TROPICAL VARIABILITY IN GRAPHCAST VS GFSV16: WHAT DOES THE NEURAL NETWORK LEARN?

MARIA GEHNE maria.gehne@colorado.edu, VIJIT MAITHEL, JULIANA DIAS, BRANDON WOLDING

CIRES Cooperative Institute for Research in Environmental Sciences

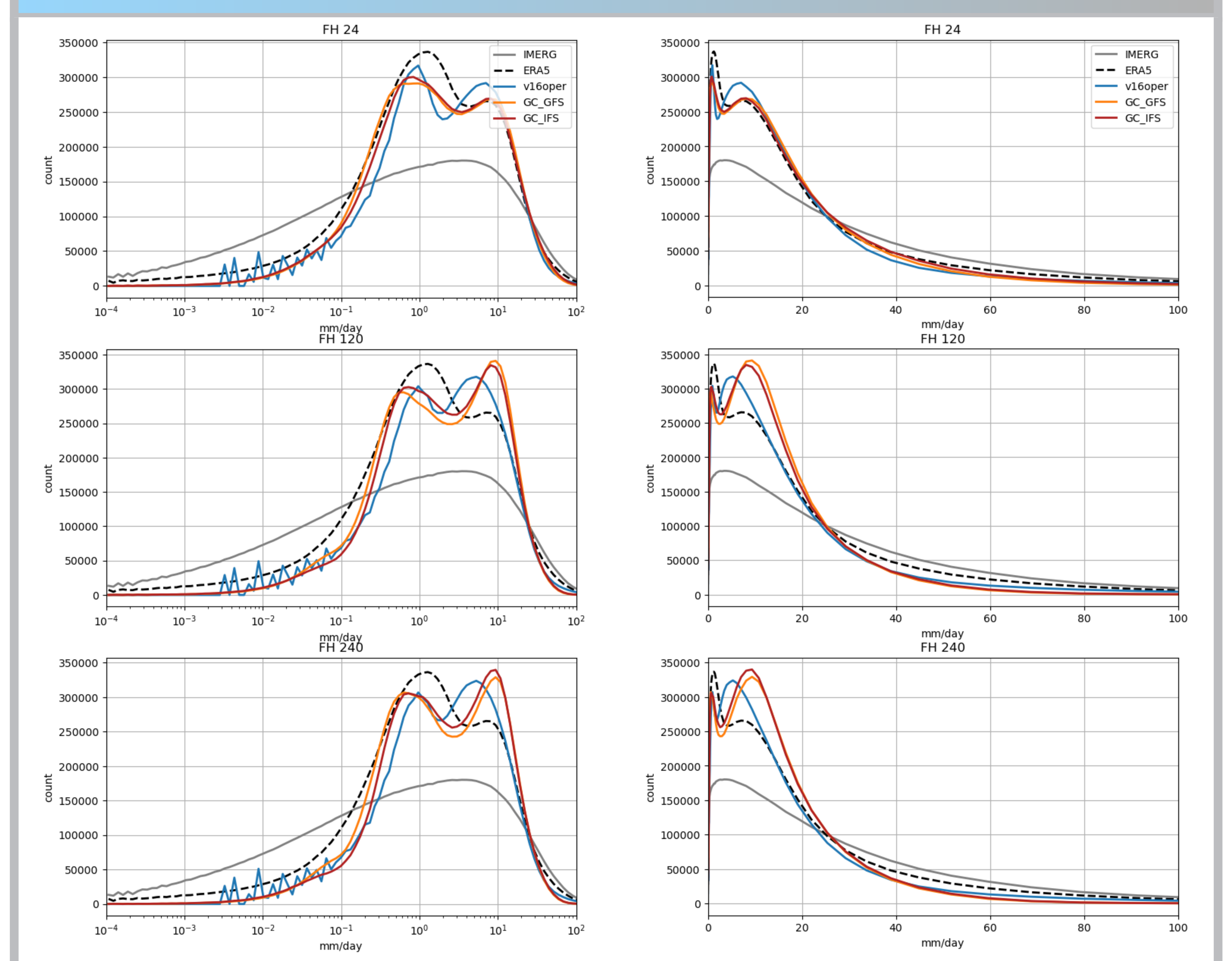
SUMMARY

Machine Learning based weather prediction (AINWP) is becoming increasingly common. AINWP relies on reanalysis, and thus traditional numerical weather prediction (NWP), for training data sets, but has clear advantages over NWP in that forecasts are very cheap to run once the model has been trained. Initial evaluation of AINWP shows that it can provide forecasts at comparable or even higher accuracy than NWP.

Here we investigate a set of 10 day forecasts from the GFSv16 operational model and compare those to forecasts from GraphCast with two sets of initial conditions (GraphCast (GFS) and GraphCast (IFS)) for the time period January 2022 - September 2024. Daily initializations during this time allow for skill evaluation by lead time.

We show that these models have very different underlying behavior in their tropical large-scale convection and coupling to dynamics. GraphCast forecasts have similar precipitation biases to ERA5, lower occurence of higher precipitation rates and stronger precipitation-circulation coupling at CCEW scales. The latter should, in theory, translate to improved teleconnections to and potentially enhanced subseasonal skill in mid-latitude precipitation forecasts. This is currently under further investigation.

PRECIPITATION RATE DISTRIBUTION



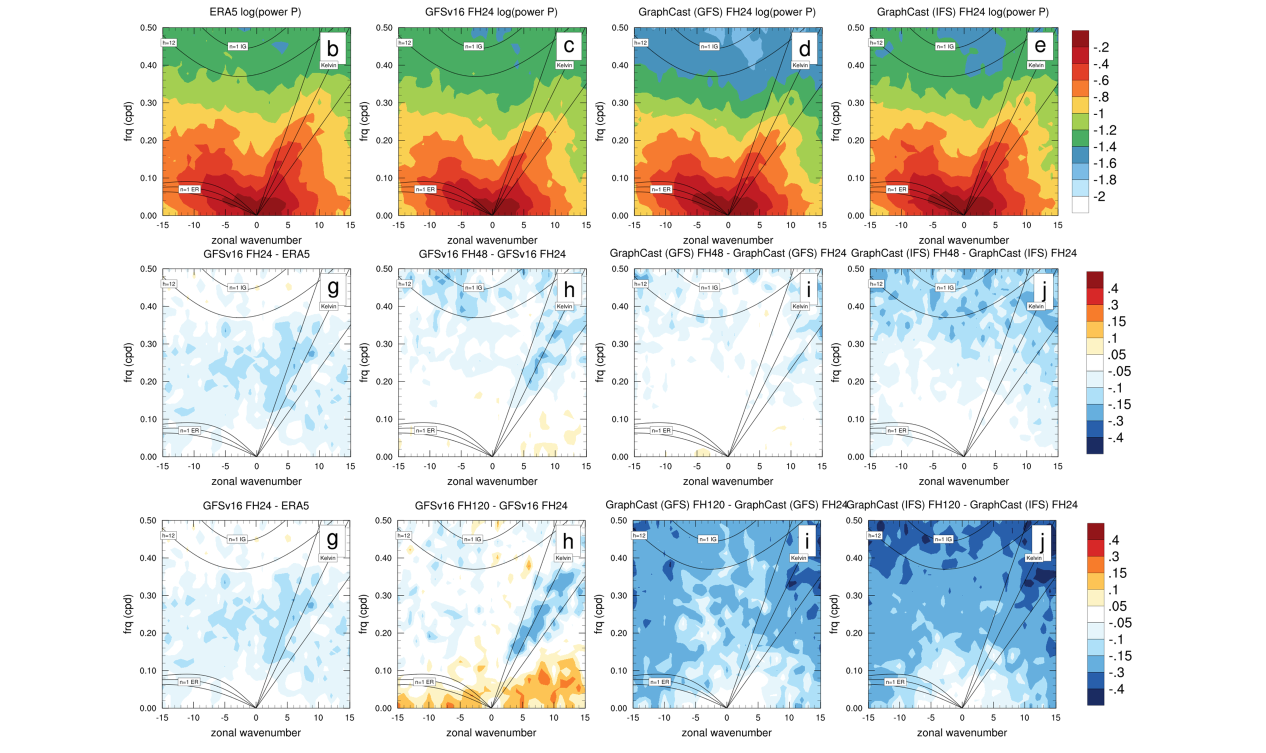
Distribution of precipitation rates using logarithmic bin sizes, plotted on log x axis.

Distribution of precipitation rates using logarithmic bin sizes, plotted on linear x axis.

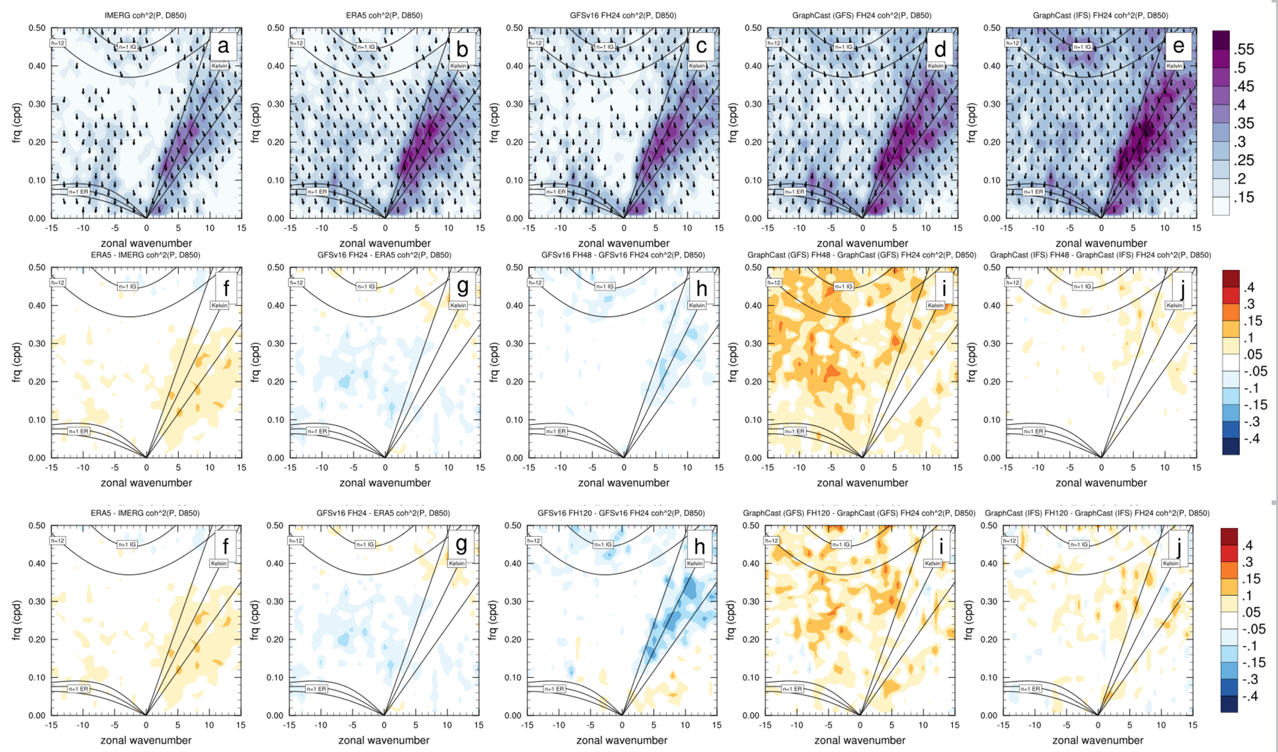
Models and **ERA5** precipitation show erroneous peak in low precipitation rates. At 24h lead time models match **ERA5** pretty well at low rates and GFSv16 underestimates the occurence of higher precipitation rates the most.

At later lead times GraphCast forecasts underestimate the occurence of higher precipitation rates the most. GFSv16 most common precipitation rates are the closest match to the **IMERG** peak location.

SPACE-TIME POWER AND COHERENCE SPECTRA



Precipitation power spectra at 24h lead time (top). Difference between precipitation power at 48h and 24h lead time (middle), and difference between precipitation power at 120h and 24h lead time (bottom).

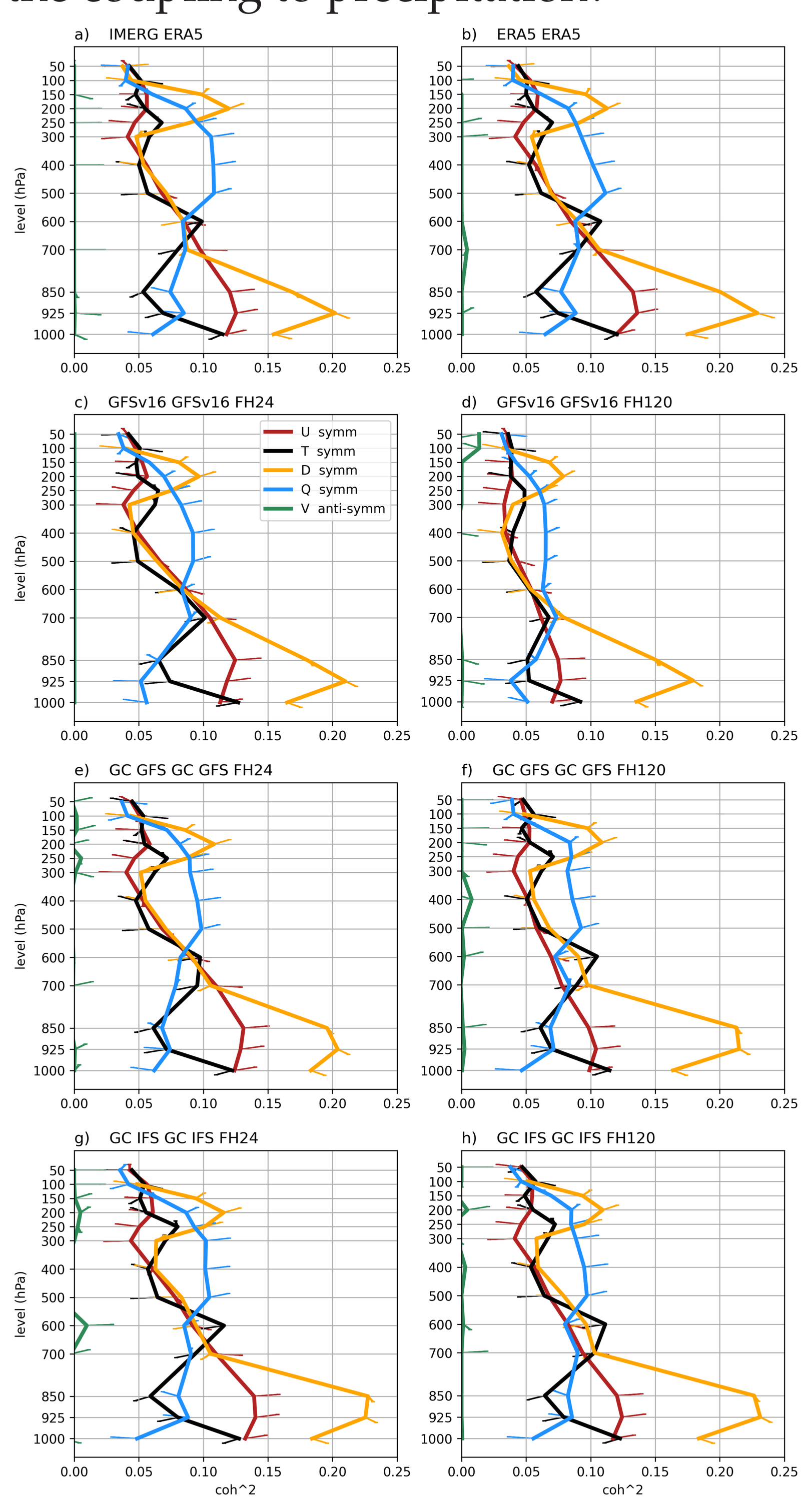


Coherence-squared spectra for precipitation and divergence at 850hPa ($\mathrm{coh}^2(P, D850)$) at 24h lead time (top). Difference between $\mathrm{coh}^2(P, D850)$ at 48h and 24h lead time (middle), and difference between $\mathrm{coh}^2(P, D850)$ at 120h and 24h lead time (bottom).

GraphCast shows decreasing variance at high frequencies and wave numbers early in the forecast, while GFSv16 shows decreasing variance in regions of CCEWs activity. Coherence between precipitation and divergence in GraphCast increases with lead time independent of scale, while GFSv16 coherence decreases mainly in regions of CCEWs activity.

VERTICAL COHERENCE CCE KELVIN WAVE

How well do the models represent the vertical structure of the coupling to precipitation?



Observed coherence in **Kelvin** wave band.

GFSv16 at lead time 24h and 120h.

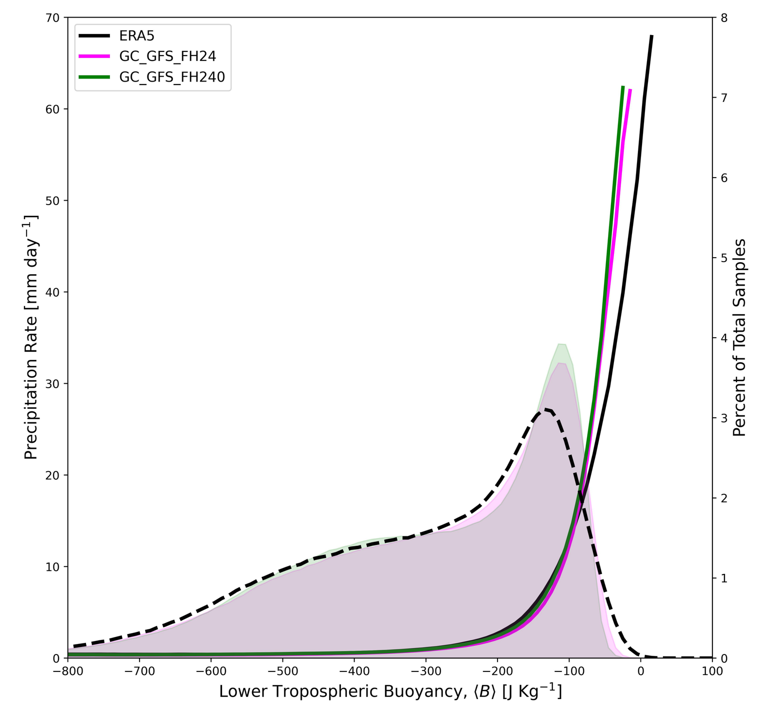
GraphCast (GFS) at lead time 24h and 120h.

GraphCast (IFS) at lead time 24h and 120h.

GFSv16 initializes with sufficient strength in the coherence between precipitation and dynamics, but looses the coherence strength quickly with lead time. GraphCast is able to keep the coherence strength (maybe due to loss of small scale variability).

THERMODYNAMICS

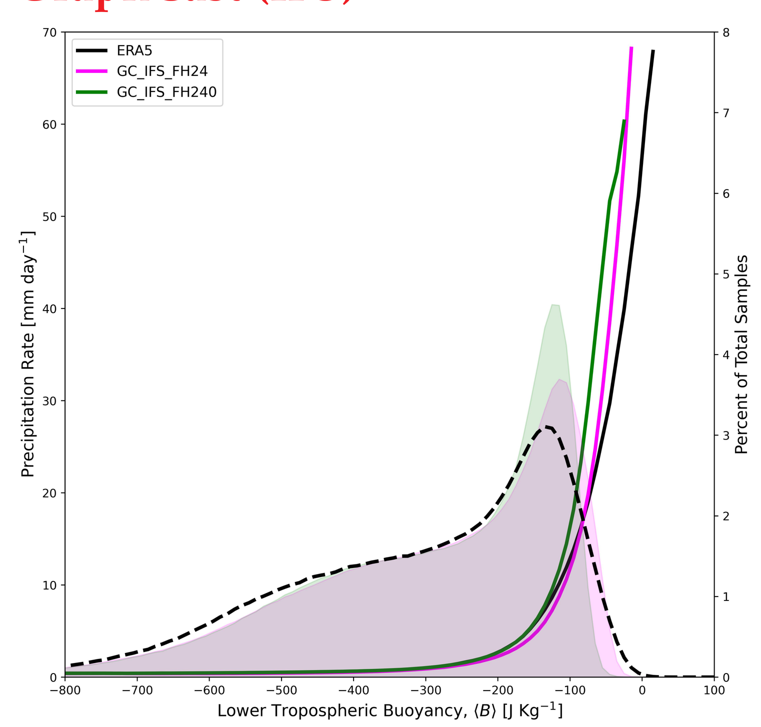
GraphCast (GFS)



Lower tropospheric buoyancy distribution and precipitation rate binned by buoyancy.

- Only small changes in distribution and precipitation rate with lead time.

GraphCast (IFS)



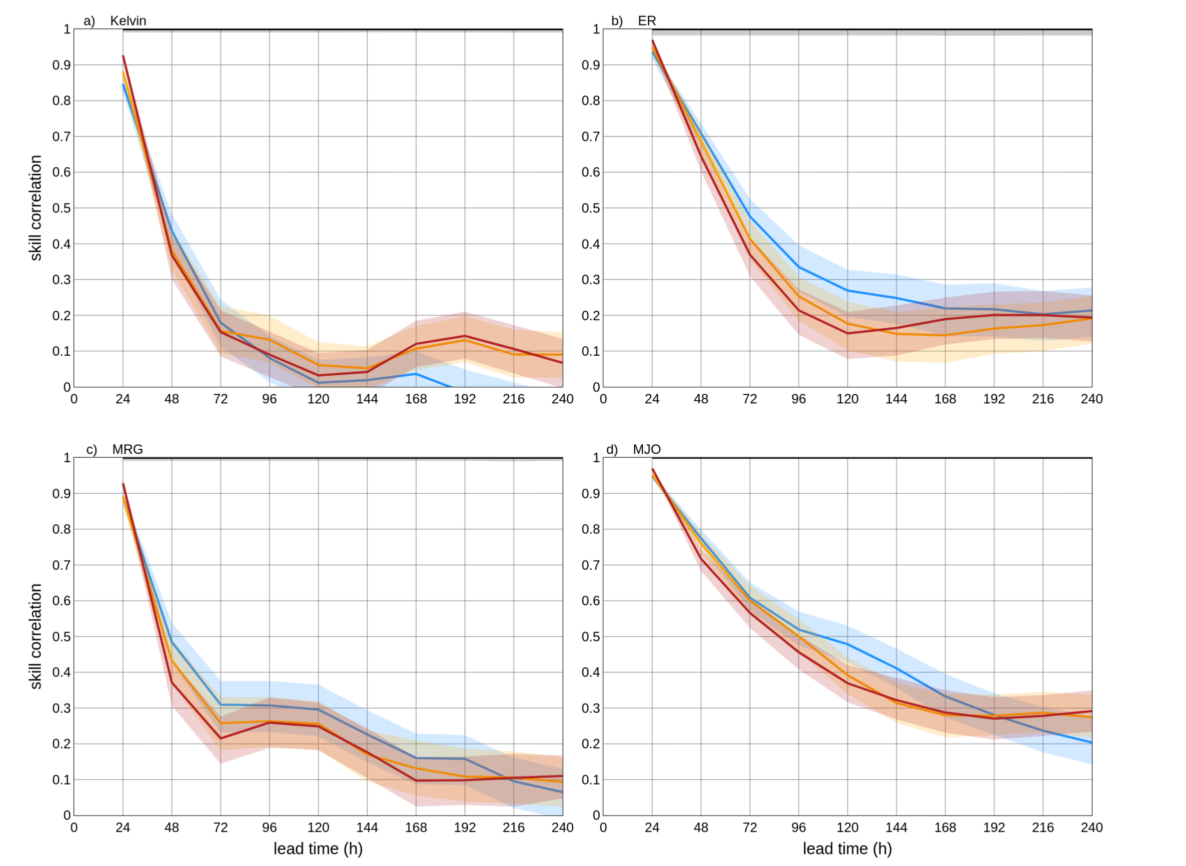
Lower tropospheric buoyancy distribution and precipitation rate binned by buoyancy.

- Shift toward lower buoyancy with lead time.
- Higher precipitation rates occur at lower buoyancy at later lead times.

Impact of the initial conditions lasts at least 10 days into the forecast. Both GraphCast (GFS) and GraphCast (IFS) underestimate the buoyancy at high precipitation rates compared to **ERA5**.

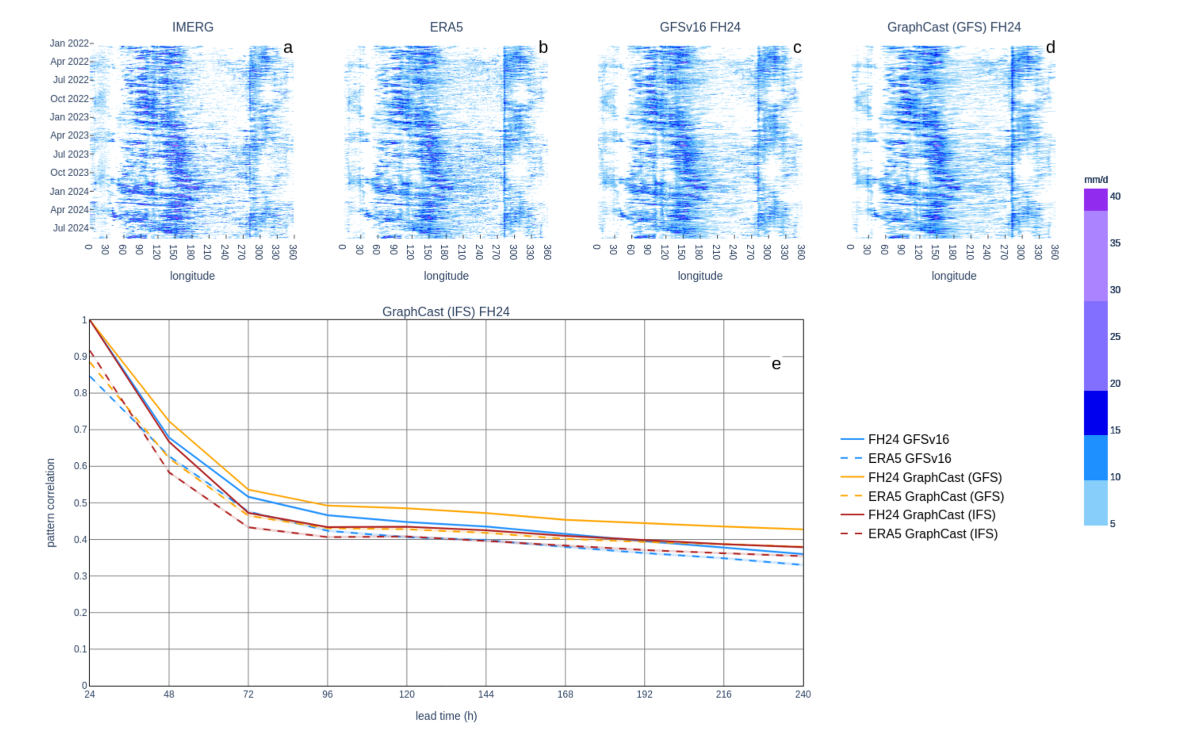
PRECIPITATION SKILL

CCEW activity skill



Comparable skill for GFSv16 and GraphCast. Better skill for GFSv16 at lead time 96h for ER?

Hovmoeller Pattern Correlation



ERA5 verification: Better skill for GraphCast at lead time 24h, comparable skill for GFSv16 and GraphCast (GFS) at longer leads.

ACKNOWLEDGEMENTS

Citations

Gehne et al., 2022: Diagnostics of Tropical Variability for Numerical Weather Forecasts. Wea. Forecasting, 37, 1661-1680, https://doi.org/10.1175/WAF-D-21-0204.1.

Wolding,et al., 2023: Atmosphere-Ocean Coupled Energy Budgets of Tropical Convective Discharge-Recharge Cycles. J. Atmos. Sci., 81, 3-29, https://doi.org/10.1175/JAS-D-23-0061.1.

Lam et al. (2023) Learning skillful medium-range global weather forecasting.Science382,1416-1421.DOI:10.1126/science.adi2336

Data

European Centre for Medium-Range Weather Forecasts. 2019, updated monthly. ERA5 Reanalysis (0.25 Degree Latitude-Longitude Grid). Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/BH6N-5N20. Accessed 12 Mar 2024.

IMERG precipitation: Huffman et al. (2023), GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V07, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [2025/01/14], 10.5067/GPM/IMERG/3B-HH/07

GraphCast Data: Radford et al. (2025) Accelerating Community-Wide Evaluation of AI Models for Global Weather Prediction by Facilitating Access to Model Output. Bull. Amer. Meteor. Soc., 106, E68-E76, https://doi.org/10.1175/BAMS-D-24-0057.1.