tech Park (Bulgaria). This research was supported by the GATE project, funded by the H2020 WIDESPREAD-2018-2020 TEAMING Phase 2 programme, under grant agreement no. 857155, and the programme “Research, Innovation and Digitalisation for Smart Transformation” 2021-2027 (PRIDST) under grant agreement no. BG16RFPR002-1.014-0010-C01.

We would like to express our appreciation to Professor Hiroyuki Kusaka, Mr. Ryota Karube, Mr. Evgeny Shrinyan, and Ms. Tereza Trendafilova for their cooperation and support during the initial stage of this research.

References

Baño-Medina, J., Manzanas, R., Gutiérrez, J.M., 2020. Configuration and intercomparison of deep learning neural models for statistical downscaling. Geosci. Model Dev. 13, 2109–2124. https://doi.org/10.5194/gmd-13-2109-2020

Damiani, A., Ishizaki, N.N., Sasaki, H., Feron, S., Cordero, R.R., 2024. Exploring super-resolution spatial downscaling of several meteorological variables and potential applications for photovoltaic power. Sci. Rep. 14, 13125. https://doi.org/10.1038/s41598-024-64055-y

Dennis Shea, 2014. The Climate Data Guide: Regridding Overview [WWW Document]. NCAR. URL Retrieved from https://climatedataguide.ucar.edu/climate tools/regridding-overview (accessed 7.9.26).

Dong, C., Loy, C.C., He, K., Tang, X., 2014b. Learning a Deep Convolutional Network for Image Super-Resolution. pp. 184 199. https://doi.org/10.1007/978-3-319-10593-2_13

European Space Agency, Airbus Defence and Space GmbH, 2022. Copernicus DEM - Global and European Digital Elevation Model [WWW Document]. Copernicus DEM. https://doi.org/10.5270/ESA-c5d3d65

Kingma, Diederik P, Ba, J., 2014. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980 1–15. https://doi.org/10.48550/arXiv.1412.6980

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

National Centers for Environmental Prediction, N.W.S.N.U.S.D. of C., 2000. NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. Boulder, Colorado. https://doi.org/10.5065/D6M043C6

Onishi, R., Sugiyama, D., Matsuda, K., 2019. Super-Resolution Simulation for Real-Time Prediction of Urban Micrometeorology. SOLA 15, 178–182. https://doi.org/10.2151/sola.2019-032

OpenStreetMap [WWW Document], 2026. URL https://www.openstreetmap.org/ (accessed 7.21.26).

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Liu, Z., Berner, J., Wang, W., Powers, J.G., Duda, M.G., Barker, D.M., Huang, X.-Y., 2019. A Description of the Advanced Research WRF Model Version 4.

Sofia Municipality, 2024. Geographical Characteristics [WWW Document]. Sofia Municipality. URL https://www.sofia.bg/web/sofia-municipality/geographical-feature (accessed 7.8.26).

Sofia Municipality, 2009. General Master Plan of Sofia Municipality [WWW Document]. URL https://sofiaplan.bg/Documents/OUP-Sofia_2009_withfonts.pdf (accessed 7.21.26).

Visconti, G., 2021. Downscaling, Regional Models and Impacts, in: Climate, Planetary and Evolutionary Sciences. Springer International Publishing, Cham, pp. 31–99. https://doi.org/10.1007/978-3-030-74713-8_2

Vitanova, L., Petrova-Antonova, D., Shirinyan, E., 2025. Urban digital twin for assessing and understanding urban Heat Island impacts. Urban Clim. 62, 102530. https://doi.org/10.1016/j.uclim.2025.102530

Vitanova, Lidia L., Peev, R., Petrova-Antonova, D., Trendafilova, T.K., Roman, D., 2026. Spatially optimised sensor networks for efficient urban temperature monitoring and prediction. Discover Cities 3, 124. https://doi.org/10.1007/s44327-026-00308-x

Vitanova, L. L., Petrova-Antonova, D., Shirinyan, E., Khanh, D.N., Trendafilova, T.K., Subasinghe, S., Khan, A., Doan, Q. V., 2026. Fusing urban informatics and weather modelling in digital twins for city-scale heat-health mapping. Computational Urban Science 6, 32. https://doi.org/10.1007/s43762-026-00268-3

Yasuda, Y., Onishi, R., Hirokawa, Y., Kolomenskiy, D., Sugiyama, D., 2022. Super-resolution of near-surface temperature utilizing physical quantities for real-time prediction of urban micrometeorology. Build. Environ. 209, 108597. https://doi.org/10.1016/j.buildenv.2021.108597