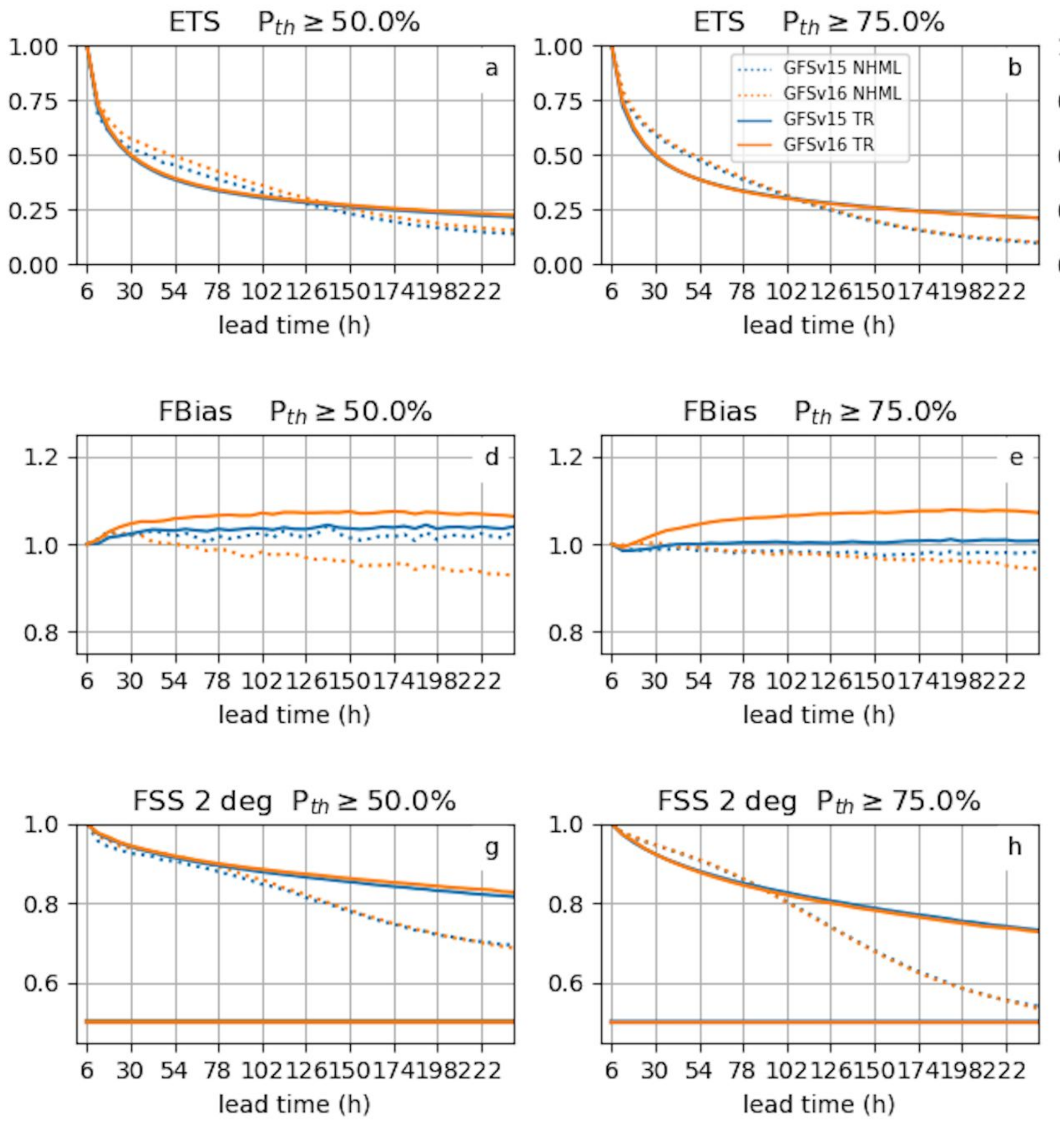
# Tropical Variability in the Unified Forecast System

**Maria Gehne[1,2], Brandon Wolding[1,2], Juliana Dias[2], George Kiladis[2]**

University of Colorado Boulder

NOAA

1 CIRES University of Colorado, Boulder
2 NOAA Physical Sciences Laboratory




NWP models tend to perform better in mid-latitudes than in the Tropics for lead times <4 days.

- The underlying dynamics are different in the Tropics and mid-latitudes.
- Convection is main driver of precipitation in the Tropics.
- Convective parameterization has a larger impact on precipitation in the Tropics.

There is evidence that better forecast skill in the Tropics can lead to improved forecasts in mid-latitudes.

# Evaluating tropical convection in NWP

It is not very well understood which processes in the Tropics are most important to mid-latitude forecast skill.

There are, however, well-known sources of predictability beyond a few days in the tropical atmosphere such as the MJO and Convectively Coupled Equatorial Waves (CCEWs).

## Consider metrics and diagnostics specifically for NWP in the Tropics:

- Better understanding of NWP model behavior with respect to tropical convection.
- Identify forecast error sources in the Tropics related to moisture-convection coupling, CCEWs and the MJO.
- We will look at variability and not biases in this presentation, although biases can be substantial at later lead times.

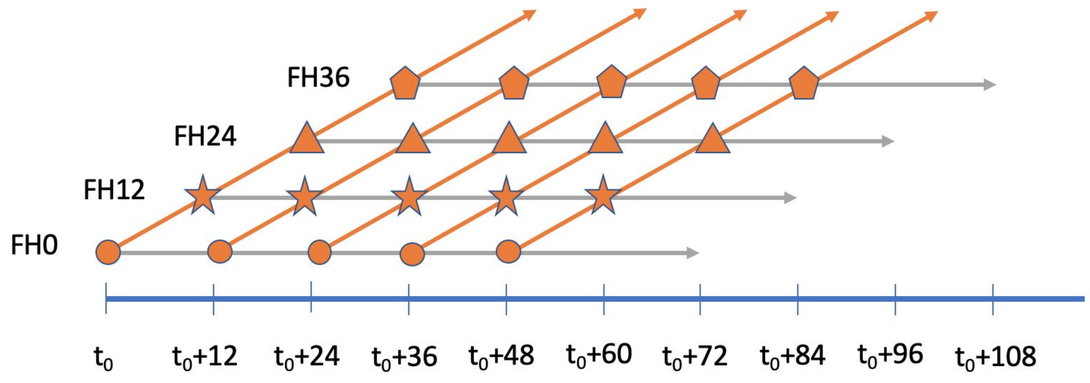
## NWP evaluation presents different challenges than climate model evaluation.

- Forecasts are shorter: days-weeks.
- Model versions change frequently.
- It is rare to have long (multi-year) time series of operational model runs.

## Consider diagnostics as a function of lead time.

If certain phenomena are initialized correctly, how long is the model able to keep that information?

# Diagnostics




Hovmoeller diagrams and pattern correlation
(zonal propagation)

Space-time coherence spectra
(scales of coupling to moisture)

Vertical structure of coherence between precipitation and dynamical fields
(vertical structure and phase relationship within CCEWs)

Convectively coupled wave activity and skill
(CCEW propagation)

Moisture - convection coupling
(coevolution of precipitation and column saturation fraction)

**Model output needed:**

- gridded 2D fields of precipitation, surface pressure, land-sea mask
- gridded 3D fields of temperature, specific humidity, winds

# Model runs

UFS coupled prototype (P5,7,8) runs - 168 initializations, every 1st and 15th of the month between 20110401 and 20180315.

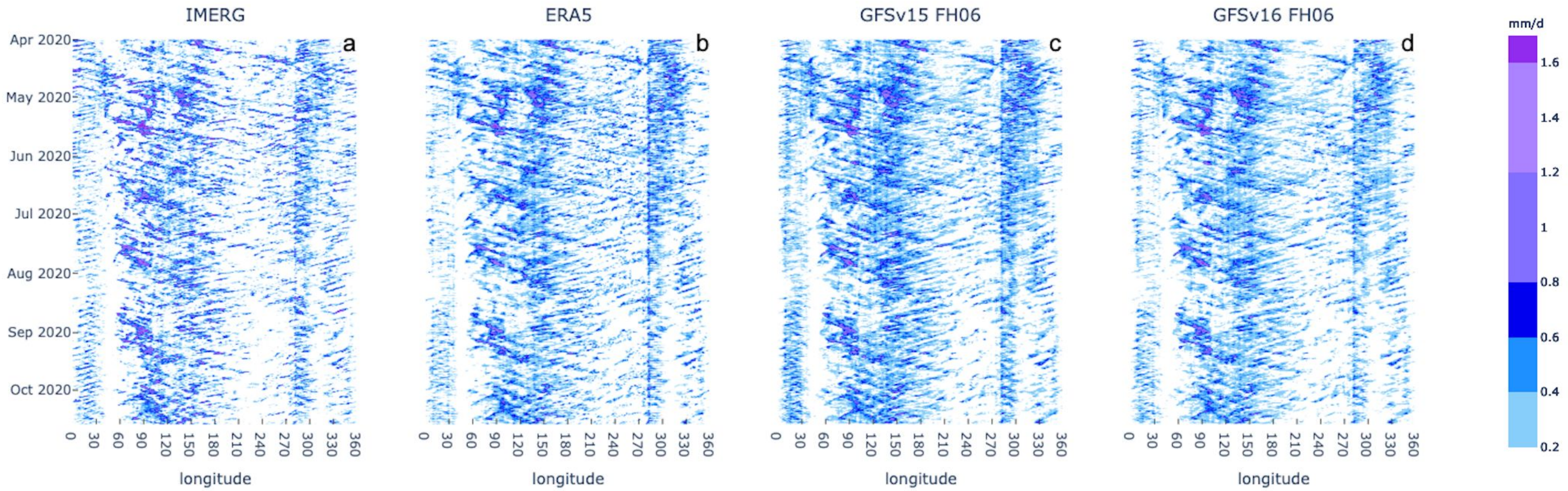
ECMWF S2S (2021 model version) (EC2021) database runs - only initializations within +-2 days of the UFS initializations.

FV3GFS V15 operational (GFSv15) and FV3GFS V16 parallel (GFSv16) runs initialized 6 hourly from April through October 2020 and run out to lead time 240h.

These are uncoupled forecasts.

More details on the GFS v15 and v16: https://www.emc.ncep.noaa.gov/emc/pages/numerical_forecast_systems/gfs.php
More details on the UFS prototypes:
https://registry.opendata.aws/noaa-ufs-s2s/#:~:text=The%20UFS%20prototypes%20are%20the,weather%20prediction%20system%20from%20NWS.
More details on the ECMFS S2S: https://confluence.ecmwf.int/display/S2S/ECMWF+model+description

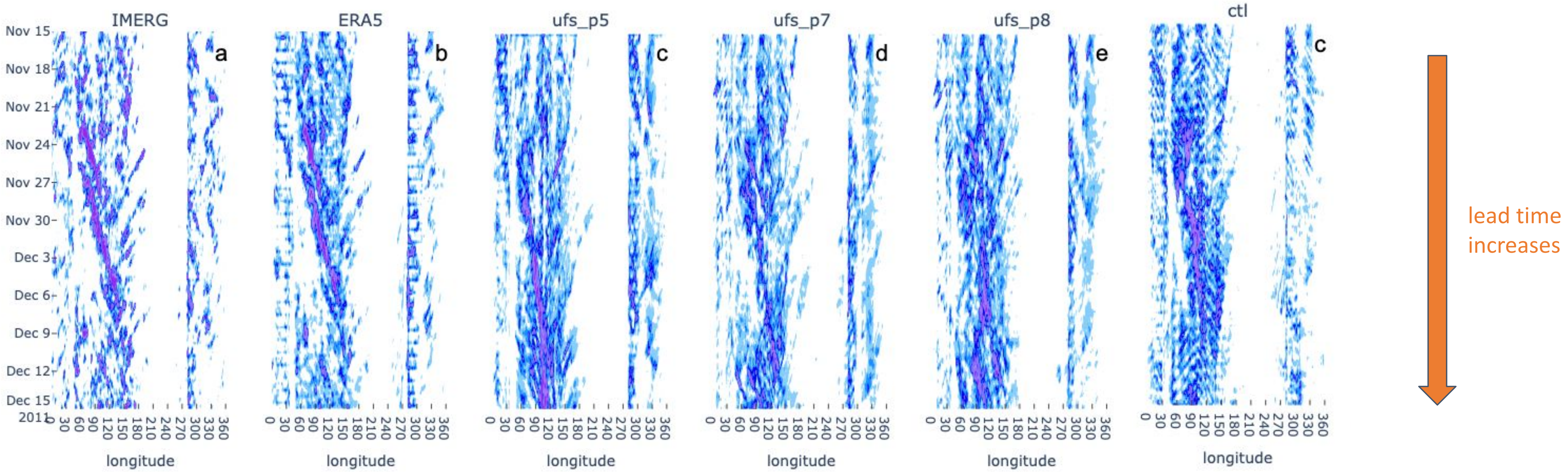
# Hovmoeller and Pattern Correlation




Assess the **zonal** propagation of convective features.

Pattern correlation between forecast and ‘truth’ can be used as a skill score.

# Hovmoeller and Pattern Correlation




30 day period in November - December 2011

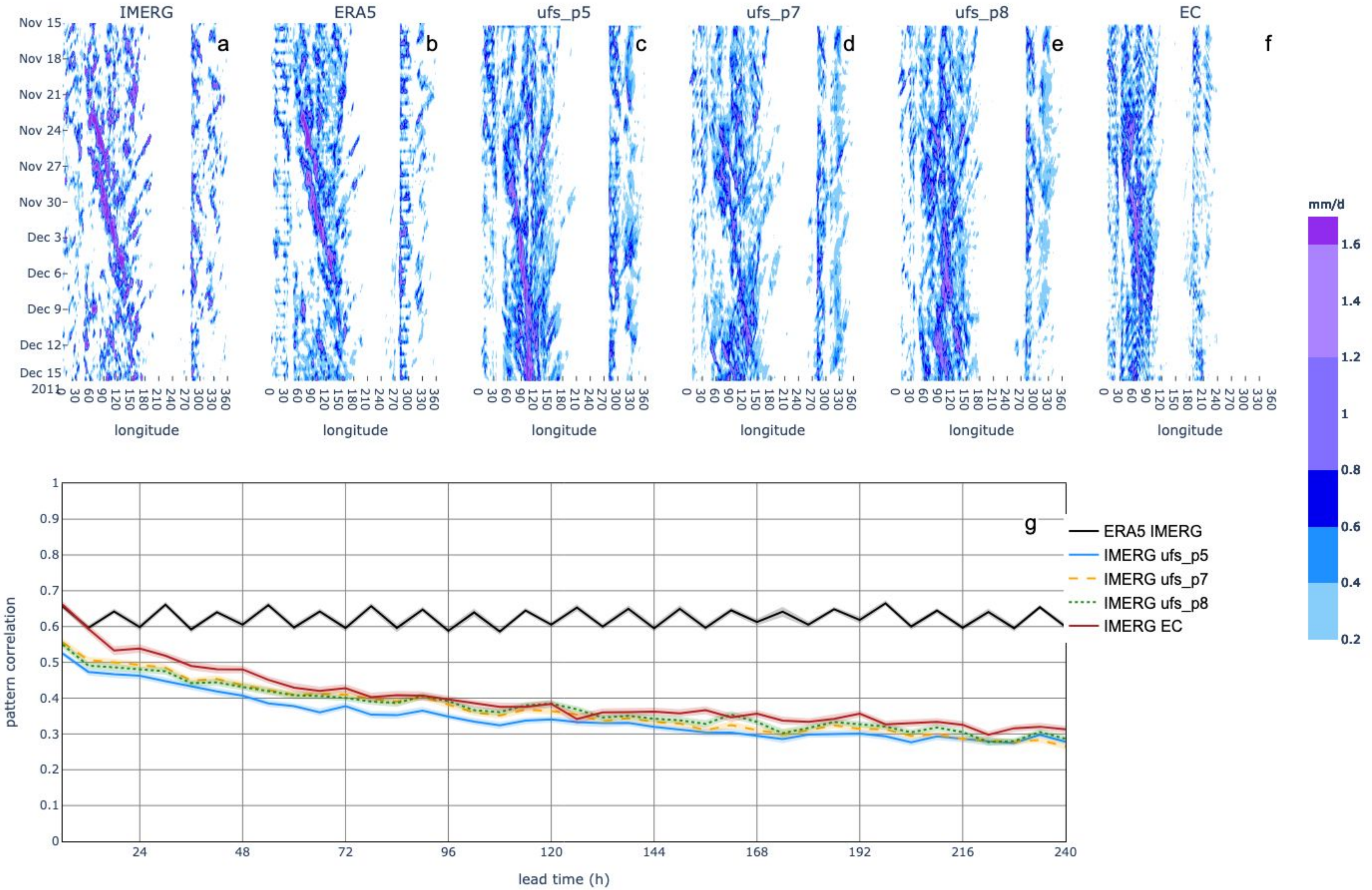
IMERG and ERA5 show the MJO event observed during DYNAMO starting around 11/22.

Model precipitation is plotted along the forecast instead of at a single lead time.

Model forecasts vary widely between models and ensemble members after a few days.

Some forecasts have an indication of enhanced convection during the observed MJO period and others don't.

# Hovmoeller and Pattern Correlation



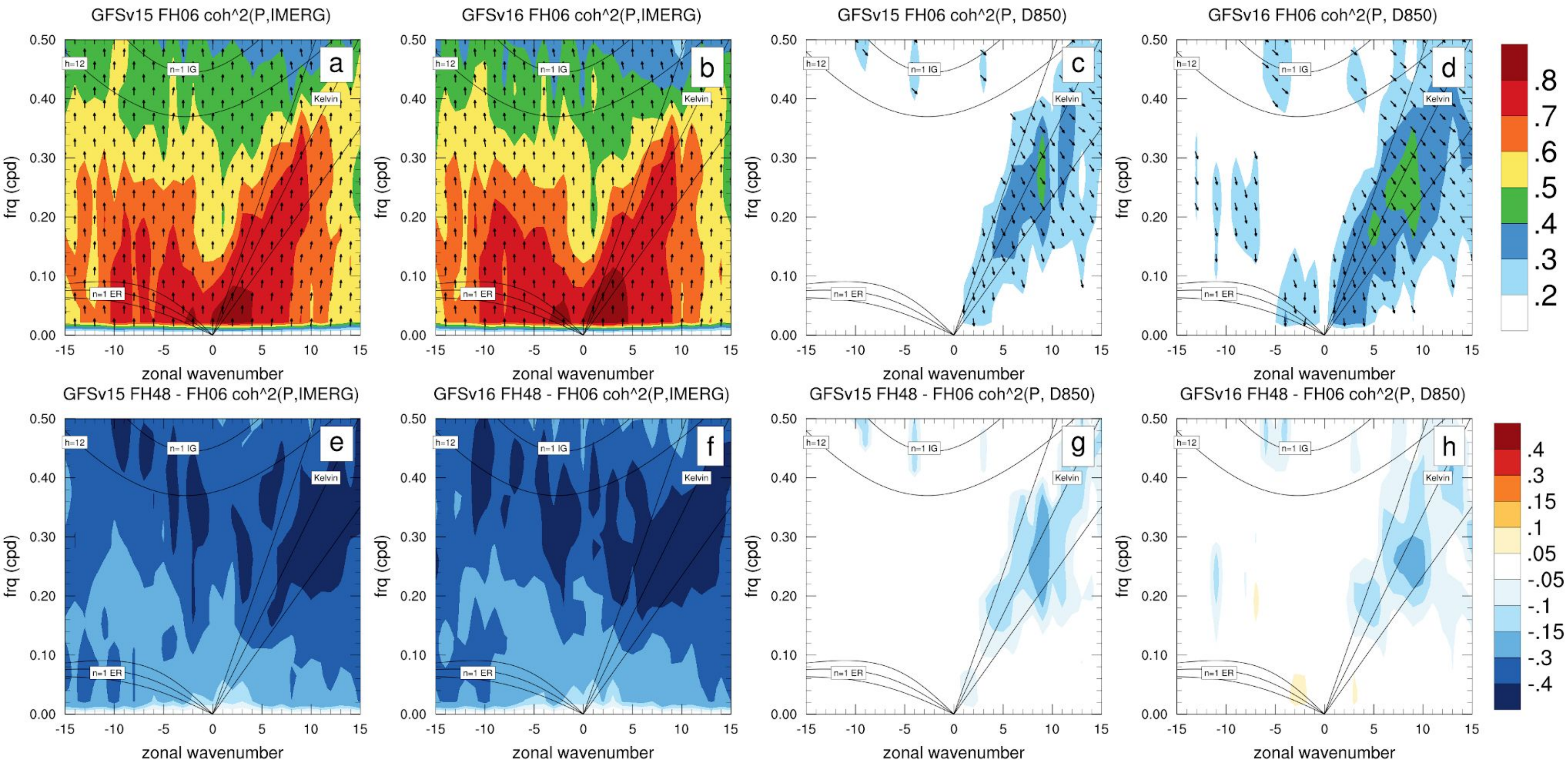

Comparing model 6h average precipitation rate to IMERGv6b precipitation. Correlations with ERA5 tend to be slightly higher.

Initial pattern correlation higher for P7 and P8 than P5.
About 6-12h improvement when verifying against ERA5 and IMERG.

Noise in the correlation is possibly due to the small sample size.

EC pattern correlation for EC control is comparable to ERA5-IMERG, skillful for 24h longer than UFS prototypes.

# Space-time coherence-squared spectra

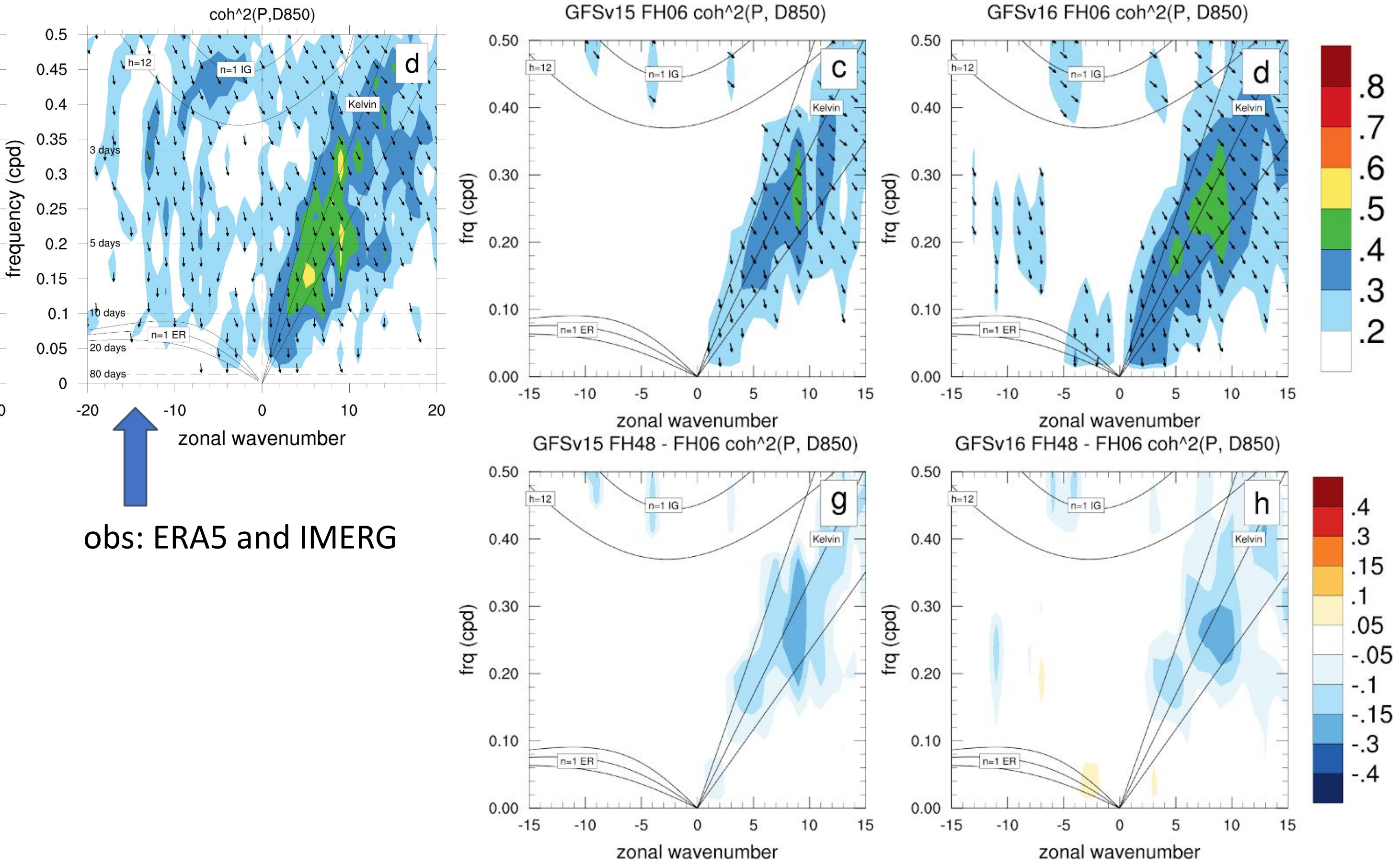



- How well do models initialize and propagate CCEWs?
- Coherence spectra show space-time regions of tropical variability without having to estimate a background.

Evaluate the consistency in variability between modeled and observed precipitation at a range of spatial and temporal scales.

It is possible to evaluate precipitation – dynamics relationship strength and how it changes with lead time.

# Space-time coherence spectra




There are distinct peaks in coherence along CCEW dispersion curves, but overall the model coherence tends to be lower than observed. Models tend to have peaks at slightly higher frequencies than the reanalysis and observations

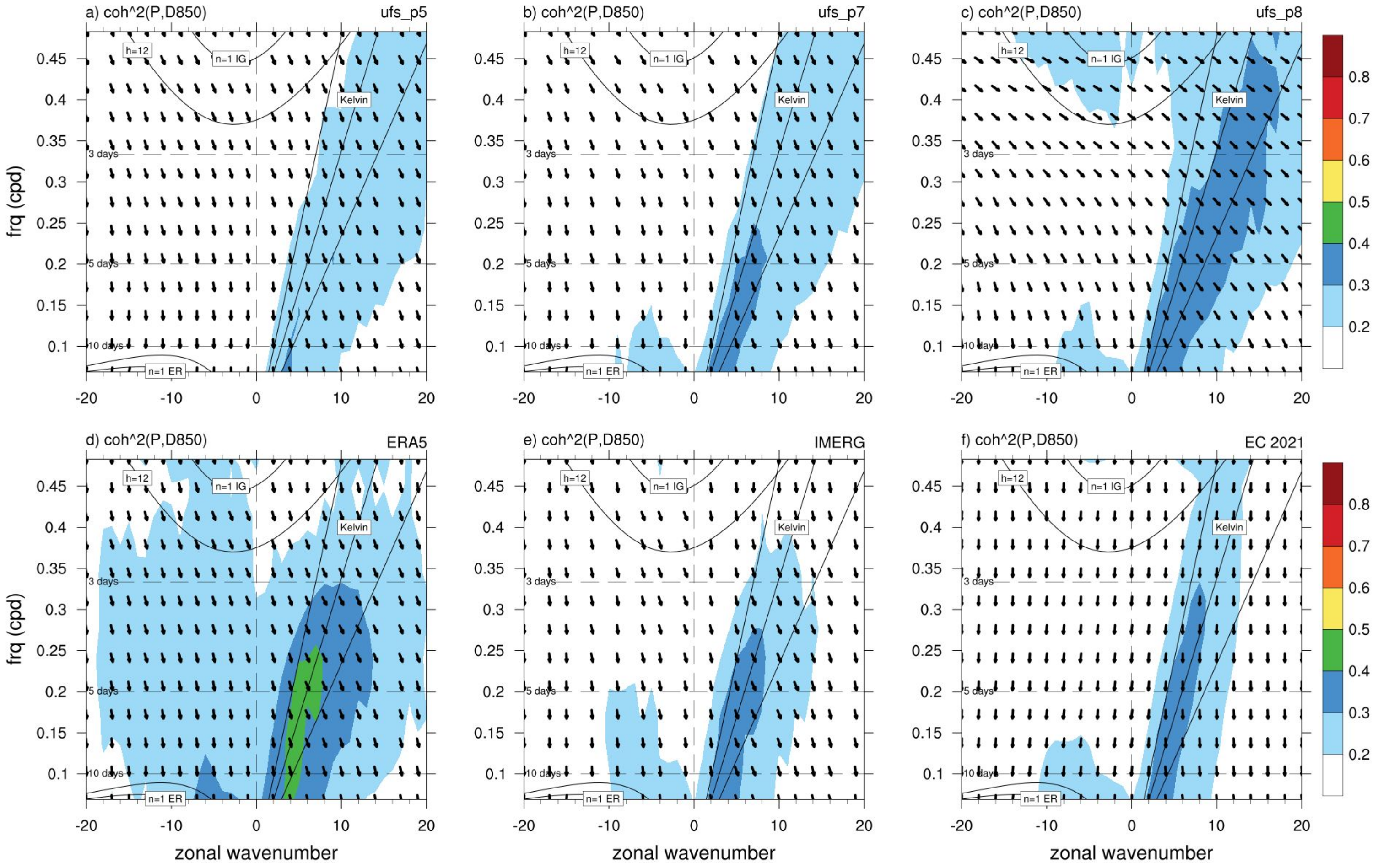
By 48h lead time coherence between precipitation and 850 hPa divergence at the peaks in the Kelvin wave band has decreased by 50-75% (GFSv15) and 30-50% (GFSv16).

Both model versions are able to initialize CCEWs, the coupling between moisture and dynamics is too weak even at initial time.
At longer lead time precipitation is not coupled strongly to the near-surface dynamics, although this is improved in GFSv16.

There is almost no coherence at very high frequencies.

Variability at higher frequencies and wavenumbers does not contribute much to S2S predictability although this activity could still be a source of feedback to the larger scales.

# Space-time Coherence Spectra

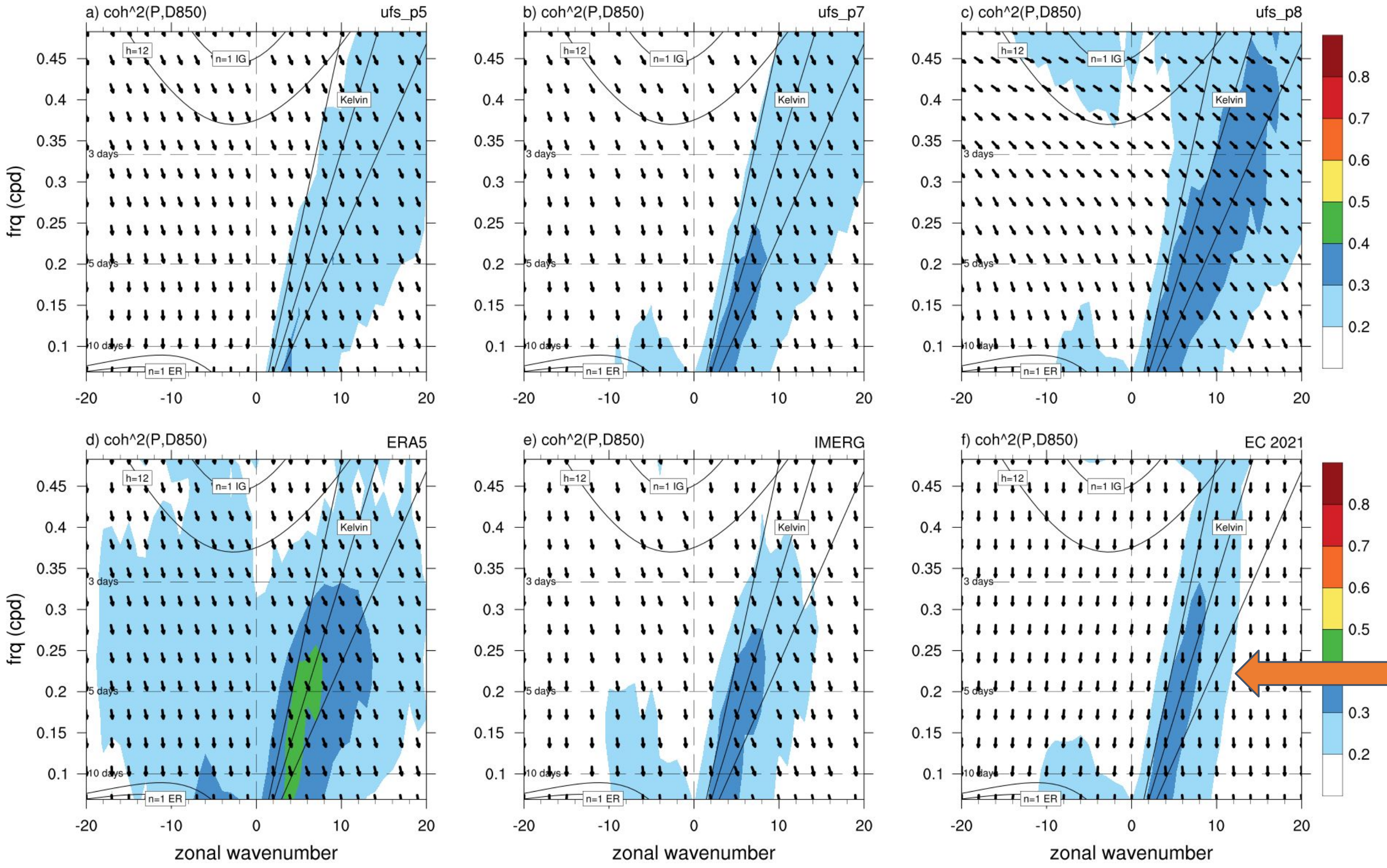



Coherence and phase spectra for precipitation and D850 from different sources for lead time 1- 30 days.

Coherence between low level convergence and precipitation for P7 is more confined.

P8 has stronger coherence and stronger dispersion. Maybe too much coherence at higher frequencies?

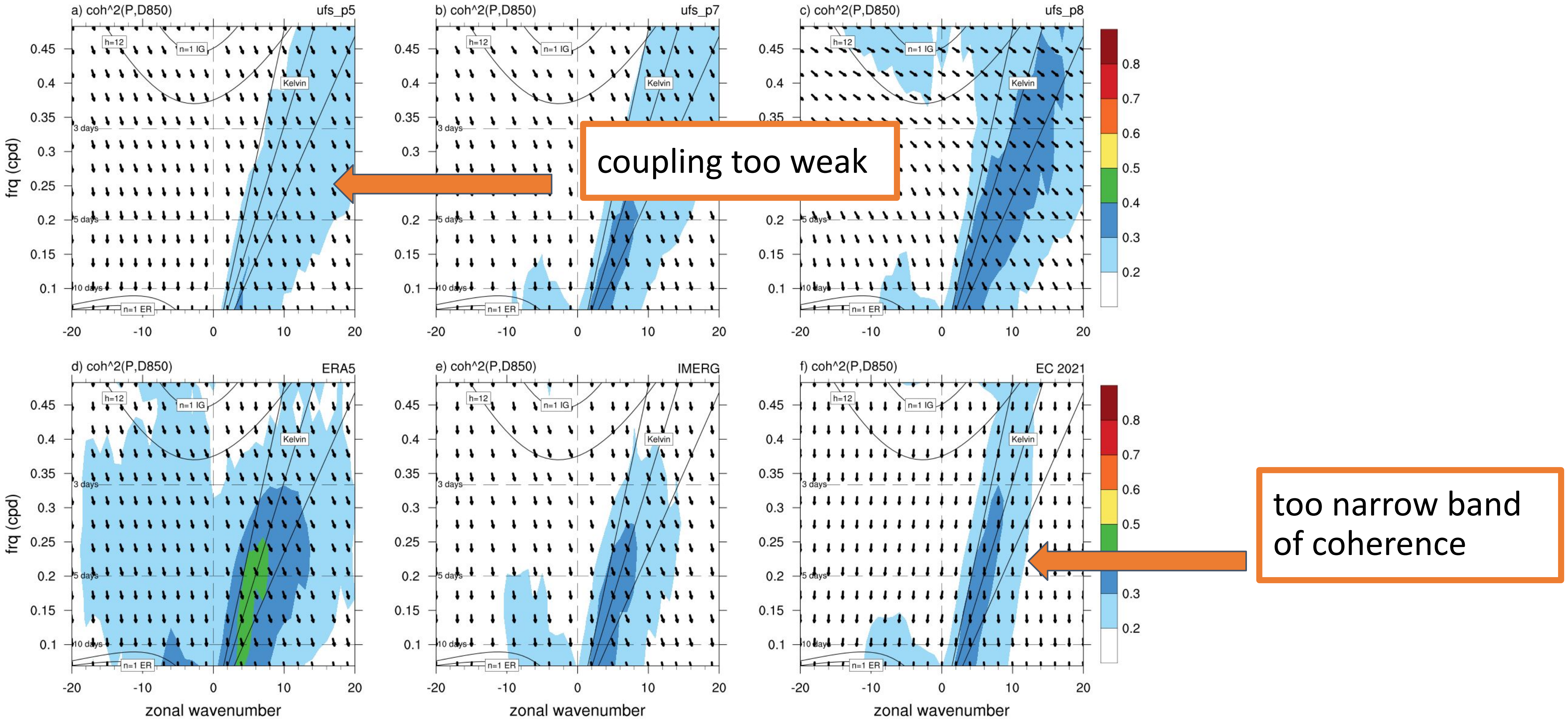
# Space-time Coherence Spectra




too narrow band of coherence

Coherence and phase spectra for precipitation and D850 from different sources for lead time 1- 30 days.
Coherence between low level convergence and precipitation for P7 is more confined.
P8 has stronger coherence and stronger dispersion. Maybe too much coherence at higher frequencies?

# Space-time Coherence Spectra

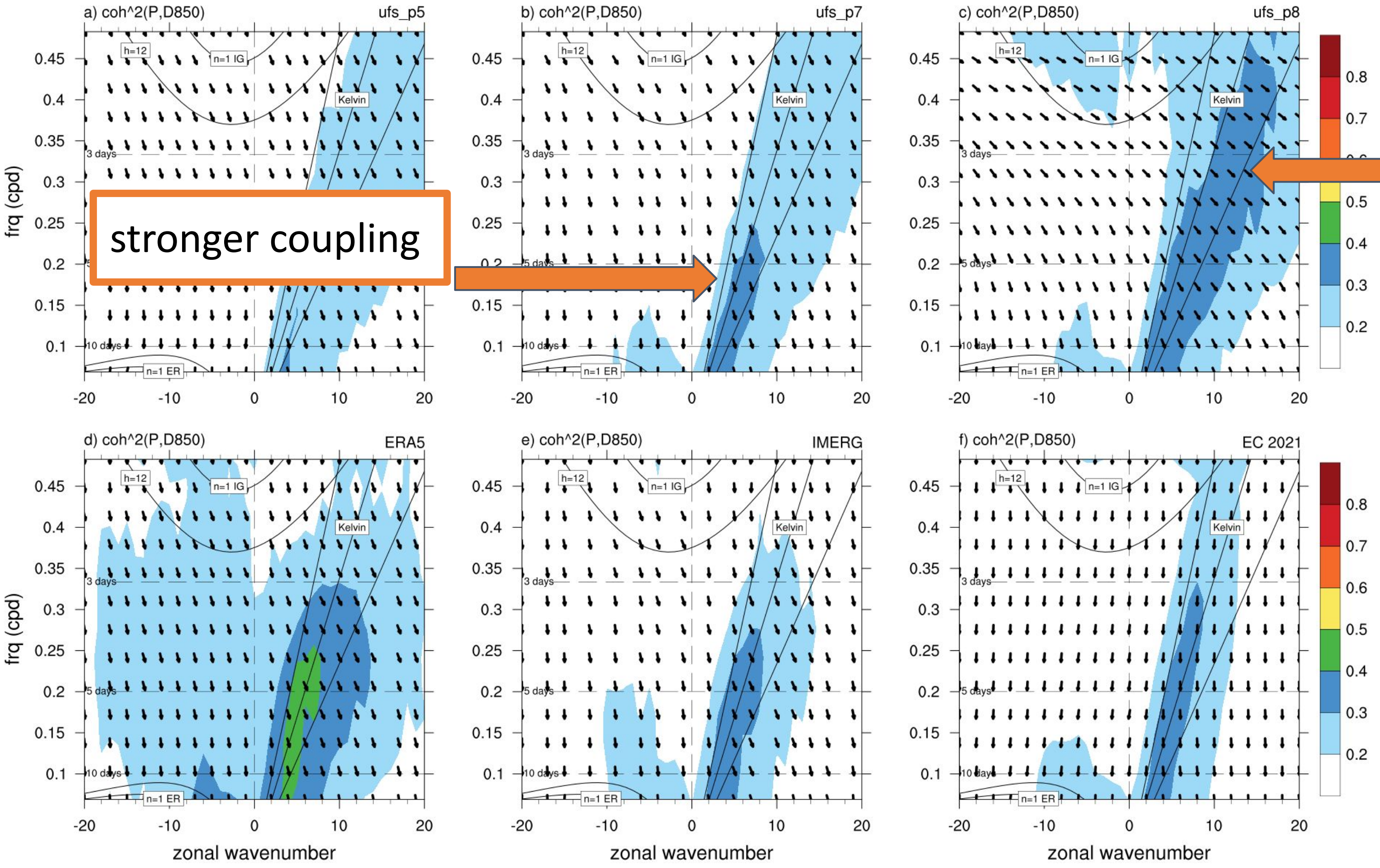



Coherence and phase spectra for precipitation and D850 from different sources for lead time 1- 30 days.

Coherence between low level convergence and precipitation for P7 is more confined.

P8 has stronger coherence and stronger dispersion. Maybe too much coherence at higher frequencies?

# Space-time Coherence Spectra




more dispersion and stronger coupling

Coherence and phase spectra for precipitation and D850 from different sources for lead time 1- 30 days.

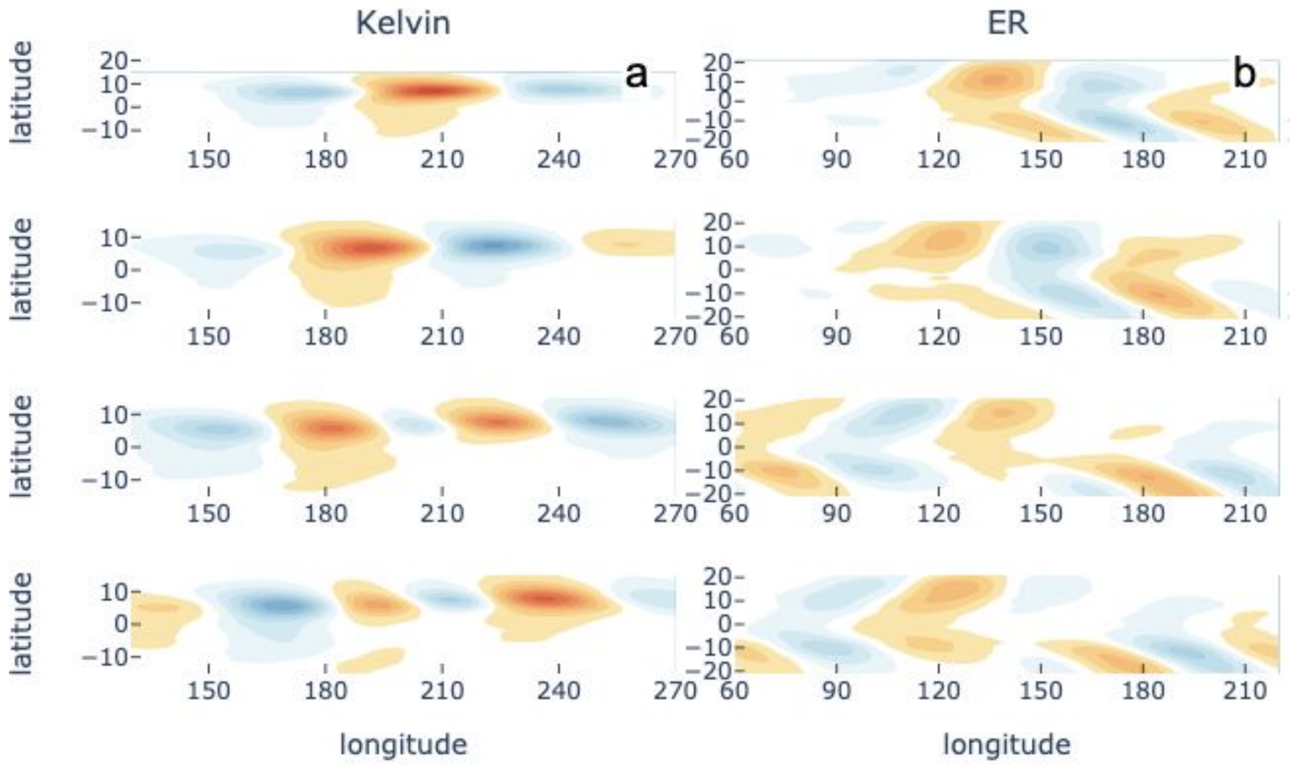
Coherence between low level convergence and precipitation for P7 is more confined.

P8 has stronger coherence and stronger dispersion. Maybe too much coherence at higher frequencies?

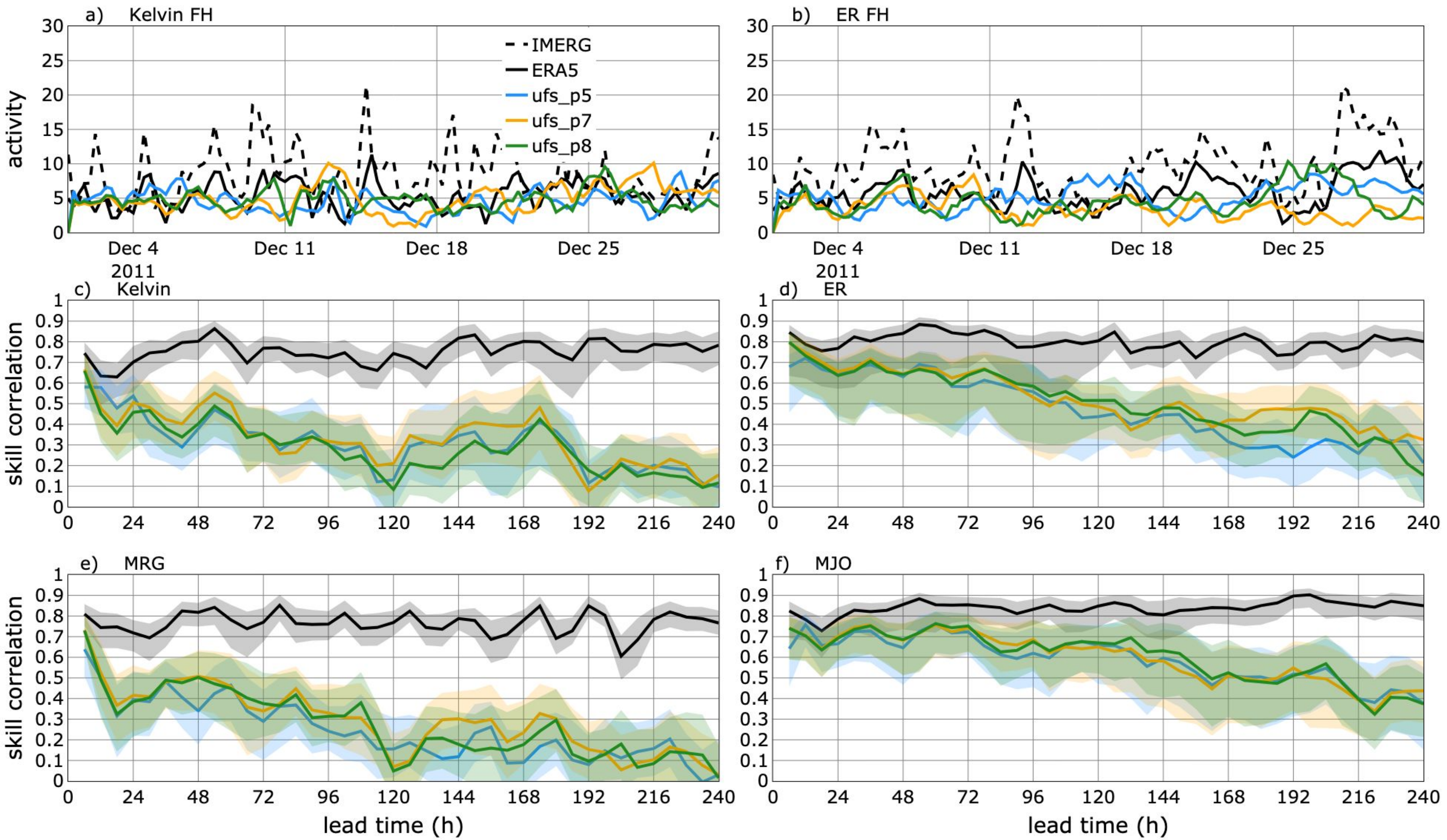
# CCEW activity skill in the UFS

How long and how well can the model predict CCEWs?

1. Use long time series (30+ years) of observed filtered precipitation to compute EOFs describing CCEW signal.
2. Project the model precipitation at each forecast hour onto these EOF patterns and compute a CCEW activity index.
3. Compute anomaly correlation between the observed and model index.



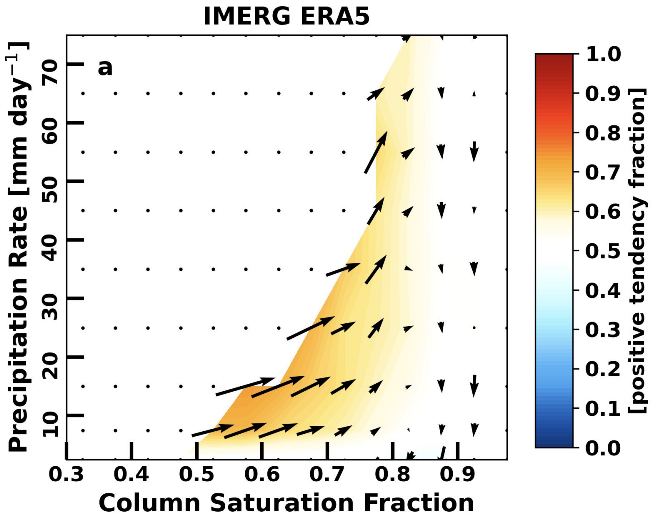
# CCEW activity skill in the UFS




Similar skill among the coupled forecasts, noise because of small sample size.

ER skill is retained for longer in the coupled forecasts than uncoupled, but not for Kelvin, MRG or the MJO.

# Moisture convection coupling

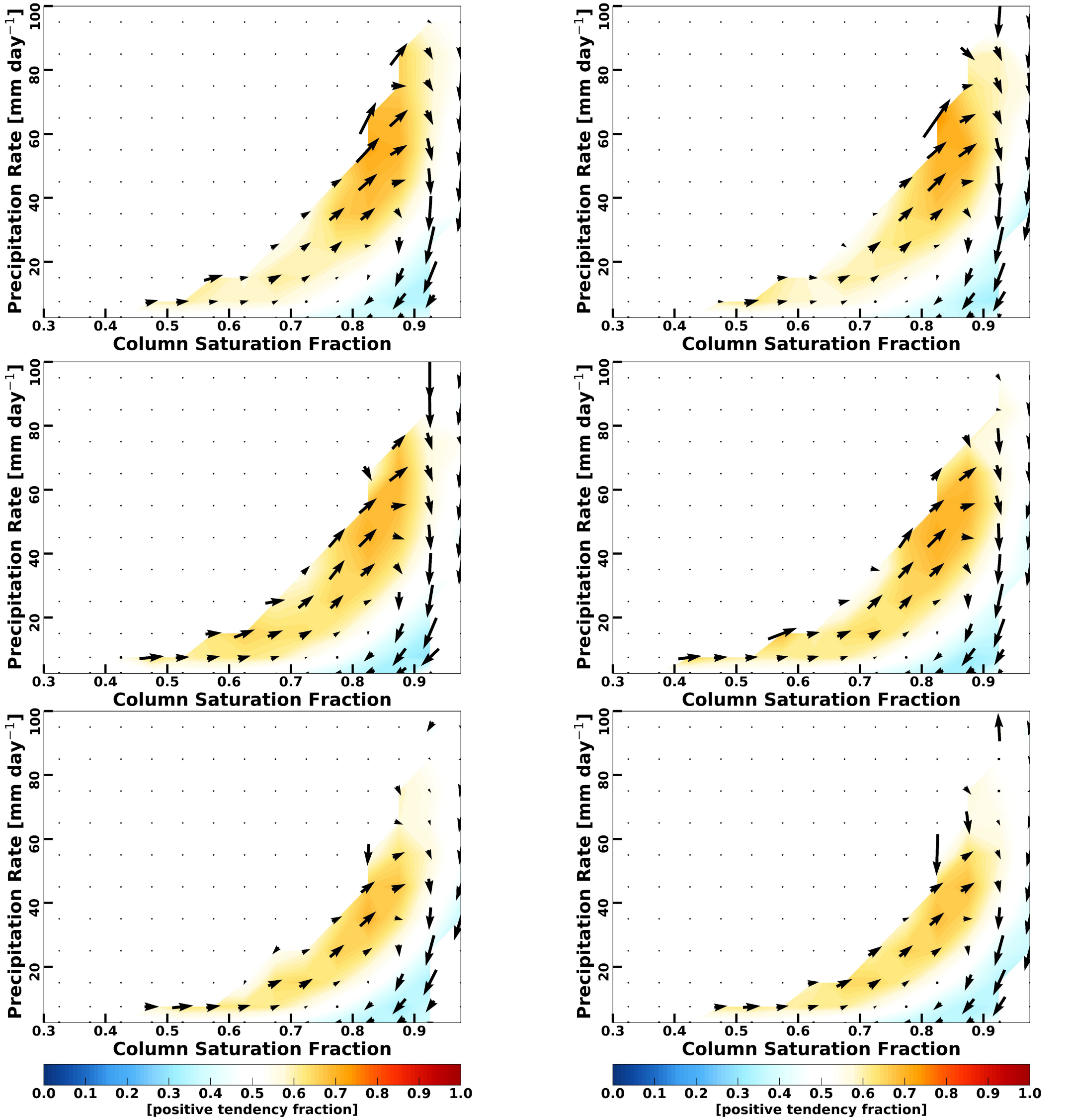



The clockwise rotation can be thought of as the evolution of precipitation and moisture through the convective life-cycle.

Moistening and increasing precipitation rates on the upper left, decreasing precipitation rates at very high moisture levels at the right, and drying and very low precipitation rates on the lower edge.

FH 120

FH 240

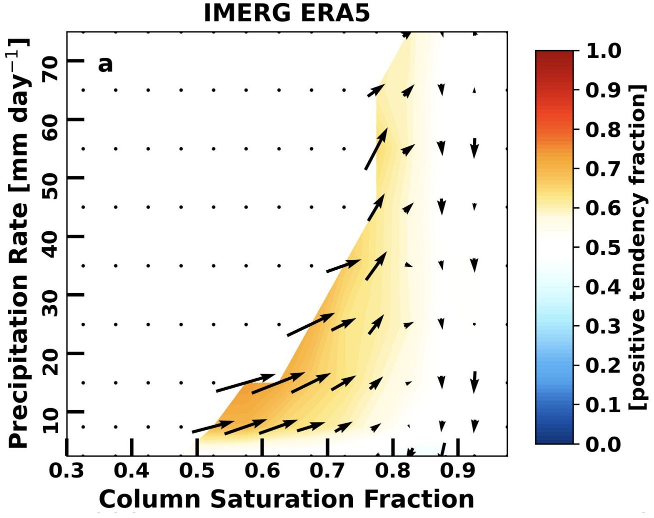


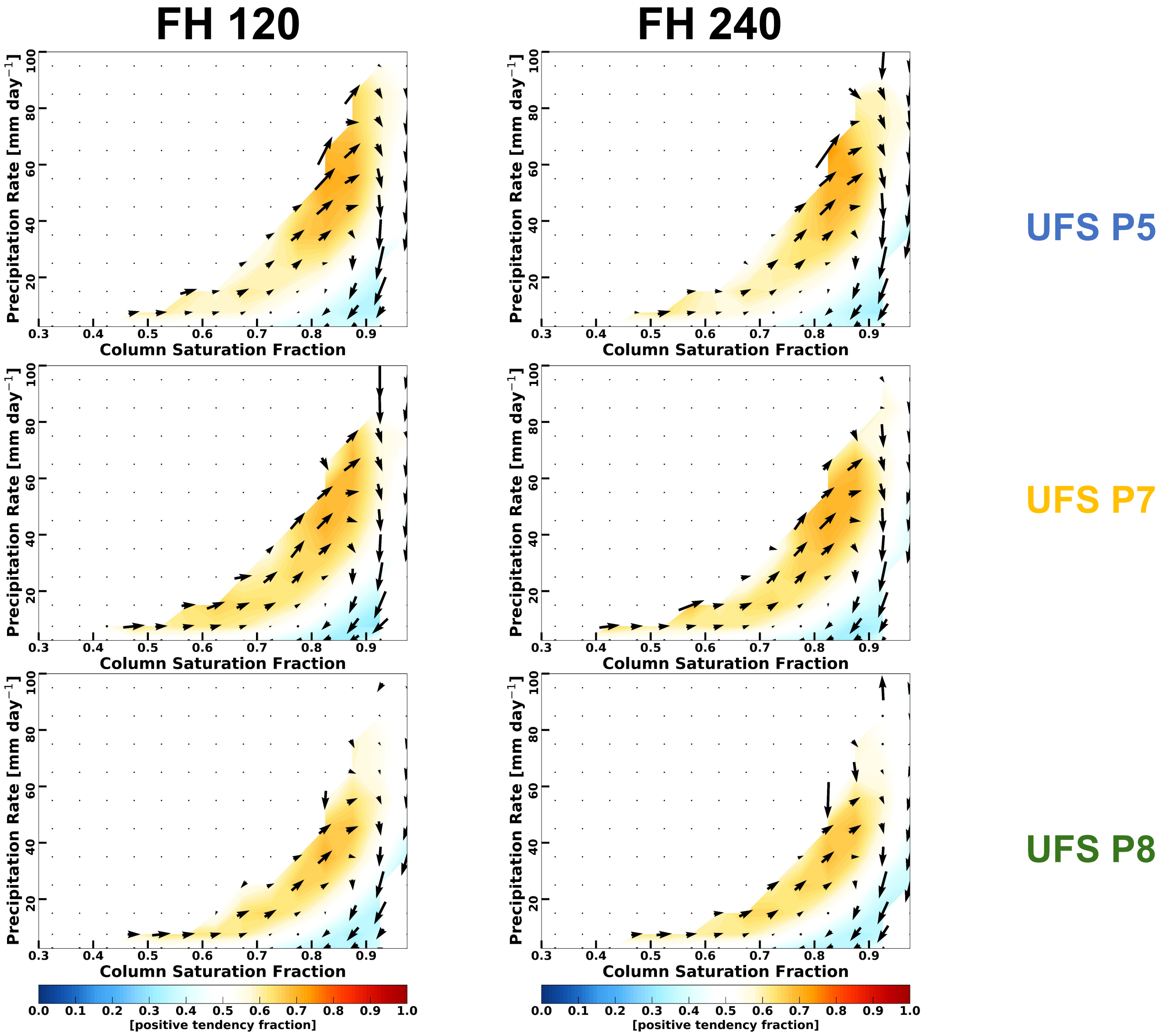

UFS P5

UFS P7

UFS P8

# Moisture convection coupling



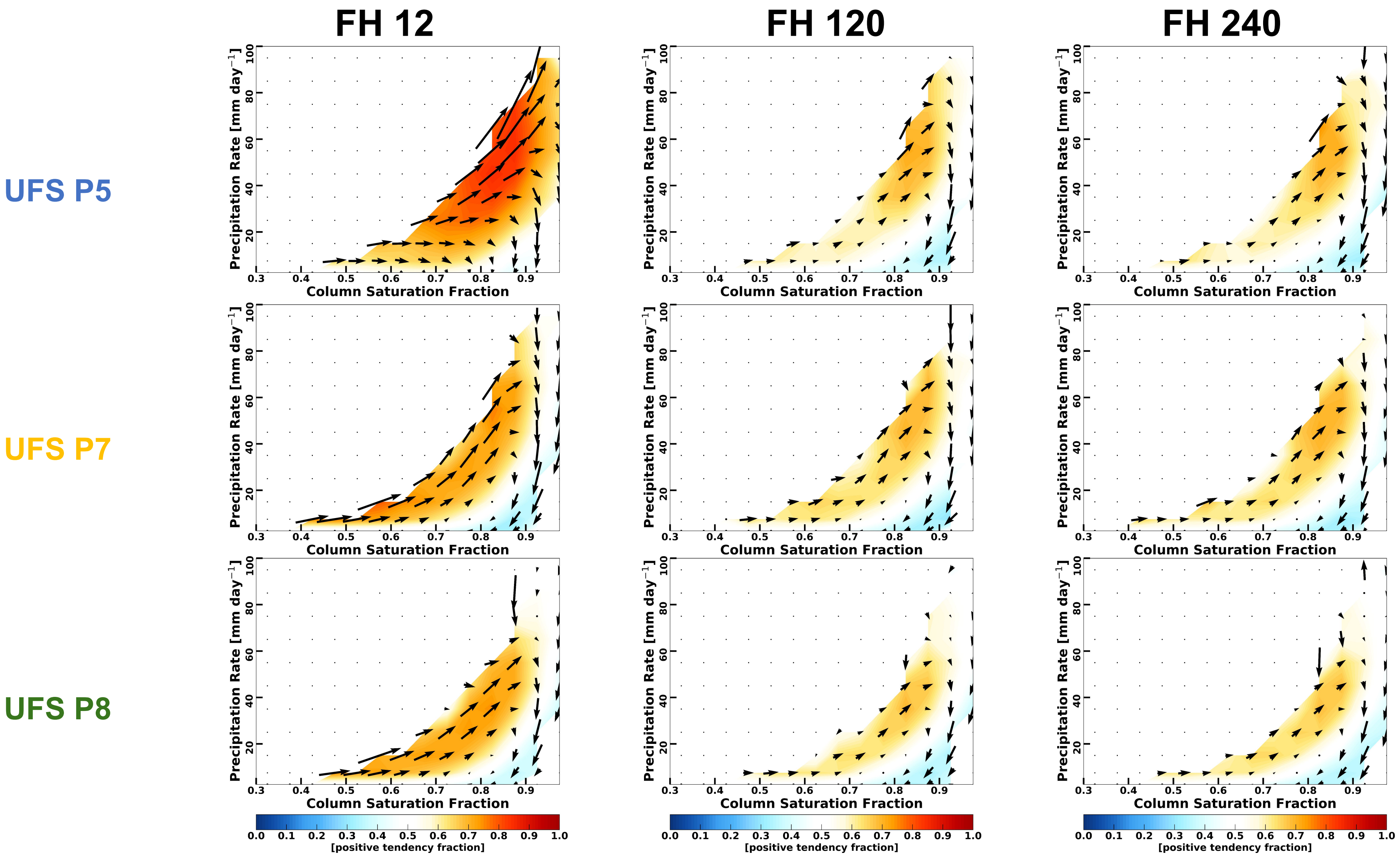




Final state is very similar for P5 and P7. Weaker in P8.

The precipitation CSF coevolution obtained from reanalysis and observed precipitation also shows a clockwise rotation.

All prototypes have stronger moistening at higher precipitation rates and CSF values than reanalysis.

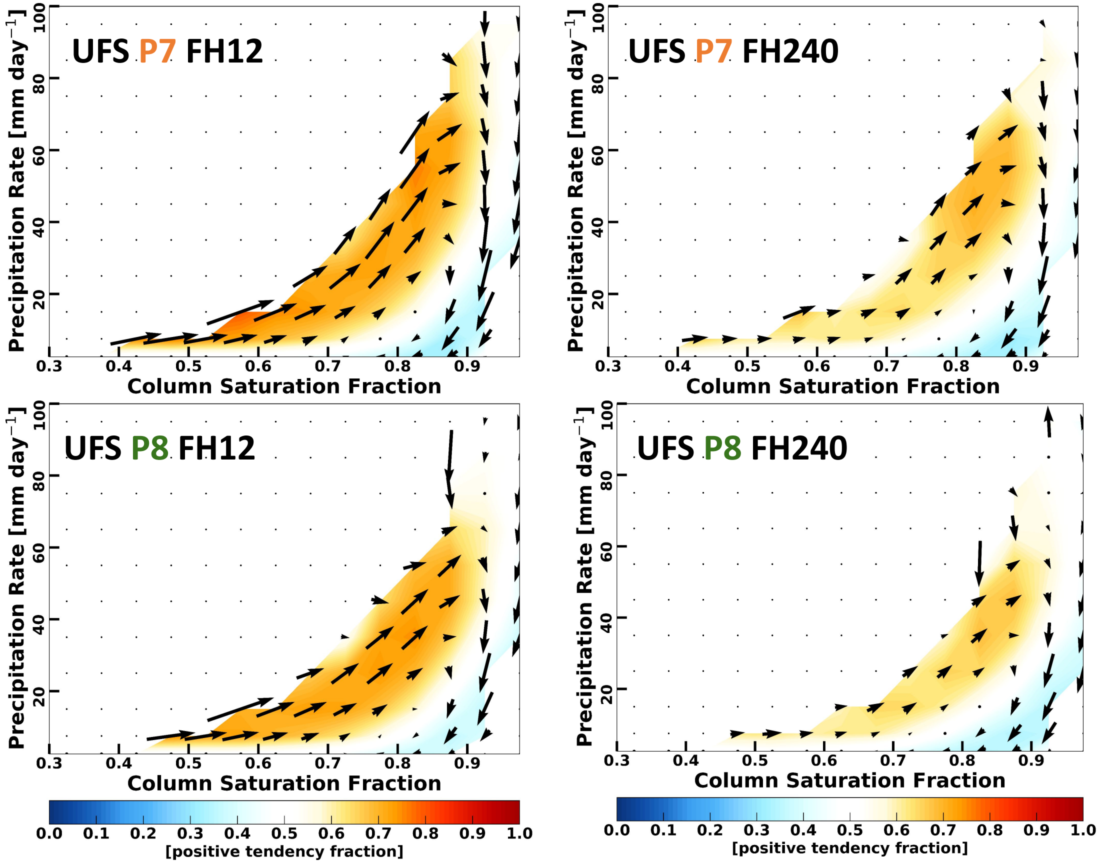
# Moisture convection coupling

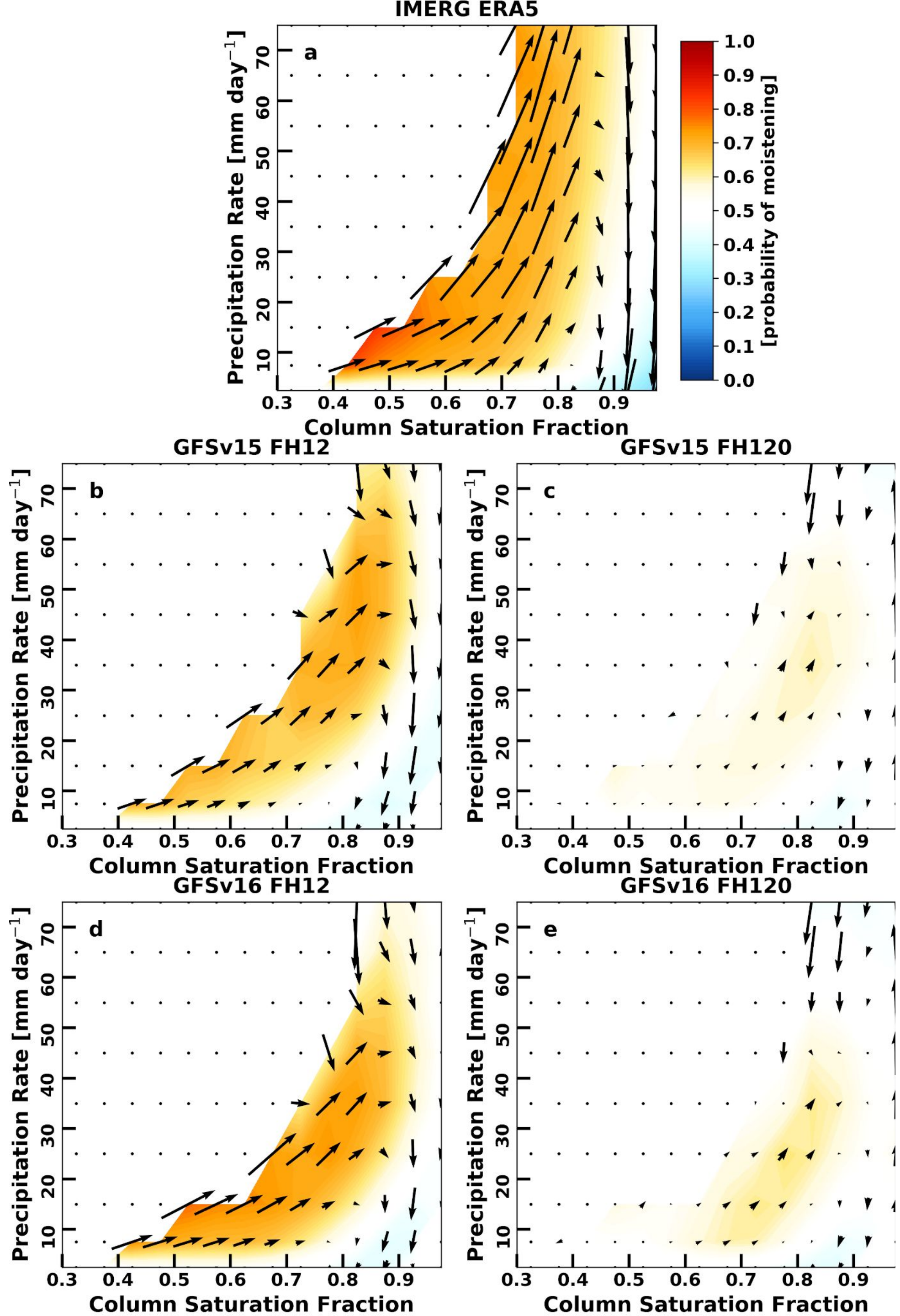


# Moisture convection coupling

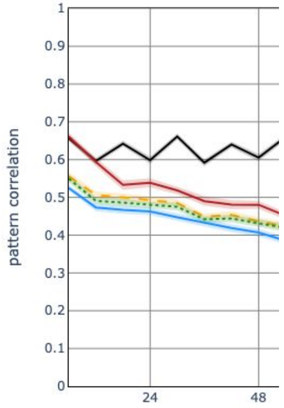
In the coupled system we see similar decrease in coupling strength with lead time.

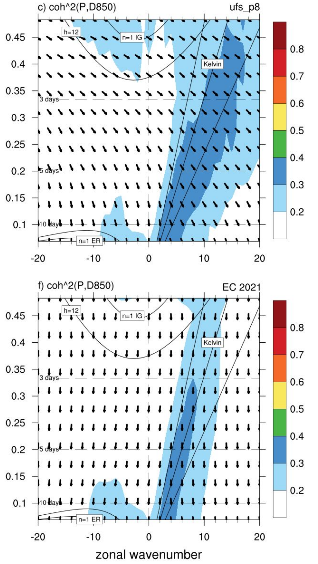
Decrease is slower and that could be due to either the coupling or model improvements in the coupled system.

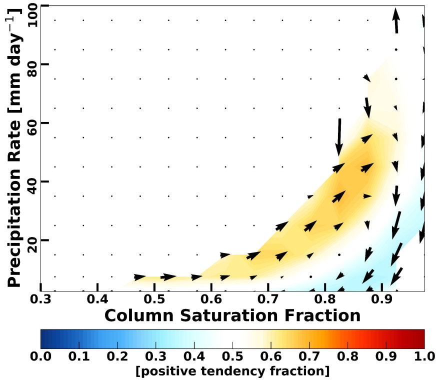


# Summary




- Consider skill metrics for tropical convection mostly based on precipitation.

- Much precipitation skill is lost in the hours immediately following initialization.

- Coupling between convection and the circulation is improved (in terms of scales and strength) in the UFS coupled prototypes, but decreases rapidly with lead time.

- To look at the changes in coupling strength in wavenumber frequency space for the coupled system a larger sample size with regular initializations is needed.

- Moisture convection coupling is also improved in the coupled system. Possibly coupling to the ocean or other model improvements are responsible.

# Summary

- A stand-alone python **GitHub** repo for these diagnostics (and more) exists (tropical_diagnostics) and a release is public for testing.

- Several of these diagnostics were included in the November beta release of METplotpy and METcalcpy of METplus. A recording of the presentation on METplus Use Cases for UFS P5 and P7 output can be found here (https://dtcenter.org/events/2022/2022-dtc-metplus-workshop/agenda-recordings)

More details on the diagnostics can be found in:

Gehne M., B. Wolding, J. Dias and G. N. Kiladis (2022). Diagnostics of Tropical Variability for Numerical Weather Forecasts, *Weather and Forecasting* (https://doi.org/10.1175/WAF-D-21-0204.1)

METplus GitHub

# Extra Slides

# Prototype Details

- Initialized each month on the 1st and 15th from 2011/04/01 through 2018/03/15 (7x24 = **168 forecasts**)
- 6 hourly output out to **30 days** (720h)
- Comparison of UFS coupled prototype 5, 7 and 8

## ufs-weather-model

https://github.com/ufs-community/ufs-weather-model

**Atmosphere**
- FV3 dynamical core
- GFS Physics with GFDL microphysics
- CCPP physics driver
- C384 (~25km), 127 levels (P6 onward)

**Ocean**
- MOM6 Modular Ocean Model
- ¼ degree tripolar grid, 75 hybrid levels
- OM4 Set up [Adcroft, 2019]

**Waves**
- WAVEWATCH III
- ½ degree regular lat/lon grid
- ST4 Physics [Ardhuin, 2010]

**Ice**
- CICE6 Los Alamos Sea Ice Model
- ¼ degree tripolar grid (same as ocean)
- 5 thickness categories
- Mushy thermodynamics on (P7 onward)

**Driver/Mediator**
- NEMS driver
- CMEPS mediator

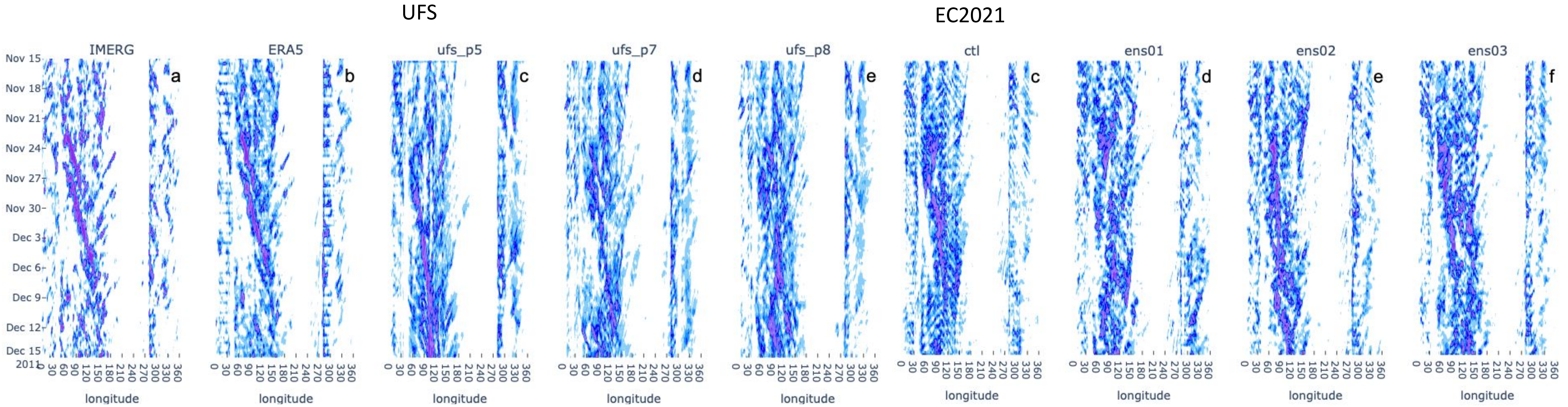
## Prototypes Overview

*For more details see the spreadsheet here.*

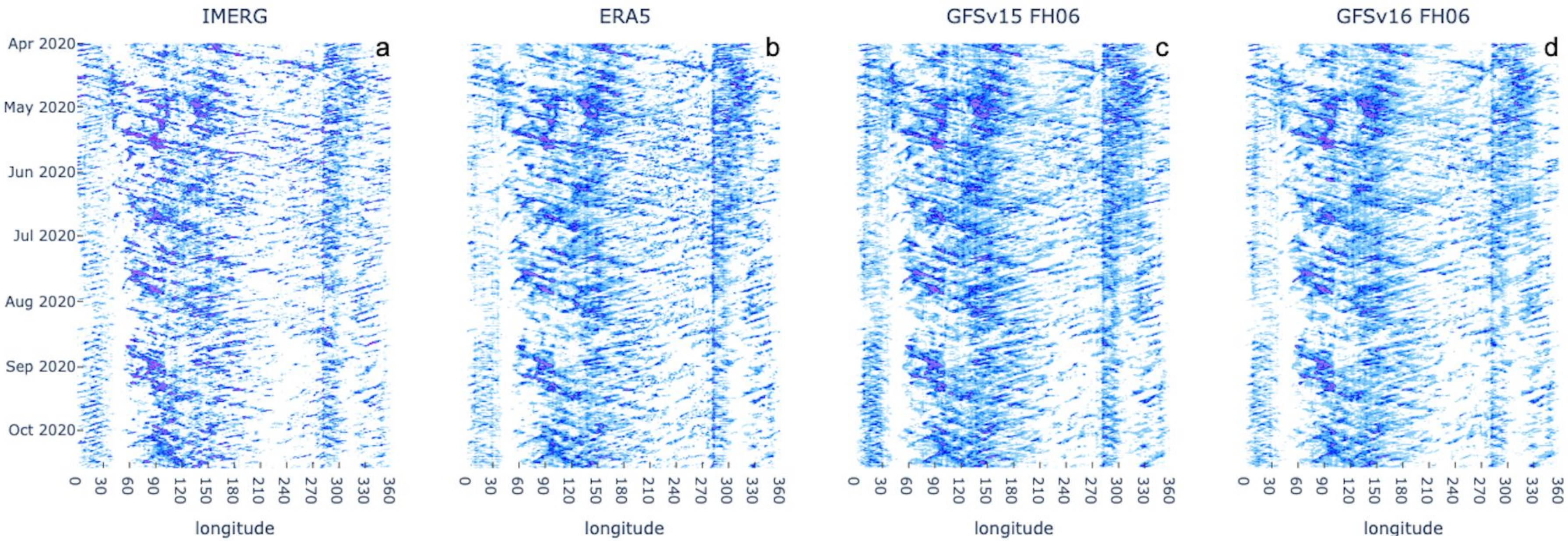
| | Initial Conditions | | | | | |
|---|---|---|---|---|---|---|
| | **FV3 GFS** | **MOM6** | **CICE** | **WW3** | **Ice Model** | **Mediator** |
| **UFS_P1** | CFSR | CFSR | CFSR | n/a | CICE5 | NEMS |
| **UFS_P2** | CFSR | **CPC 3Dvar** | CFSR | n/a | CICE5 | NEMS |
| **UFS_P3.1** | CFSR | CPC 3Dvar | **CPC ice analysis** | n/a | CICE5 | NEMS |
| **UFS_P4** | CFSR | CPC 3Dvar | CPC ice analysis | **Generated with CFS forcings** | CICE5 | NEMS |
| **UFS_P5** | CFSR | CPC 3DVar | CPC ice analysis | Generated with CFS forcings | **CICE6** | **CMEPS** |
| **UFS_P6** | **CFSR Frac grid** | CPC 3DVar | CPC ice analysis | Generated with CFS forcings | CICE6 | CICE6 |
| **UFS_P7** | **GEFS, NOAH-MP land, Frac grid** | CPC 3DVar | CPC ice analysis | **Generated with GEFS forcings** | CICE6 | CICE6 |

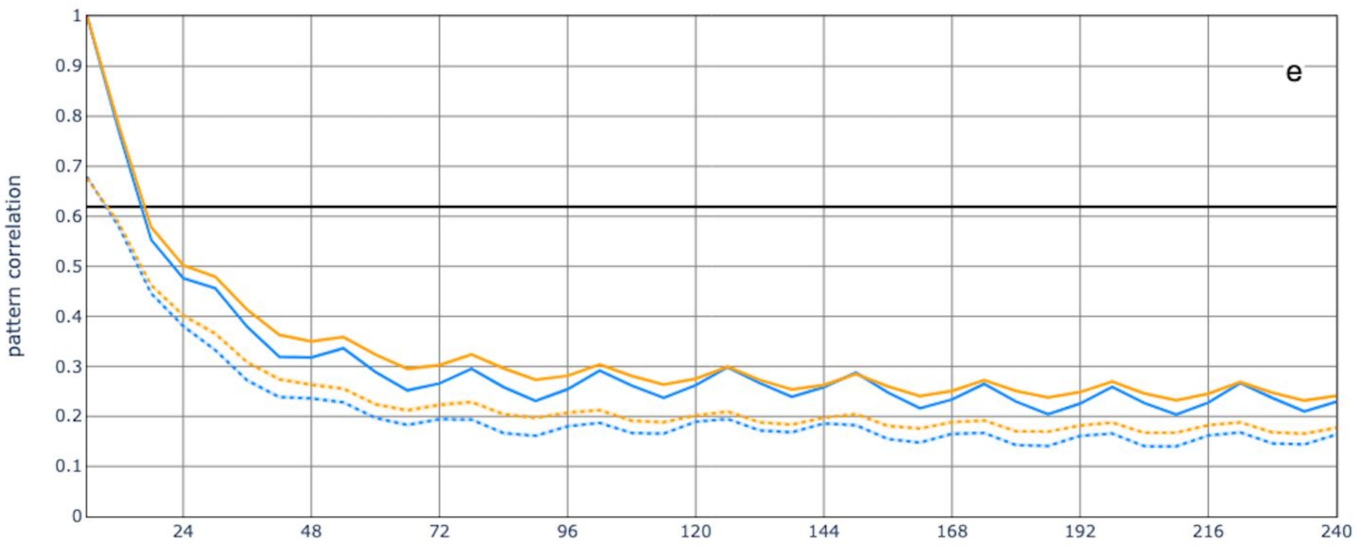
| Prototype | Atm ICs (gfs* files) | Soil ICs (in sfc* files) | Snow ICs (in sfc* files) |
|---|---|---|---|
| P7 | From GEFSv12 Reanalysis | Spun up with Noah-MP, using GDAS forcing | From GEFSv12 Reanalysis |
| P8a P8b | No change | Spun up with updated Noah-MP, using NASA GLDAS forcing (GSWP3/GDAS for Antarctica) | Spun up with updated Noah-MP, using NASA GLDAS forcing (GSWP3/GDAS for Antarctica) |
| P8 | No change | Spun up with updated Noah-MP using P8a forcing and new land/lake mask | Spun up with updated Noah-MP using P8a forcing and new land/lake mask |

# Hovmoeller and Pattern Correlation



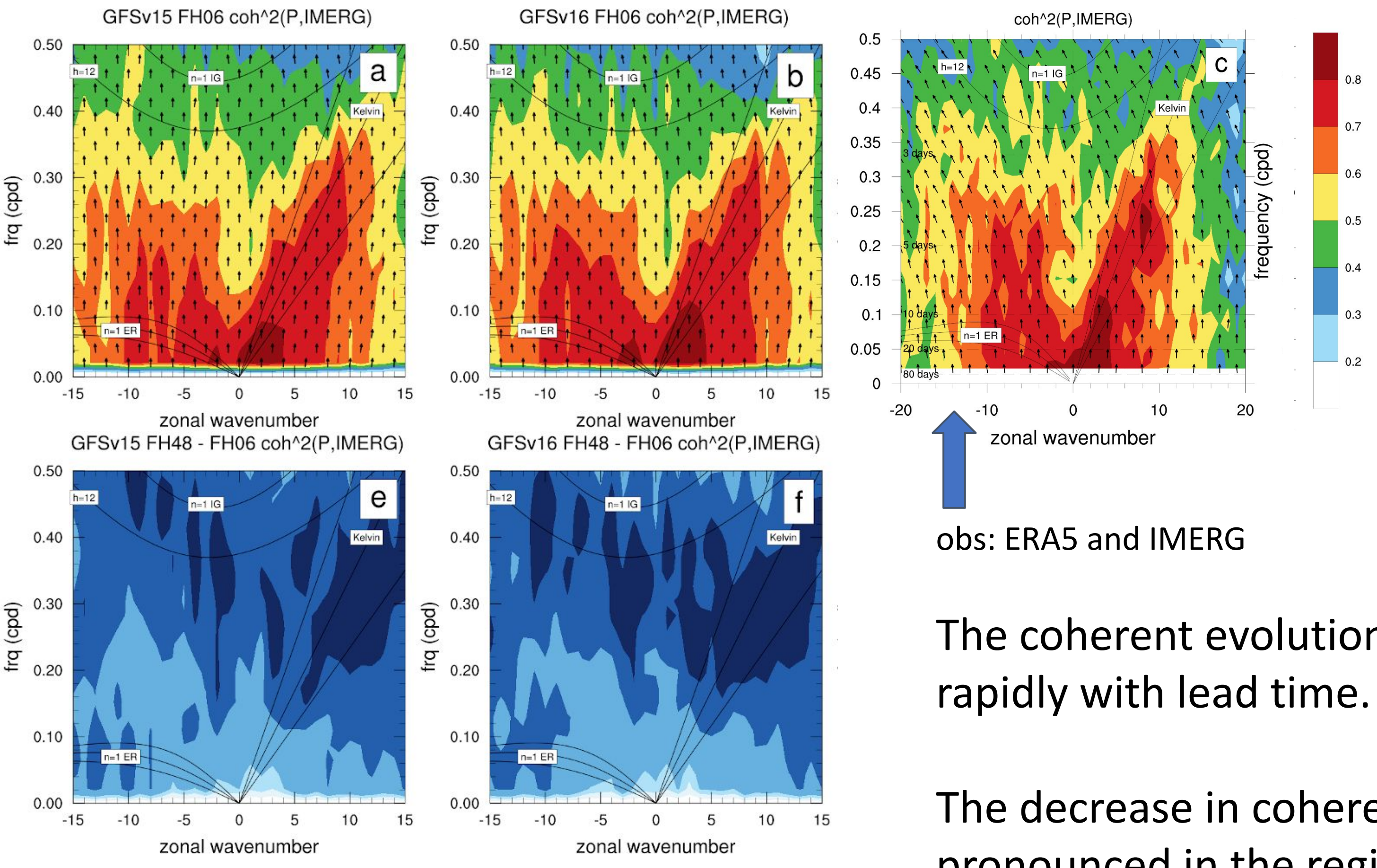
# Hovmoeller and Pattern Correlation




- GFSv15 operational vs GFSv16 parallel shows only minor differences with GFSv16 slightly outperforming GFSv15.
- Correlation with IMERG is higher initially (<FH12) than correlation between IMERG and ERA5.
- Much potential skill in precipitation forecasts is already lost during the first few hours after initialization.

# Space-time coherence spectra




obs: ERA5 and IMERG

Initially larger coherence values tend to be located near CCEW dispersion curves and at lower frequencies and larger spatial scales.

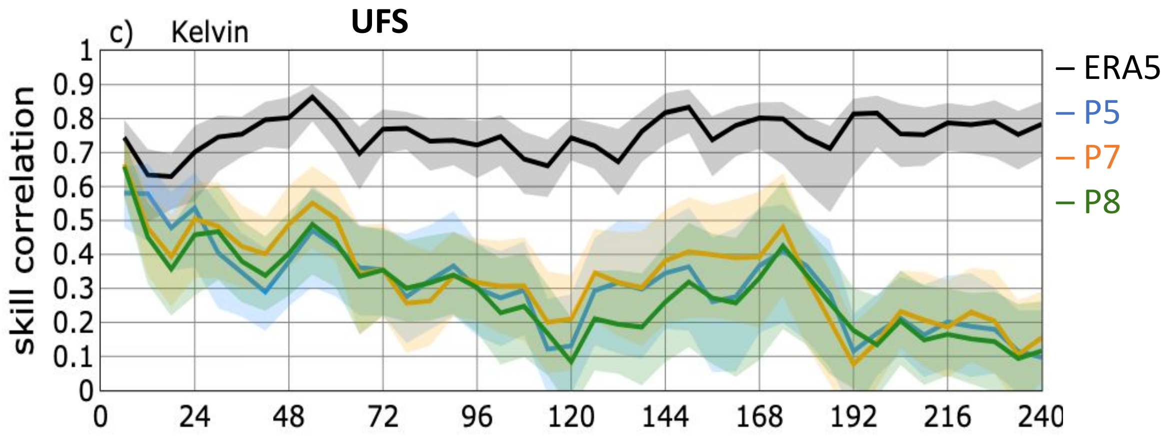
Precipitation in both GFSv15 and GFSv16 in the first 12 - 24h past initialization is largely able to initialize and maintain large scale CCEW events

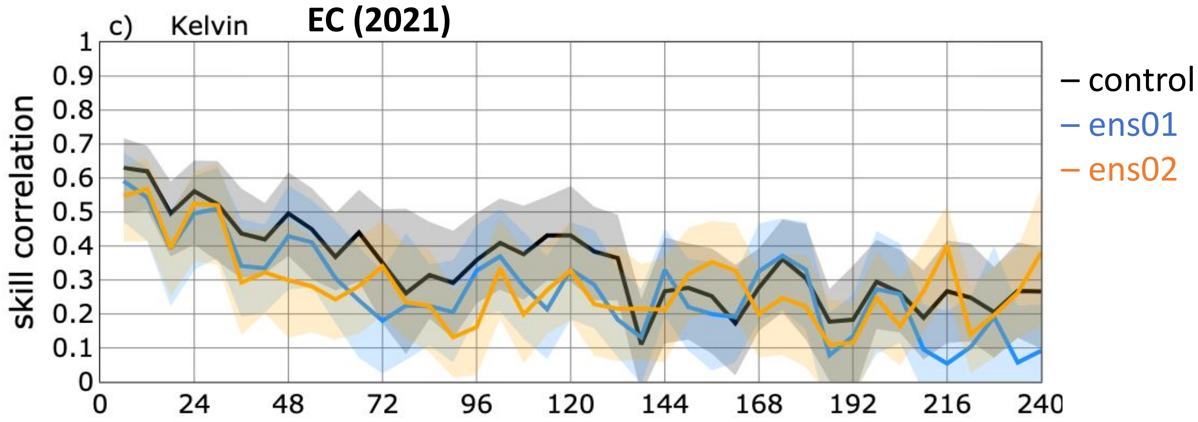
The coherent evolution of observed and modeled precipitation decreases rapidly with lead time.

The decrease in coherence squared from 6h to 48h lead time is most pronounced in the regions of CCEW dispersion curves and higher frequencies and wavenumbers.

The coherence decay rate is related to the wave lifecycle and the model is able to propagate waves present in the IC, but spontaneous initialization of CCEWs is much harder.

# CCEW activity skill comparison to EC 2021



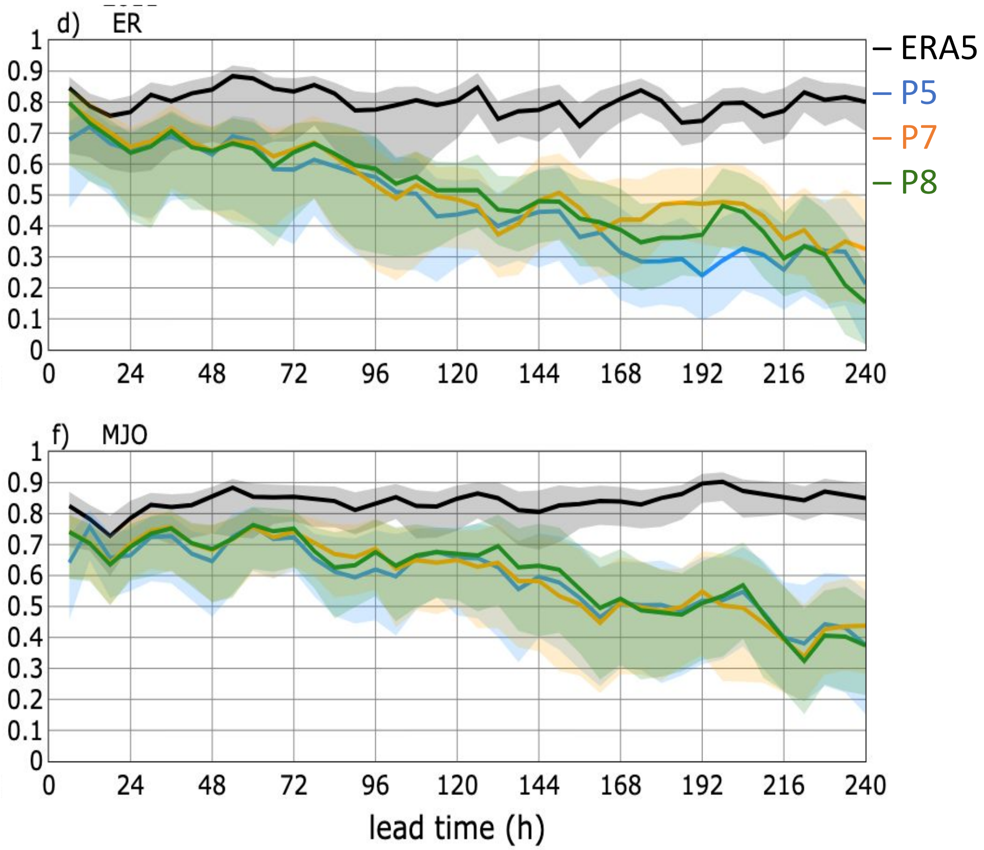



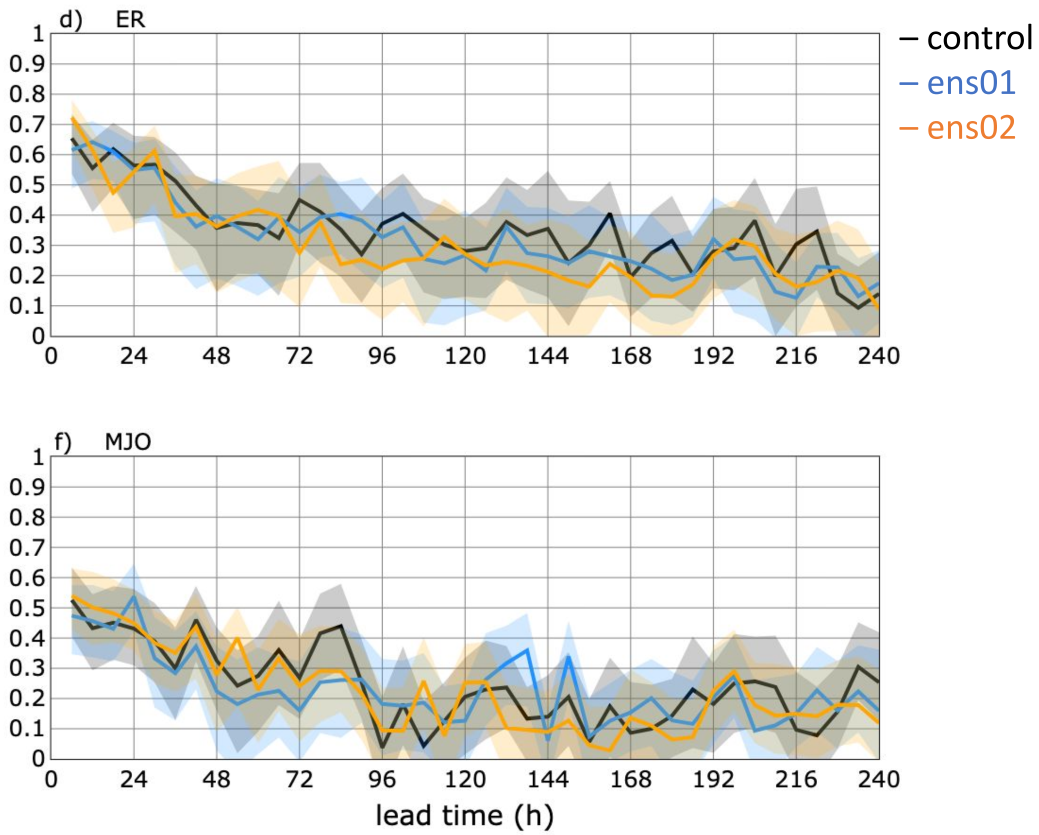

In general, UFS prototypes have comparable skill to the EC S2S ensemble.

UFS initial (at 6h lead time) Kelvin skill for P7 and P8 is slightly higher than in the EC, although difference is not significant.

EC skill at 12h lead time is still above 0.5 correlation while P7 and P8 have dropped below the 0.5 threshold.

# CCEW activity skill comparison to EC 2021



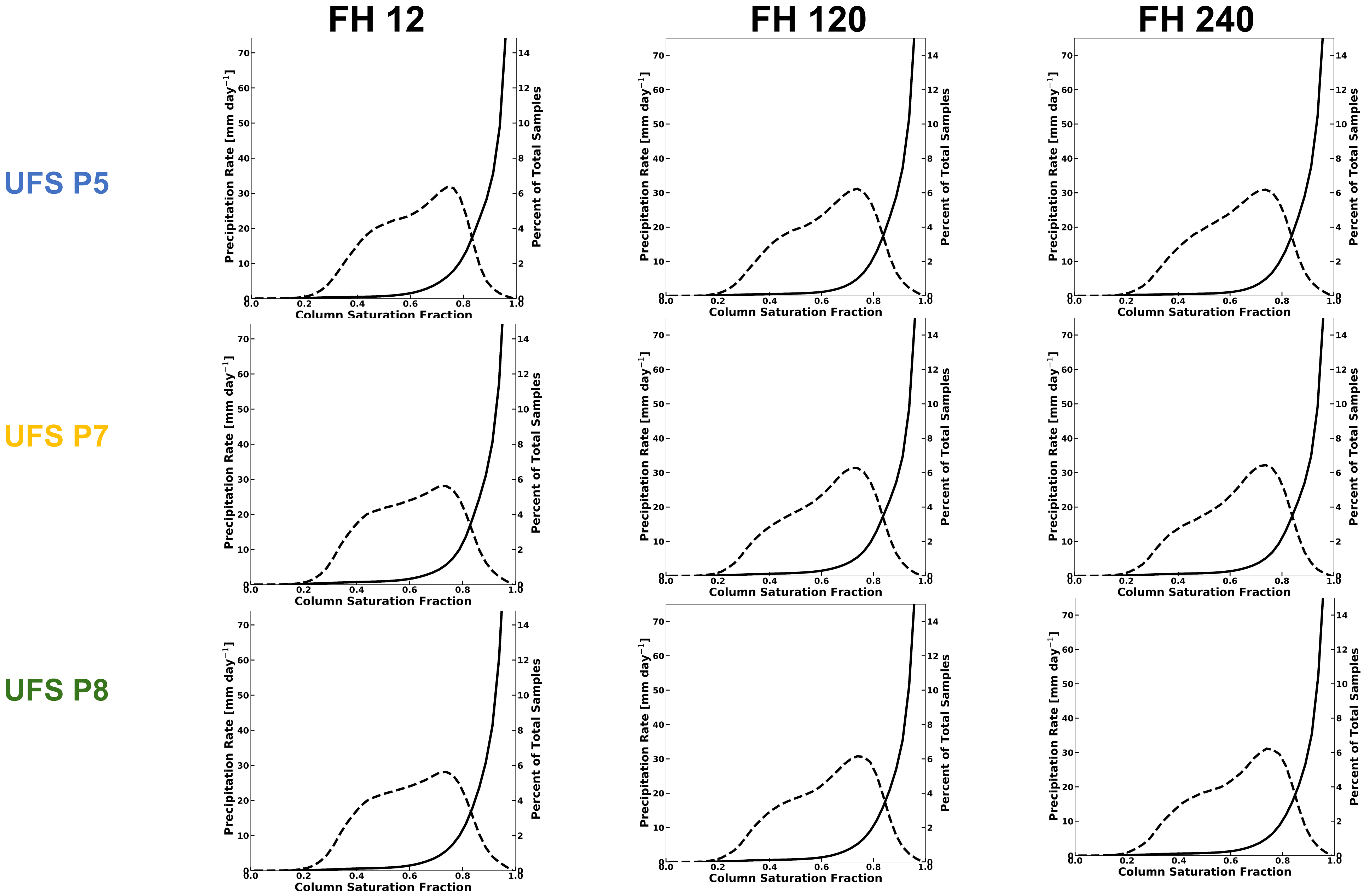

Initial ER skill is comparable between UFS prototypes and EC ensemble forecasts.

UFS prototypes have ER skill correlation above 0.5 until 96h lead time, while the EC skill correlation drops below 0.5 before 48h lead time.

Initial MJO skill is significantly higher in P7 and P8 then the EC forecasts.

EC MJO skill drops below 0.5 in the first 24h, while the UFS MJO skill stays above 0.5 correlation until 144h lead time.

# Precipitation - CSF coevolution



# Moisture convection coupling

Coevolution of precipitation and column saturation fraction (CSF).

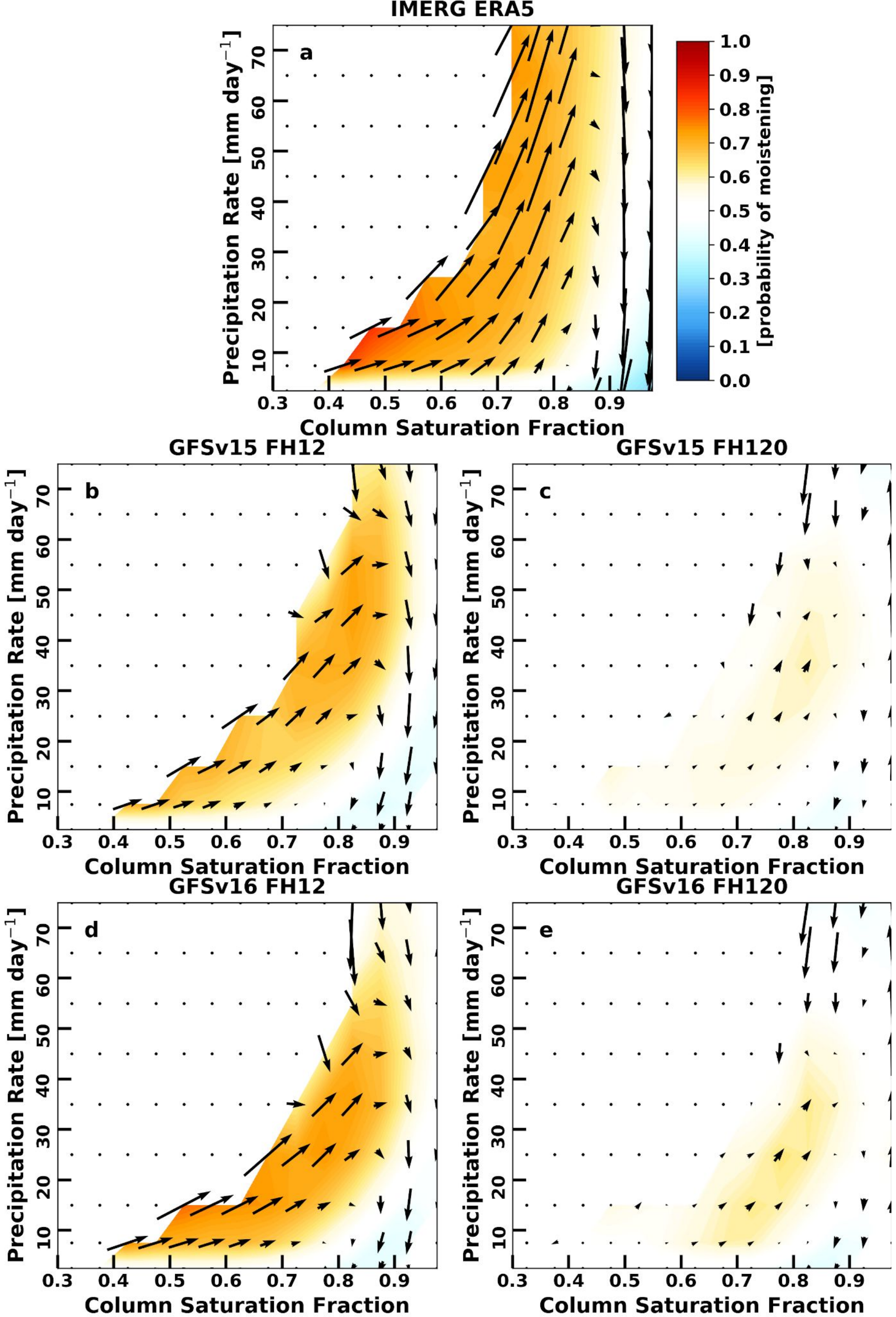
IMERG-ERA5 coevolution shows gradual moistening and increasing precipitation for intermediate precipitation rates. At very high CSF precipitation drops off and drying occurs.

At FH 12 GFSv15 shows similar evolution to IMERG-ERA5 although weaker moistening and less vigorous coevolution.

GFSv16 has slightly stronger moistening compared to GFSv15 at FH12.

In both model versions the coevolution weakens with lead time. By FH120 there is only a hint of the original coevolution left.

Caveat: ERA5 may also not be showing the "real" evolution as there are differences between reanalyses and radiosonde data. This is being investigated in more detail at the moment.

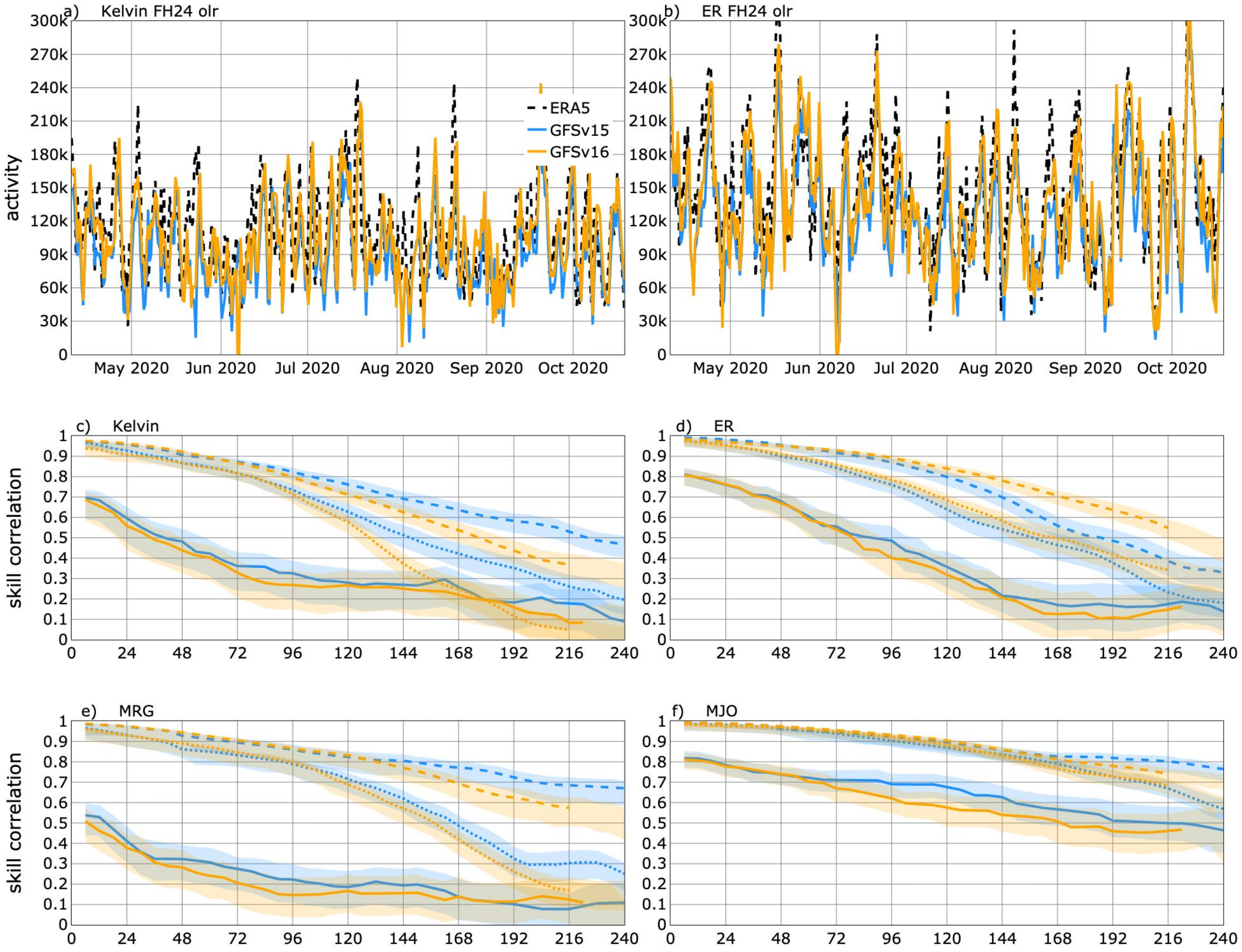


# CCEW activity

Regress OLR and winds on precipitation PCs to find EOFs.

Skill for OLR is higher than precipitation for Kelvin (by 0.2) and ER (by 0.15) and comparable for MRG and MJO.

Wind skill is much higher than OLR or precipitation skill.



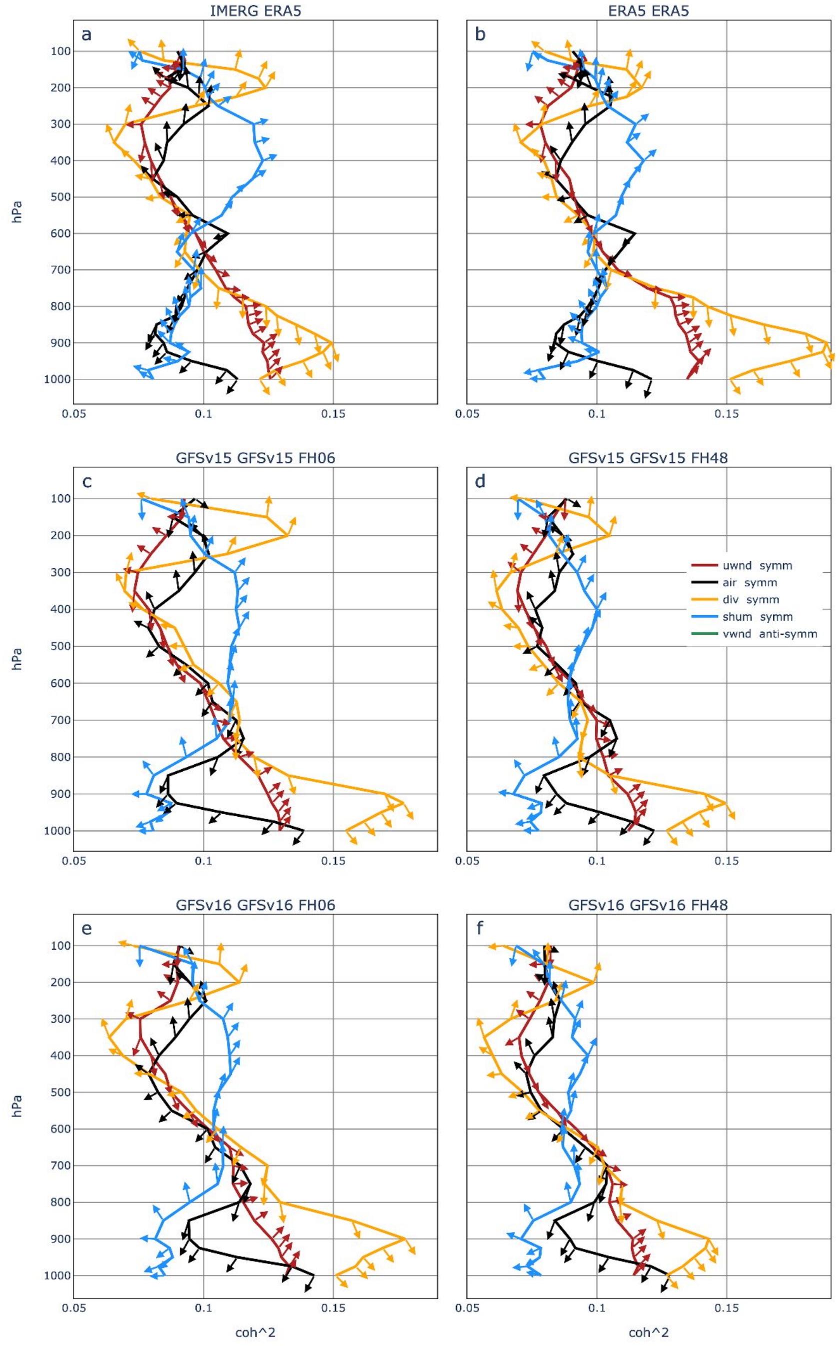
# Vertical structure of coherence

- Proxy for vertical profile of latent heating associated with deep convection.
- Filtered P is used to compute coherence with dynamical variables at all vertical levels.

Results point to several issues in the coupling between large-scale dynamics and convection.

The low-level divergence coherence peak appears too weak and at slightly lower levels than observed and decreases with lead-time.
Mid-level peak in temperature coherence-squared is lower in GFSv15 and GFSv16.
Coherence with specific humidity does not show a well defined peak between 550 and 250 hPa in either model version.

Coherence weakens with lead time at all levels.

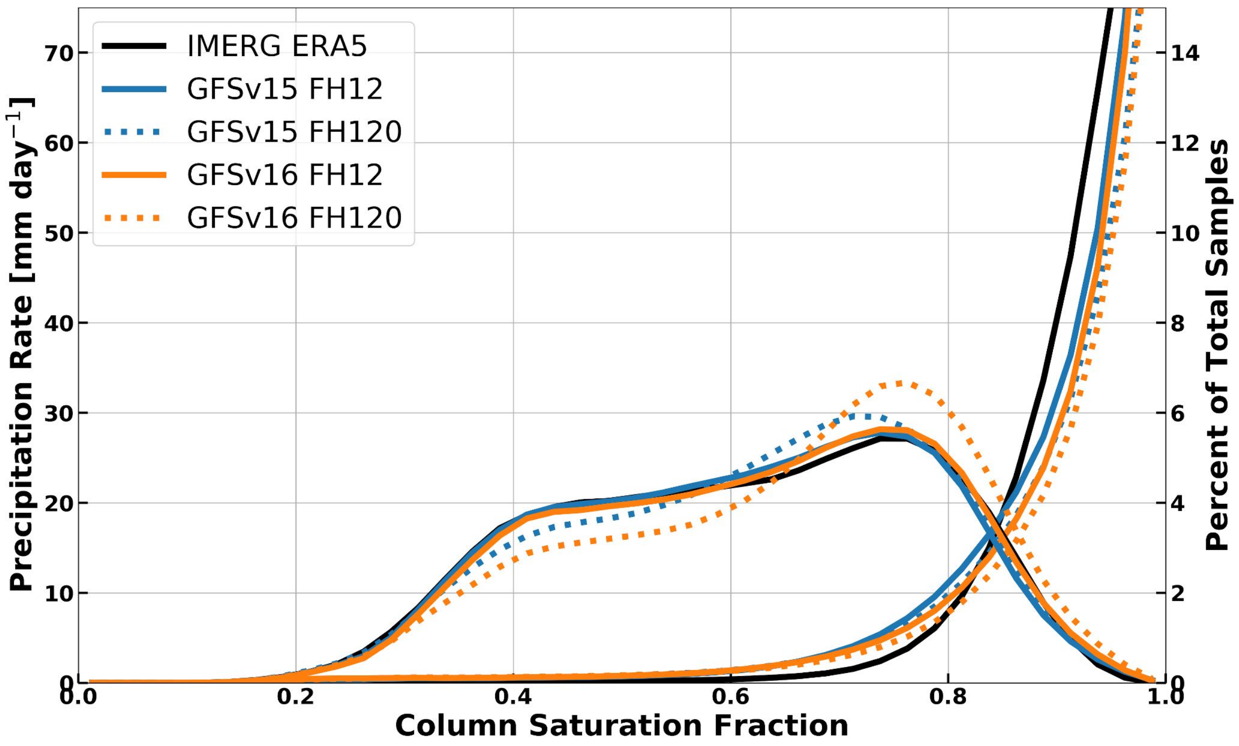


# Moisture convection coupling

Column saturation fraction (CSF) distribution and precipitation rate conditioned on CSF. IMERG/ERA5 shows exponential pick-up of precipitation rate with CSF.

GFSv16 shows slight improvement over GFSv15 in the precipitation pick-up. This is sustained with lead time.

GFSv16 has larger shift in CSF distribution with lead time toward an increase in the occurrence of larger CSF values.




Convective adjustment time is slightly longer for both GFSv15 and GFSv16 than IMERG and ERA5. Shorter convective adjustment time indicates increased sensitivity of precipitation to atmospheric moisture.
Drift with lead time for GFSv16 convective adjustment time scale?