

Accelerating Aerosol Hygroscopicity Prediction with Deep Learning and HPC Deployment

Radek Halfar 1), Billy Braithwaite 2), Juha Tonttila 2), Antti Vartiainen 2), Arup Sarkar 3), Harri Kokkola 3,4), Anton Laakso 3)

1. IT4Innovations, VSB - Technical University of Ostrava, Czech republic
2. CSC - IT Center for Science Ltd., Espoo, Finland
3. FMI - Finnish Meteorological Institute, Helsinki, Finland
4. University of Eastern Finland



Background & Motivation

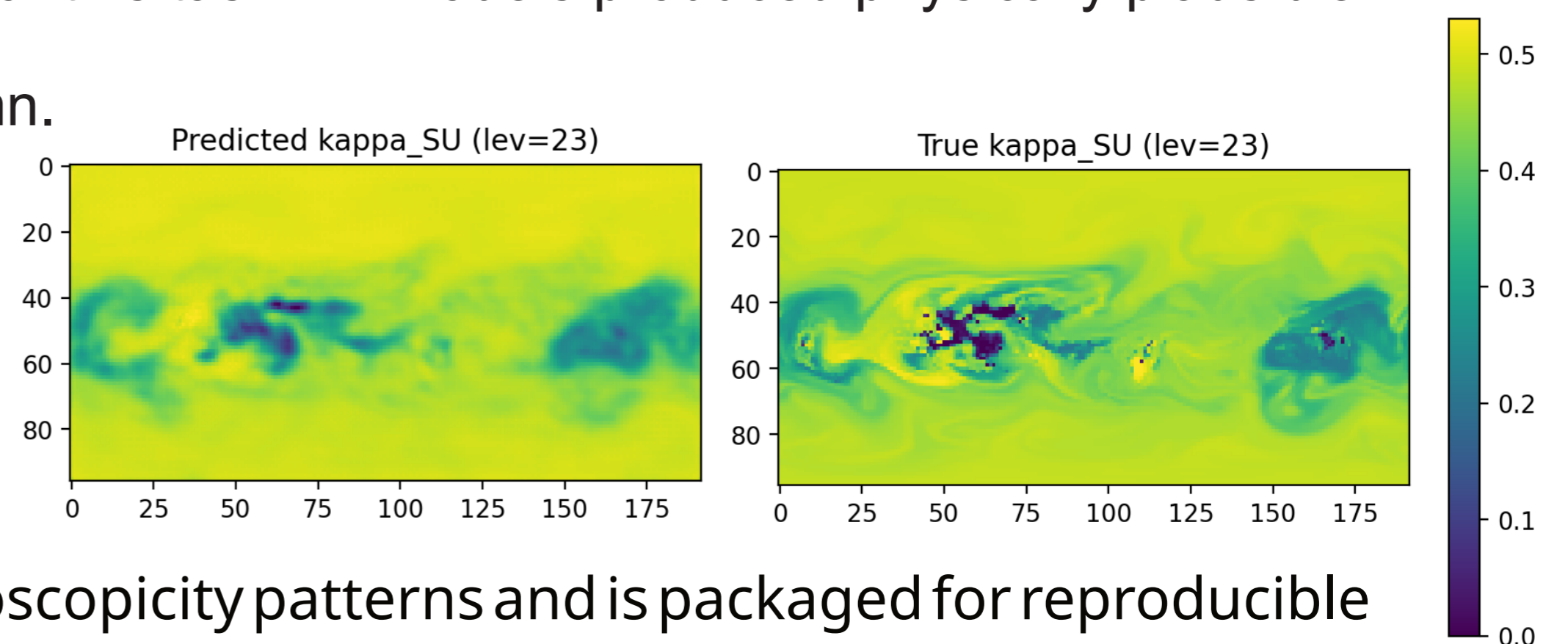
Aerosol hygroscopicity strongly influences clouds, radiation, and air quality, but physics-based simulations are computationally expensive. This project develops a machine-learning emulator to predict K_{SU} and K_{CA} directly from atmospheric composition fields, as a step toward an AI-based aerosol process component for integration into high resolution global climate models within the DestinE Climate Digital Twin initiative.

Data & Methodology

The emulator is trained on NetCDF spatiotemporal data using 12 input variables covering aerosol concentrations, dynamical variables, surface temperature, and emissions for four aerosol species. Inputs are subject to per-channel z-score normalisation. Three deep learning architectures are evaluated: a 3D U-Net with temporal attention fusion (UNet3D-Temporal), a Convolutional LSTM encoder-decoder (ConvLSTM), and a temporal Vision Transformer with 3D patch embeddings and learnable spatial-temporal positional encodings (ViT-Temporal).

Results

All three architectures achieved consistent convergence. ConvLSTM reached the lowest validation error, followed closely by ViT-Temporal, while UNet3D-Temporal suggests that sequence-aware recurrent and attention-based models capture temporal dependencies more effectively for this task. All models produced physically plausible kappa field distributions across the full atmospheric column.



Conclusions & Outlook

The emulator reproduces realistic spatiotemporal hygroscopicity patterns and is packaged for reproducible HPC deployment with Singularity. The near term outlook is to mature and validate the standalone aerosol AI component through longer-sequence training, larger pretraining datasets, and robust testing, while supporting implementation of the new aerosol model in IFS atmospheric model.

