Forecasting the interferometric phase stability at the VLA with machine learning

Brian Svoboda
Ken Minschwaner, Bryan Butler, Paul Demorest,
Daniel Faes, Kyle Massingill

July 14, 2026

Atmospheric water vapor and phase

- Over-densities of water vapor in the troposphere change the index of refraction along the line of sight and delay the wavefront by varying amounts for each antenna. At >10 GHz, this causes the phase to vary on short timescales.
- Uncalibrated phase errors → **decorrelation**
- Phase referencing can be used to correct, but nearby calibrators are not always available and overheads can be prohibitive.
- At the VLA the *Atmospheric Phase Interferometer* (API) measures the phase RMS in real time for dynamic scheduling.

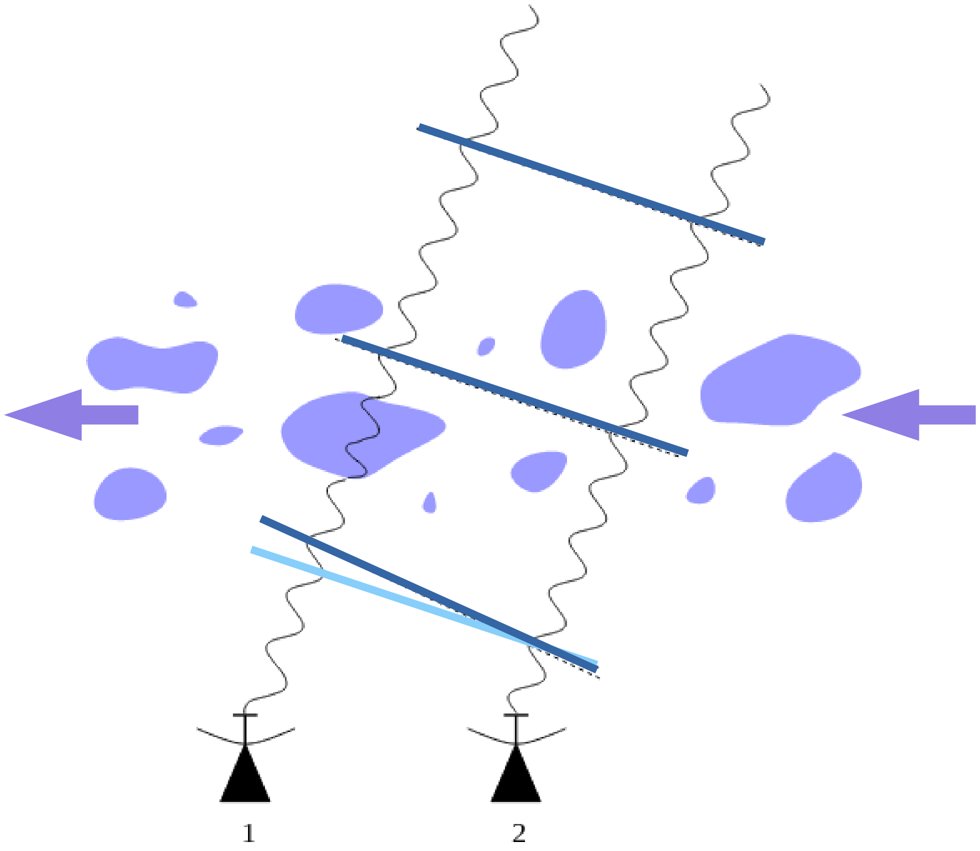
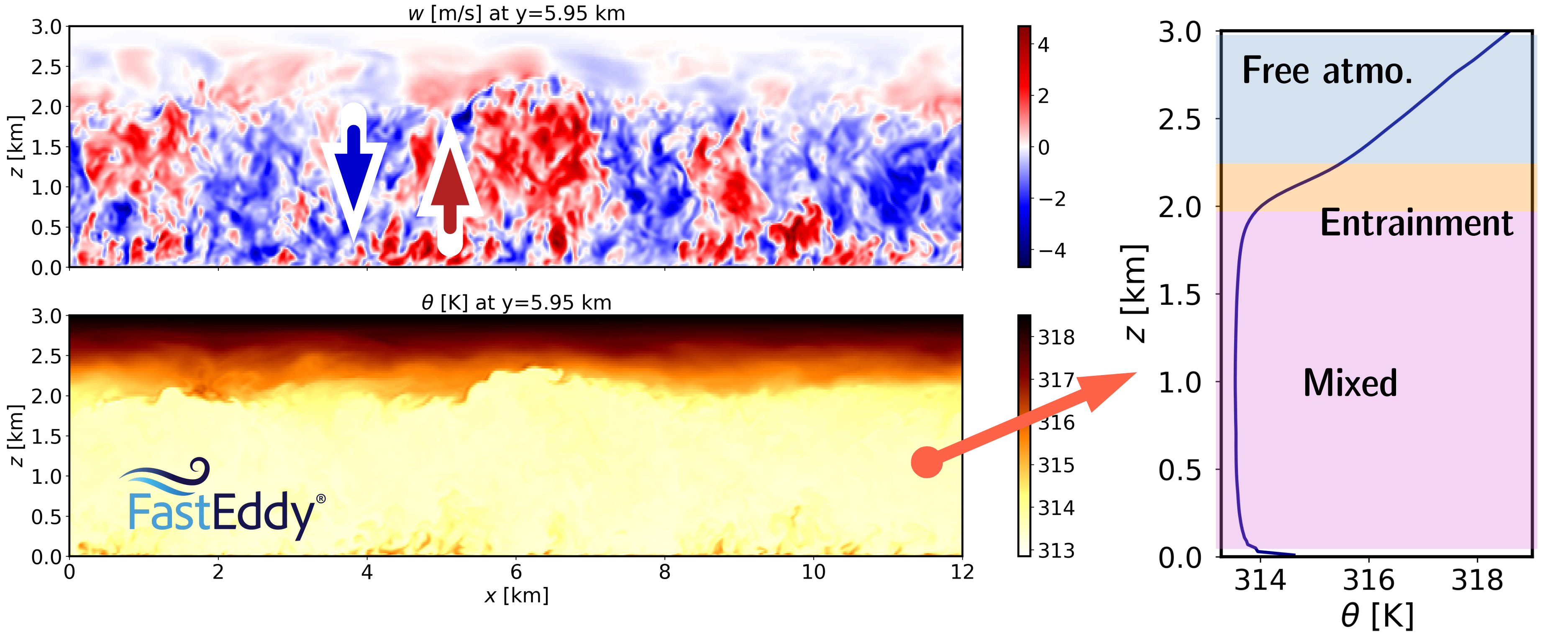
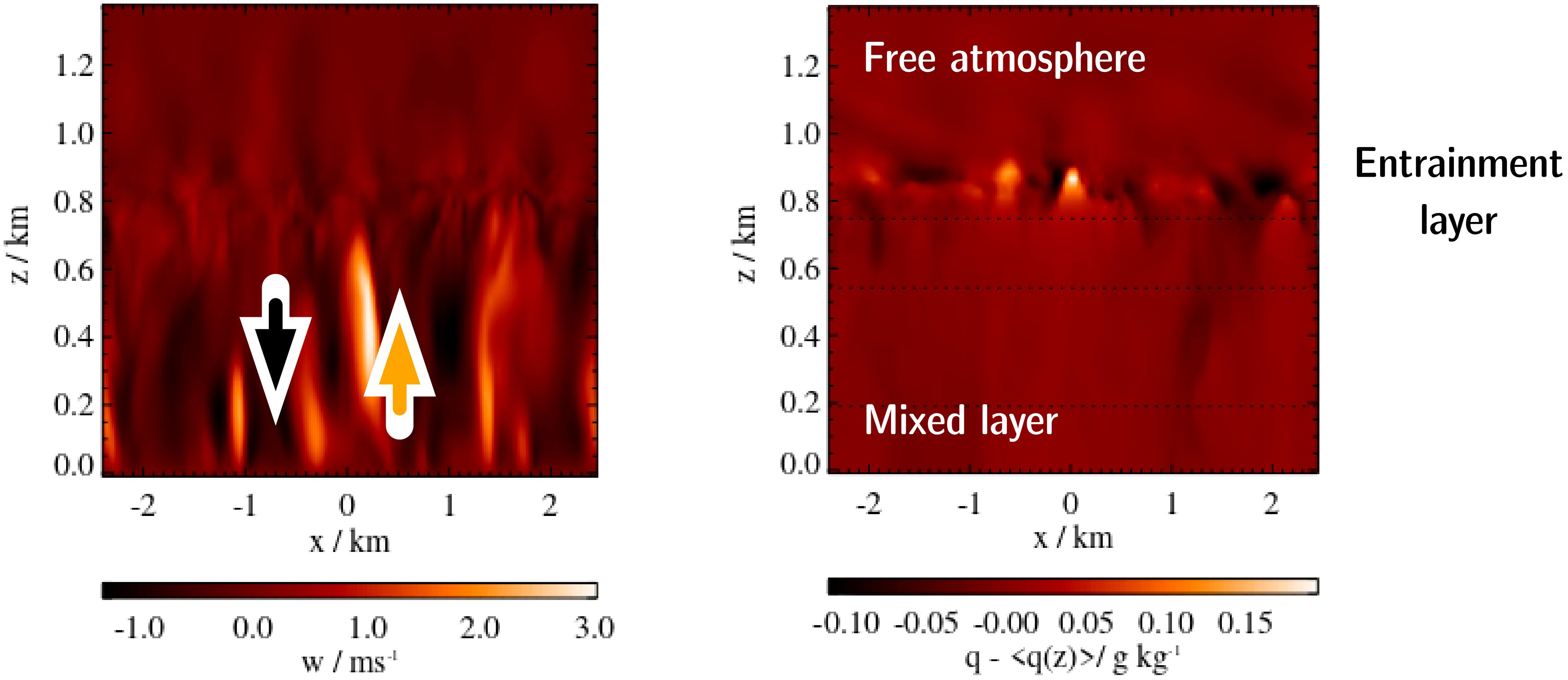


Image credit: Sault et al. (2006), figure originally by M. Wiedner (thesis)



"Large eddy" numerical fluid simulations of a dry, convective boundary layer (daytime). Convection creates a well-mixed layer and at the upper turn-over (2km) *entrains* air from the free atmosphere. This entrainment layer has strong fluctuations in density and water vapor.

Image credit: Sauer and Munoz-Esparza (2020), UCAR/NCAR, FastEddy documentation

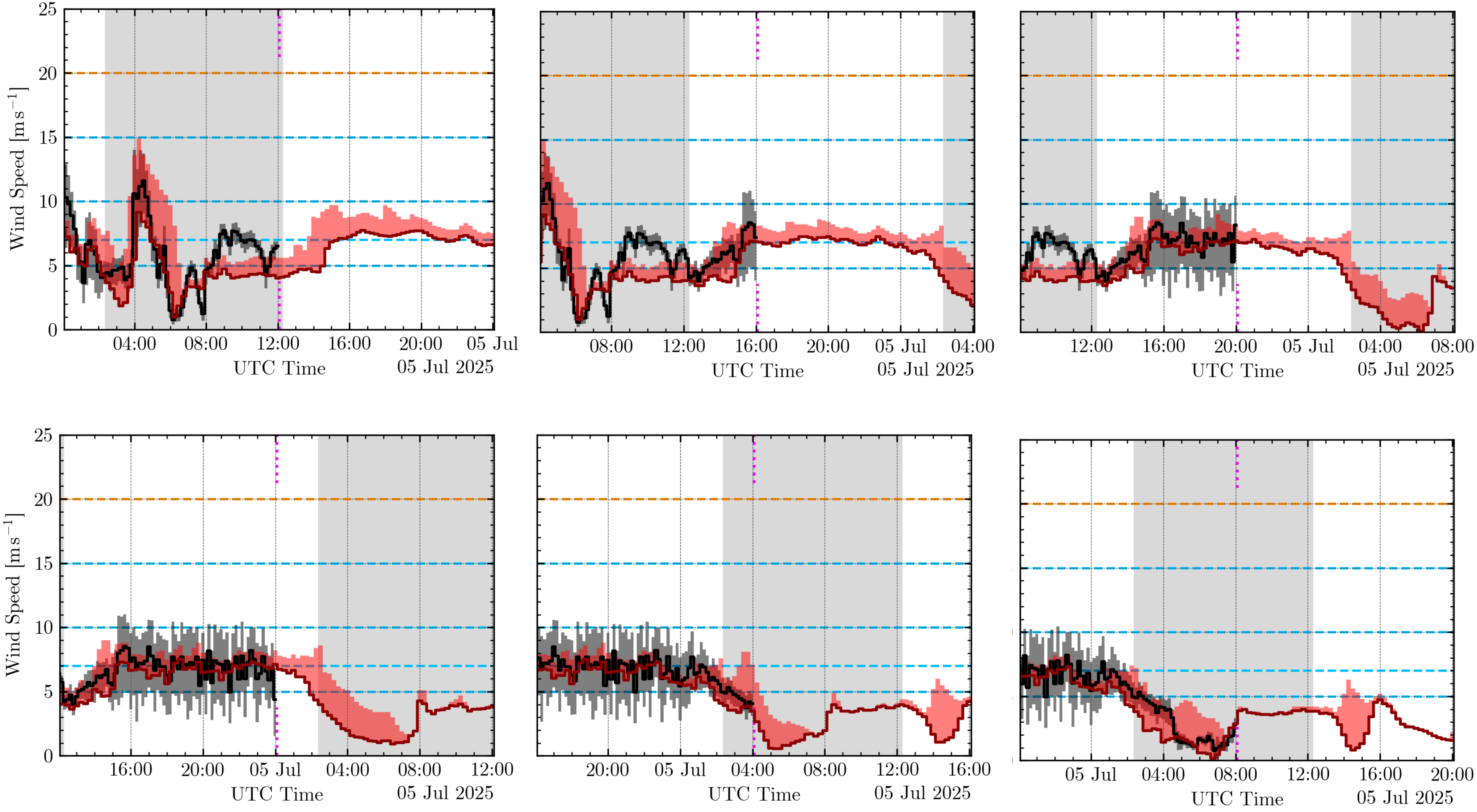


Stirling et al. (2005) performed large eddy simulations to support the ALMA site testing. Above left shows the vertical velocity and convective cells. Above right shows the strong turbulent fluctuations in water vapor densities in the entrainment layer.

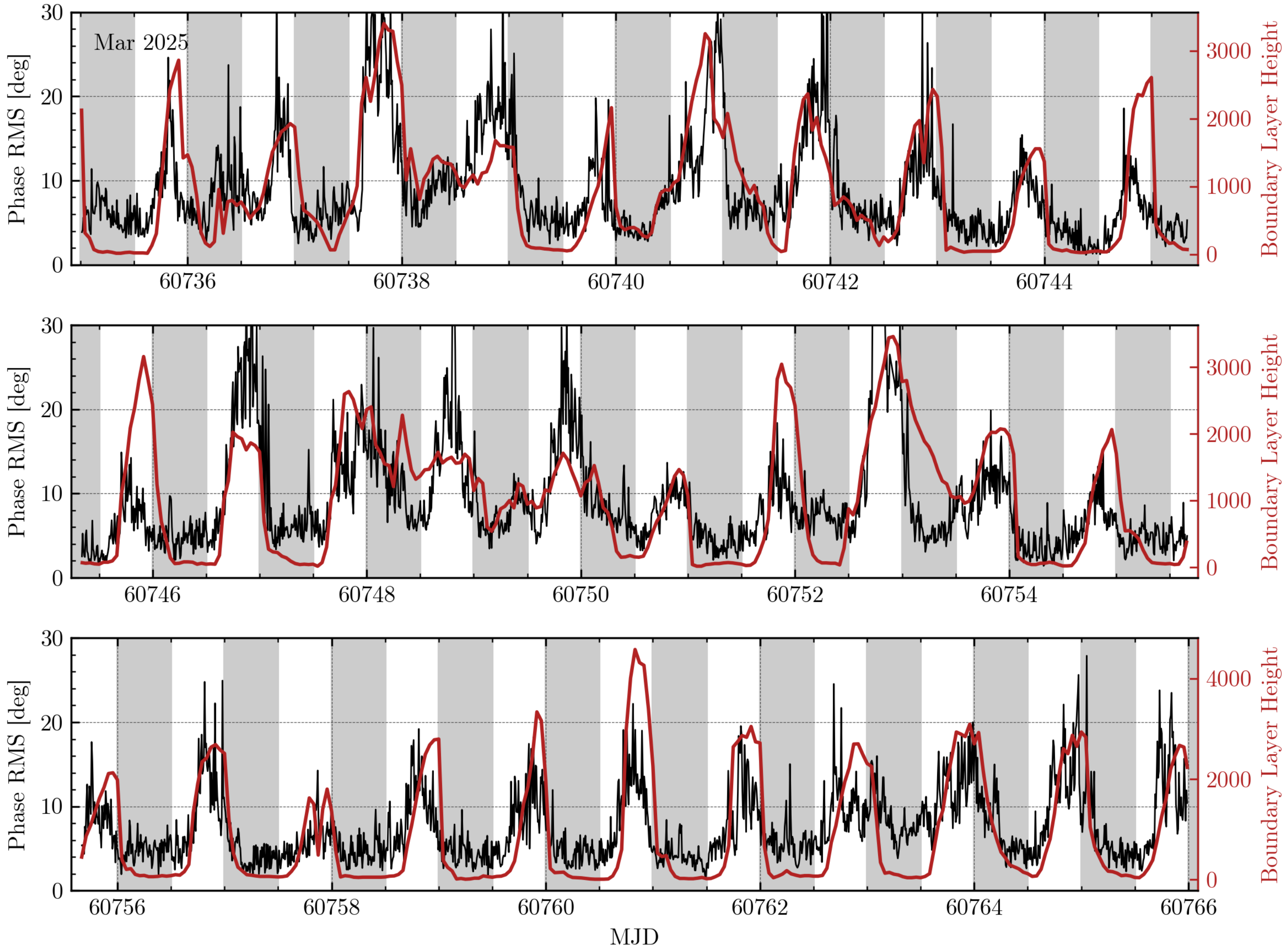
Image credit: Stirling et al. (2005; ALMA #517); Fig. 2 & 5.

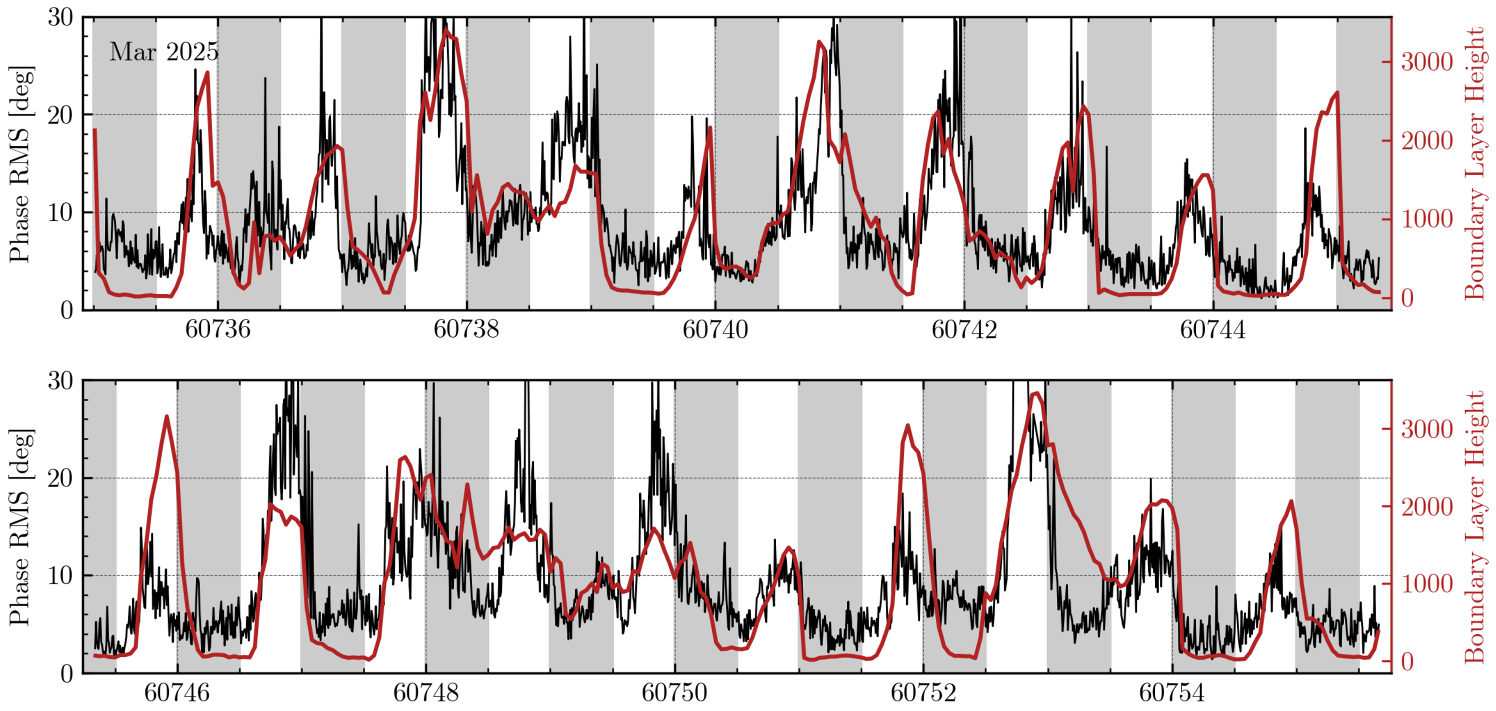
Motivation

- High-frequency VLA observations are valuable but are limited by weather constraints: *phase stability*, *wind speed*, and *clouds*.
- Currently only forecasted wind speed is used in automated scheduling.
- Forecast guidance can help:
  - improve scheduling efficiency in changing conditions,
  - make better use of daytime,
  - estimate cloud liquid column density day or night,
  - aid scheduling of geographically distributed arrays (VLBA, ngVLA).
- Quality of numerical weather prediction (NWP) models has dramatically improved over the last few decades and are easy, free to access.
- NCEP HRRR: best-in-class, short-term forecast for US (18h hourly).
- **Can we forecast the phase RMS from NWP data?**



Wind Speed [ms^{-1}]
25
20
15
10
5
0
04:00
08:00
12:00
16:00
20:00
05 Jul
UTC Time
05 Jul 2025


Mar 2025
Phase RMS [deg]
Boundary Layer Height
MJD
60736
60738
60740
60742
60744
60746
60748
60750
60752
60754
60756
60758
60760
60762
60764
60766



Planetary Boundary Layer Height (BLH) from HRRR captures the diurnal trend in phase RMS from convection well, along with nocturnal wind shear. This suggests that the parametrized convection in NWP models contains relevant information for predicting the phase RMS. Workable to train a machine learning model?

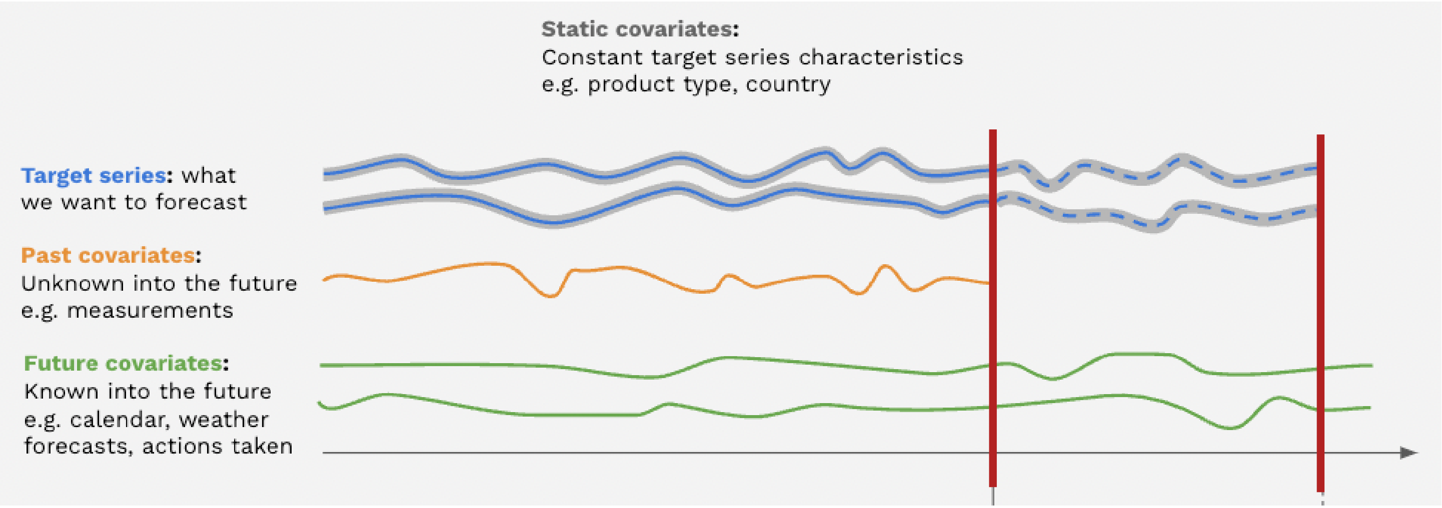
Machine learning model fitting

- Modeling a combination of:
  - auto-regression of measured *target*: phase RMS.
  - multivariate, non-linear regression of *covariates*: weather.
  - time series dependence, or effects from *lags*.

Phase RMS

Measured wind speed

Forecasted wind speed



start horizon

Image credit: darts documentation.

Machine learning model fitting

- Split into a 6 year (2017-2023) training data set and cross-validated on a 1 year (2024-2025) hold-out dataset.
- Applied a variety of time series machine learning techniques using the Python `darts` library.
- The gradient-boosting decision tree method **LightGBM** provided the best MAE/RMS on the validation set.
- GBDT's are an ensemble technique that is similar to a random forest algorithm. Fast and easy to use (viz. neural network).

Regression Models (GlobalForecastingModel)					
SKLearnModel: wrapper around any scikit-learn-like regression model		✅✅	✅✅✅	🔴🔴	✅
LinearRegressionModel		✅✅	✅✅✅	✅✅	✅
RandomForestModel		✅✅	✅✅✅	🔴🔴	✅
LightGBMModel		✅✅	✅✅✅	✅✅	✅
XGBModel		✅✅	✅✅✅	✅✅	✅
CatBoostModel		✅✅	✅✅✅	✅✅	✅
PyTorch (Lightning)-based Models (GlobalForecastingModel)					
RNNModel (incl. LSTM and GRU); equivalent to DeepAR in its probabilistic version	DeepAR paper	✅✅	🔴✅🔴	✅✅	✅
BlockRNNModel (incl. LSTM and GRU)		✅✅	✅🔴🔴	✅✅	✅
NBEATSModel	N-BEATS paper	✅✅	✅🔴🔴	✅✅	✅
NHiTSModel	N-HiTS paper	✅✅	✅🔴🔴	✅✅	✅
TCNModel	TCN paper, DeepTCN paper, blog post	✅✅	✅🔴🔴	✅✅	✅
TransformerModel		✅✅	✅🔴🔴	✅✅	✅
TFTModel (Temporal Fusion Transformer)	TFT paper, PyTorch Forecasting	✅✅	✅✅✅	✅✅	✅
DLinearModel	DLinear paper	✅✅	✅✅✅	✅✅	✅
NLinearModel	NLinear paper	✅✅	✅✅✅	✅✅	✅
TiDEModel	TiDE paper	✅✅	✅✅✅	✅✅	✅
TSMixerModel	TSMixer paper, PyTorch Implementation	✅✅	✅✅✅	✅✅	✅

Machine learning model fitting

- Split into a 6 year (2017-2023) training data set and cross-validated on a 1 year (2024-2025) hold-out dataset.
- Applied a variety of time series machine learning techniques using the Python `darts` library.
- The gradient-boosting decision tree method **LightGBM** provided the best MAE/RMS on the validation set.
- GBDT's are an ensemble technique that is similar to a random forest algorithm. Fast and easy to use (viz. neural network).

- `(6.33, 9.65) MAE, RMS Global Naive Seasonal.`
- `(3.95, 6.27) Linear Regression with target lags and future covariates.`
- `(3.48, 5.80) LGBM using only hour and day of year.`
- `(3.23, 5.33) LGBM using target lags, hour, and day of year.`
- `(2.80, 4.71) LGBM with target lags and future covariates.`

Forecasted quantities being used:

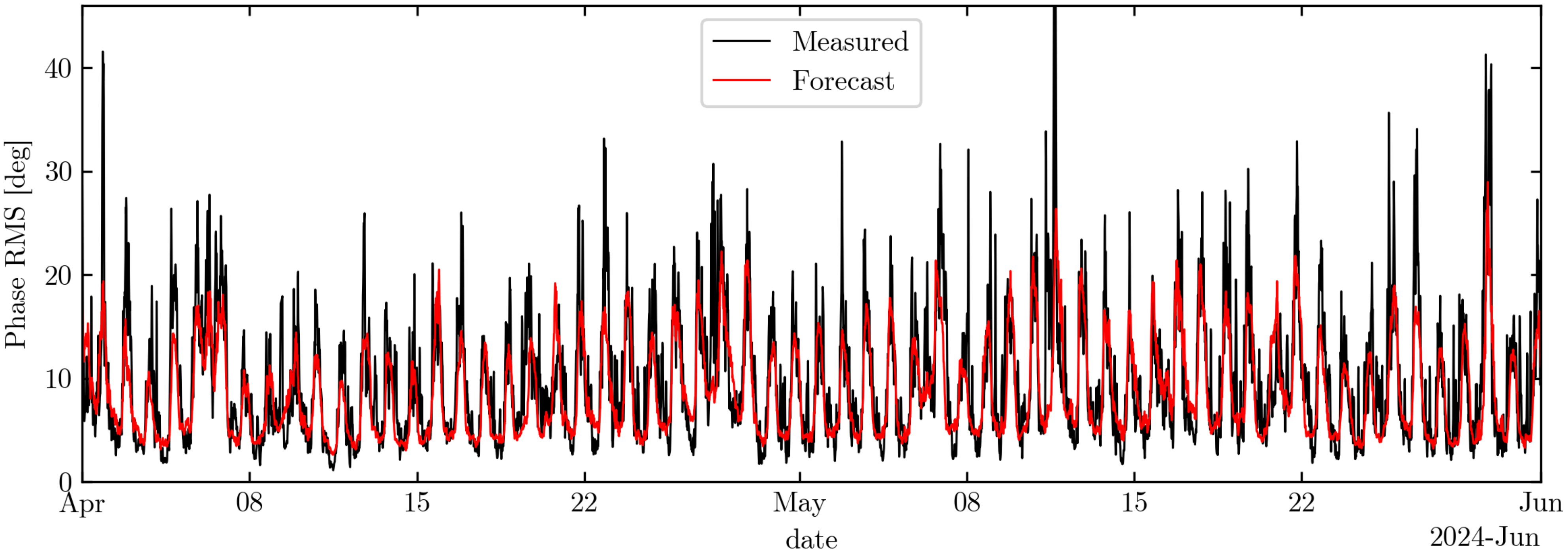
Sub-hourly (15-minute) quantities:

temperature, relative humidity, dewpoint at 2m
wind speed & direction at 10 and 80m
wind gusts at 10m
direct & diffuse solar radiation (i.e., sun angle plus cloud cover).
convective available potential energy (CAPE)
precipitation
visibility
freezing level height

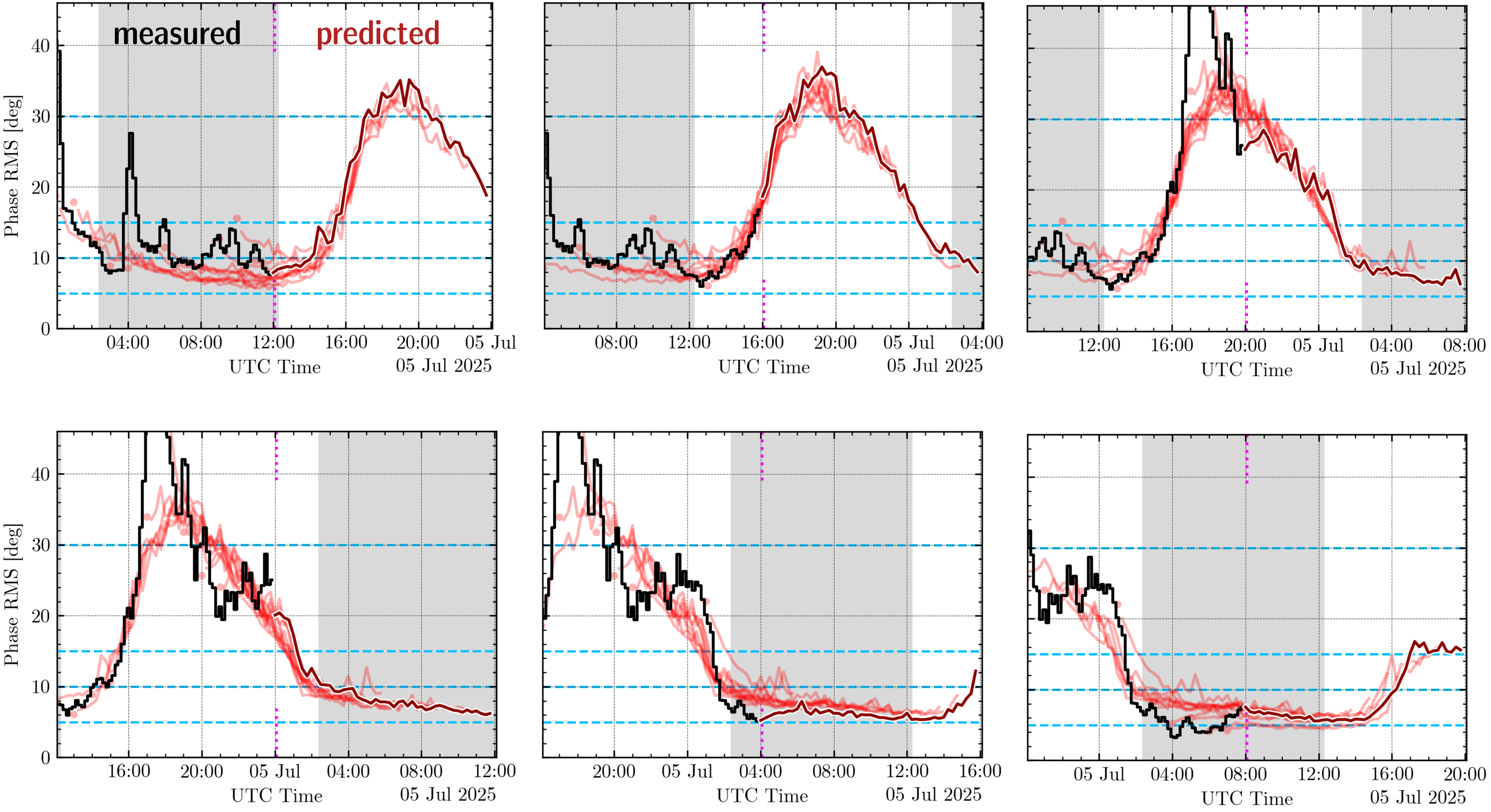
Hourly quantities:

lifted index
convective inhibition
surface pressure
cloud cover at low, medium, and high altitudes (plus total)
rain, showers, snowfall, precipitation probability

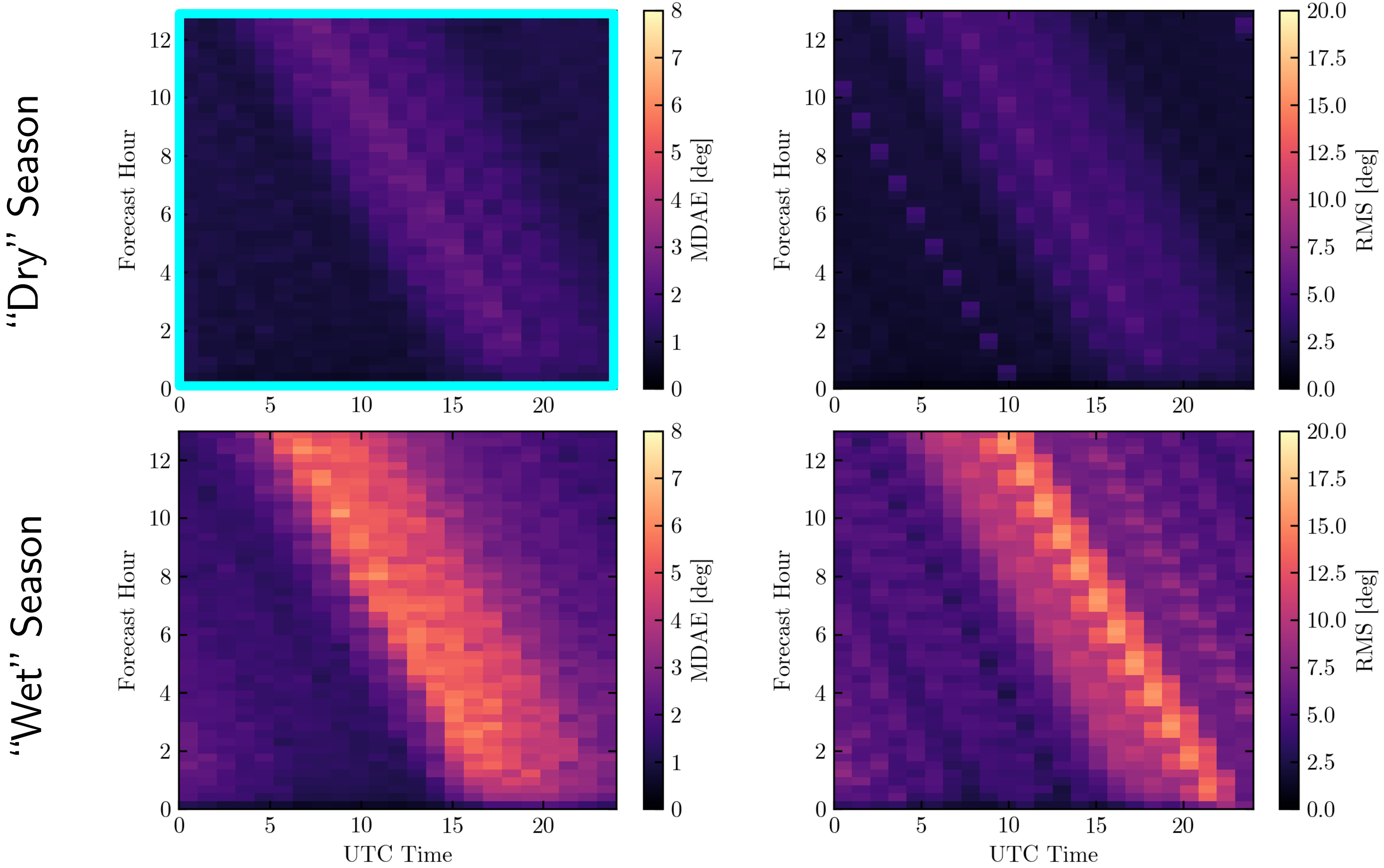
Note: more model quantities exist, but are not available through open-meteo: low-level wind shear, vertical velocity, helicity, and vorticity, among others. Can download current values relatively easily, but difficult to download a large volume for training.



Forecasts applied to the validation dataset, or “backtests.” The above shows the measured phase RMS with 6-hour forecasts strung together for two months. Weather data used as covariates not shown. Generally quite effective but typically under-predicts large excursions in phase RMS.



measured
predicted
Phase RMS [deg]
UTC Time
05 Jul 2025

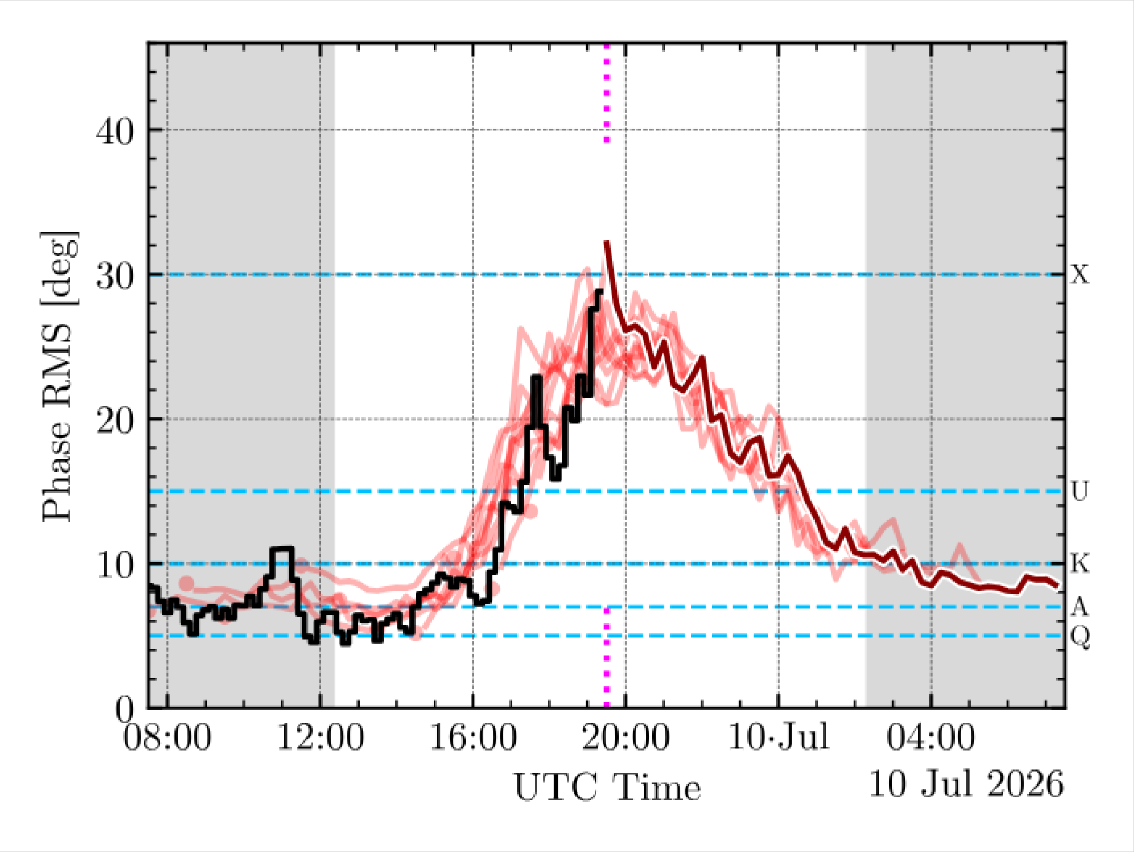


"Dry" Season
"Wet" Season
Forecast Hour
UTC Time
MDAE [deg]
RMS [deg]

Help! VLA Phase RMS (12 hr) ← Prev. Next → Latest

YYYY: 2026 MM: 07 DD: 09 HH (UTC): 19 Load Time

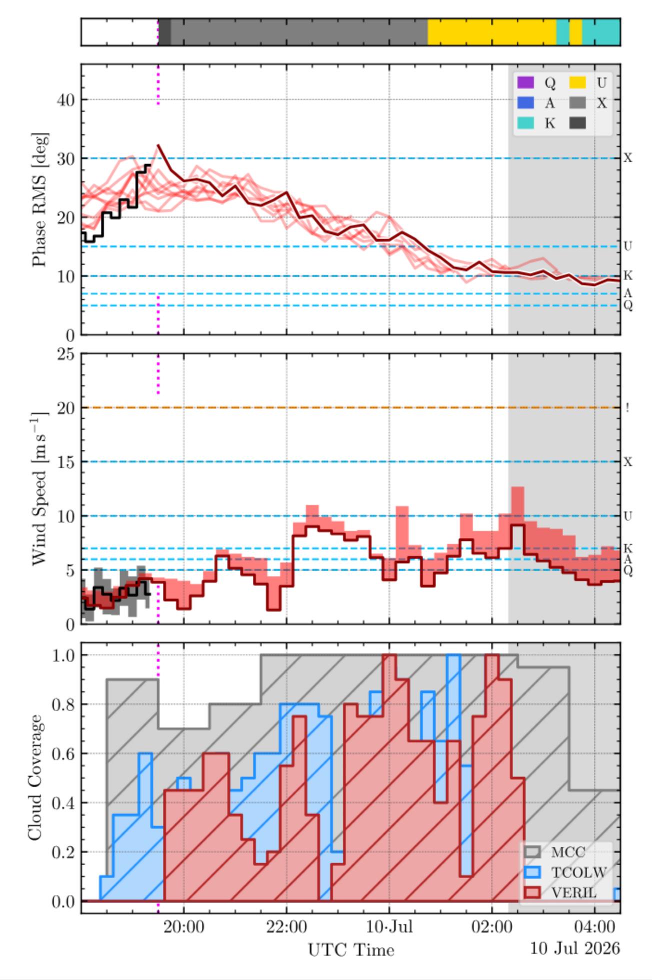
Displaying: 2026-07-09 19:00 UTC



Help! VLA Summary (9 hr) ← Prev. Next → Latest

YYYY: 2026 MM: 07 DD: 09 HH (UTC): 19 Load Time

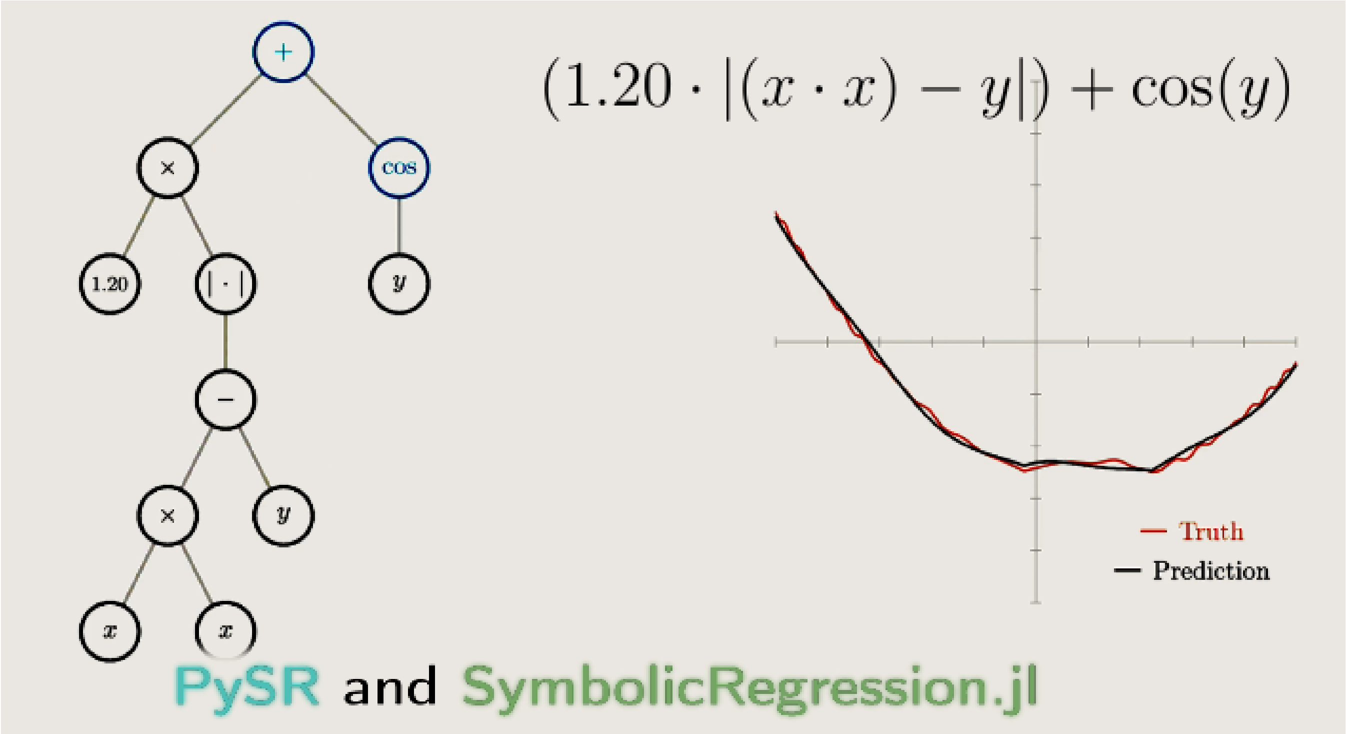
Displaying: 2026-07-09 19:00 UTC



Forward modeling phase RMS

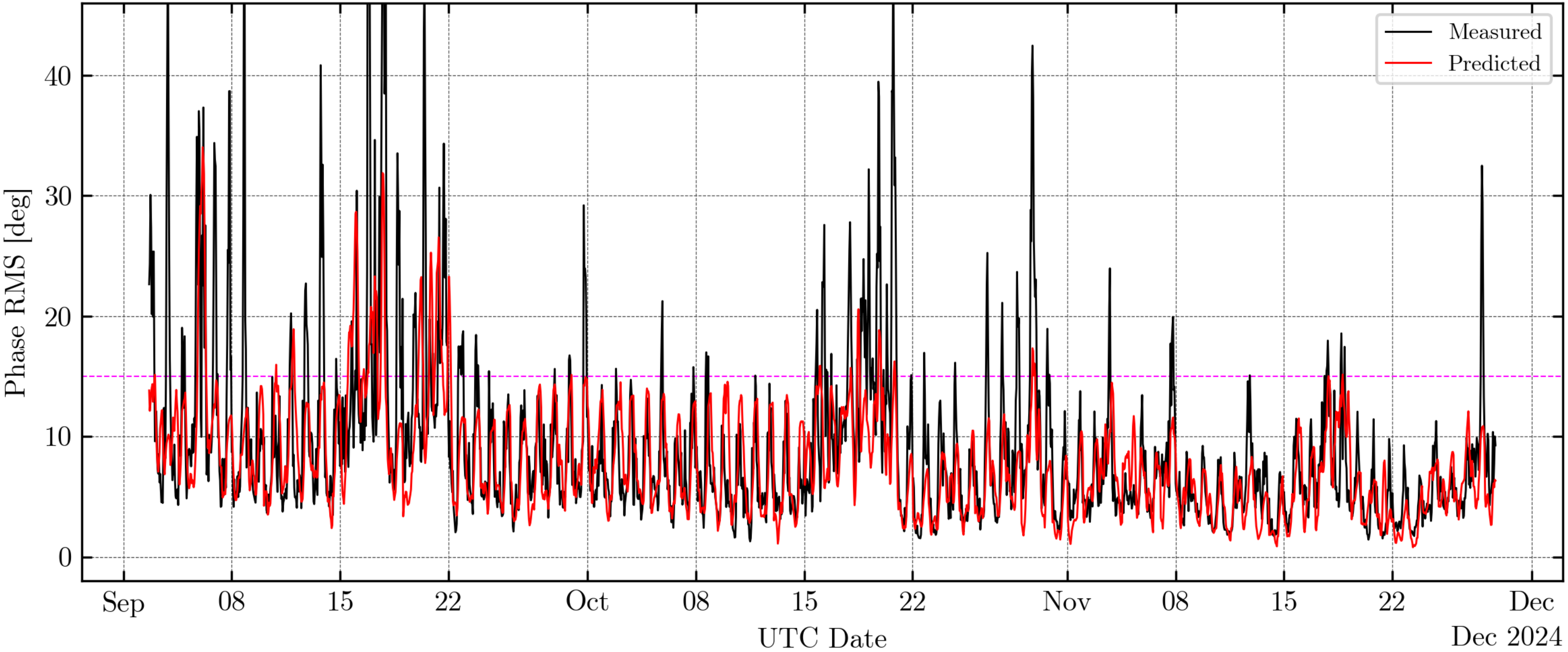
- Supervised machine learning approach is effective for forecasting the phase stability at the VLA site and can likely adapt to other sites in Southwest with Kernel Mean Matching with limited data.
- Dissimilar sites lack sufficient training data (i.e., an API).
- Forward modeling the phase RMS directly with FastEddy is possible but difficult (c.f., Stirling et al. 2005; ALMA Memo 517).
- **Can we develop a simple, physically motivated expression?**
- Advantages:
  - Can more effectively generalize to other sites (VLBA, GBT, EHT, ngVLA) because the NWP models embed the relevant physics.
  - Can perhaps capture more difficult to predict "events" by using full vertical profiles of model-level quantities (TKE, shear).

Symbolic regression and equation discovery



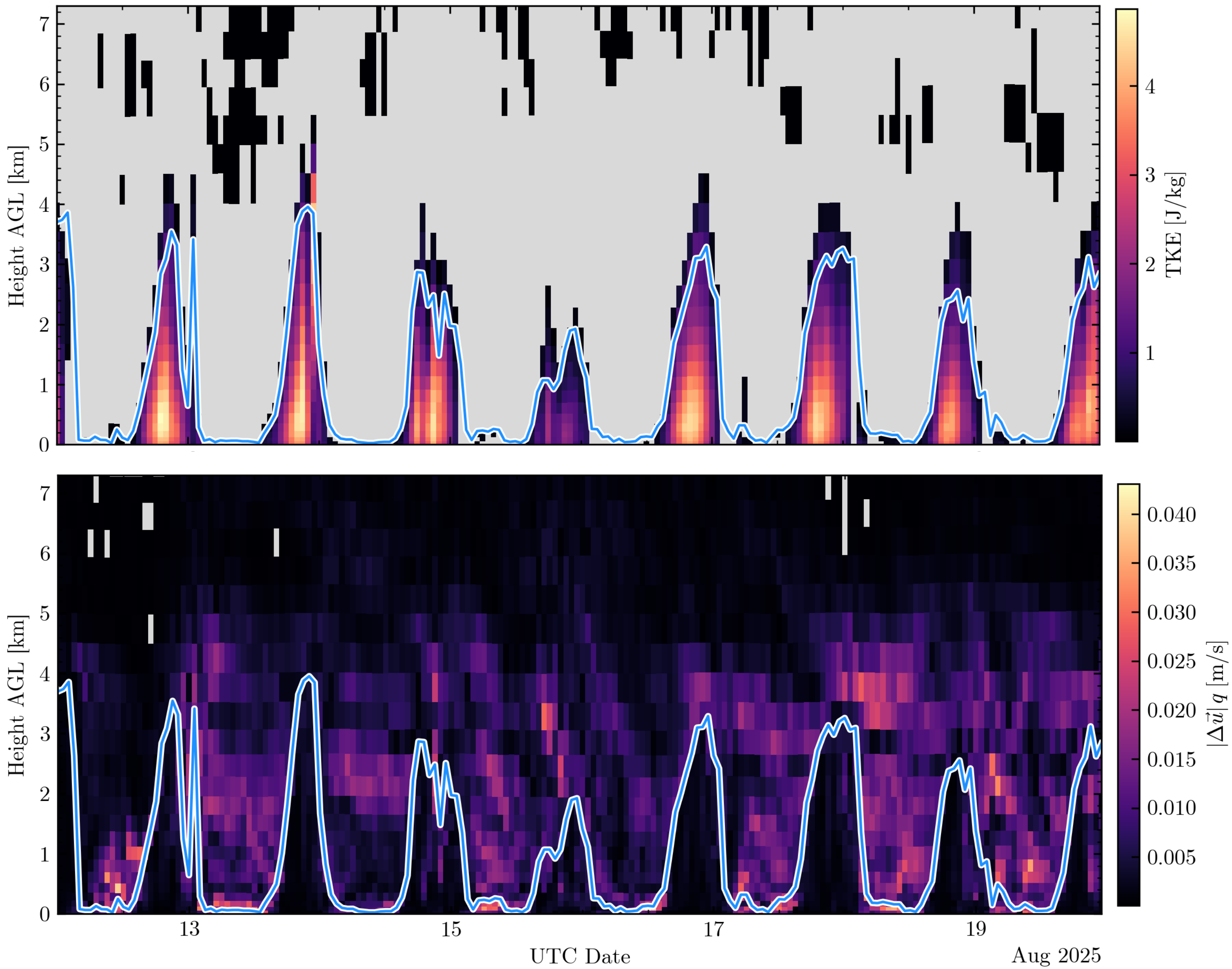
Simple expressions on a small number of physically motivated quantities are the most well-suited to generalizing to different locations where the VLA site training data may be unsuitable. Use *symbolic regression* to find such expressions.

Image credit: https://github.com/MilesCranmer/PySR



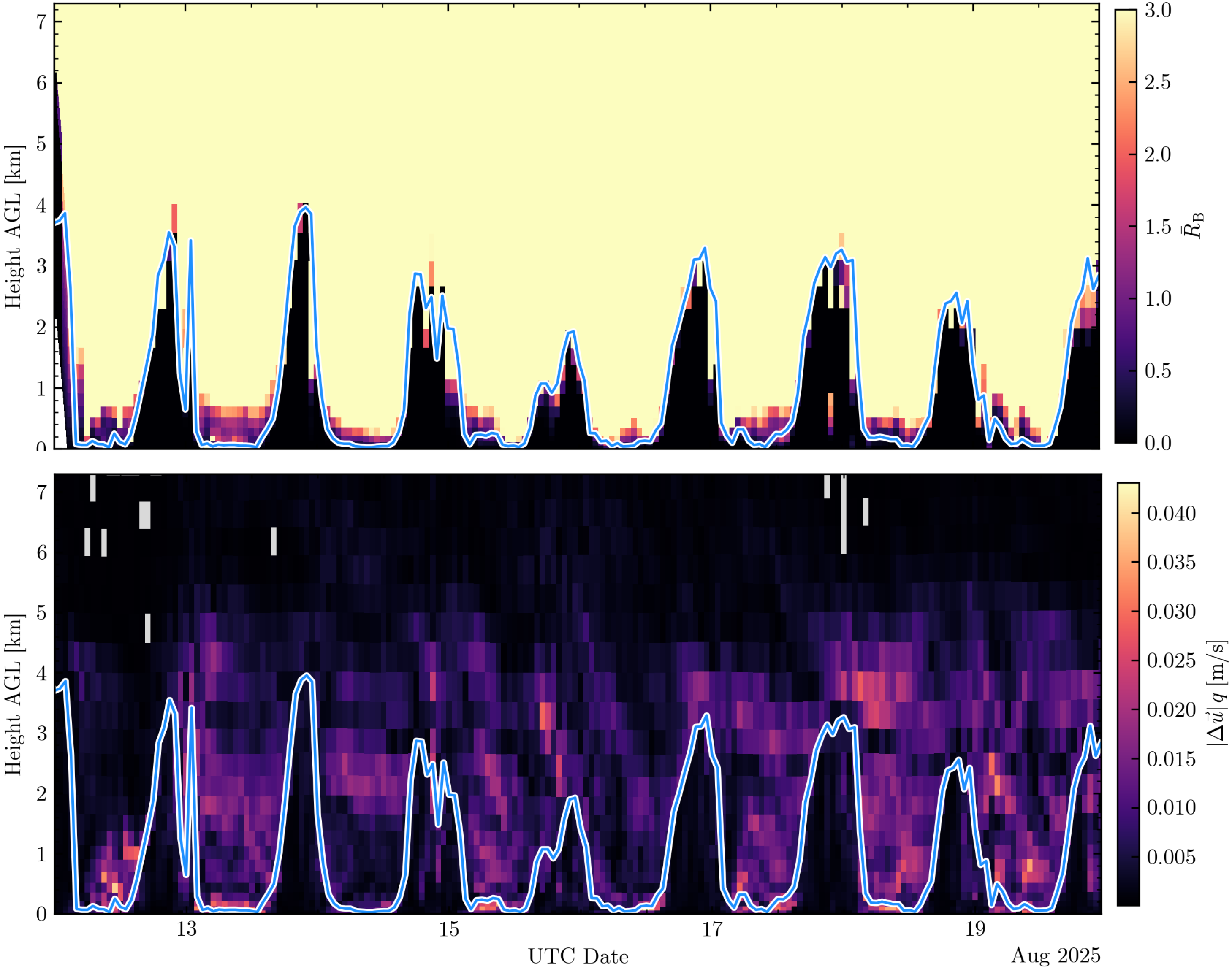
Application of equation discovery through symbolic regression. Results in a “pretty good” prediction of phase RMS using just **three** quantities in an expression of low complexity.

$$\sigma_\phi \propto \max(\sigma_{\rm pwv},\ [h_{\rm pwv} + h_{\rm bl}^{1/2}]_{\rm day})$$



Vertical profile of **turbulence kinetic energy** (TKE) with boundary layer height in blue.

Vertical profile of **wind shear magnitude** weighted by water vapor mixing ratio.



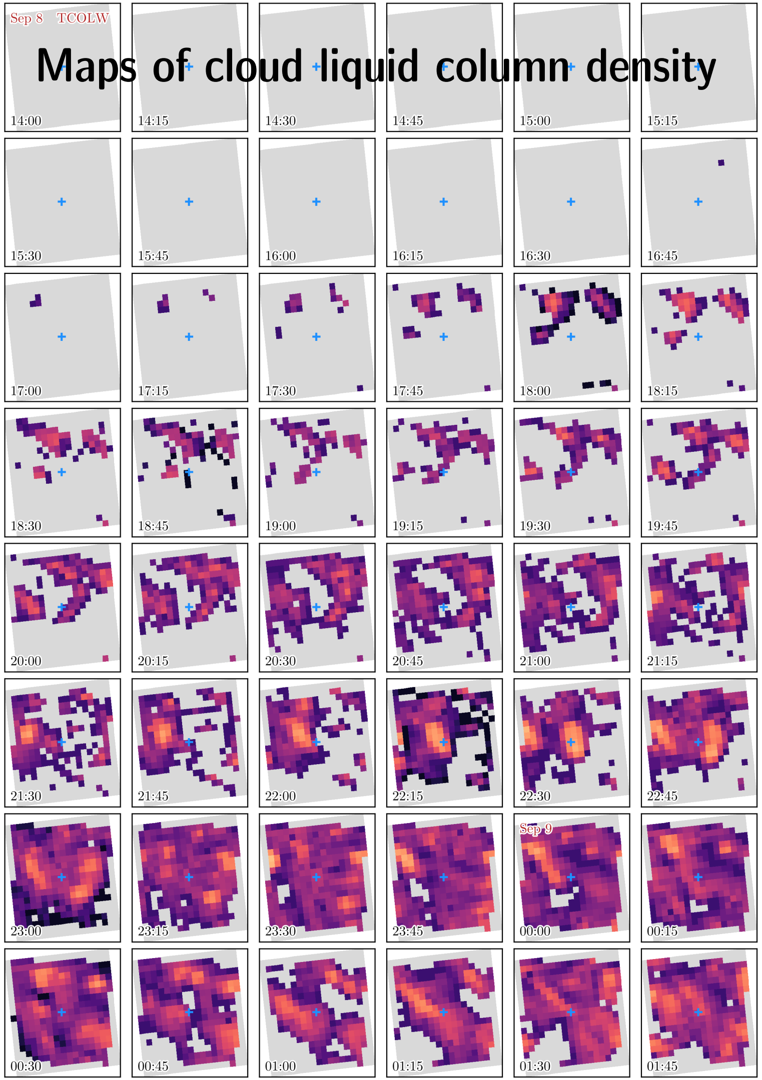
Vertical profile of **bulk Richardson's number**, predicts conditions unstable to turbulence less than 1.5

Vertical profile of **wind shear magnitude** weighted by water vapor mixing ratio.

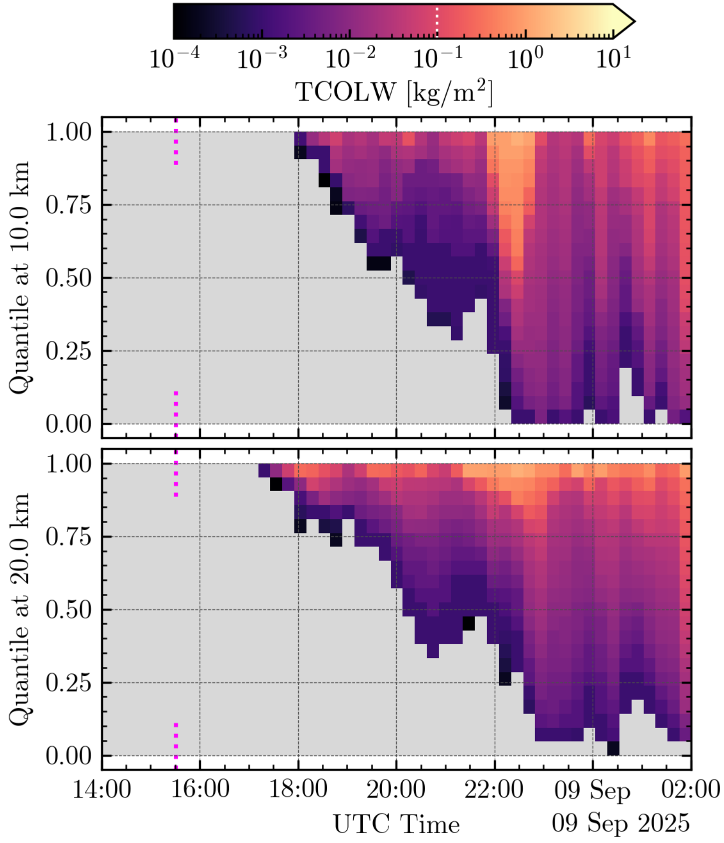
Summary and outlook

- Forecast guidance for the phase RMS is important for scheduling high-frequency observations (along with wind and opacity).
- We find that the phase RMS can be reasonably well predicted from numerical weather model quantities.
- We train a gradient boosting decision tree model using LightGBM/darts to "now-cast" the phase RMS and create longer term forecasts.
- System is running operationally: generates new forecast every hour.
- Work on a simpler, interpretable expression appears promising.
- Next steps are to engineer more physically motivated features from the model-native, vertical profile data. Aim to develop a system that can robustly predict sites other than the desert Southwest. Can validate on limited quantities of VLBA/GBT data.

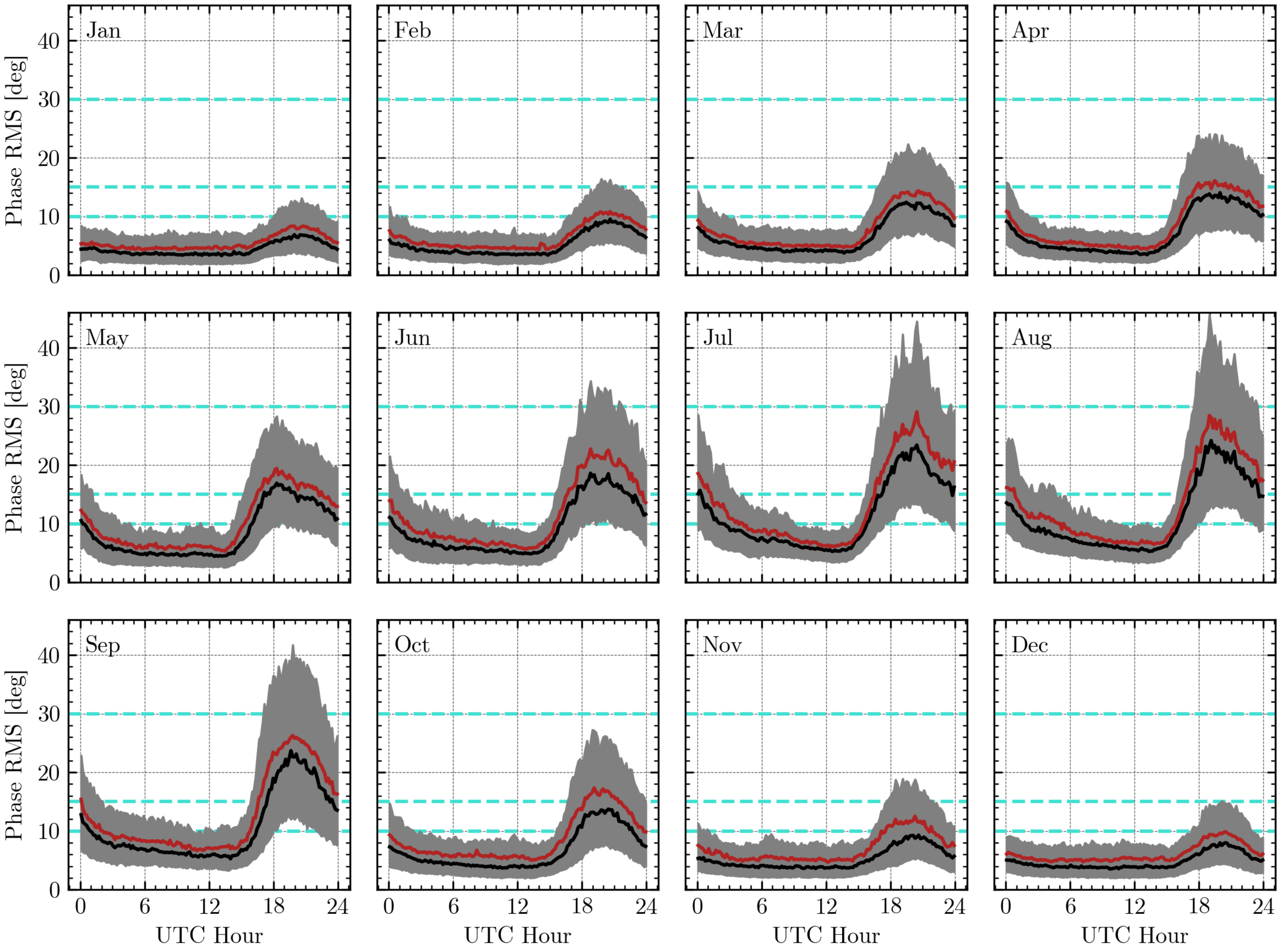
Maps of cloud liquid column density

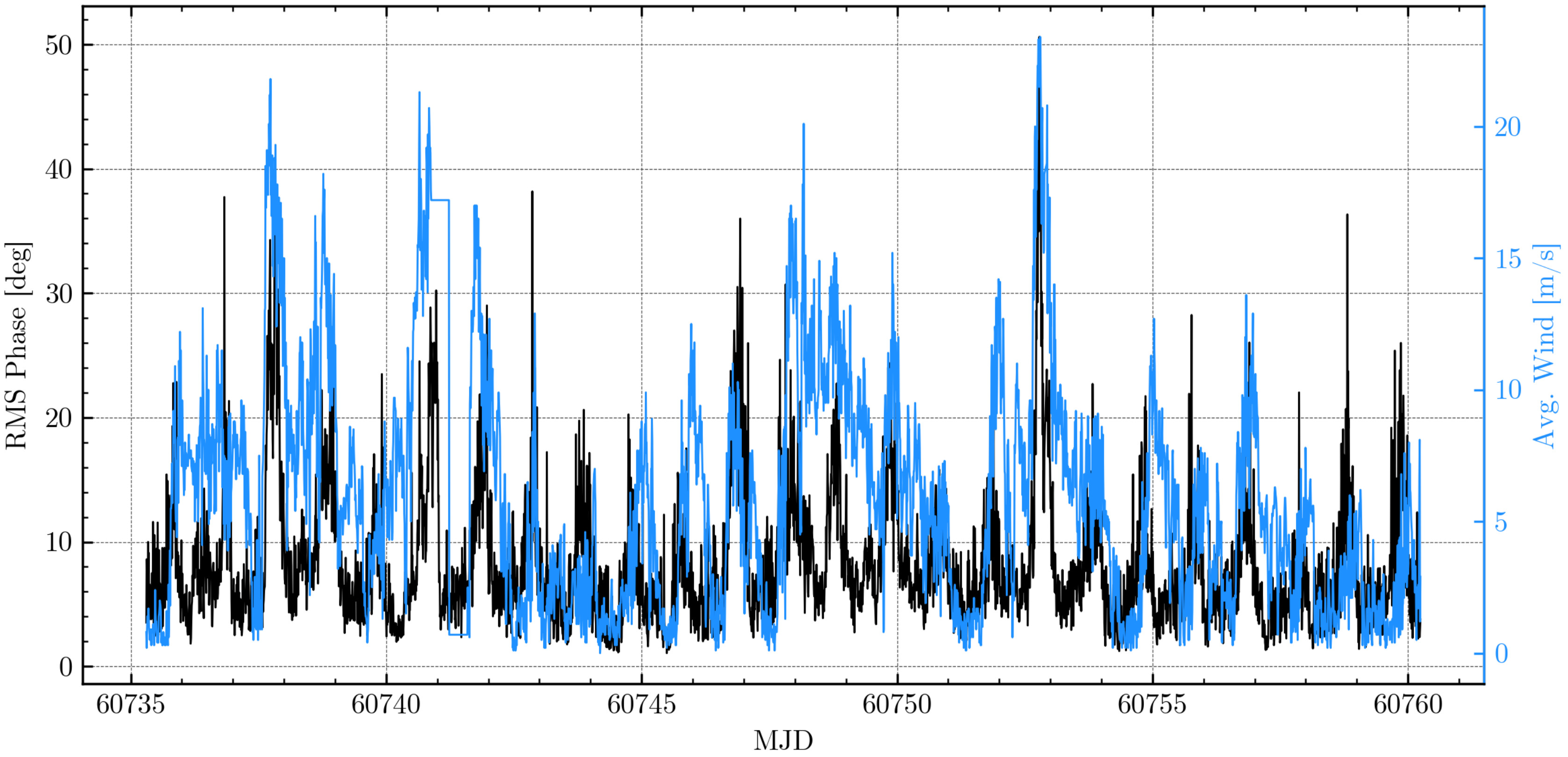


CDF of cloud column by time

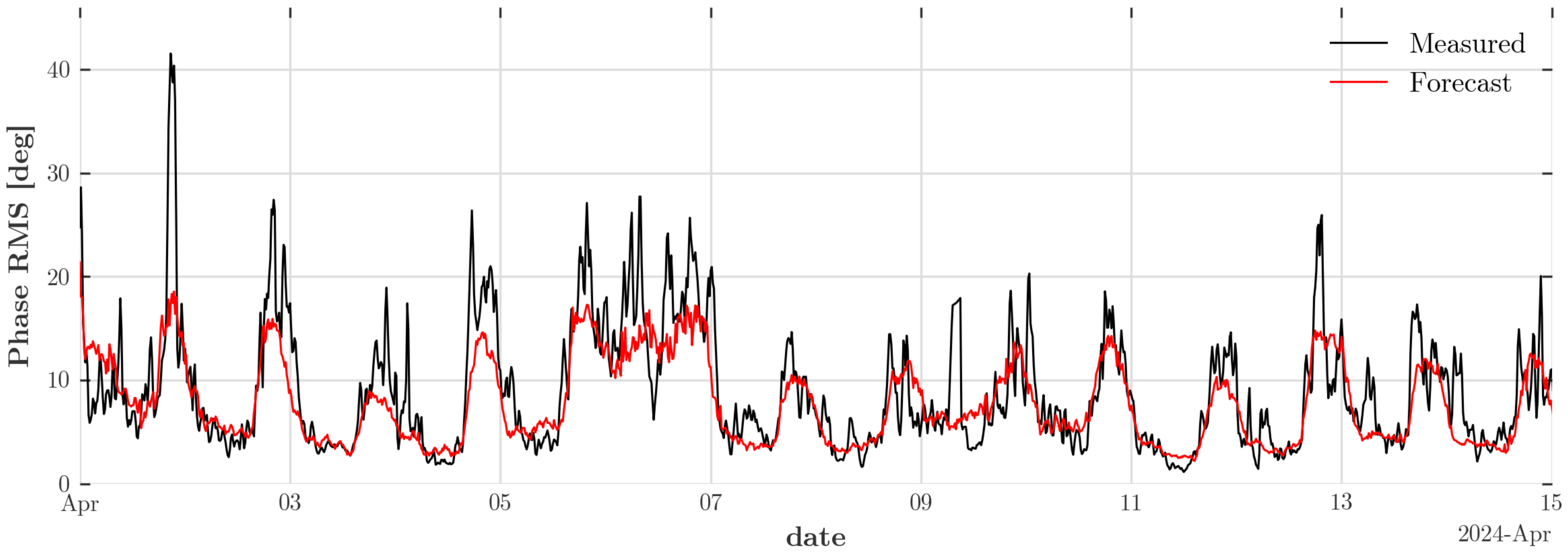


60 km, 15min

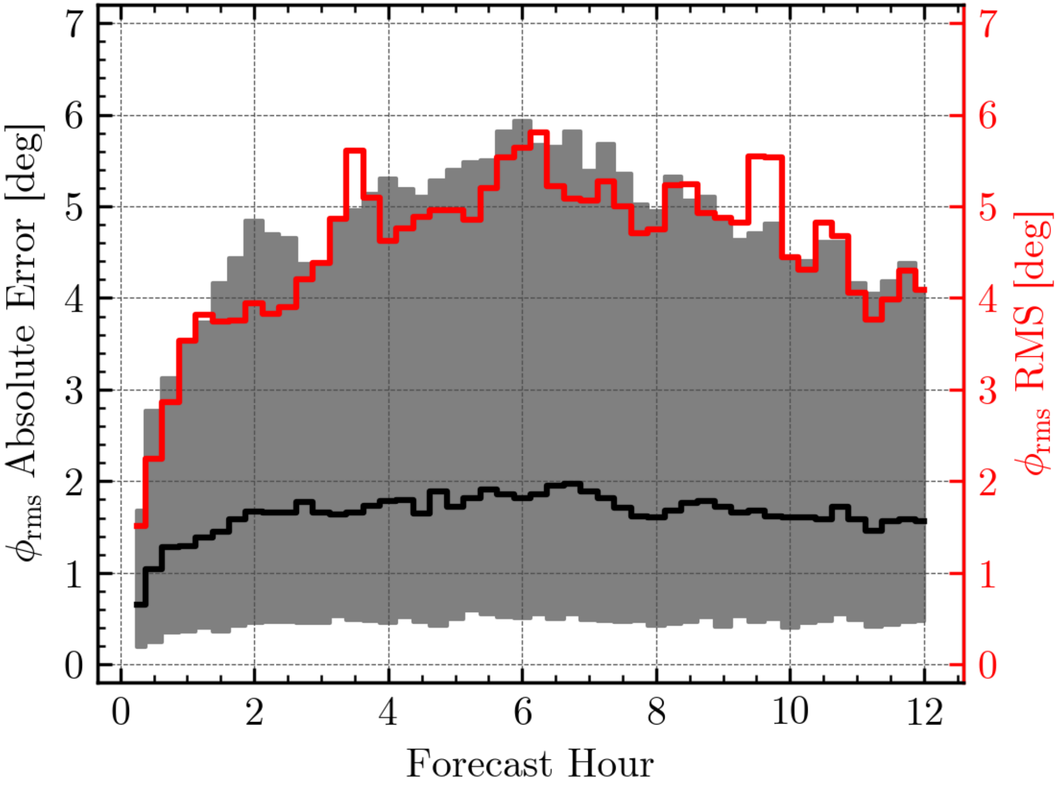

Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Phase RMS [deg]
0
10
20
30
40
0
6
12
18
24
UTC Hour



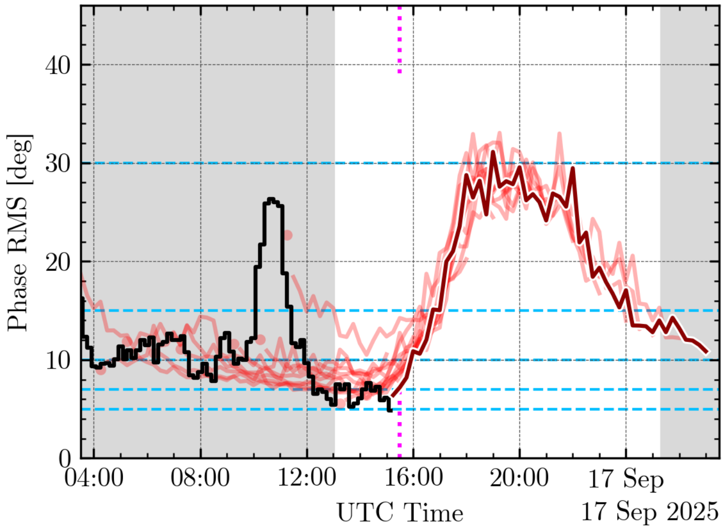
The above shows 25 days of API phase RMS data (10 min) and wind speed from the WXT520 (1 min). Absolute wind speed is generally a poor predictor. Although there is a diurnal trend to both wind and phase RMS, it corresponds especially poorly at night.

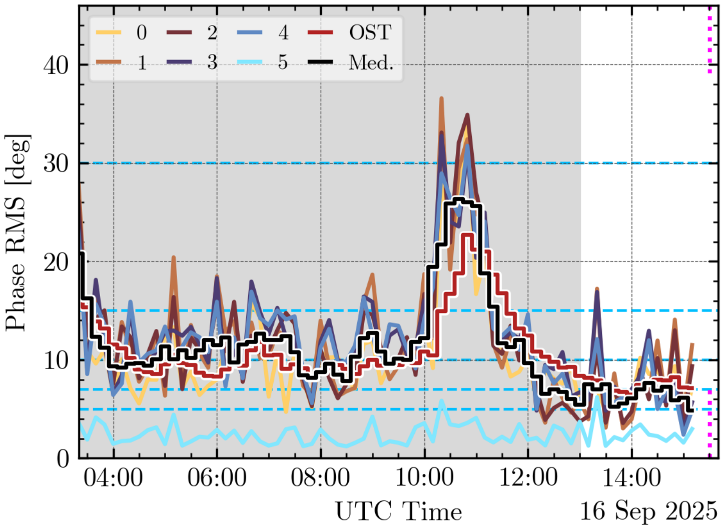


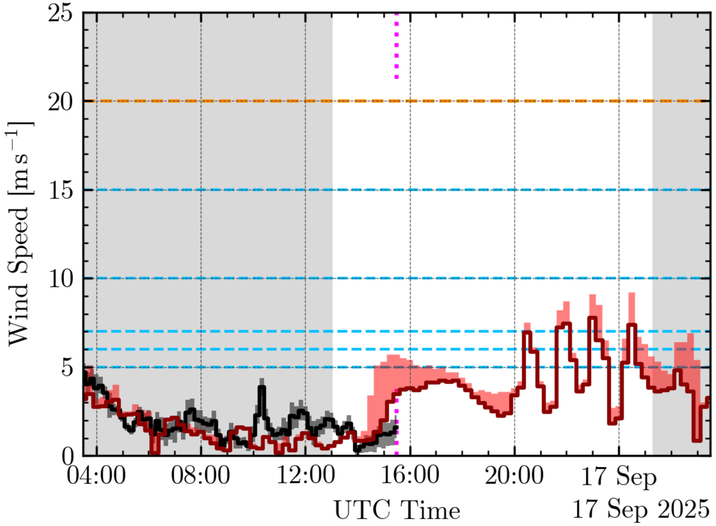
Forecasts applied to the validation dataset, or “backtests.” The above shows the measured phase RMS with 6-hour forecasts strung together for two months. Weather data used as covariates not shown. Generally quite effective but typically under-predicts large excursions in phase RMS.

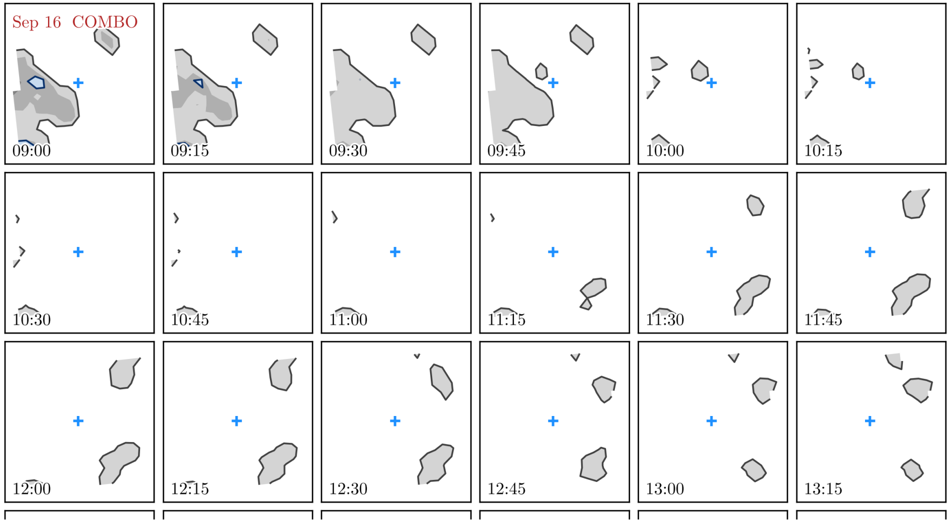


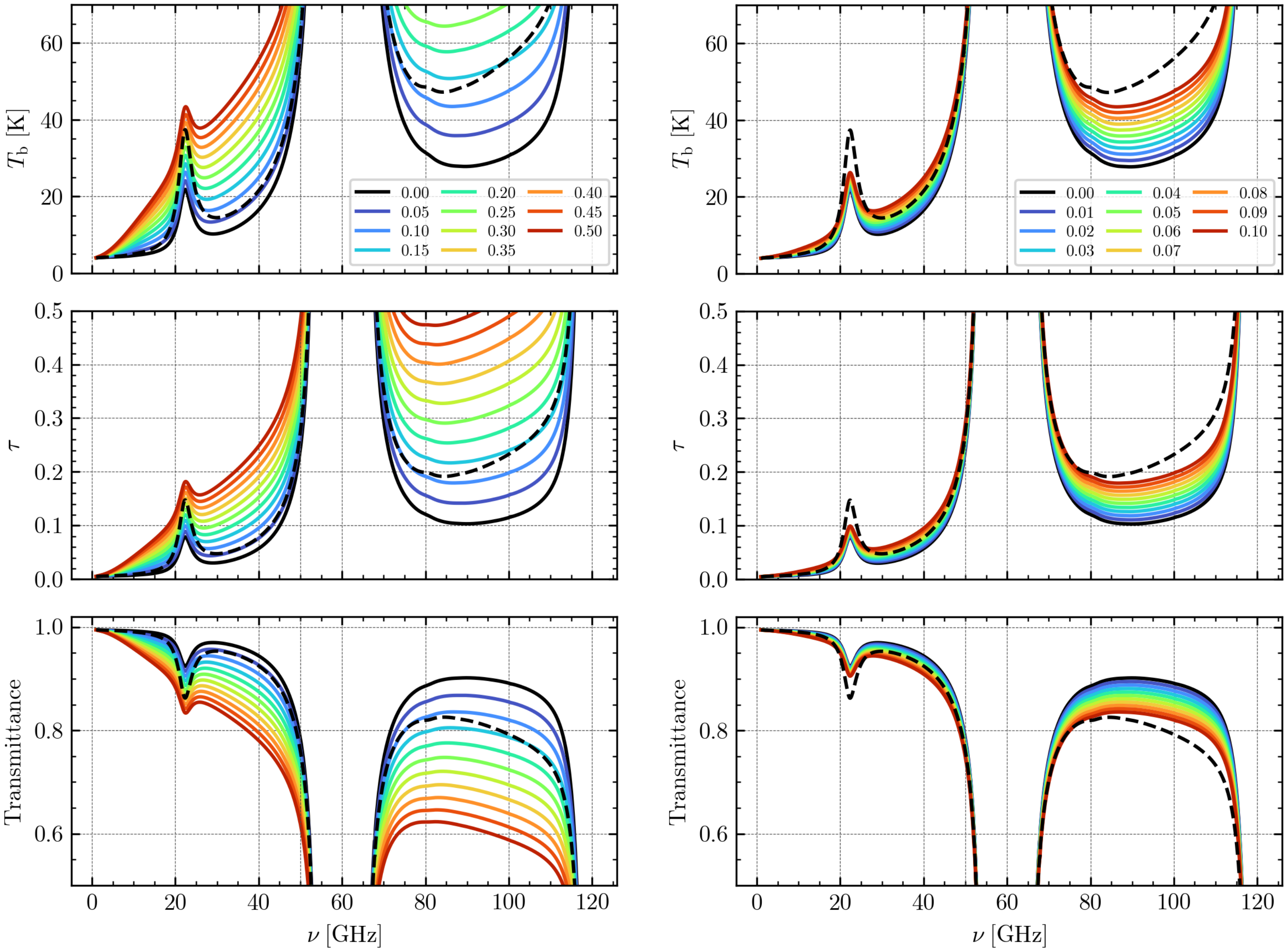
MAE and RMS of the forecasted phase RMS as a function of forecast hour. Near timescales of 1-2 hours benefit from "anchoring" against the measured phase RMS, but on longer timescales converges to a typical error reflecting the prediction from the weather data. Values aggregated for all times are shown, but errors are likely to be different for dry and wet seasons.


Phase RMS [deg]
0
10
20
30
40
04:00
08:00
12:00
16:00
20:00
17 Sep
UTC Time
17 Sep 2025


0
1
2
3
4
5
OST
Med.
Phase RMS [deg]
0
10
20
30
40
04:00
06:00
08:00
10:00
12:00
14:00
UTC Time
16 Sep 2025


Wind Speed [m s^{-1}]
0
5
10
15
20
25
04:00
08:00
12:00
16:00
20:00
17 Sep
UTC Time
17 Sep 2025


Sep 16 COMBO
09:00
09:15
09:30
09:45
10:00
10:15
10:30
10:45
11:00
11:15
11:30
11:45
12:00
12:15
12:30
12:45
13:00
13:15



T_b [K]
τ
Transmittance
ν [GHz]
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.10