Post-Processing High-Resolution Weather Forecasts with Deep Learning to Create Operational Ensembles

Tyler McCandless, Luke Conibear, Allison Reed Harris, Ashley Payne, Steven Brey, Stylianos Flampouris, Luke Peffers, Kushal Keshavamurthy, Max Green, Forest Cannon

Tomorrow.io

ECMWF User's Conference
8 June 2023

Goal of this work: Provide forecasts to our customers that are **more accurate at the local scale than standalone operational NWP,** and **provide probabilities** to support their decision process

Key Question:
What is the added value of private sector forecasts?

Bringing it all together to provide concise, accurate and actionable weather intelligence forecasts

TM

Tomorrow.io Platform

Multi-Task Neural Network (MTNN) Pipeline

CONUS MTNN Model

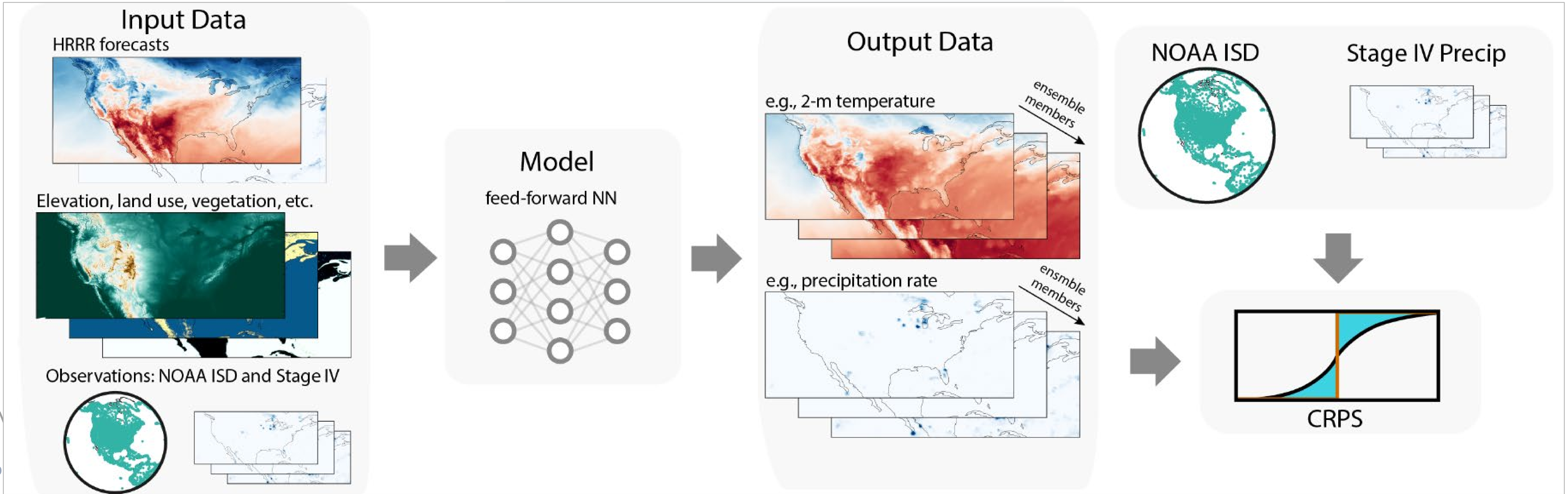
Artificial Neural Network model to predict multiple outputs and create probabilistic forecasts based on HRRR at 3-km resolution

Model Verification

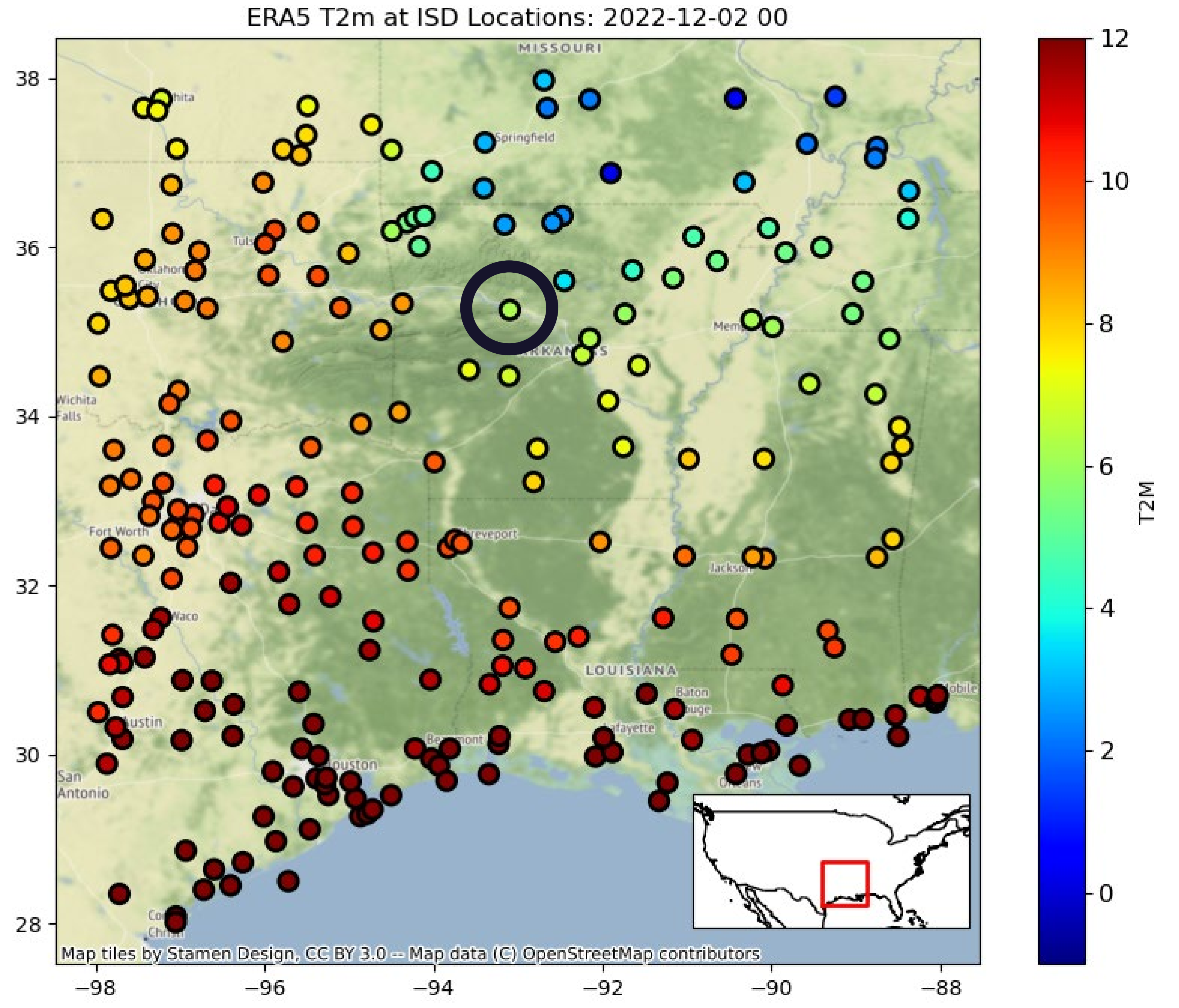
4-year dataset split into 2 year training, 1 year test and 1 year validation. Evaluate error metrics on scorecard to determine whether to move to production

Inference

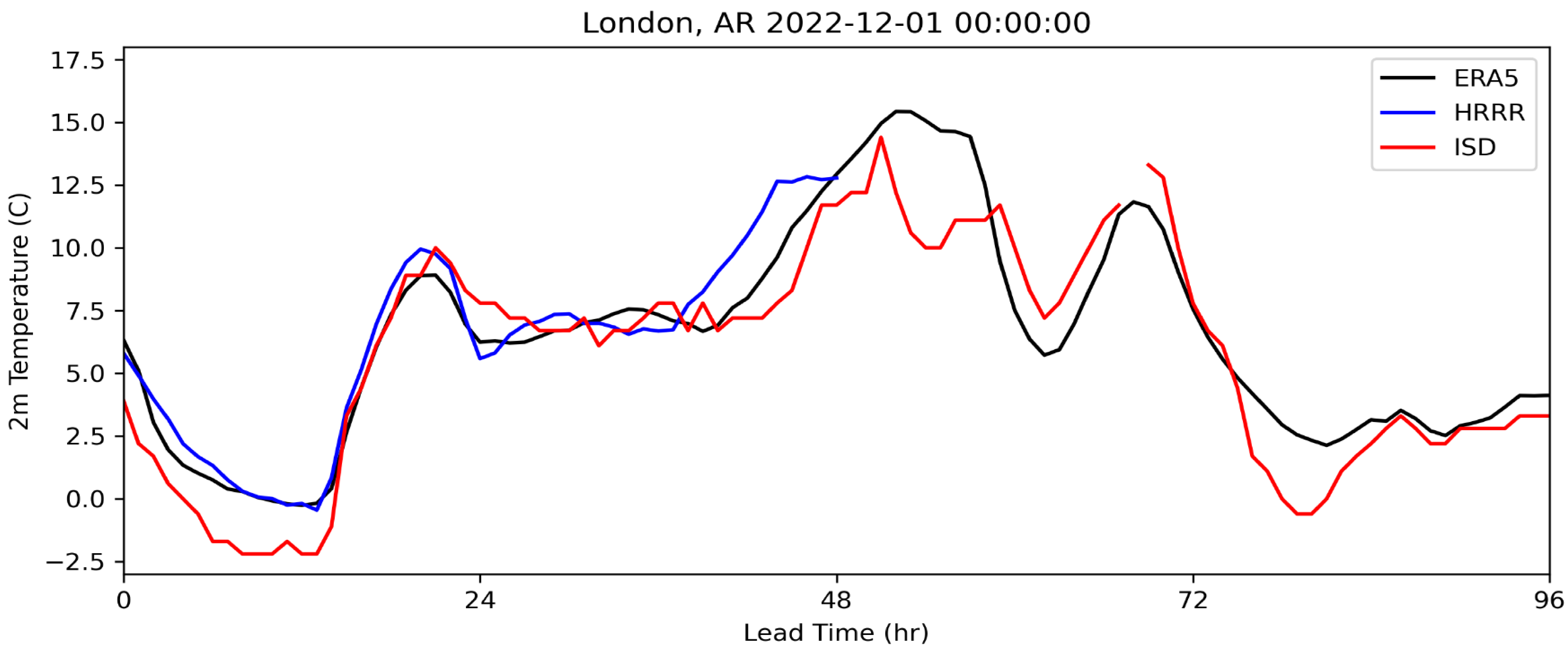
Operational, containerized, pipeline to prepare feature inputs from hourly HRRR initializations and to output predictions in a format that is ingested by the platform & API



Target Selection: Reanalyses or Stations?



- ERA5 is desirable as it is (relatively) high resolution, spatially-continuous, and many publications have quantified its exceptional performance relative to other reanalyses
- High Resolution Rapid Refresh (HRRR – Nation's operational high-res forecast system) is used as feature input to the NN, and generally matches HRRR at initialization
- Station data (from the Integrated Surface Database (ISD)) is more representative of the conditions that users experience – Decision to target improving HRRR predictions for customers.

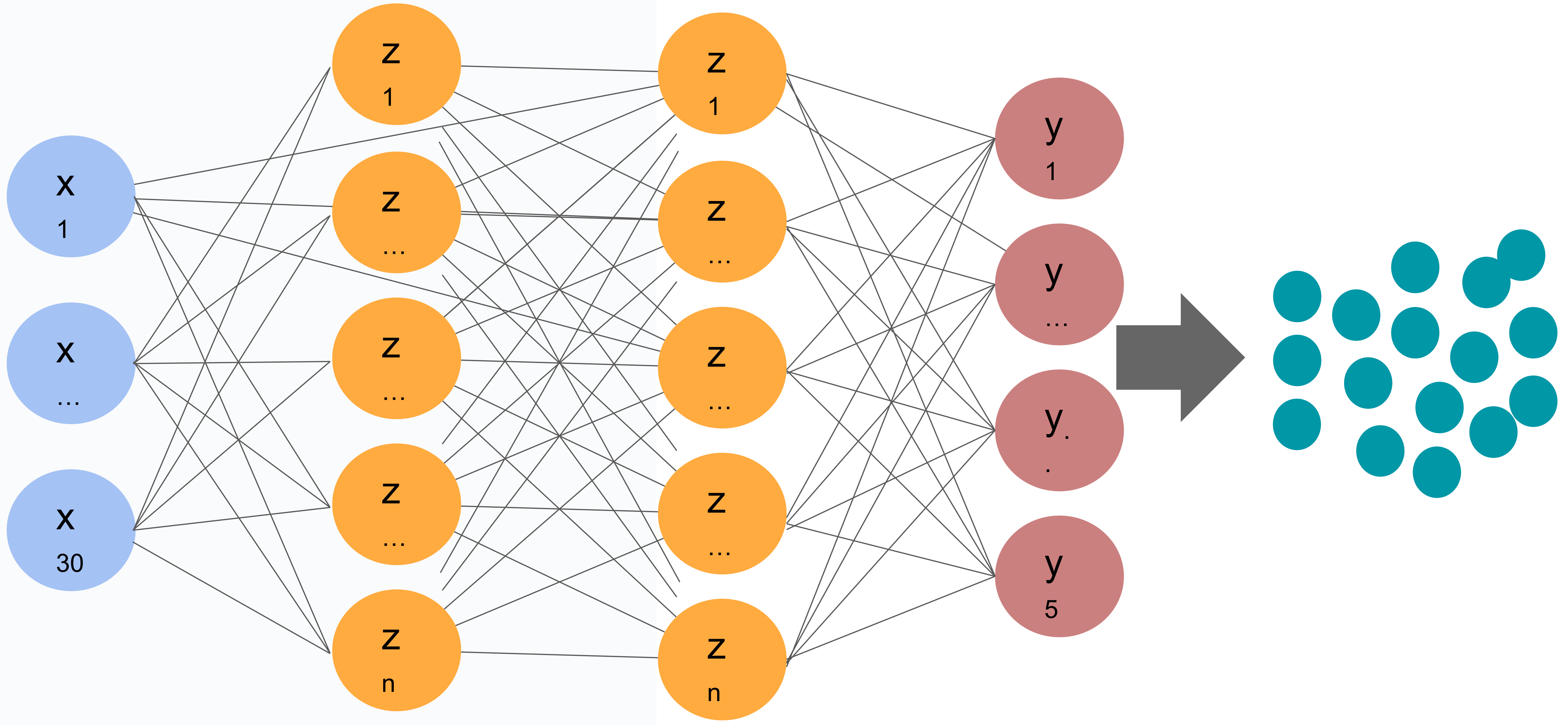


tomorrow

MTNN Summary

ANN (TensorFlow)

- CRPS loss function
- Multiple target variables from ISD observations
- Trains in 3 hours on 1x NVIDIA T4 (16GB) - 4 year dataset (2 years for training)
- Key Hyperparameters:
  - Ensemble members per target =21
  - Learning Rate = 0.001
  - Epochs = 200
  - Batch Size = 4096
  - Activation function = elu
  - Nodes per hidden layer = 48
  - Hidden layers = 6
  - Overfitting control (next slide)



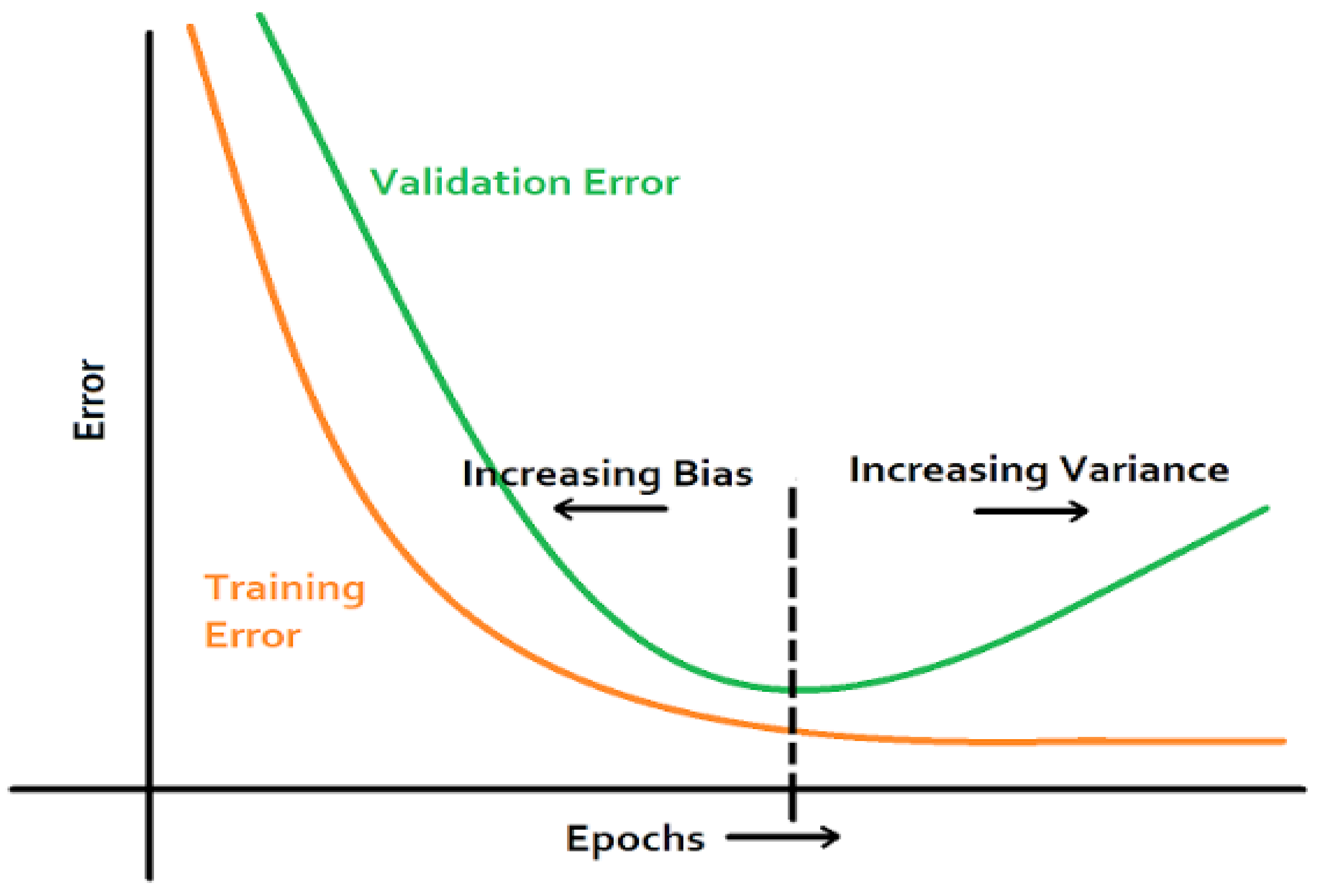
LC

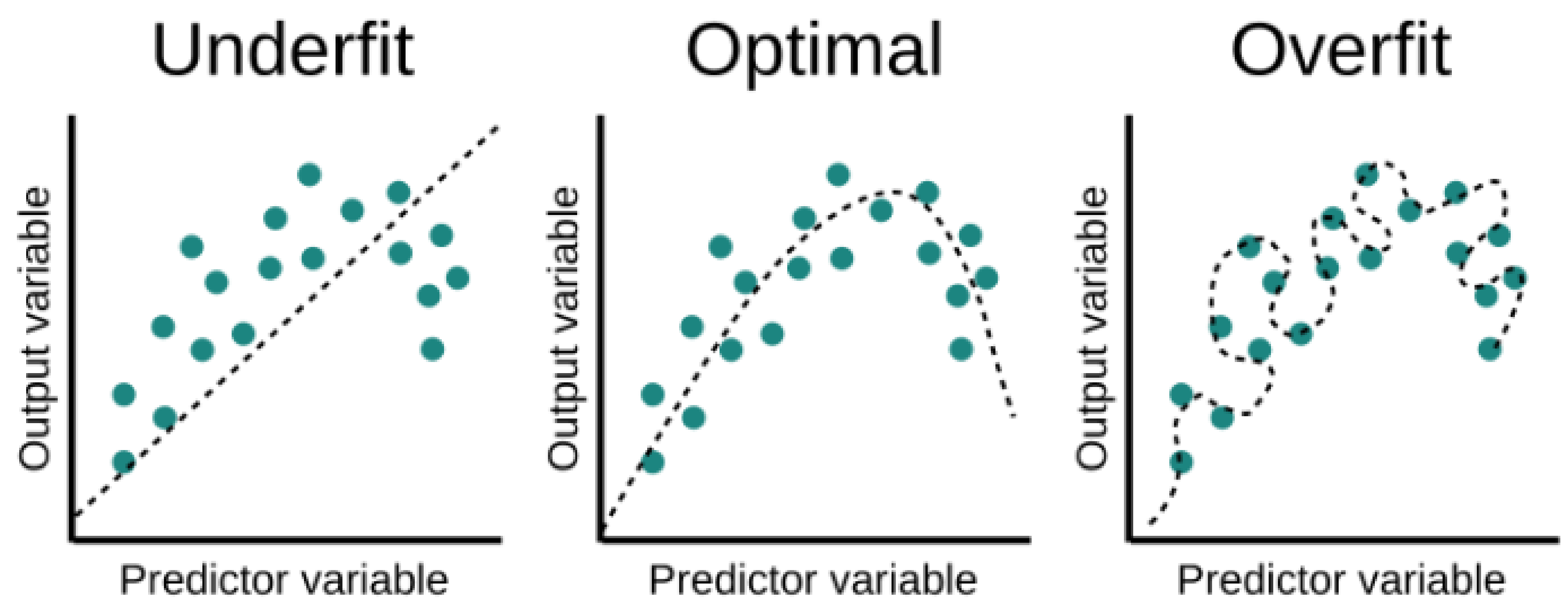
tomorrow.io

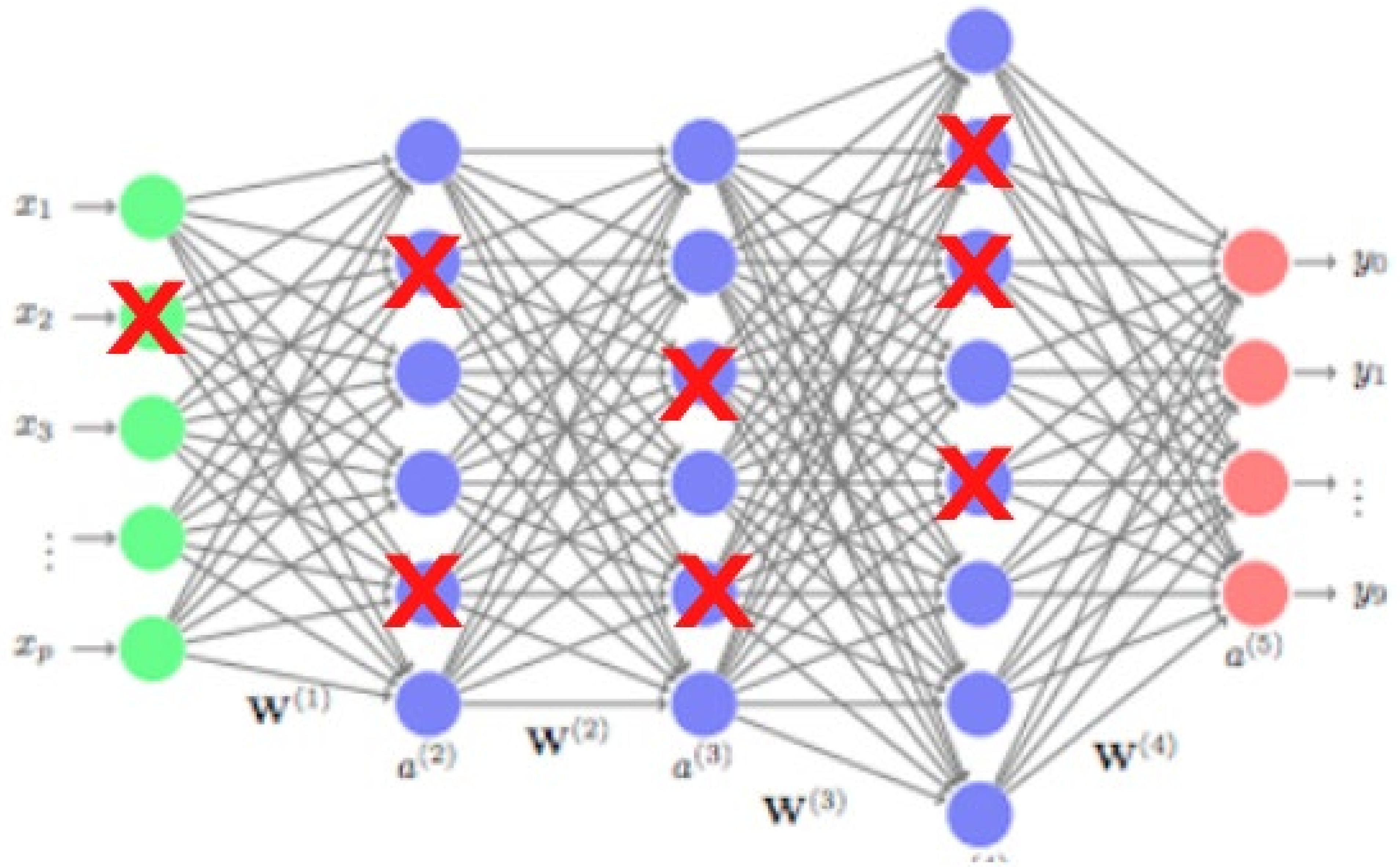
MTNN Summary

Avoiding overfitting

- L1 Regularization
- L2 Regularization
- Batch normalization
    - Helps coordinate update of multiple layers
- Early Stopping
- Dropout
    - 10%, 15%, 10%, 15%, 10%, 15%

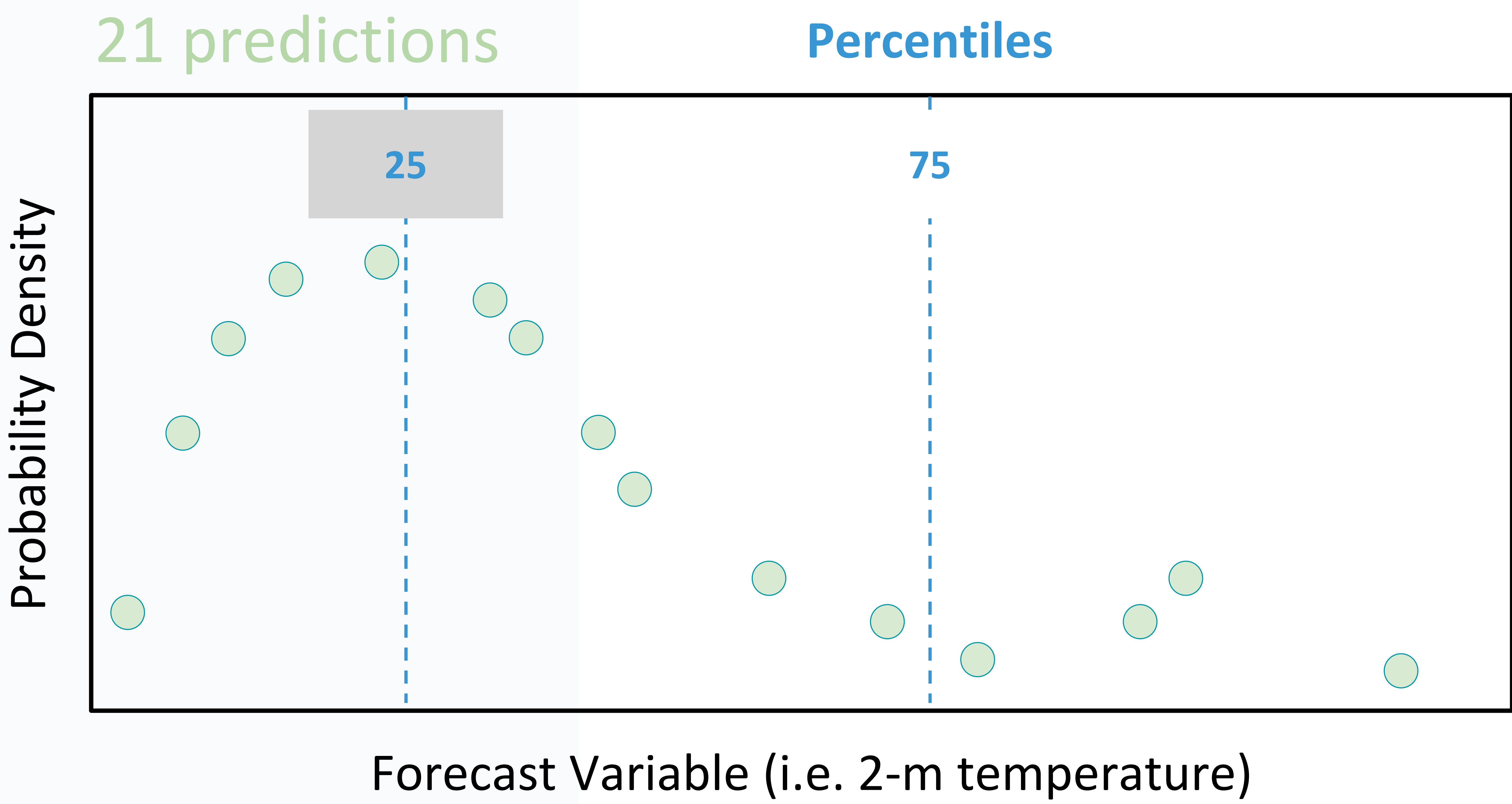






MTNN Output

- Predictions hourly out to 48-hr
- Forecast Variables:
  - 2-m temperature
  - 2-m dewpoint temperature
  - 10-m wind speed and direction
  - surface pressure
  - mean sea surface pressure
  - relative humidity
  - precipitation rate
- Probabilistic (forecast percentiles)
  - 5, 10, 25, 50, 75, 90, 95

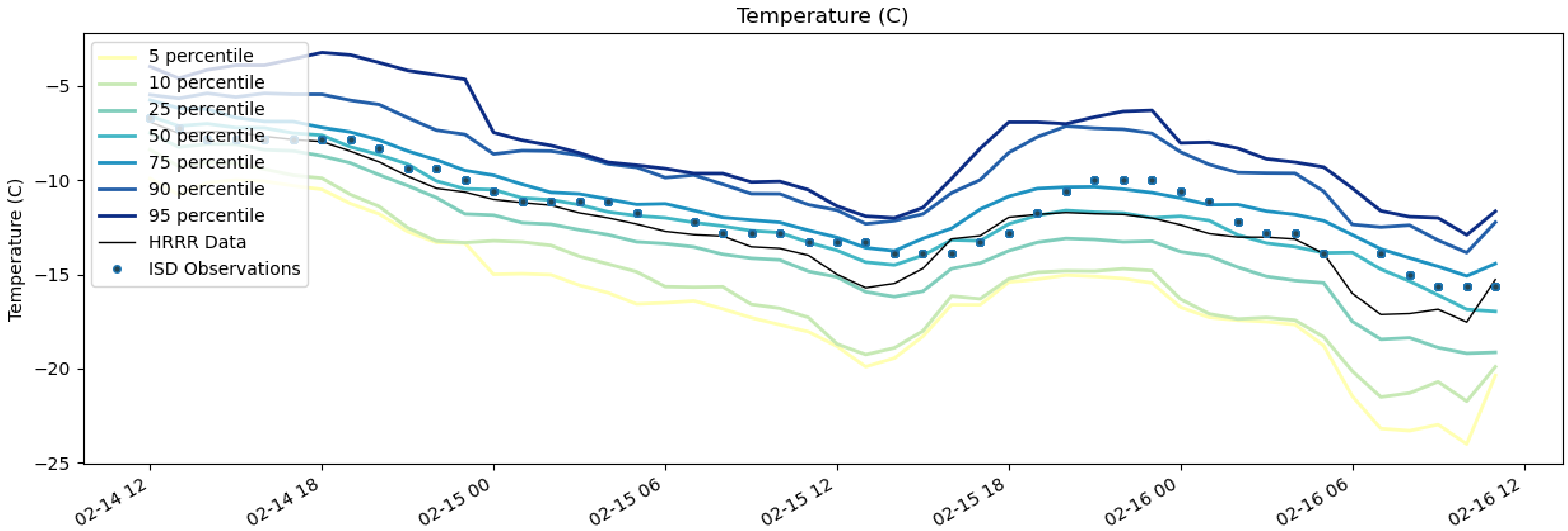


Percentiles are used to translate information from each variable's probability distribution into useable information

AP

MTNN Event Example: Winter Storm Uri

- Winter Storm Uri marked by a prolonged cold-air outbreak in Texas, with catastrophic impacts on energy and water infrastructure.
- Ensemble Predictions Capture Extreme Low Temperature Event (Dallas, TX shown)



AP

MTNN Summary Results - Probabilistic

CRPS results demonstrate value of MTNN probabilistic relative to operational HRRR

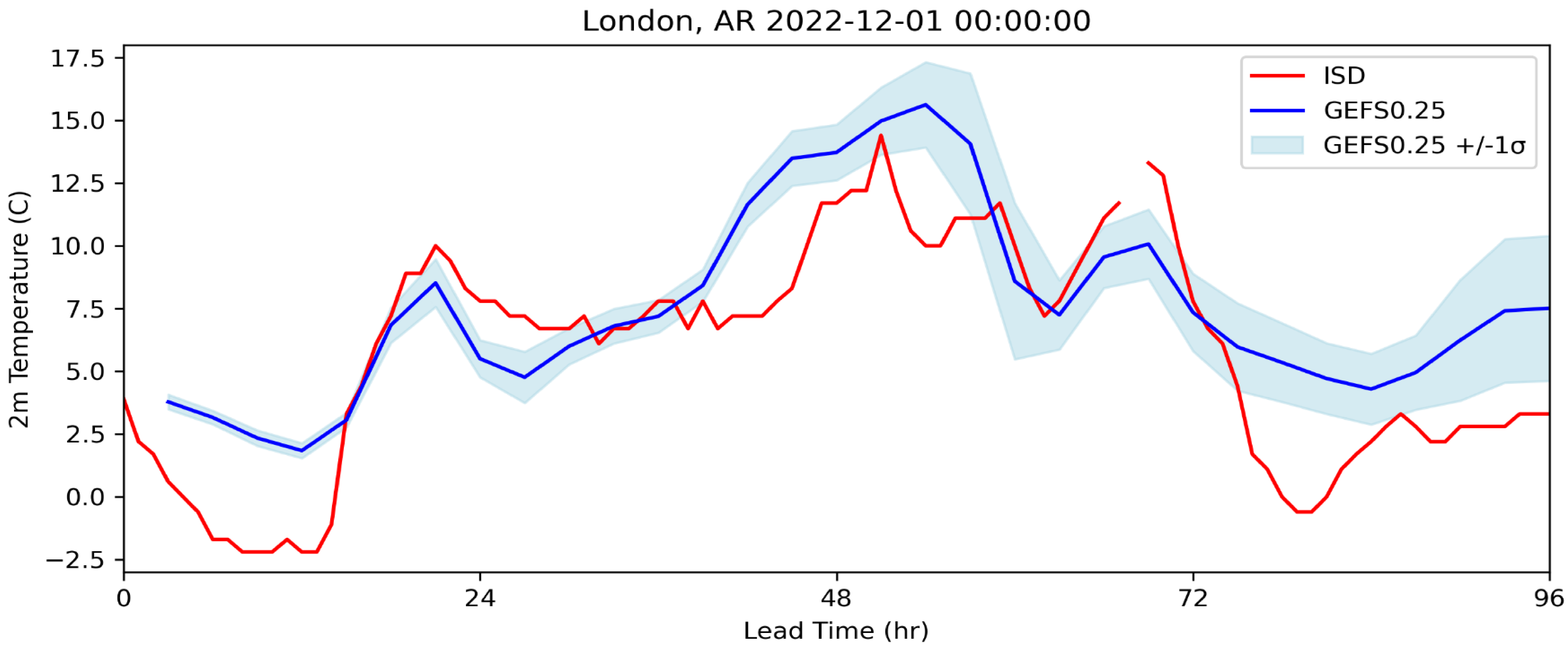
- CRPS is our loss function, so we expect improvement
- Will evaluate compared to climatology in the future

Forecast Variable (level)	HRRR CRPS (equal to MAE)	MTNN CRPS	MTNN percent improvement
2-m Temperature	1.57°C	1.11°C	29.3 %
10-m Wind Speed (U,V)	~1.5 m/s	~0.97 m/s	~35%
2-m Dewpoint	2.02°C	1.27°C	37.1 %
Surface Pressure	2.30 hPa	1.79 hPa	22.2 %
Precipitation Rate (Surface)	0.16 mm/hour	0.10 mm/hour	37.5 %

ARH

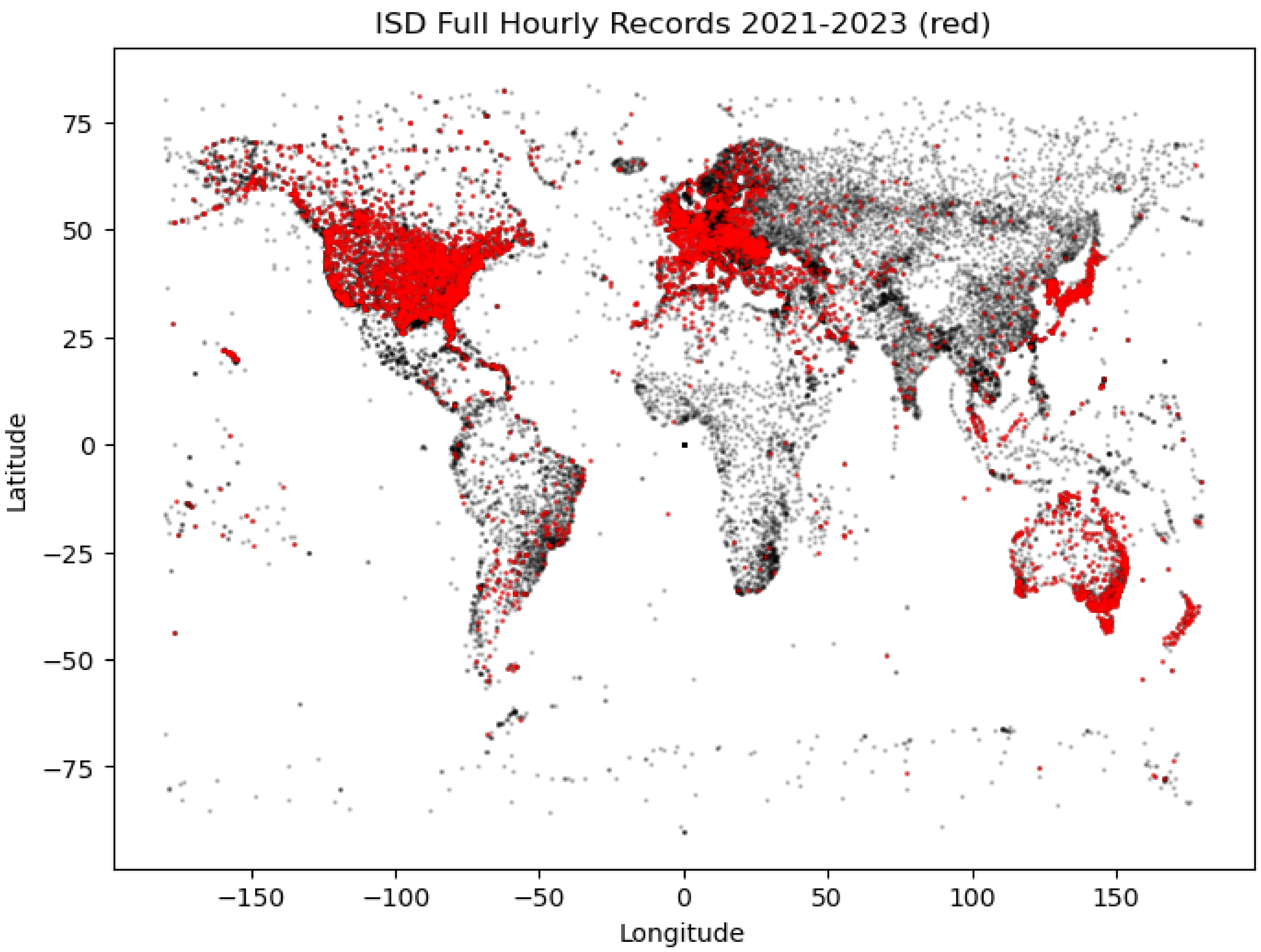
Application of MTNN to Extended Lead Times and Global Locations

- Longer lead-times require ingest of global models (e.g., ECMWF HRES, GFS, and their respective ensembles)
- Ability to handle NWP uncertainty across lead-time, condition and location as NN features (Reforecasts Needed!)
- ISD data is still a desirable target, but the coverage limitation is an issue



tomorrow

Revisiting Model Targets for Global Applications



● >95% Hourly Records 2021-2023 (~5500/35000 ISD Stations)

- ISD coverage is sparse outside of the US, Europe, Japan and parts of Australia.
- Global precipitation obs are not practical for ML use (not consistent enough in space and time, and generally not available sub-daily. Also, challenging to standardize QC)
- What are the alternatives (e.g., satellite obs?)

tomorrow.io

The Tomorrow.io Radar/Radiometer LEO Constellation

Building on decades of technology advancement in active/passive microwave instrumentation to enable the world's first ***operational*** precipitation mapping constellation comprised of ***science-quality*** payloads

Instrument Details

SmallSat Radar:

- Ka-band (35.5-36 GHz)
- Wide swath (400 km)
- High sensitivity (≤ 12 dBZ)
- High resolution (5 x 5 x .25 km)
- Adaptive, reconfigurable software-defined radar

CubeSat Sounder:

- Variant of TROPICS MWS
- 12 channels from 92-205 GHz
- 118 GHz O_2 temp. sounding
- 183 GHz H_2O sounding

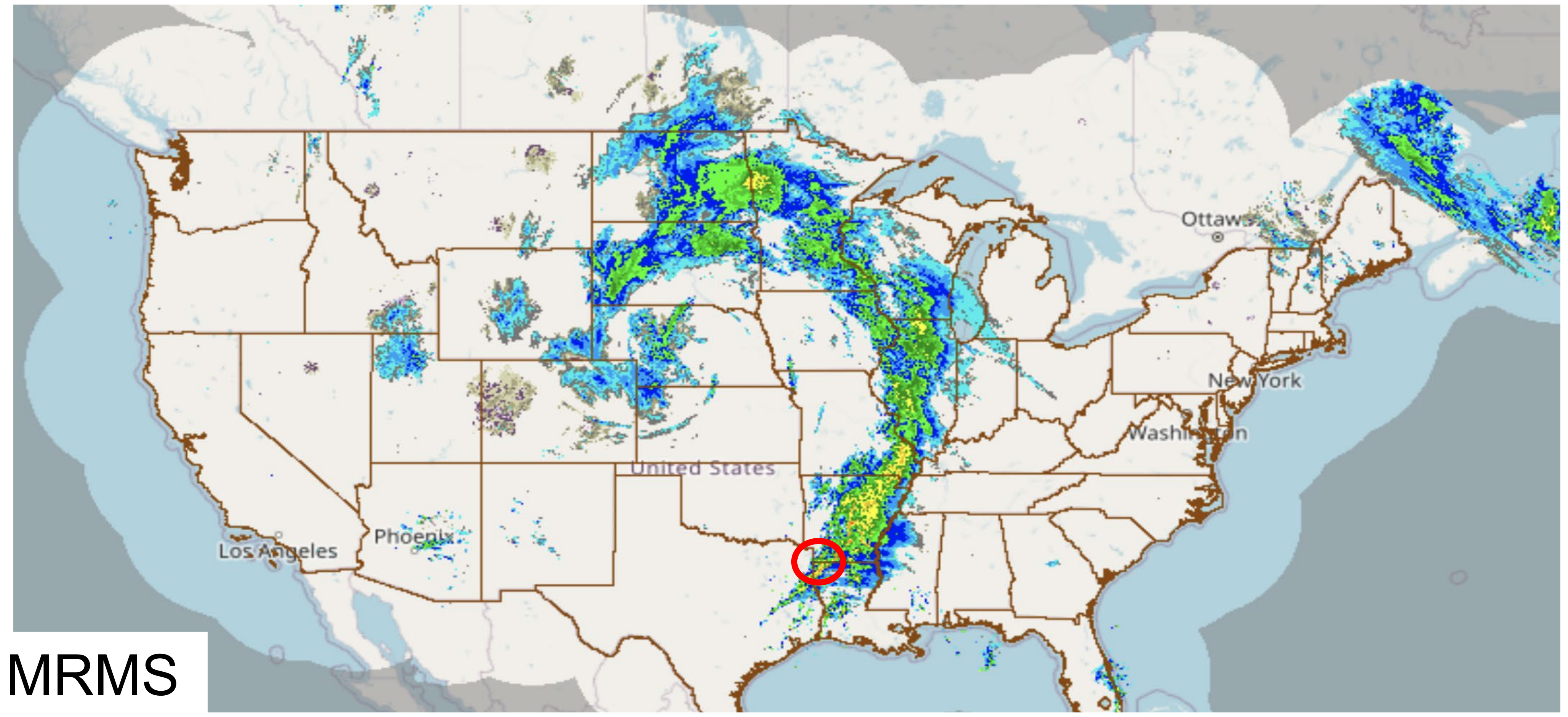
Constellation Capabilities

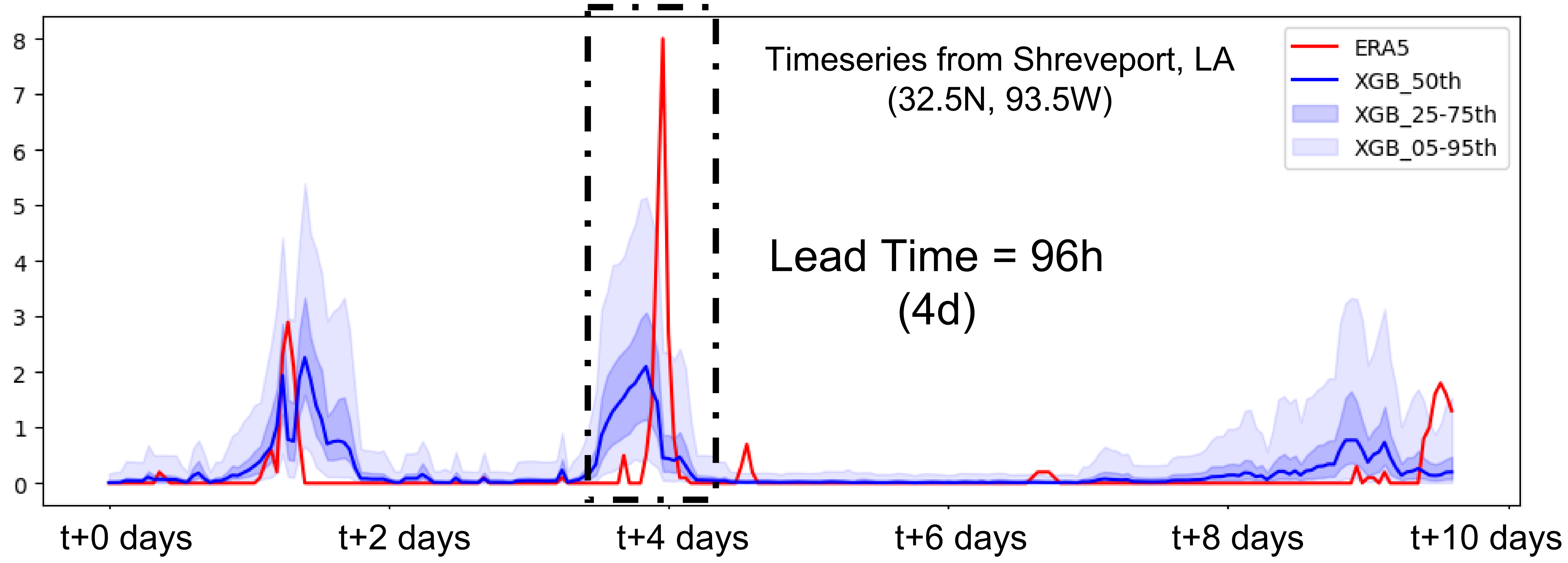
- Heterogeneity: 10 radars + 18 sounders planned
- ~hourly mean revisit over majority of globe
- 15 minute downlink latency
- Real-time precipitation, temperature, and humidity profiling
- Surface characterization from NRCS (radar) & clear-sky T_b (radiometer)

Post-processing high-resolution weather forecasts with deep learning to create operational ensembles

We present a machine learning model that post-processes high-resolution, deterministic forecasts to produce short to medium-range probabilistic forecasts for seven core weather variables. We developed and operationally implemented a multi-task neural network with a custom loss function, namely the Continuous Ranked Probability Score (CRPS). We tested this methodology using raw deterministic ECMWF HRES forecasts, ERA5 reanalysis fields as target data, and ERA5 invariant fields as supplementary features. Additionally, we tested and operationalized this model using data from the High Resolution Rapid Refresh (HRRR) model as input. This technique combines the strengths of high-resolution Numerical Weather Prediction (NWP) modeling with complex non-linear machine learning to generate more accurate deterministic forecasts alongside probabilistic forecasts, adding substantial predictive and actionable information. The results show deterministic forecast improvements in Root Mean Square Error (RMSE) over the HRRR from 1% to 12.5%. Using the CRPS as a metric to validate the probabilistic forecasts, we find a 22% to 38% improvement over the HRRR model. In addition to this increased forecast skill, the multi-task neural network approach is affordable to train and lightweight enough to run operationally on hourly forecasts. While our application used 21 ensemble members, this machine learning based approach has the flexibility to generate any number of ensemble members to best fit distribution of the forecast variables without significantly increasing computational costs.

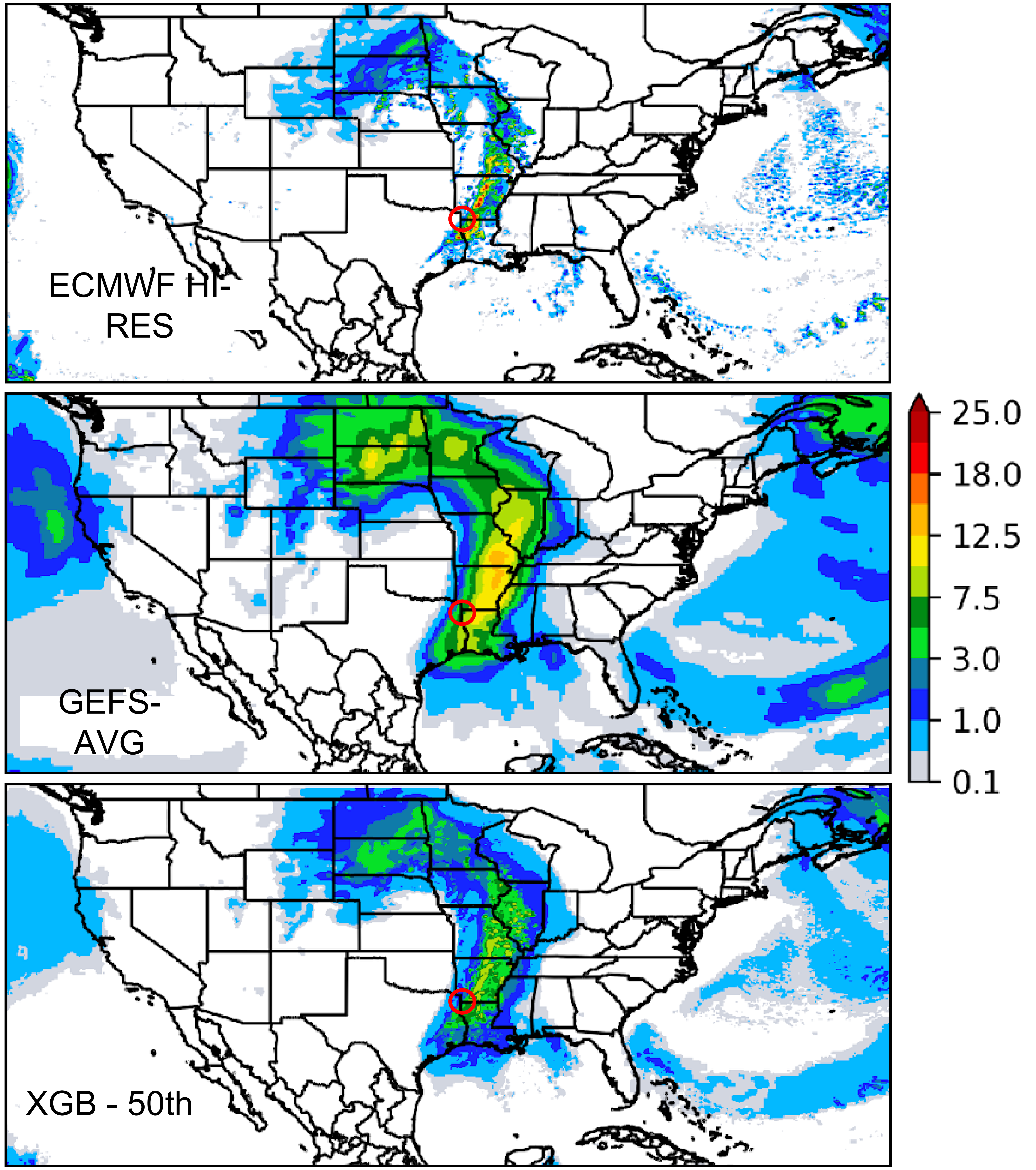
Questions that arise from a POC of global/multi-model post-processing:





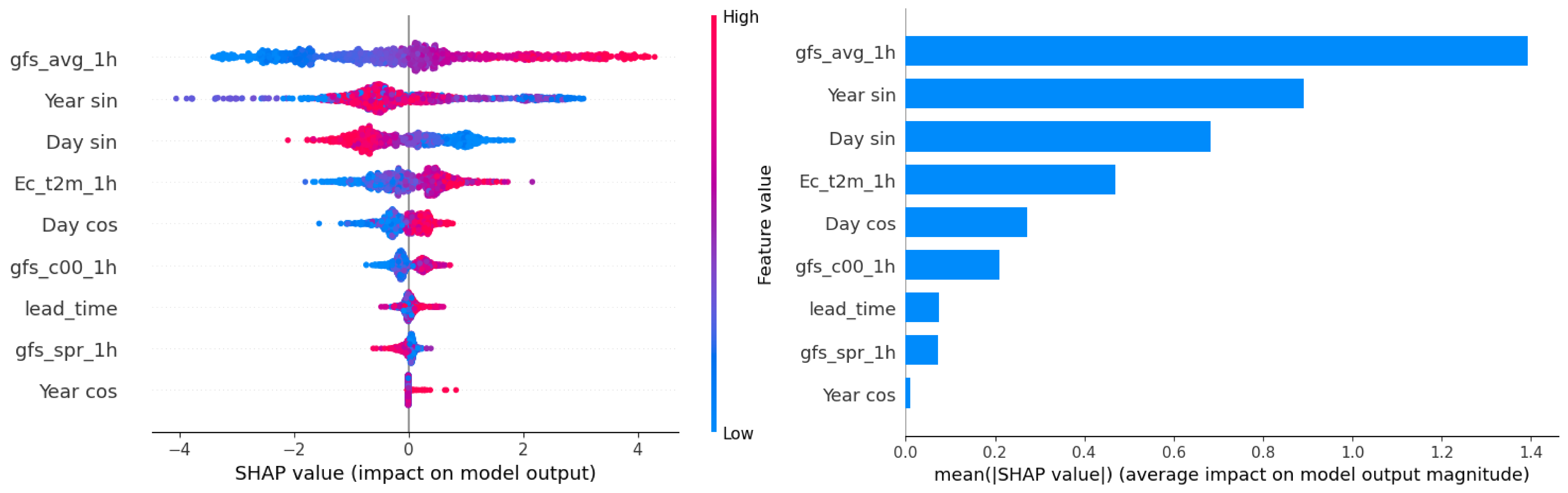
Initialized (t):
2022-12-10-00Z

Precipitation Rate (mm/h): 2022-12-14 00:00Z
Lead Time = 96h (4d)



Questions that arise from a POC of global/multi-model post-processing:

- What are the appropriate methods for evaluating the models themselves?
- How can physical constraints be baked in?



tomorrow.io