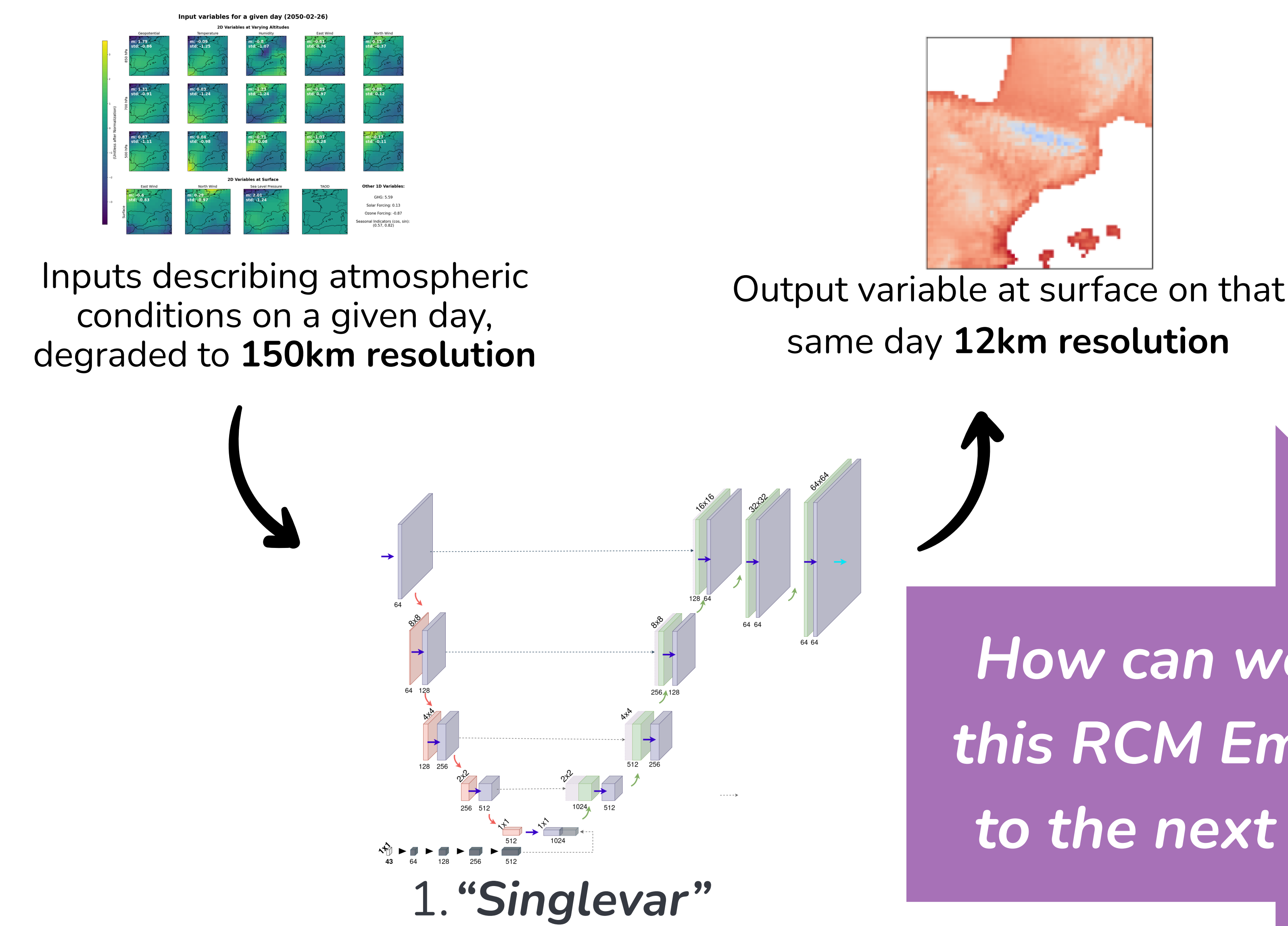


Background

Deep-learning based methods have been shown to successfully circumvent the computational cost of Regional Climate Models (RCM's) by learning how to downscale coarse resolution data into high resolution outputs

The work at present builds upon the singlevariate U-Net RCM Emulator proposed by Doury et. al (2023). In this framework, all data (input, training, testing) come from the same RCM



How can we take this RCM Emulator to the next level?

Next Steps in DL-Based RCM Emulators

Christina Carty¹, Antoine Doury¹, Julien Boé²
 contact: christina.carty@meteo.fr

1.CNRM, Météo-France, CNRS, Université de Toulouse, Toulouse, France, 2. CNRS/Cerfacs/IRD, CECI, Université de Toulouse, Toulouse, France



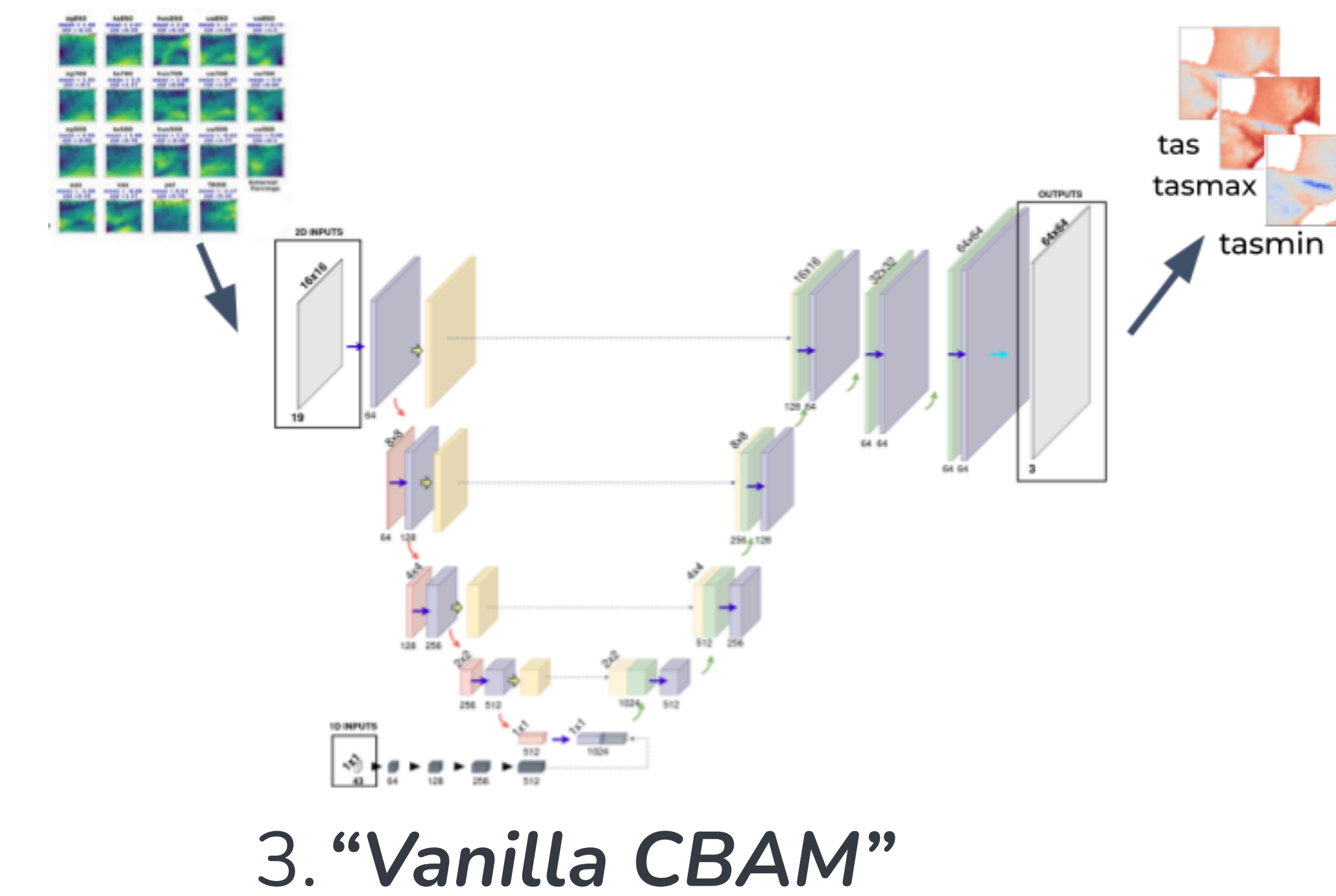
Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or [name of the granting authority]. Neither the European Union nor the granting authority can be held responsible for them.

Models

This work introduces climate model downscaling methods that are...

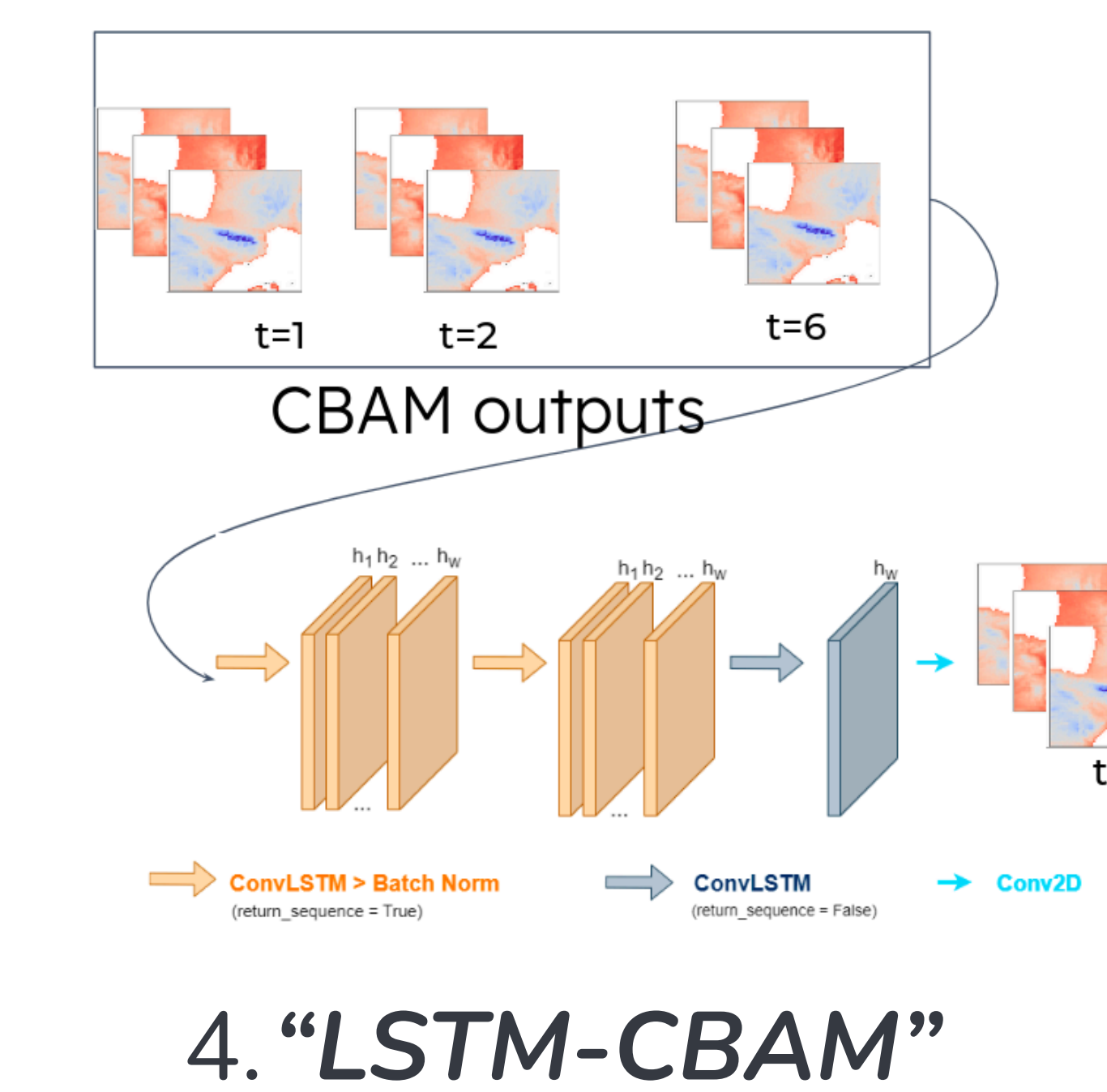
Multivariate

In addition to creating a U-Net emulator capable of outputting three variables at a time (2. "Multivar") we can also adapt the architecture by including Convolutional Block Attention Modules

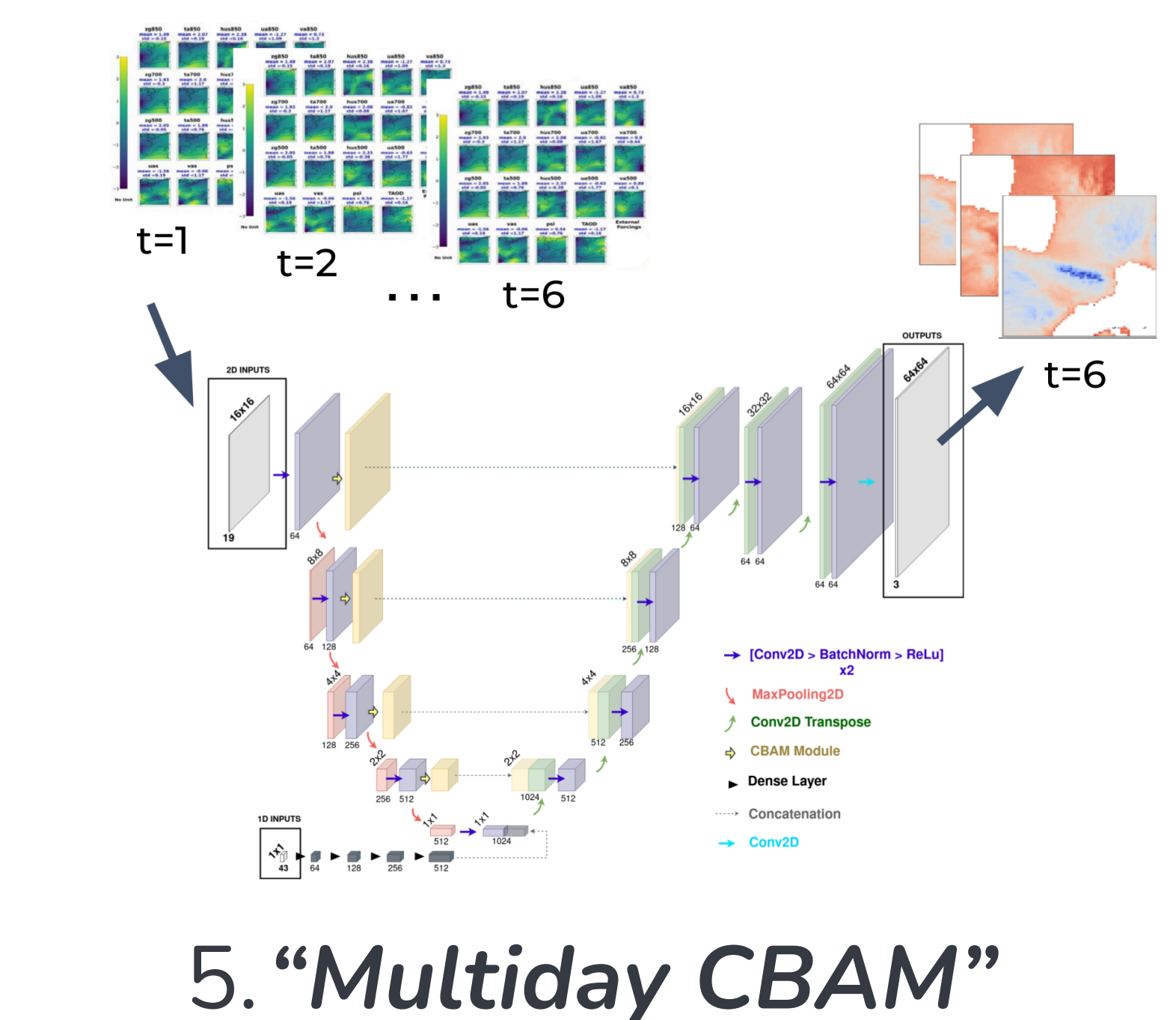


Temporal

One method of introducing temporal memory is by passing multiple days worth of CBAM outputs through a shallow model made up of ConvLSTM layers



Alternatively, we can take a Vanilla CBAM and simply give it multiple days worth of input data to predict a single day



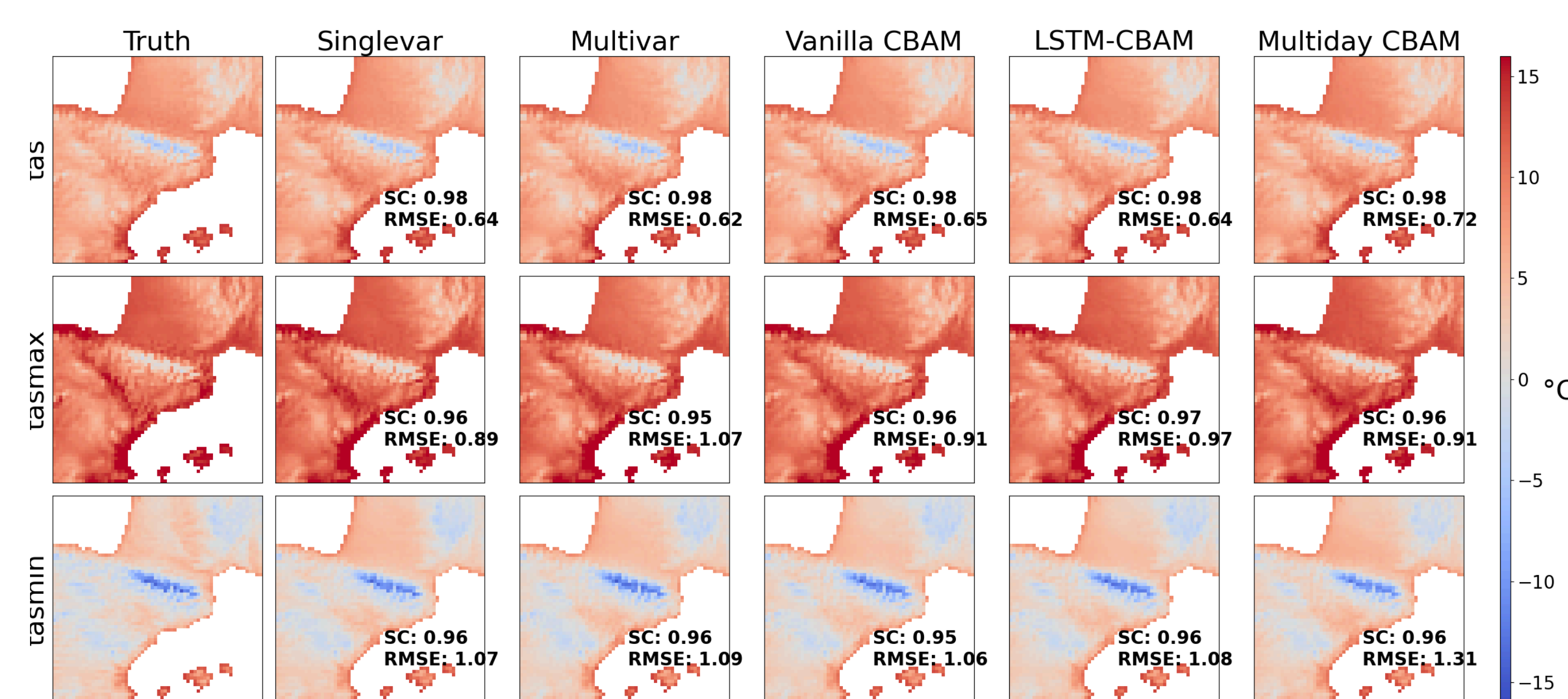
How can we meaningfully evaluate and compare these 5 models?

Results

Visually

e.g. what do the model outputs actually look like when compared to the RCM Truth?

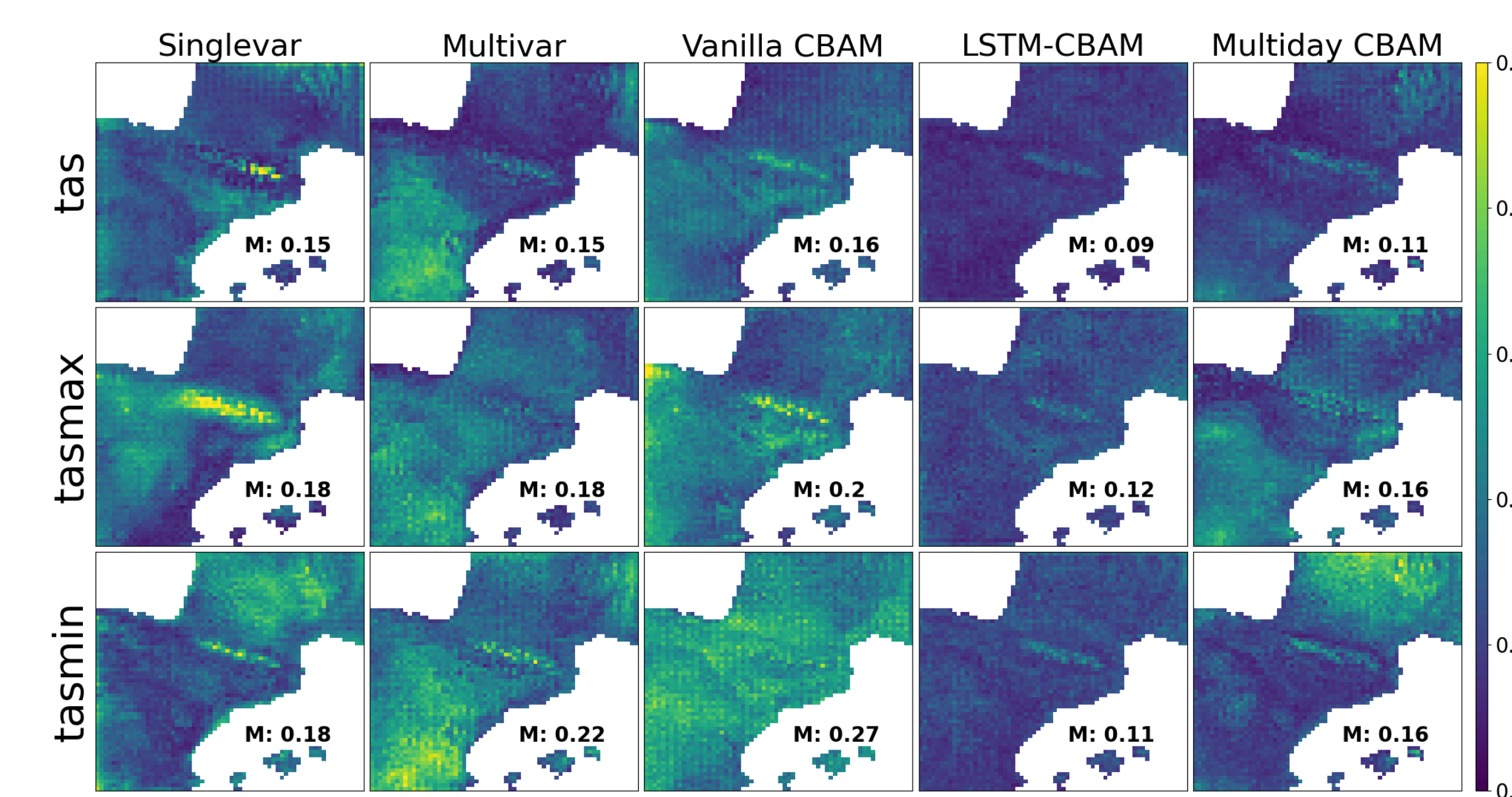
Raw Emulator Output On Single Day (2055-02-26)



Statically

e.g. what is the difference in distribution at each point between a model and the truth?

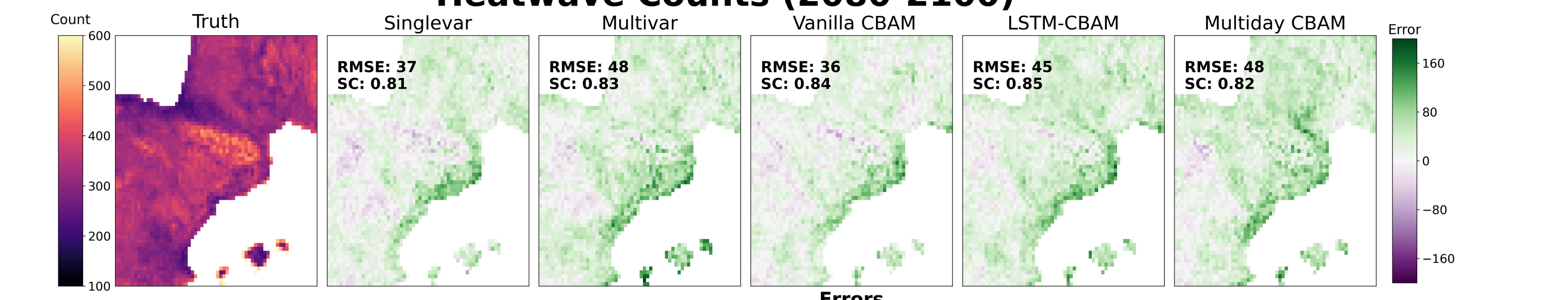
Wasserstein Distance



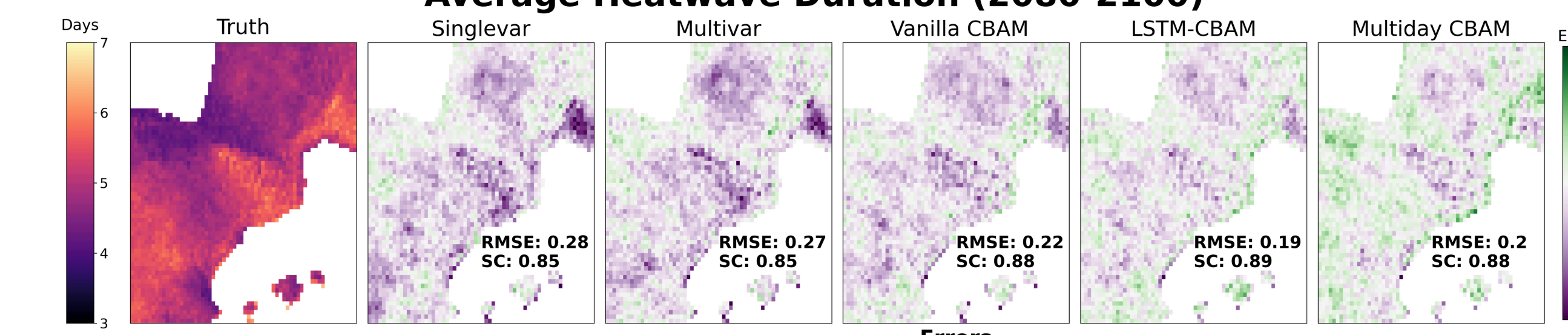
Applicably

e.g. can these models capture the presence of extreme compound weather events, like heatwaves?

Heatwave Counts (2080-2100)



Average Heatwave Duration (2080-2100)



Robustly

e.g. which models show meaningful difference in performance when randomness in training is accounted for?

Statistical Significance Testing Results: Improvements vs. Singlevar
 Via Kruskal-Wallis test w/ Dunn's post-hoc

