High-Resolution Urban Temperature Downscaling Using Machine Learning and ICON-2I Predictions

5th ECMWF-ESA Machine Learning Workshop

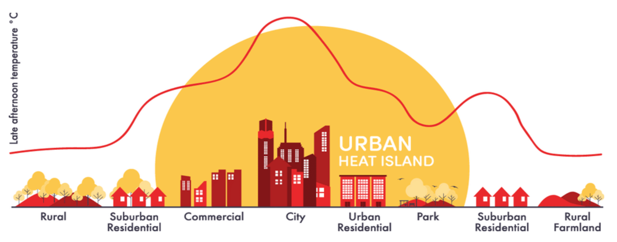
Tanguy Houget [1,2], Valeria Garbero [4], Marco Piras [1], Emmanuel Dellandrea [3], Pietro Salizzoni [1,2]

[1] Politecnico di Torino, Dipartimento di ingegneria per l'ambiente, il territorio e le infrastrutture (DIATI), Turin, Italy
[2] Ecole Centrale de Lyon, CNRS, Universite Claude Bernard Lyon 1, INSA Lyon, LMFA UMR5509, 69130, Ecully, France
[3] Ecole Centrale de Lyon, CNRS, Universite Claude Bernard Lyon 1, INSA Lyon, Université Lumière Lyon 2, LIRIS, UMR5205, 69130 Ecully, France
[4] Arpa Piemonte, Dipartimento Meteorologia, Clima e Qualità dell'Aria, Via Pio VII, 9, 10135 Torino, Italy

Contact :
tanguy.houget@polito.it

CONTEXT

a. The issue of extreme urban temperatures



The urban heat island (UHI) effect, observed in urban areas, is characterized by higher air temperatures compared to their rural surroundings.

The prediction of extreme temperatures and the development of mitigation strategies constitute a major public health challenge.

b. Limits of meteorological models

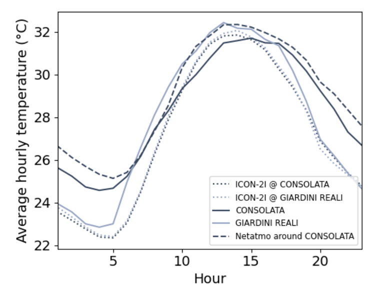


Figure 1: Comparison of the ensemble and period-averaged temperatures between the simulation and the measures at the location of CONSOLATA and Giardini Reali stations.

The ARPAE operational meteorological model ICON-2I fails to reproduce the observed nighttime temperature gradient between the park and the dense urban area due to its coarse resolution.

c. Objectives

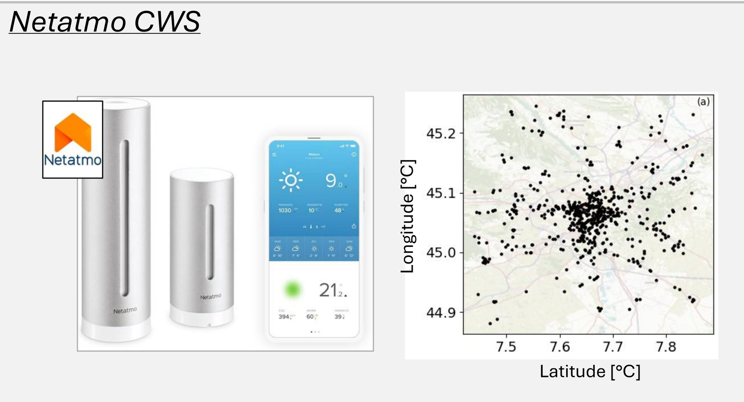
Develop a Machine Learning methodology to predict the fine-scale spatial heterogeneity of urban air temperature using Citizen Weather Station observations as training data and urban morphological parameters as predictive features.

Compare the performance of different machine learning architectures and feature sets used for model training.

DATA COLLECTION

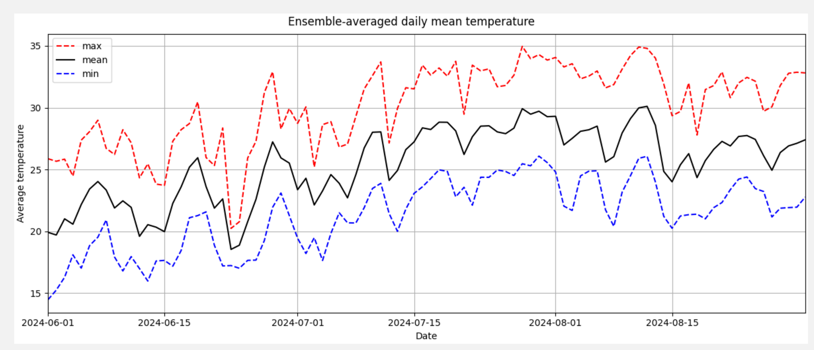
a. Temperature data

Netatmo CWS



Period considered

Heatwave of July 2024
21/07/2024 – 02/08/2024

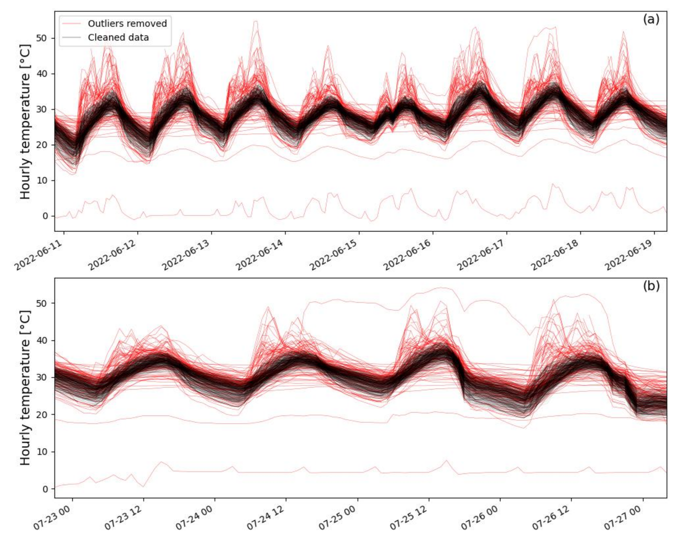


b. Predictors – Urban parameterisation

Notation	Description
abh	Building average height
alt	Ground elevation
build_sf	Fraction of built surface
imp_sf	Fraction of road surface
SVF	Sky View Factor
NDVI	Normalised Difference Vegetation Index

ICON-TU	Temperature, pressure, wind force and wind direction predicted by the ICON-TU numerical model

DATA CLEANING

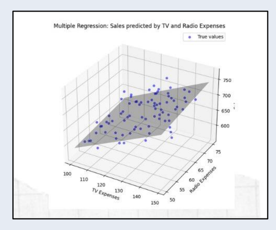


CWS temperature data are often affected by quality issues due to improper sensor installation sensors and non-compliance with WMO standards. To ensure data reliability, data cleaning is required to remove faulty stations. A quality control procedure based on Napoly et al. (2018) is therefore applied, flagging stations whose measurements deviate significantly from the overall Netatmo network trend.

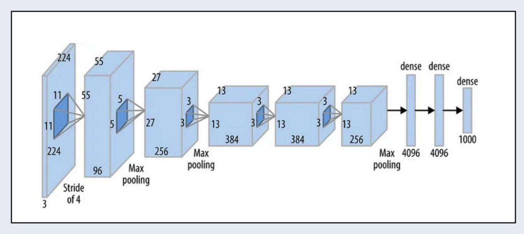
Figure 2: Time series of the temperature measured by the Netatmo sensors located in Turin in January 2022 (a) before cleaning (b) after cleaning

MACHINE LEARNING MODELS

Machine Learning models



Multi Linear Regression (MLR)



Convolutional Neural Network (CNN)

RESULTS

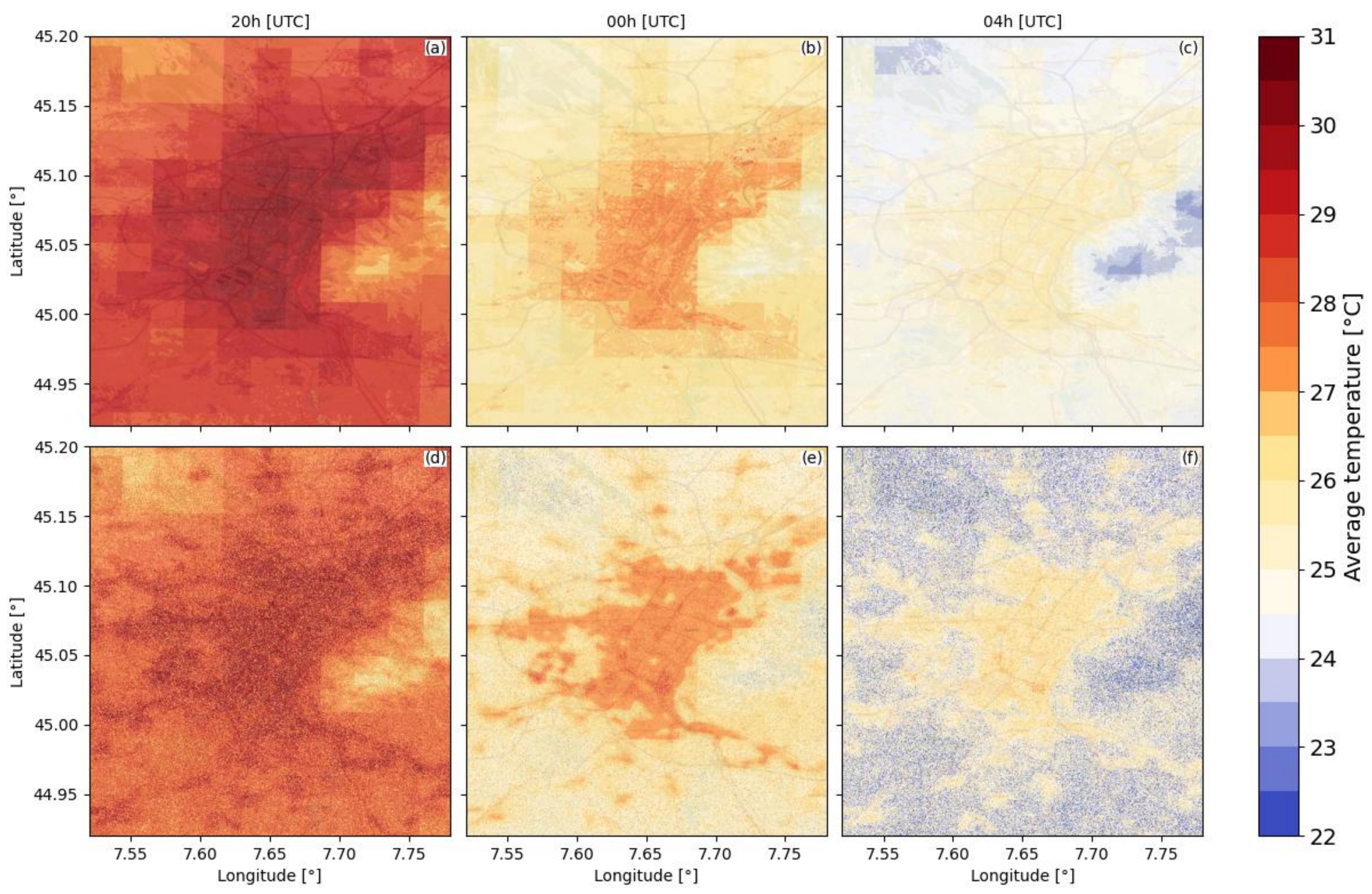


Figure 3: Maps of **period-averaged temperatures** generated using temperature from ICON-2I analysis as input feature: (a–c) results from the MLR model and (d–f) results from the CNN model. Panels correspond to (a, d) 20:00, (b, e) 00:00, and (c, f) 04:00.

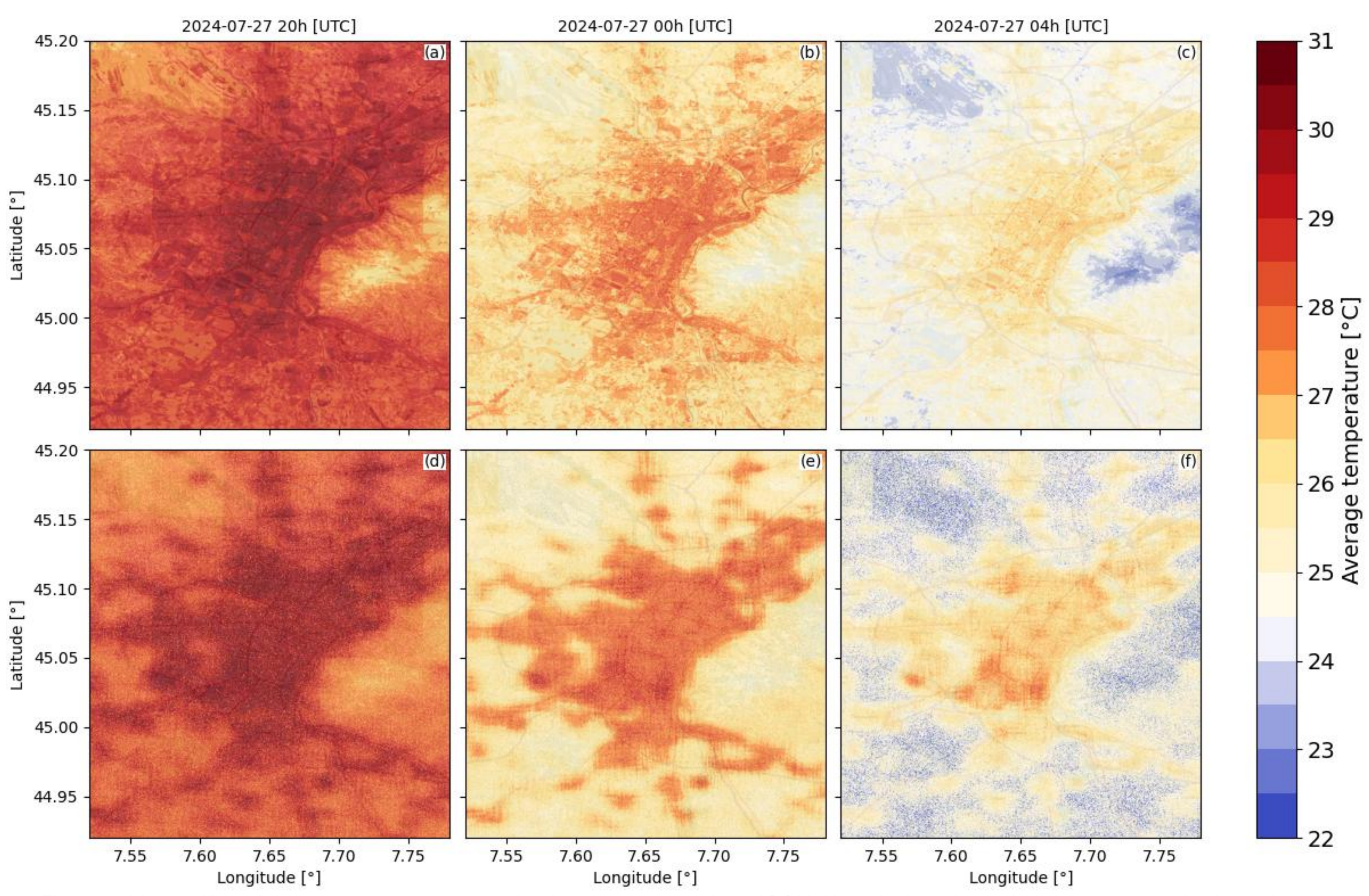


Figure 4: Maps of **hourly temperatures** generated using temperature from ICON-2I analysis as input feature: (a–c) results from the MLR model and (d–f) results from the CNN model. Panels correspond to (a, d) 20:00, (b, e) 00:00, and (c, f) 04:00.

Figure 3-4

All models, predicting period-average temperatures or hourly temperatures, MLR or CNN, reproduce the UHI effect, predicting higher temperatures in the urban center, reaching **up to +5°C** in some areas. The MLR generated maps, based on local predictors, exhibit maps with strong thermal contrasts. Conversely, the CNN incorporates spatial context, generating maps without sharp contrasts but still heterogeneous, providing a physically more realistic representation of urban temperatures.

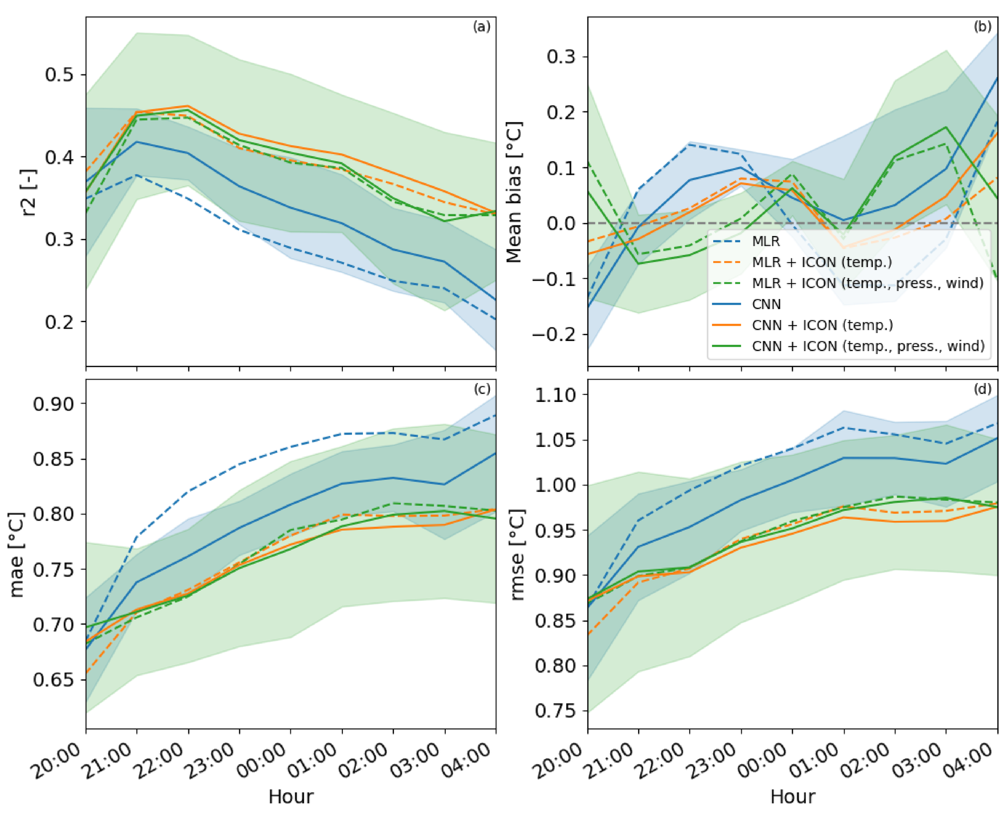


Figure 5: Performance of the MLR and CNN models at modelling period-averaged temperatures (a) R2 (b) Mean Bias (c) MAE (d) RMSE.

Figure 5

The performance of both MLR and CNN models in predicting period-averaged temperatures remains moderate ($R^2 < 0.5$), although the best configurations achieve RMSE values below 1 °C, in agreement with results reported in similar CWS-based studies (Venter et al., 2020; Vulova et al., 2020; Zumwald et al., 2021).

Incorporating ICON analysis data as predictors overall improves model performance in reproducing period-averaged temperatures during the 2024 heatwave period.

When ICON data are included, the CNN does not outperform the MLR model. In the absence of ICON features, the CNN shows slight improvements, suggesting a greater ability to capture complex spatial patterns using urban parameterization alone.

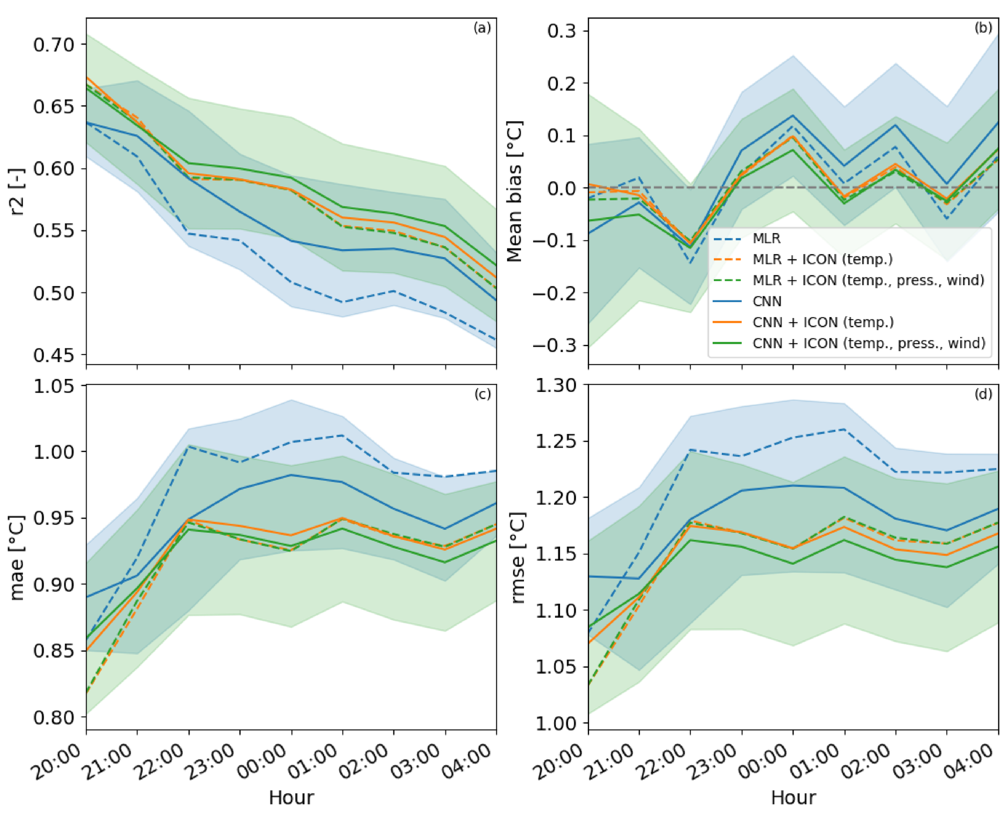


Figure 6: Performance of the MLR and CNN models at modelling hourly temperatures (a) R2 (b) Mean Bias (c) MAE (d) RMSE.

Figure 6

Predicting hourly temperatures generally reduces model performance, with **RMSE values around 1.15 °C** during most nighttime hours for the best-performing configurations.

As observed for period-averaged models, the inclusion of ICON features improves performance, while the CNN only provides added value when ICON predictors are absent; when ICON data are included, MLR and CNN models perform similarly.

Predicting hourly temperatures involves larger variability driven by daily temperature changes, leading to higher R^2 values compared to period-averaged temperature predictions. These results indicate that **the models effectively capture daily variability**; however, their accuracy, as reflected by RMSE values, is lower than for period-averaged predictions.

Discussion

Models trained to predict hourly temperatures perform less well than those trained on period-averaged temperatures. CNN models provide improved accuracy only when ICON analysis data are not included as predictors; when these data are incorporated, MLR and CNN models achieve similar performance. Overall, model precision remains moderate, potentially due to limited accuracy of temperature observations used for training or the omission of relevant features in the urban parameterization.

References :

Zumwald, M., Knüsel, B., Bresch, D.N., Knutti, R., 2021. Mapping urban temperature using crowd-sensing data and machine learning. Urban Climate 35, 100739. https://doi.org/10.1016/j.uclim.2020.100739

Napoly, A., Grassmann, T., Meier, F., Fenner, D., 2018. Development and Application of a Statistically-Based Quality Control for Crowdsourced Air Temperature Data. Front. Earth Sci. 6, 118. https://doi.org/10.3389/feart.2018.00118

Venter, Z.S., Brousse, O., Esau, I., Meier, F., 2020. Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. Remote Sensing of Environment 242, 111791. https://doi.org/10.1016/j.rse.2020.111791

Vulova, S., Meier, F., Fenner, D., Nouri, H., Kleinschmit, B., 2020. Summer Nights in Berlin, Germany: Modeling Air Temperature Spatially With Remote Sensing, Crowdsourced Weather Data, and Machine Learning. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 13, 5074–5087. https://doi.org/10.1109/jstars.2020.3019696