A Lightweight Deep Learning Approach to Spatiotemporal Heat Prediction

Euan Marney, Linus Ericsson | University of Glasgow

TLDR: Our U-Net CNN learns short-term temperature patterns using minimal data and computational resources.

Introduction

Extreme heat is an increasing public health risk, requiring accurate rate short-term temperature forecasts for early warning systems. Operational forecasting models provide high accuracy but rely on large-scale computational infrastructure.

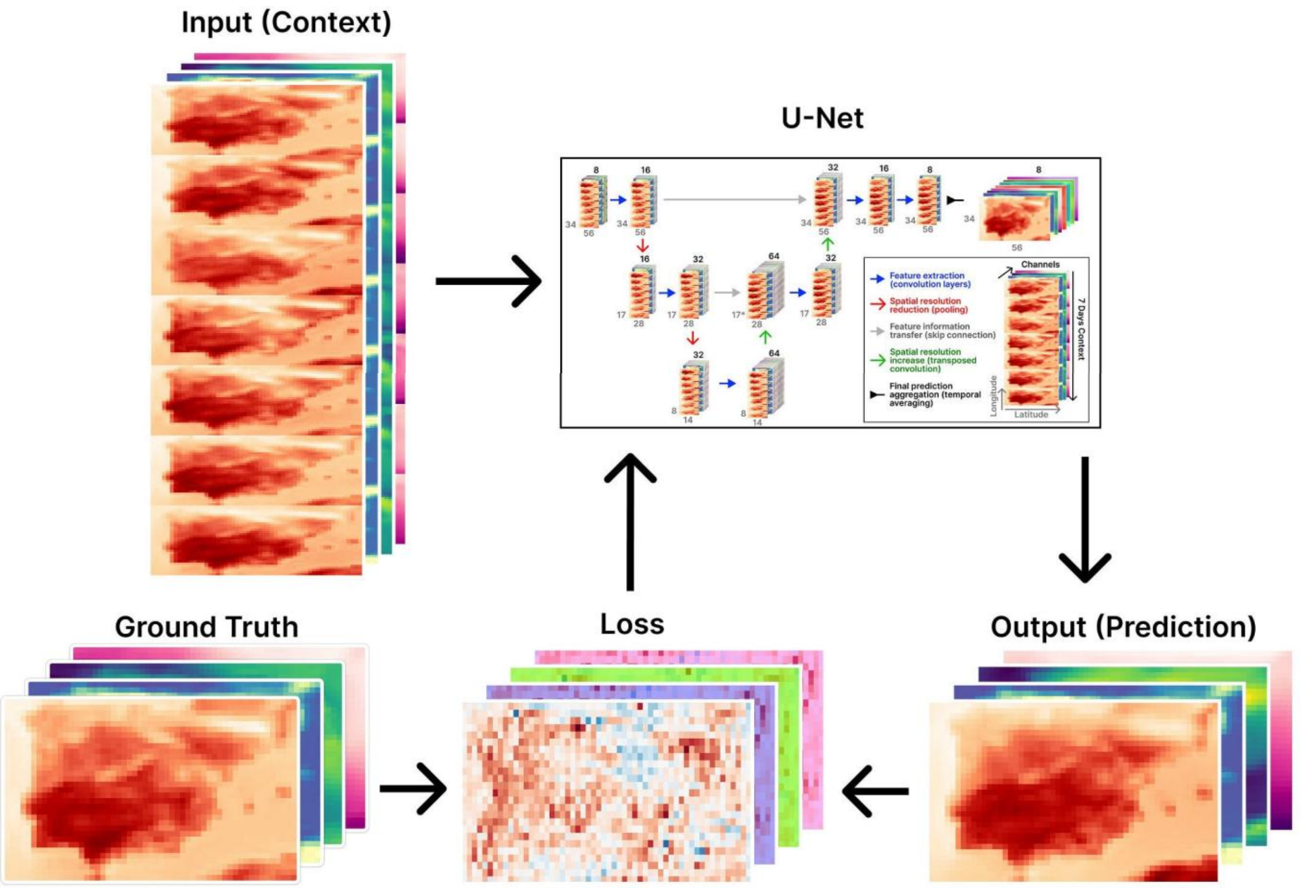
Can a lightweight deep learning approach provide useful short-range forecasts using minimal data and compute?

A 3D U-Net is trained on ERA5 reanalysis data to predict daily maximum temperature over the Iberian Peninsula.

U-Net Architecture

Models were trained using root mean squared error across all predicted variables.

The model processes inputs structured as latitude × longitude × time × channel, where each channel represents an atmospheric variable.



Experiments

RQ1: how does the length of the input context window influence model performance at different lead times?

RQ2: which input variables contribute most strongly to model performance?

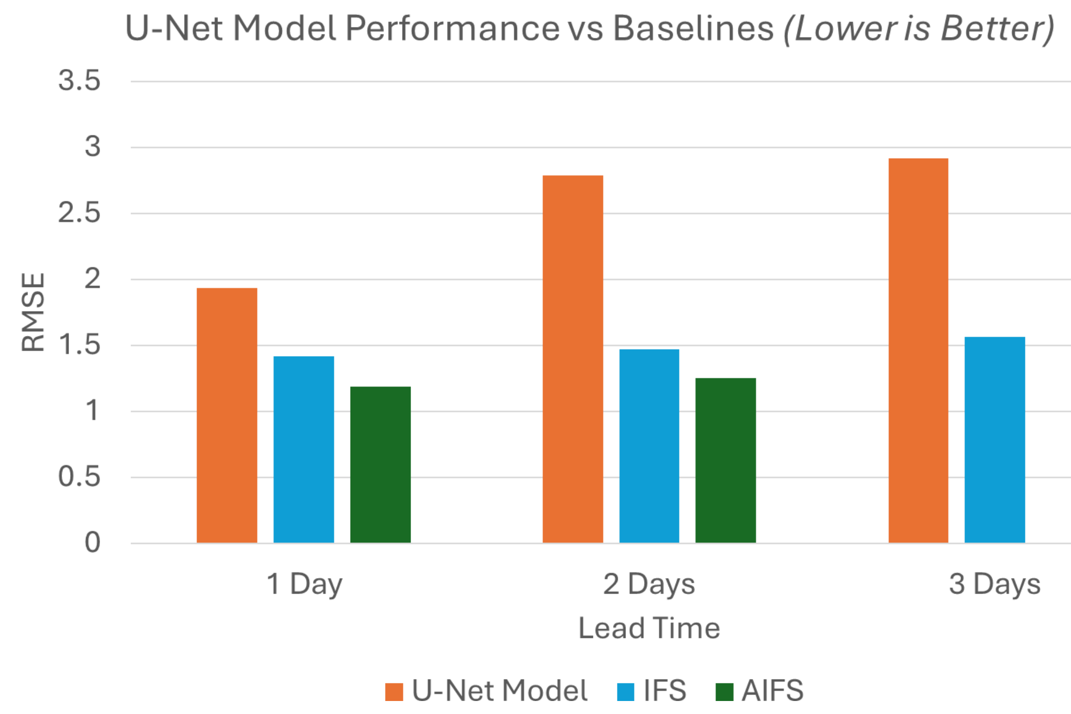
RQ3: how does the proposed U-Net-based approach compare to current operational forecast products?

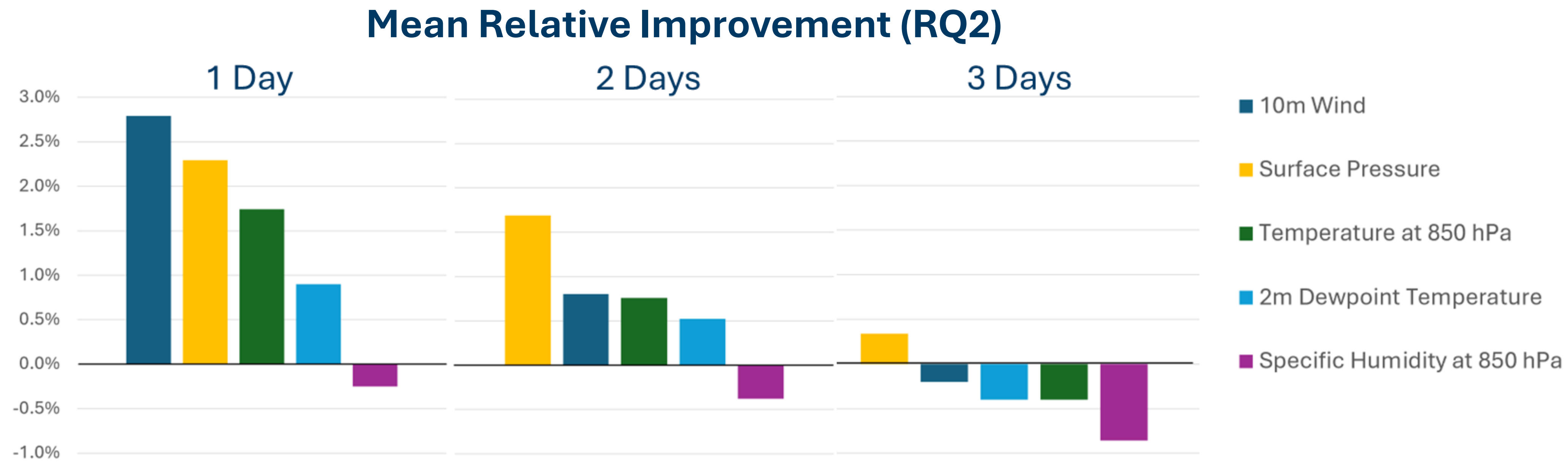
Results

RQ1: forecast accuracy declines rapidly with lead time, with strongest performance at D+1 and D+2, and minimal gains from extending context beyond seven days.

RQ2: only a small subset of variables contribute meaningfully, with surface pressure and 10 m wind providing the most consistent improvements.

RQ3: when compared to ECMWF baselines, the model is less accurate but achieves useful short-term skill with ~330k parameters and <0.5 GB GPU memory at inference.





Limitations

The model is restricted to a small set of atmospheric variables due to ERA5 data access constraints, limiting representational capacity.

Performance is evaluated over domain-wide averages, which may obscure important regional variations in predictive skill.

Conclusions

The model achieves meaningful skill at short lead times with orders-of-magnitude lower cost.

This presents a practical alternative for short-range forecasting in resource-constrained environments, such as deployment on edge devices.