The Bureau of Meteorology

Integrating data-driven global models into our evaluation framework: A Bureau of Meteorology perspective

Debbie Hudson

Catherine de Burgh-Day, Chen Li, Li Shi, Hongyan Zhu, Robin Wedd, Griff Young, Anderson Murray

Overview

- Background and setting the scene
- Initial exploration: GraphCast and SFNO (main focus sub-seasonal/multi-week)
- Moving to AIFS-based systems, starting with AIFS and weather forecasting
- Future plans

Initial exploration: GraphCast and SFNO

Evaluation of pre-trained ML weather models GraphCast and FourCastNetV2 ("SFNO") for sub-seasonal prediction

ML models:

- Pre-trained on ERA-5 data
- Uses an 6h-autoregressive timestep
- We assess forecasts out to 28-days (although they were NOT developed with S2S in-mind)
- No ensemble

Experimental Testbeds/Hindcasts

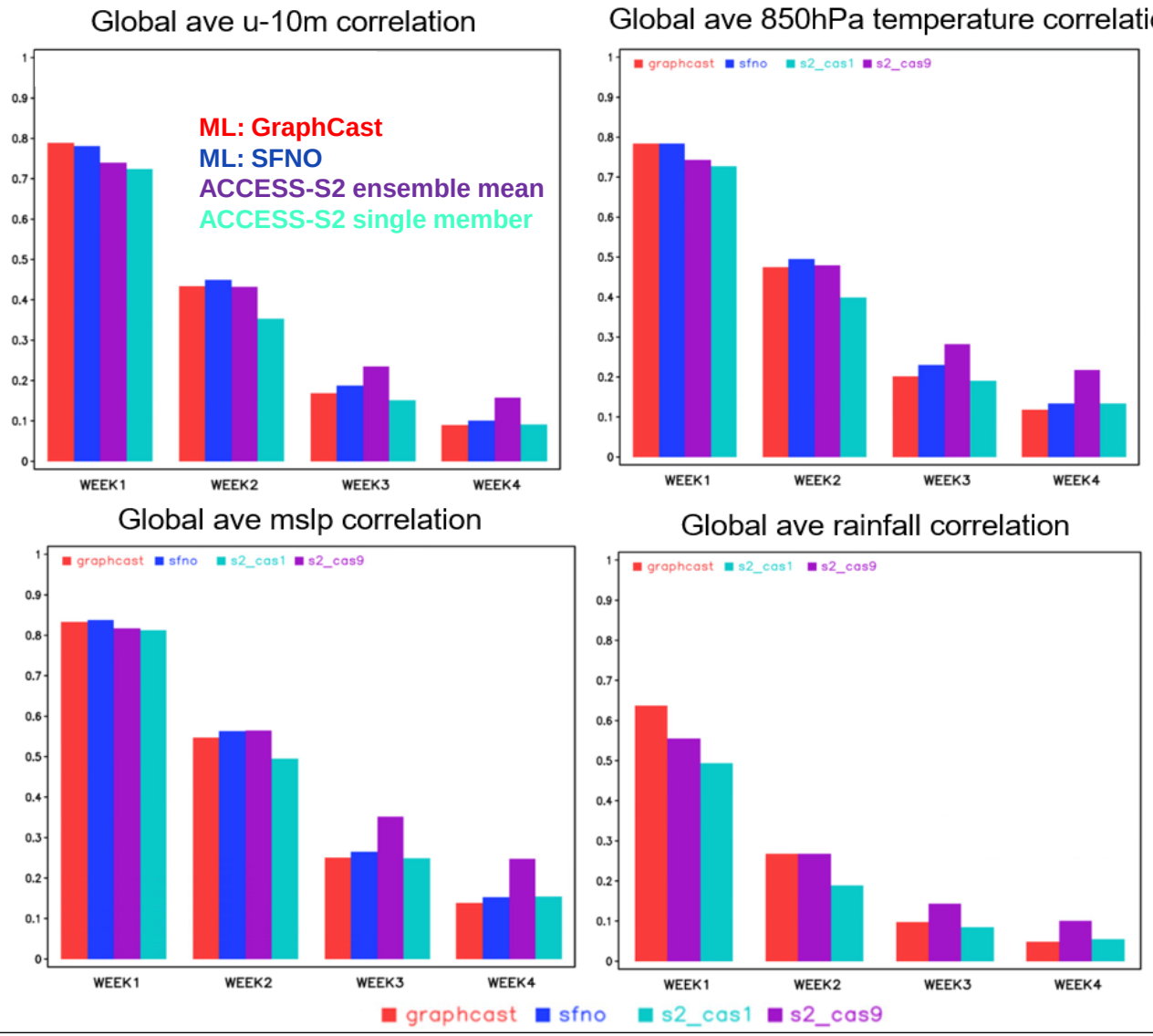
Extended hindcast:
38 years
Forecasts on 1st of each month (456 start dates)

- Same period as the ACCESS-S2 hindcast, allowing for comparison with our operational seasonal model
- However, period overlaps with ML training period
- A 9-member ACCESS-S2 ensemble is used for some comparisons
- Focus of results shown here

Mini-hindcast:
Jan 2022 – Jun 2024
1st and 16th each month (60 start dates)

- Several of the latest dynamical model versions from the Met Office (e.g., GC5) are being evaluated over this testbed
- The period is out-of-sample used for ML training
- Evaluation of anomalies? Statistical significance of results? Enough samples of climate drivers?

Multi-week skill: correlation



Compare ML models to ACCESS-S2 9-member ensemble mean and single member

- Forecasts of anomalies
- ML models slightly better in weeks 1 and 2
- For *most* variables, at weeks 3 and 4, ACCESS-S2 single member is similar to ML models
- Gains in skill to week 3 and 4 from ACCESS-S2 ensemble
- How will an ML ensemble approach perform?

From: Li Shi

Multi-week skill: correlation

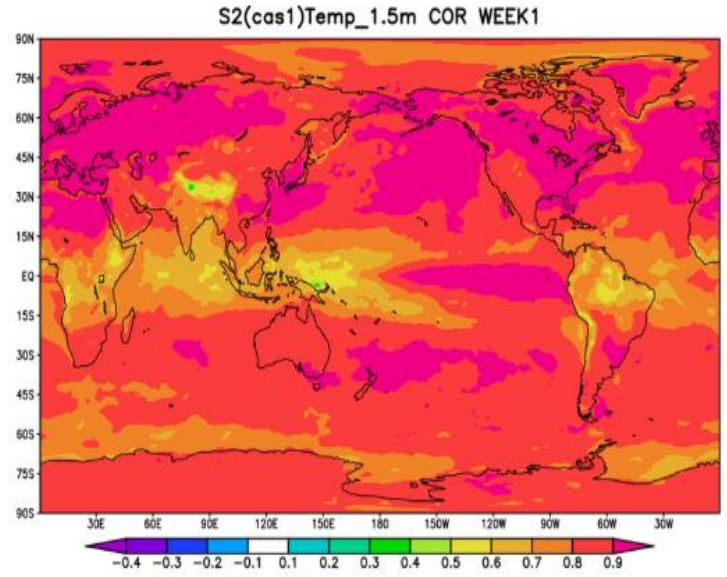
Screen-level temperature

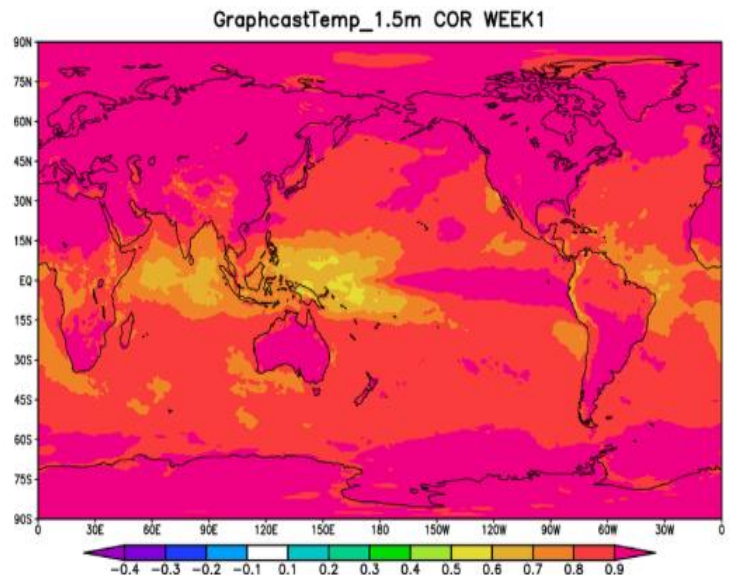
ACCESS-S2 (single member)

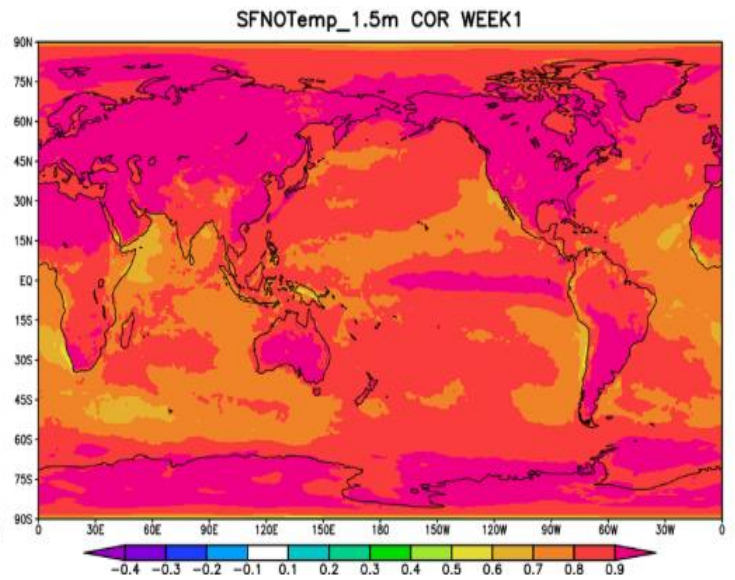
GraphCast

SFNO

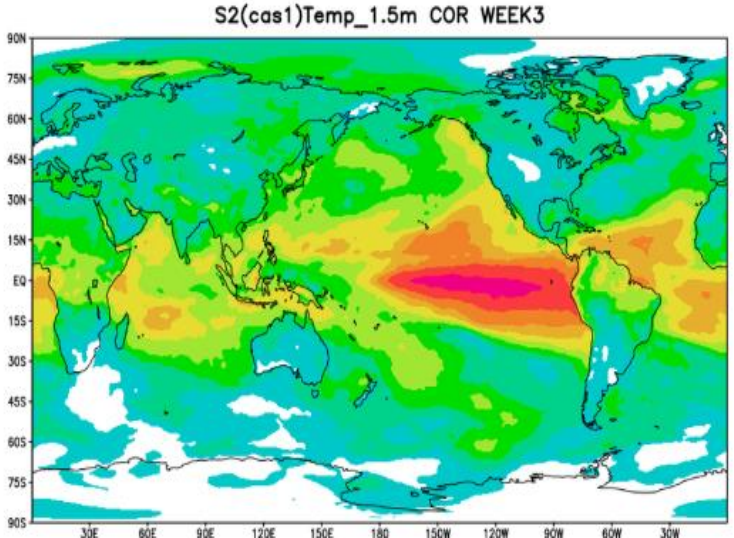
Week 1

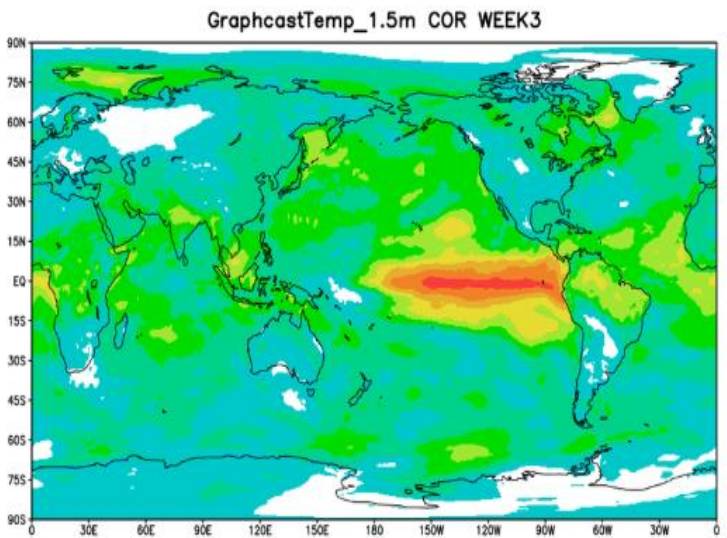


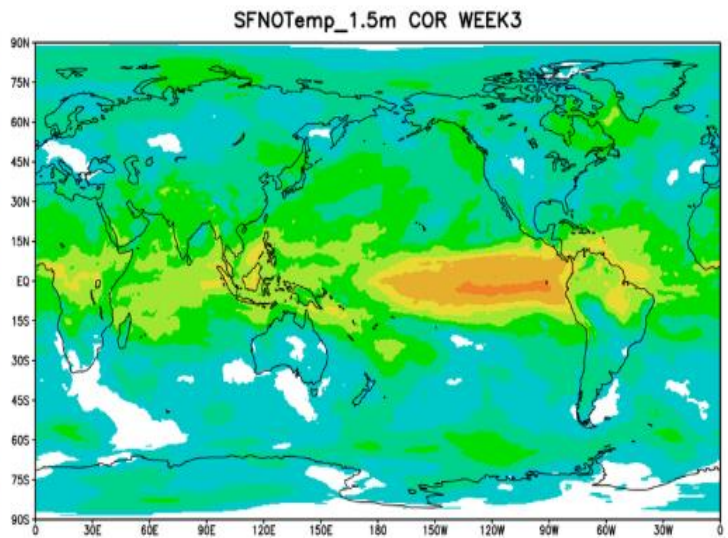




Week 3







From: Li Shi

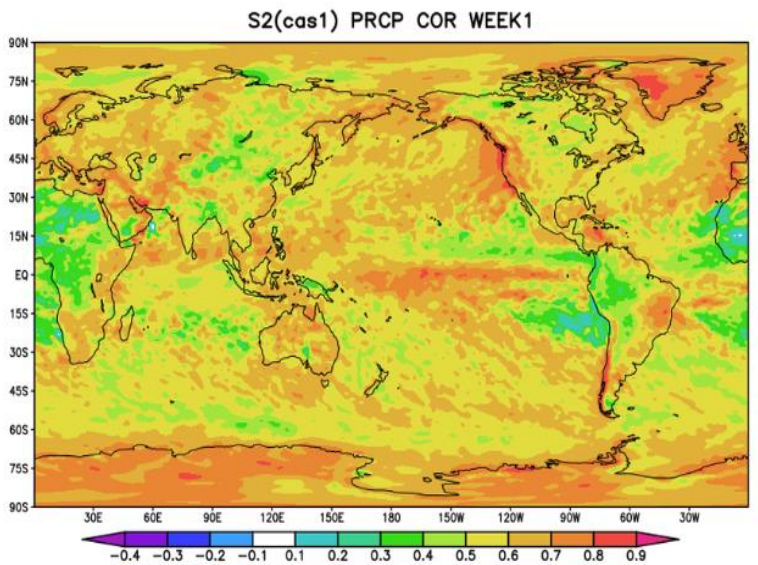
Multi-week skill: correlation

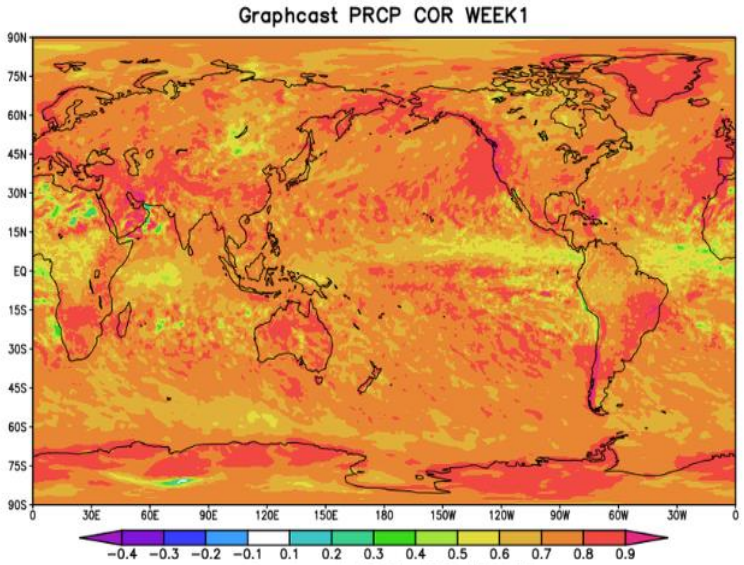
Rainfall

ACCESS-S2 (single member)

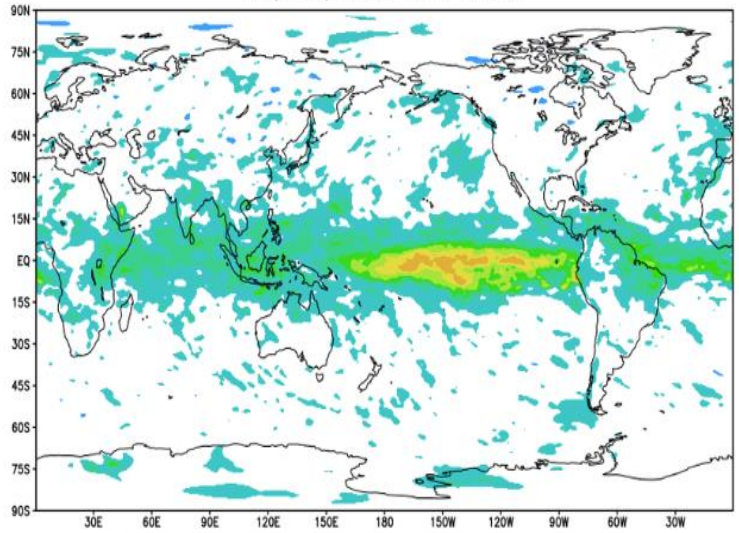
GraphCast

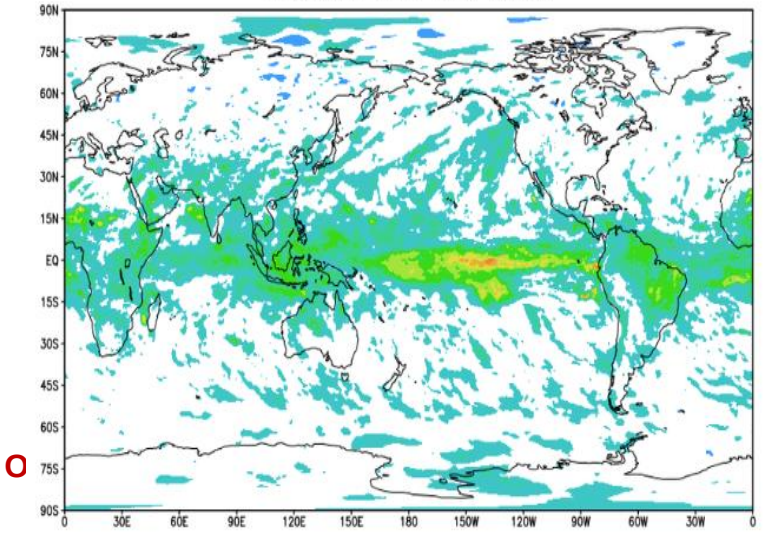
Week 1





Week 3





From: Li Shi

MJO performance in the ML models?

Two cases

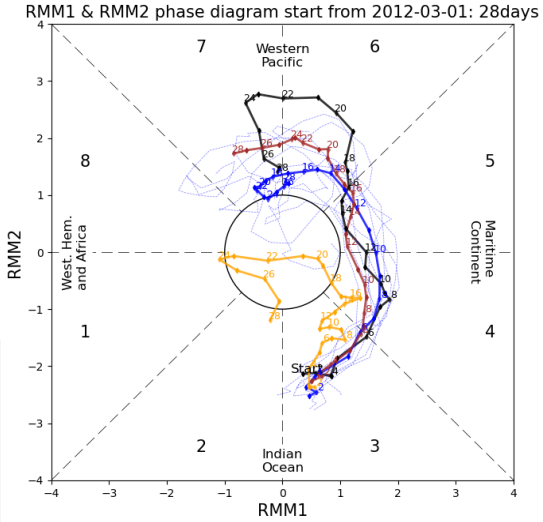
ERA5

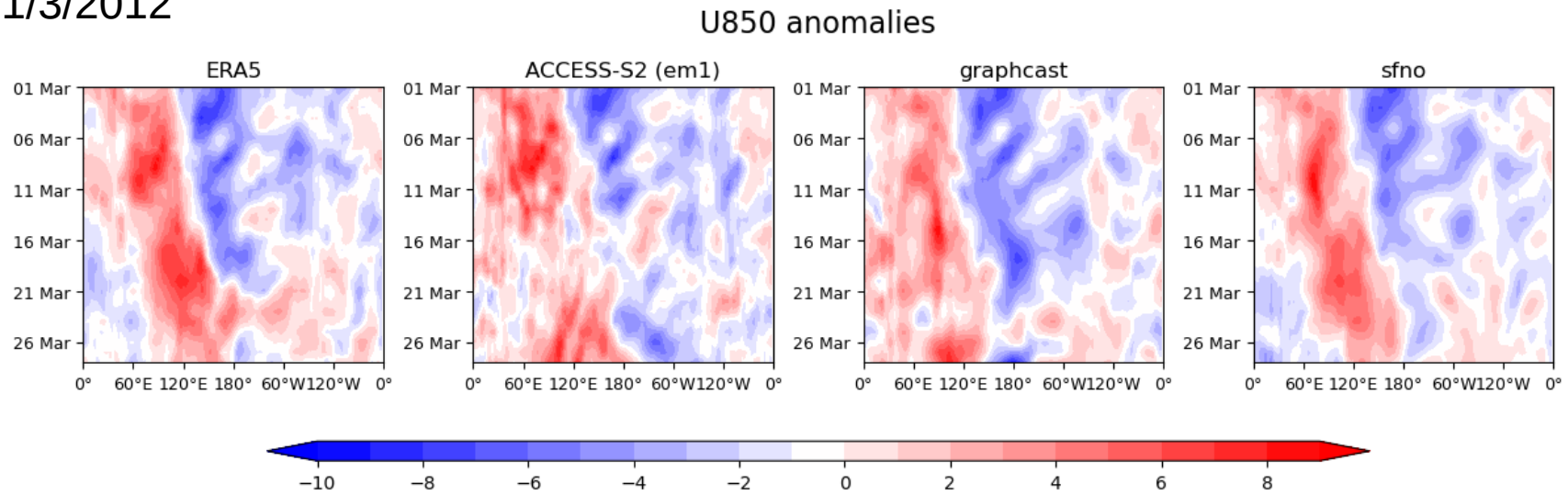
ACCESS-S2

GraphCast

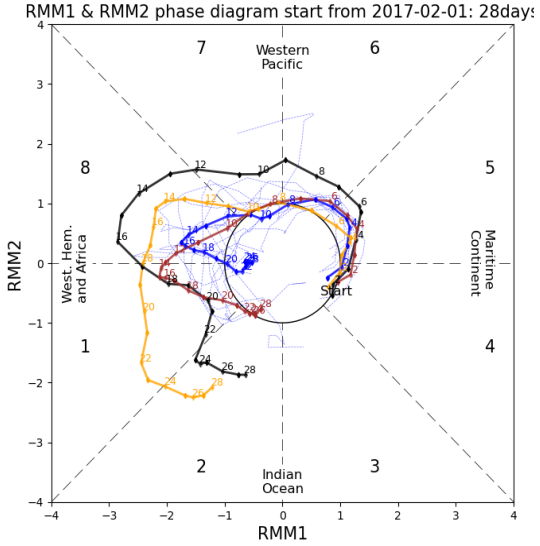
SFNO

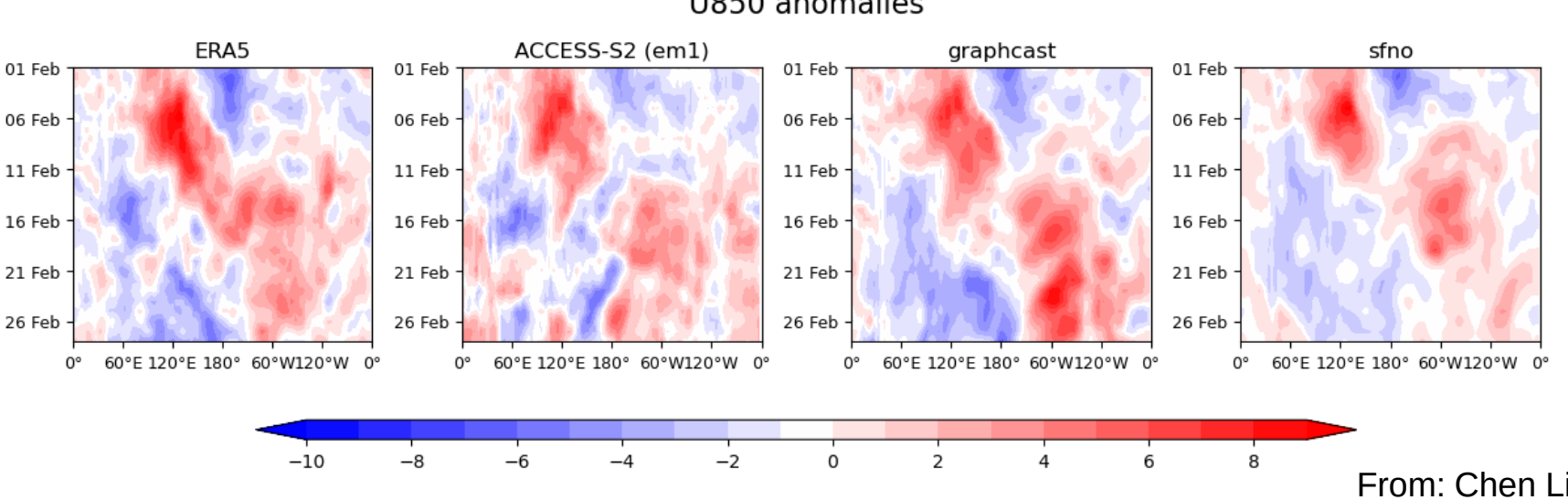
Initialised: 1/3/2012





Initialised: 1/2/2017

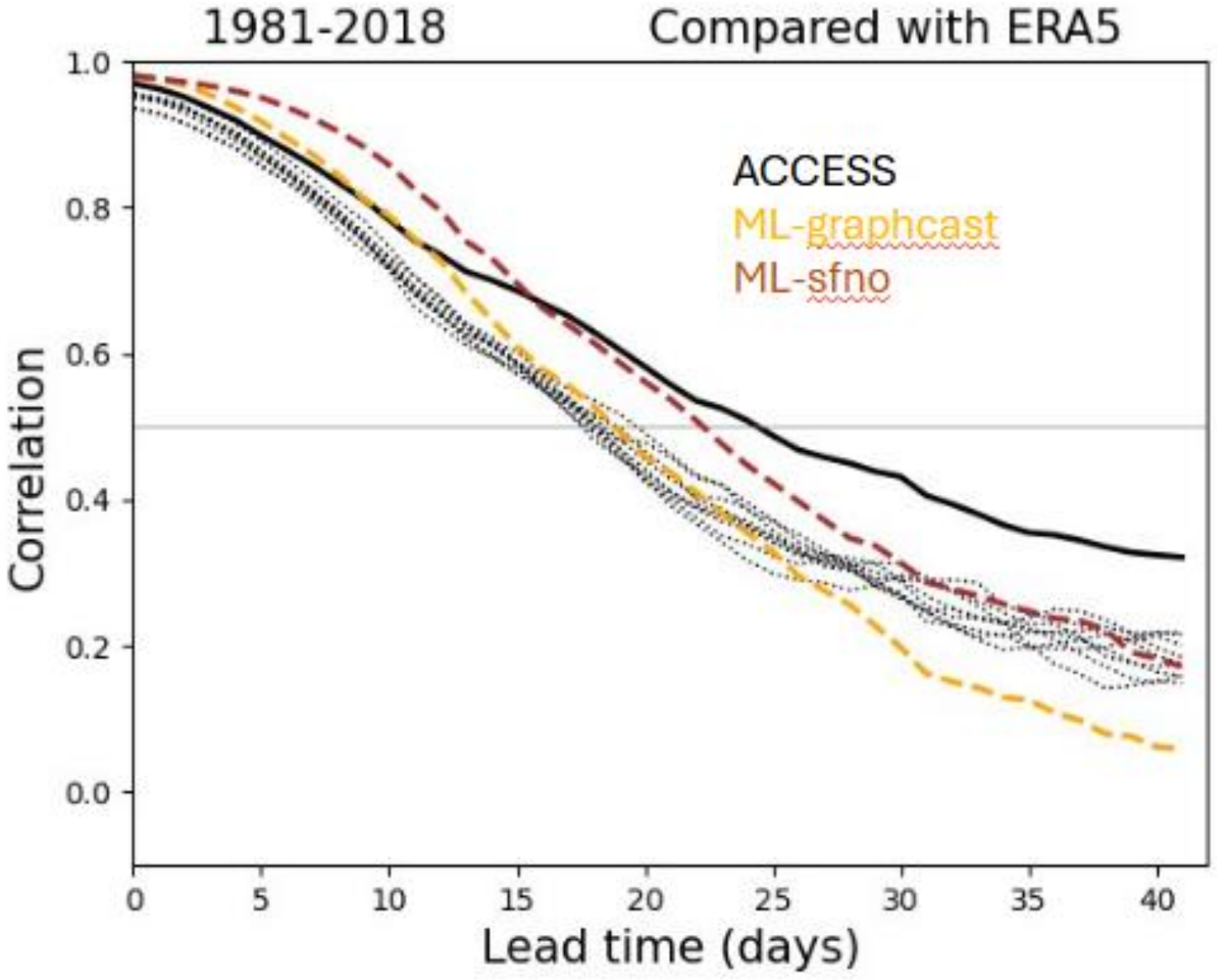




From: Chen Li

MJO skill

Correlation



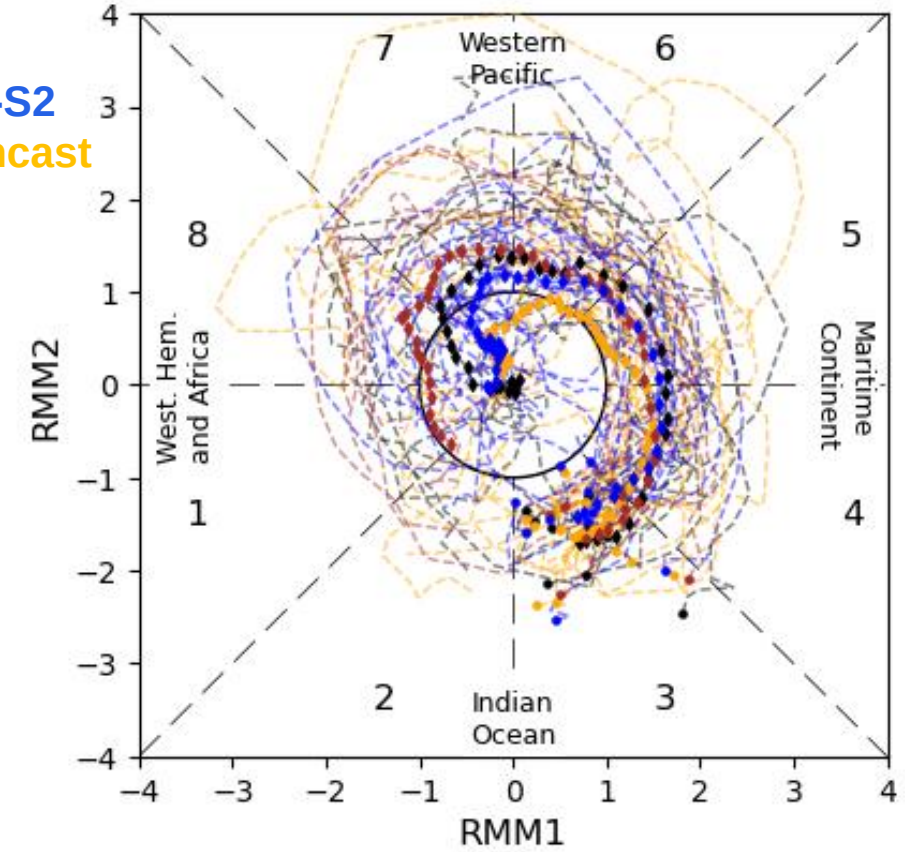
From: Chen Li

MJO: speed of propagation

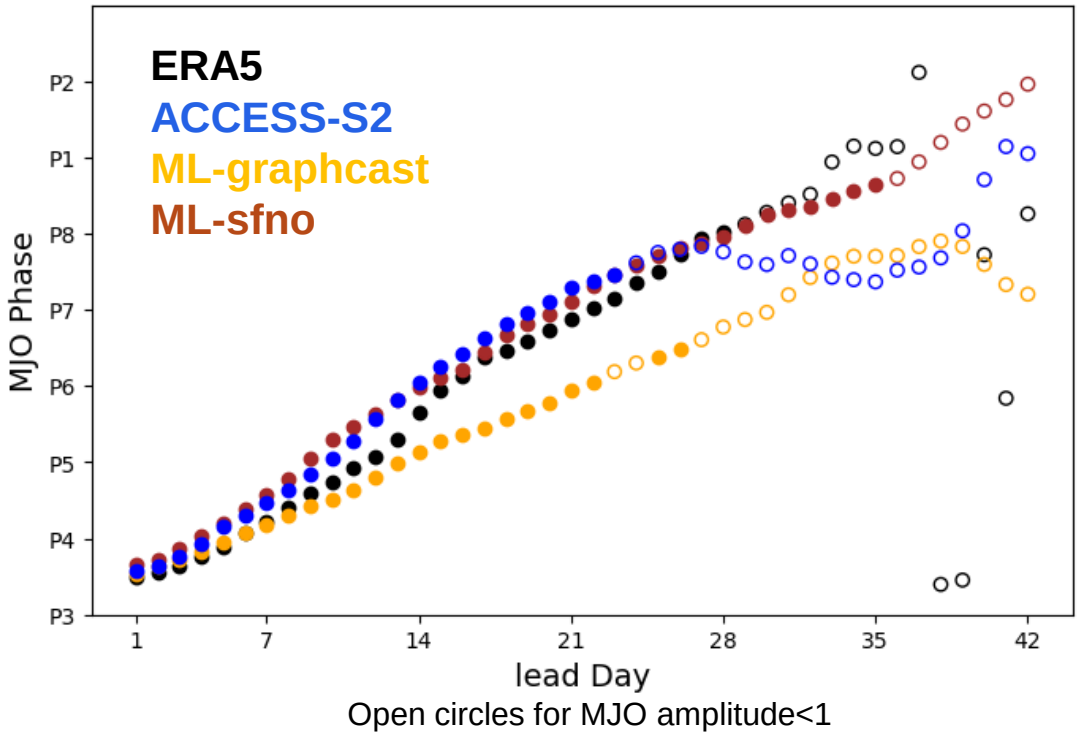
- Composites of MJO events by initial phase
- GraphCast propagation too slow; SFNO close to ERA5 (and similarly for NCEP/NCAR reanalysis)

Example for MJO's starting in Phase 3

ERA5
ACCESS-S2
ML-graphcast
ML-sfno



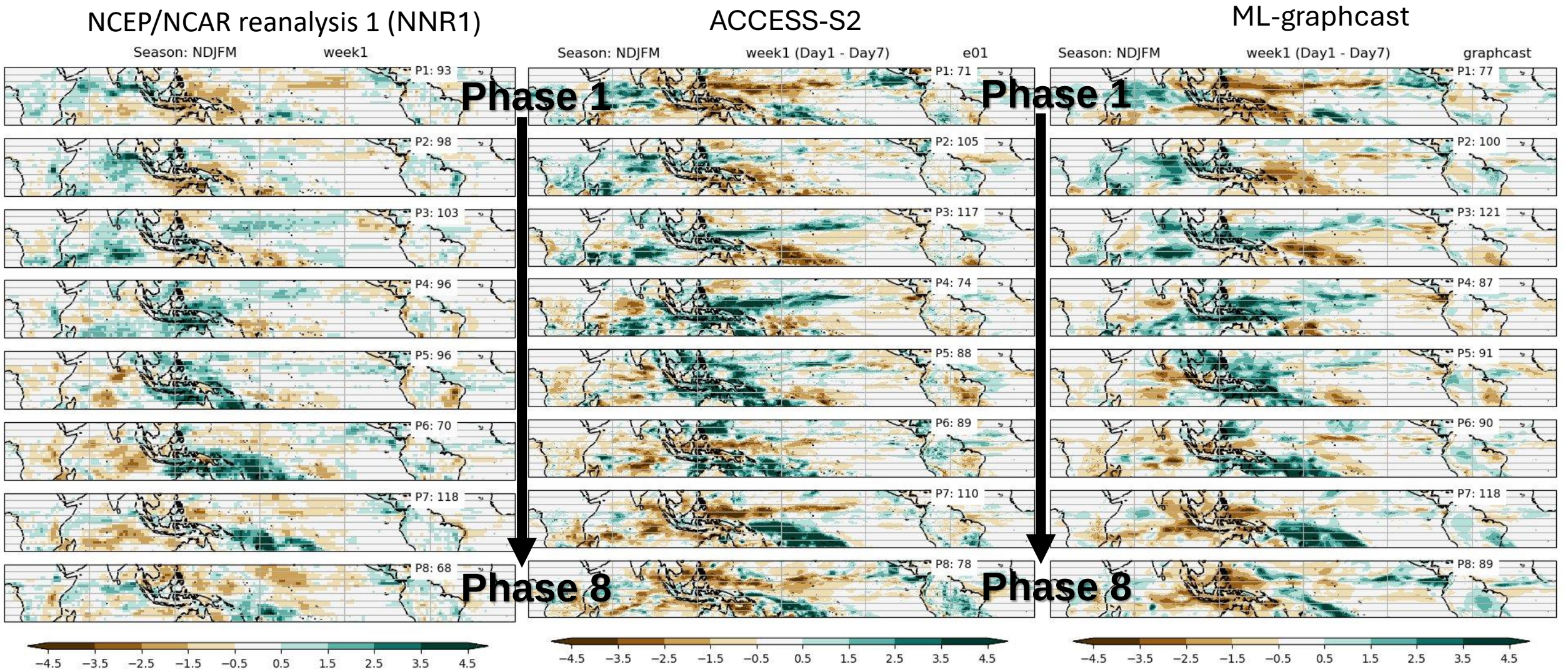
Composite MJO speed
(starting in Phase 3)



Open circles for MJO amplitude<1

MJO-rainfall composites

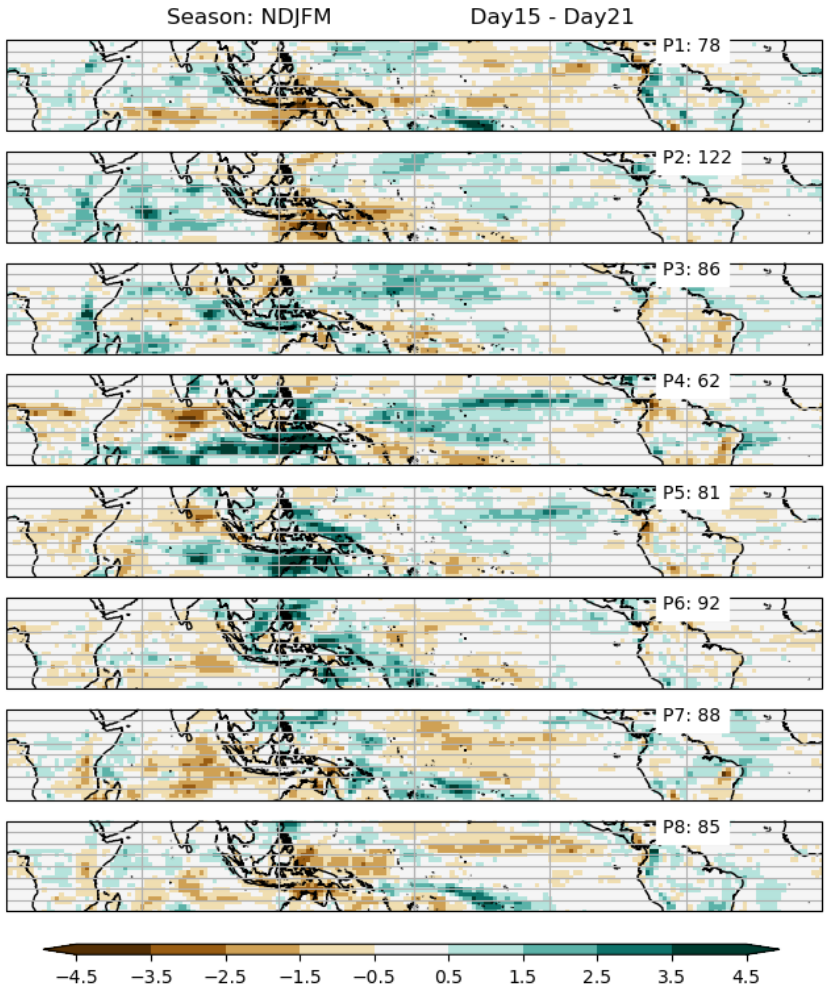
Week 1 (days 1-7)



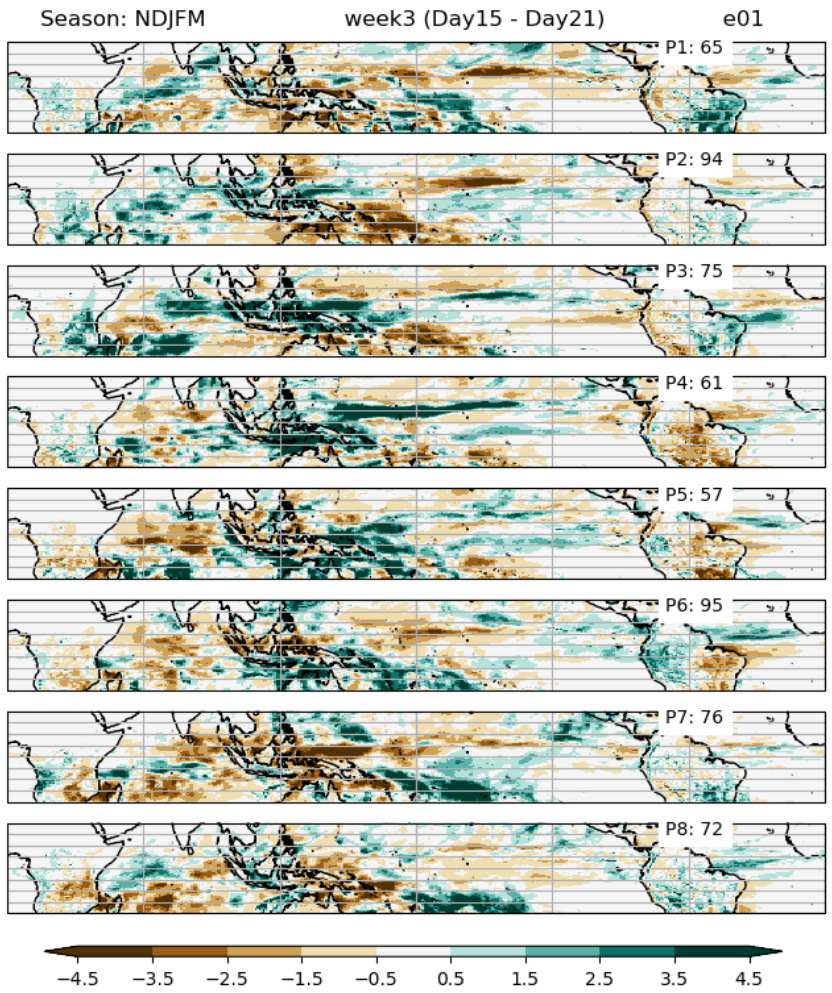
From: Chen Li

Week 3 (days 15-21)

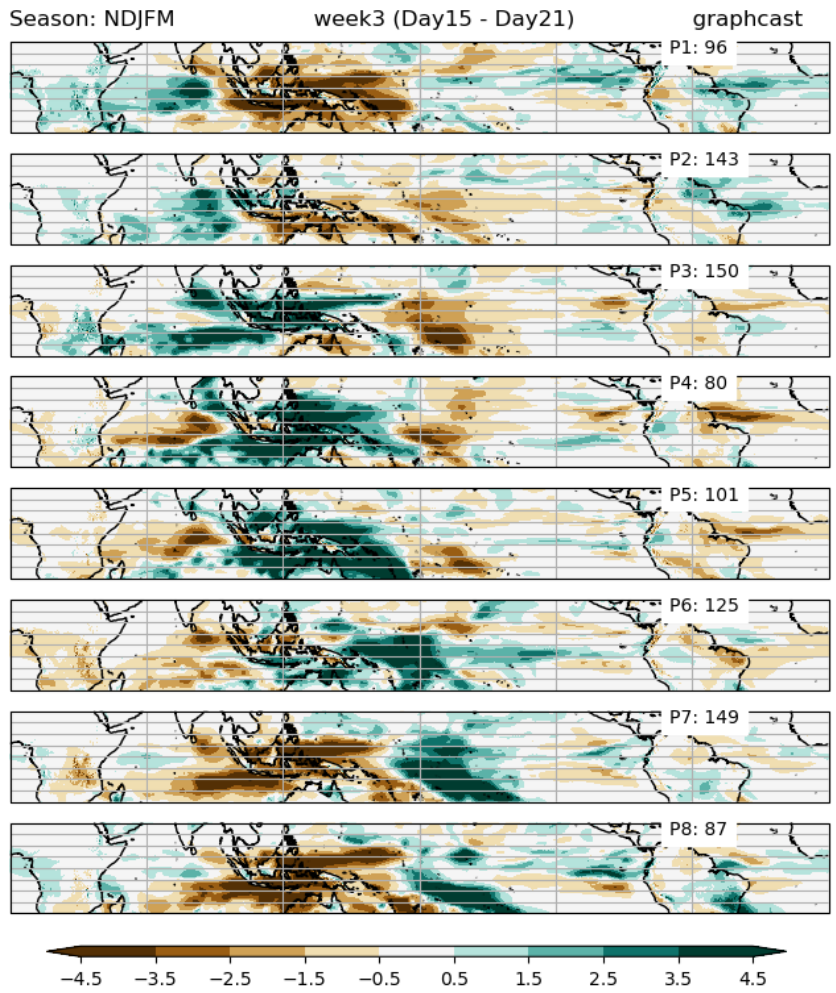
NCEP/NCAR reanalysis 1 (NNR1)



ACCESS-S2



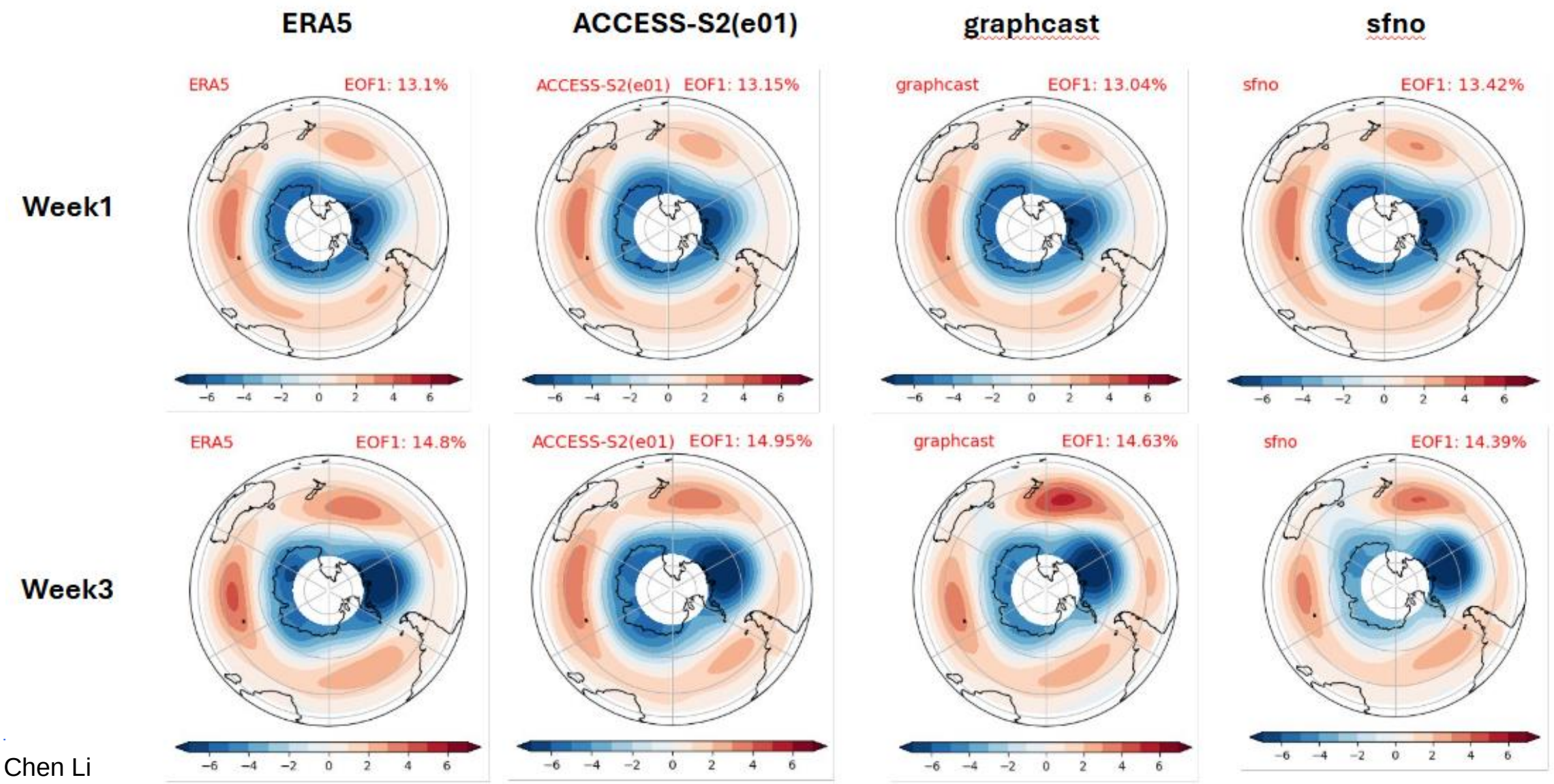
ML-graphcast



From: Chen Li

SAM performance in the ML models?

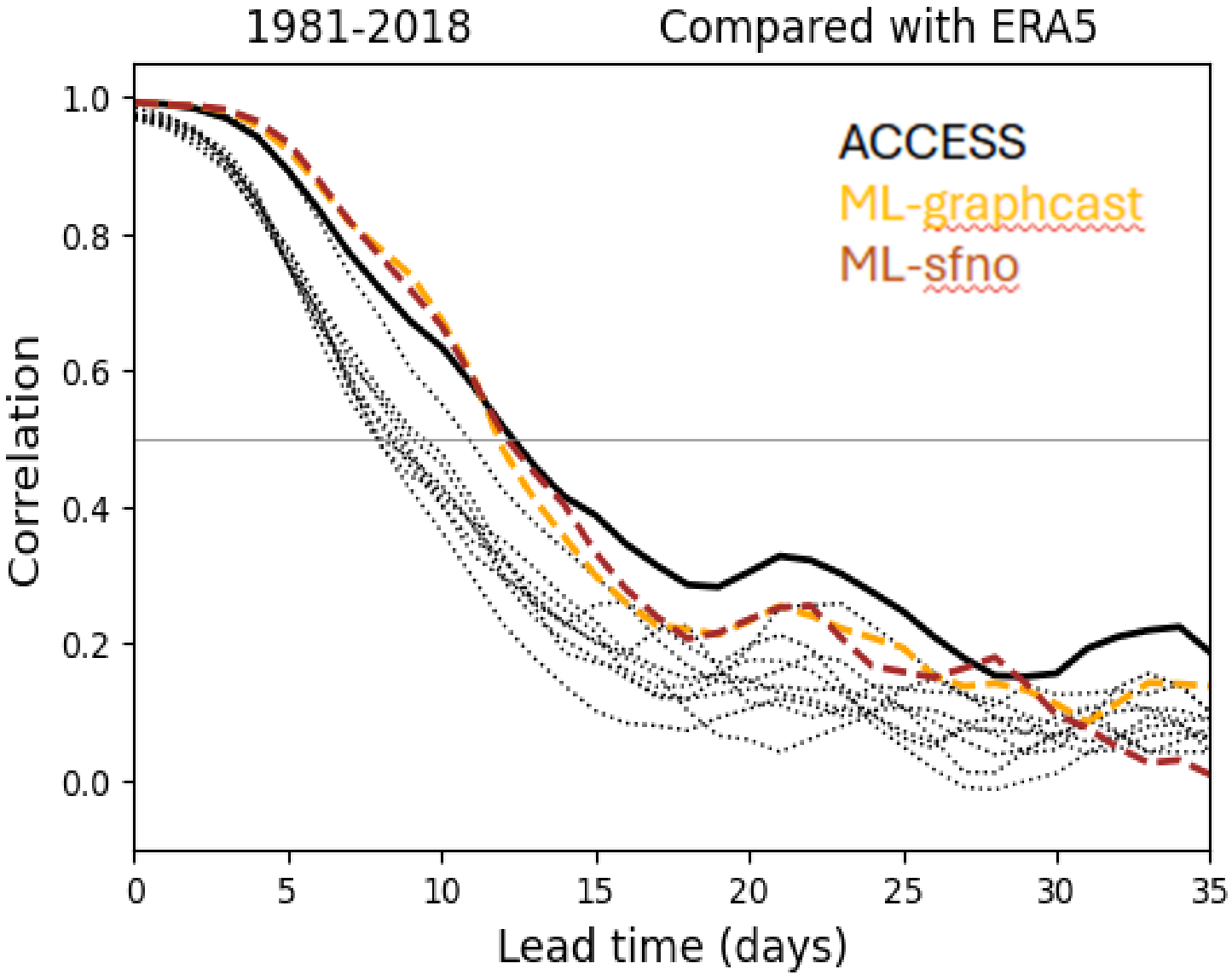
SAM pattern (EOF1 mode)



From: Chen Li

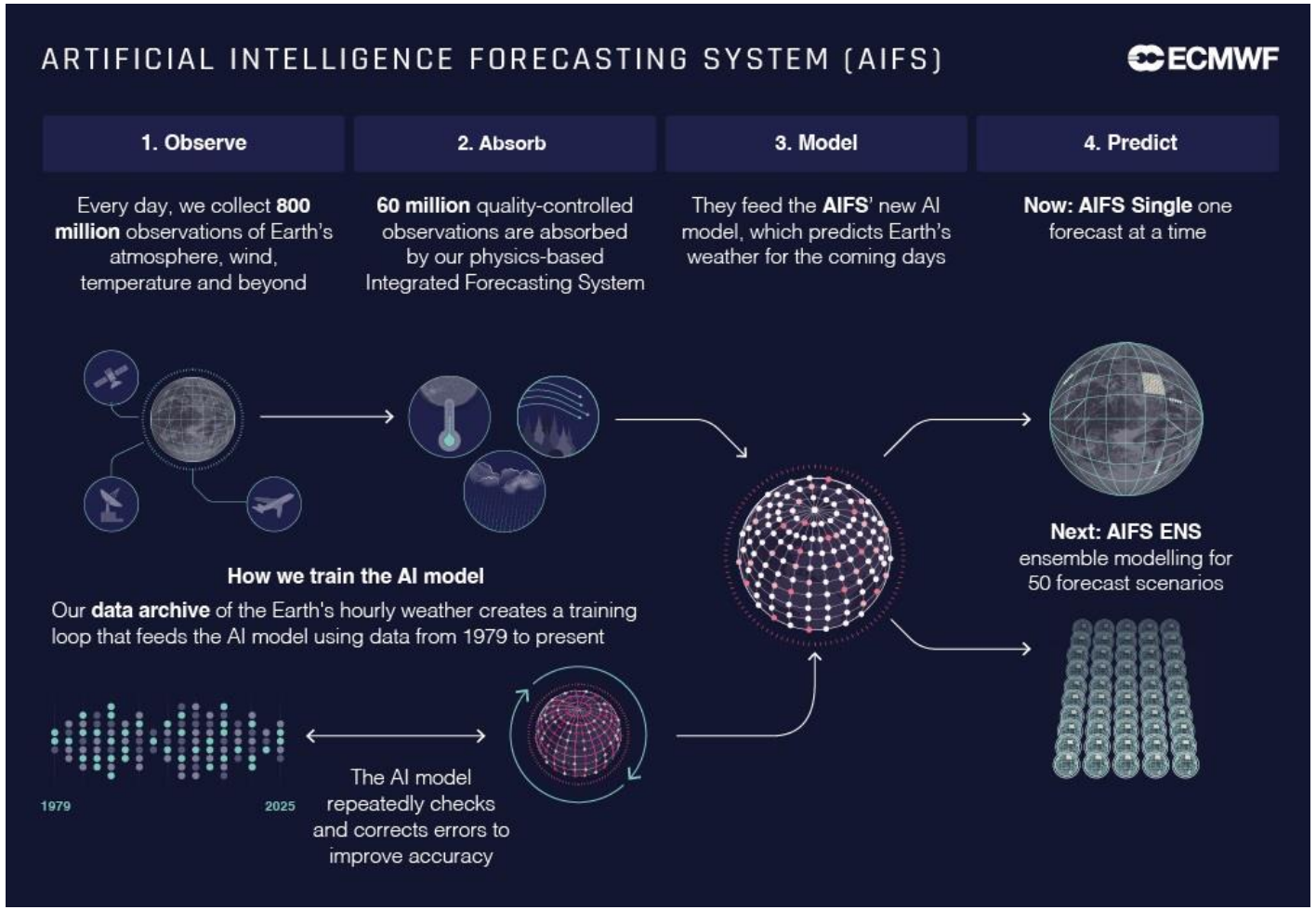
SAM skill

Correlation



From: Chen Li

Moving on to investigate ECMWF ML systems… Starting with AIFS



https://www.ecmwf.int/en/about/media-centre/news/2025/ecmwfs-ai-forecasts-become-operational

Artificial Intelligence Forecasting System (AIFS)
Operational: 25 February 2025

AIFS Single: deterministic forecasts

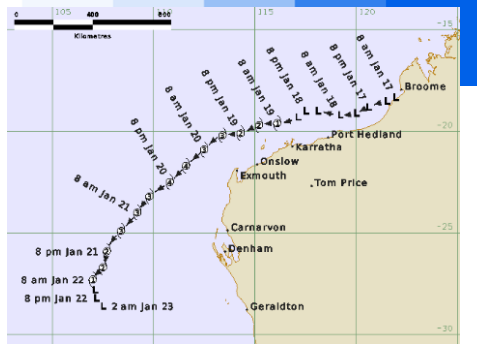
AIFS outperforms state-of-the-art physics-based models for many measures, including TC tracks, with gains of up to 20%

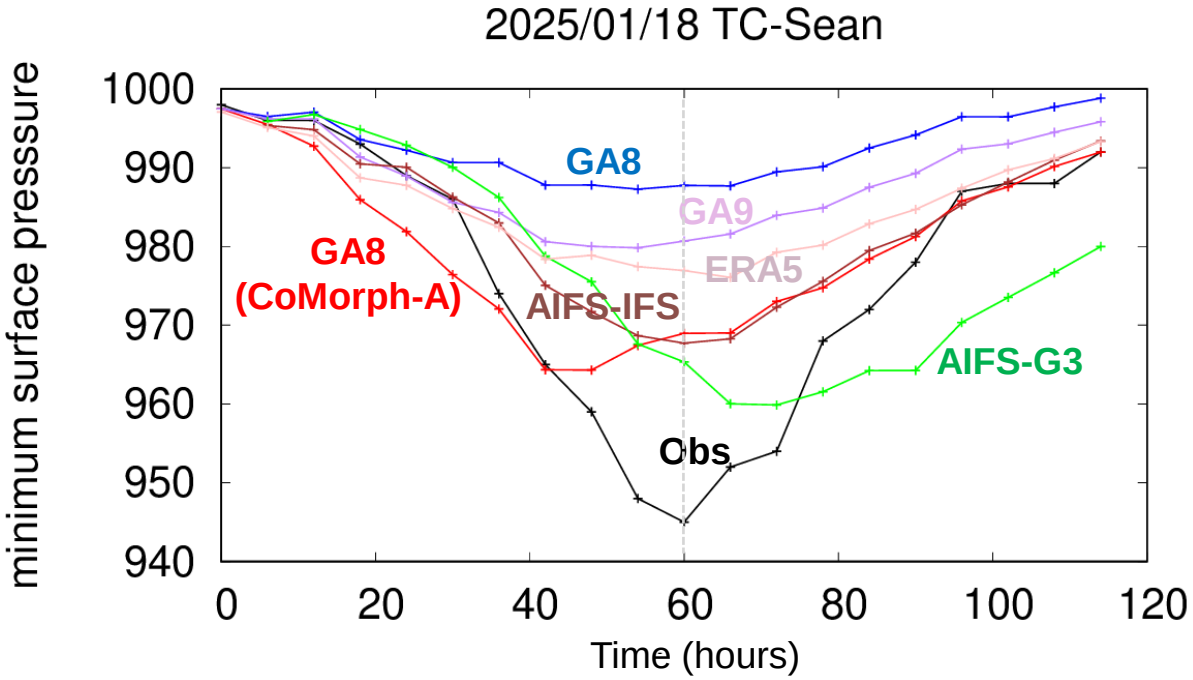
Grid spacing ~28km

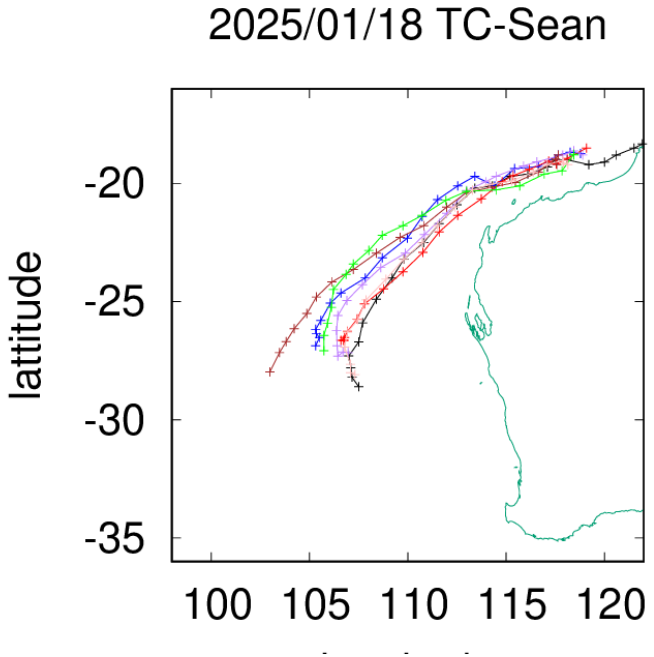
Implemented at the Bureau (NCI) to run forecasts using initial conditions from IFS, ACCESS-G and ERA5

AIFS case studies

TC Sean (Jan 2025)





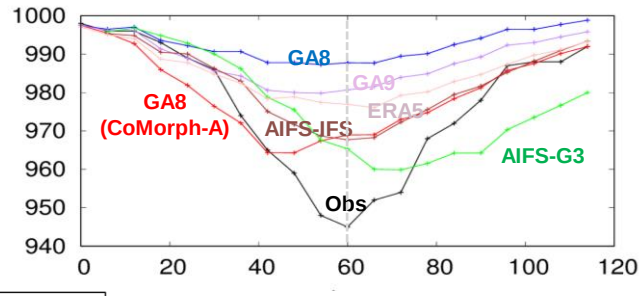


More accurate intensity forecasts with AIFS compared to GA8/GA9

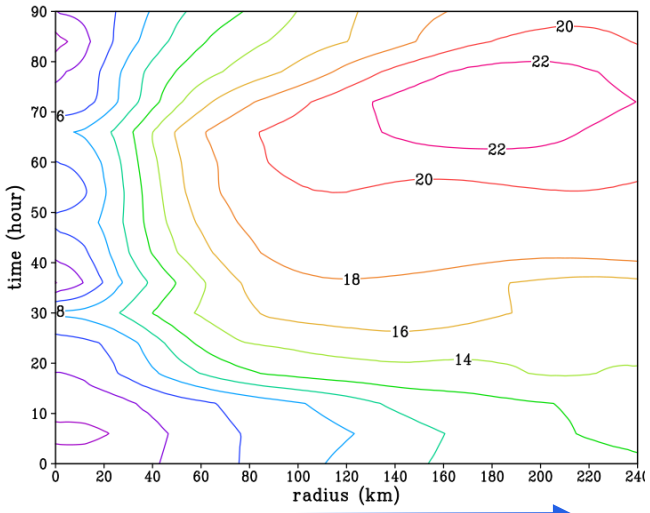
AIFS-IFS: initial conditions are IFS
AIFS-G3: initial conditions are ACCESS-G3
GA8: model used in ACCESS-G4
GA9: next generation global model
GA8-CoM: GA8 using the new CoMorph-A convection scheme

From: Hongyan Zhu

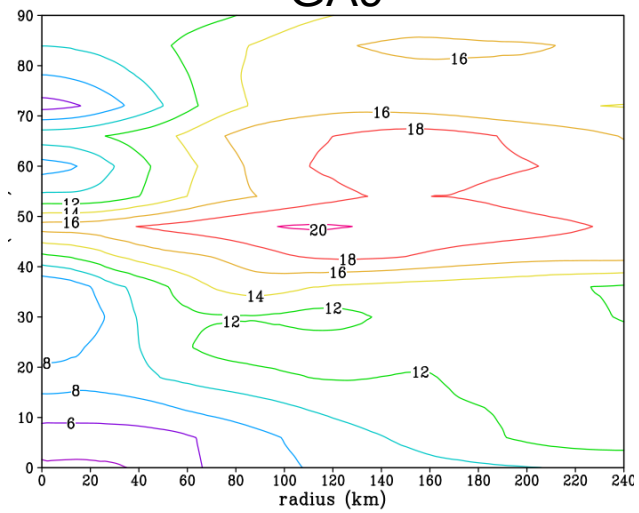
TC Sean: Azimuthal Wind (at 850hPa)



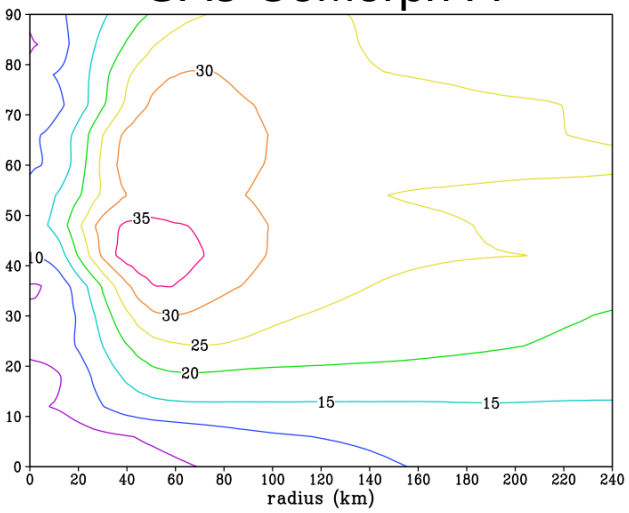
GA8



GA9



GA8-CoMorph-A

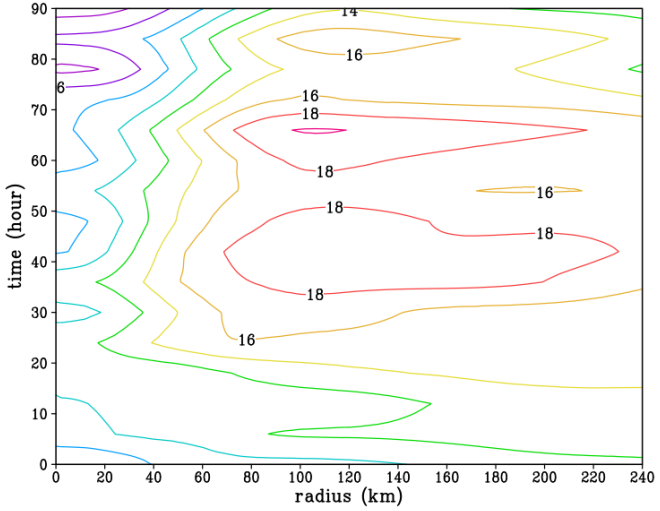


Time

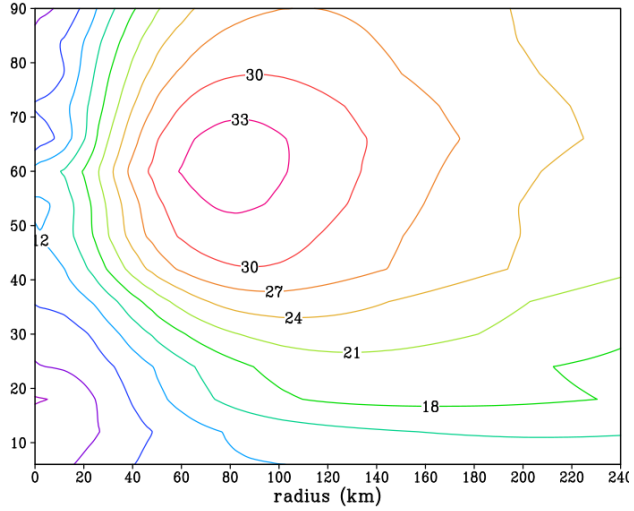
Distance from centre of the TC

Radius of max azimuthal wind:
~200 km for GA8
~120 km for GA9
~60 km for GA8-CoM

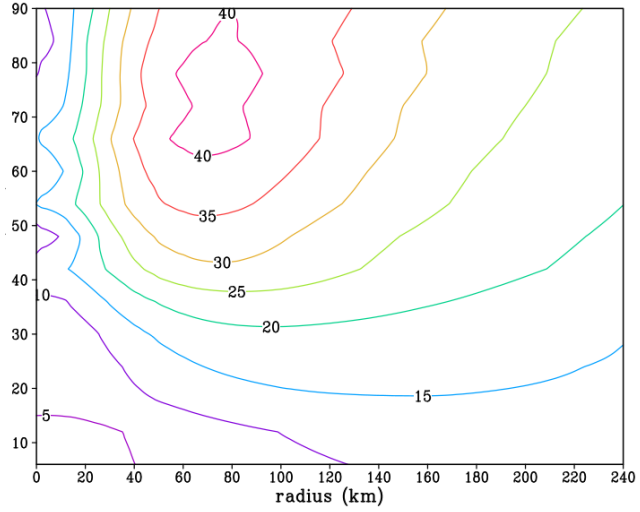
ERA5



AIFS with IFS IC



AIFS with ACCESS-G3 IC



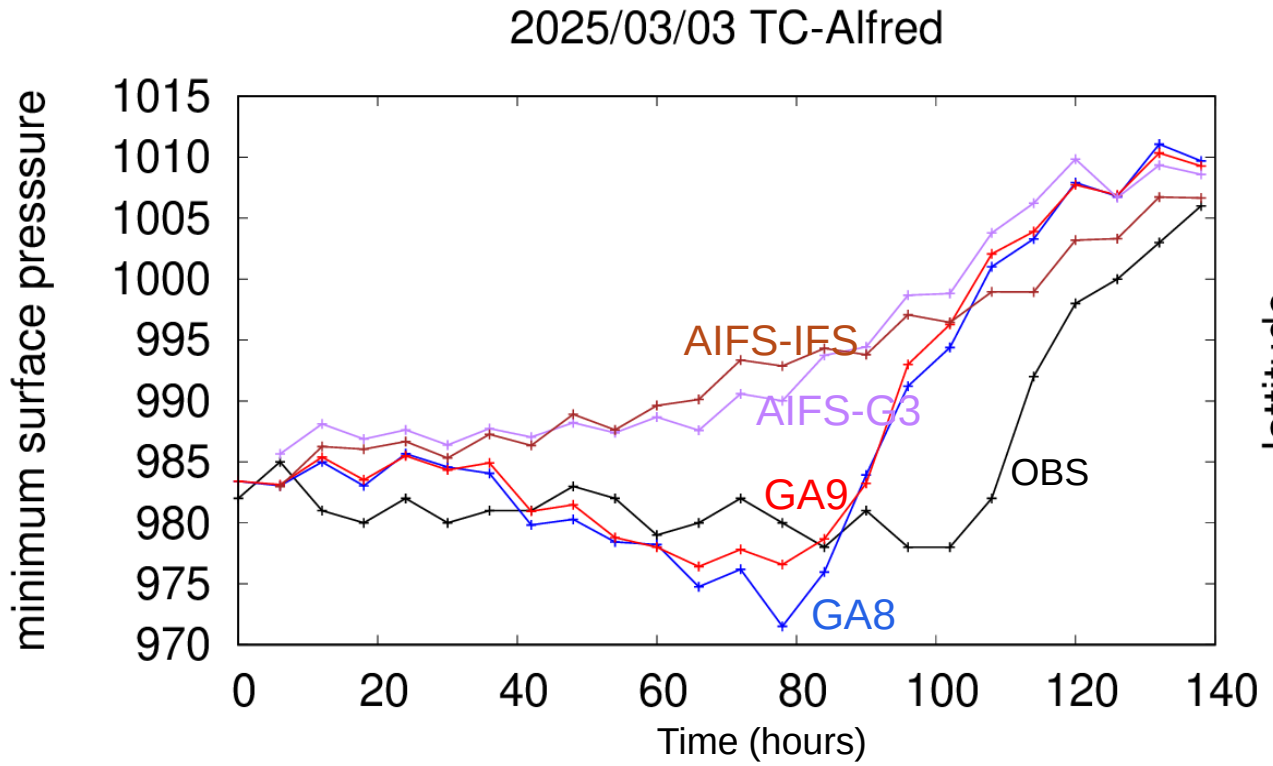
Time

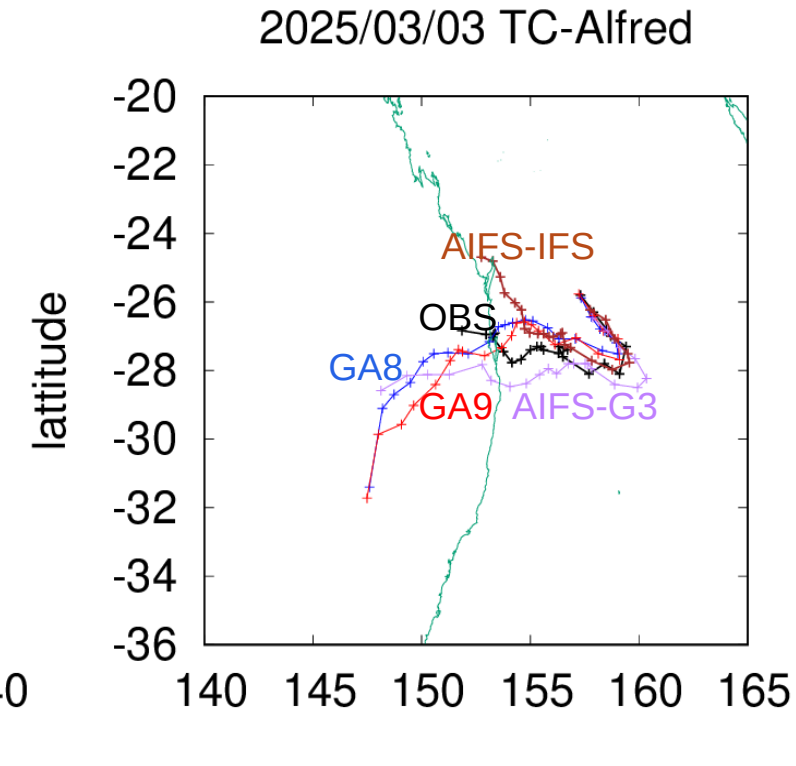
Radius of max azimuthal wind
~100 km for ERA5
~80 km for both ML versions

From: Hongyan Zhu

AIFS NWP case studies

TC Alfred (Mar 2025)





GA8 and GA9 have more accurate intensity forecasts compared to AIFS. The AIFS-G3 track forecast aligns most closely with observations, while AIFS-IFS predicts a northward movement of the TC instead of a westward movement before landfall.

AIFS-IFS: initial conditions are IFS
AIFS-G3: initial conditions are ACCESS-G3
GA8: model used in ACCESS-G4
GA9: next generation global model

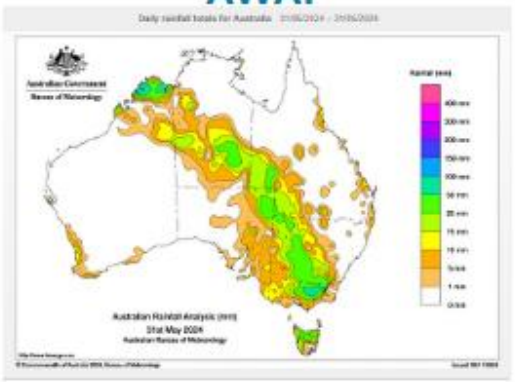
From: Hongyan Zhu

AIFS NWP case studies

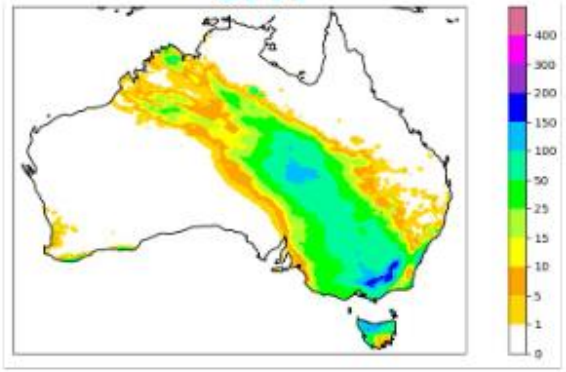
May 2024 cold front: produced damaging winds and widespread rain

Rainfall "observations": 30 May 2024

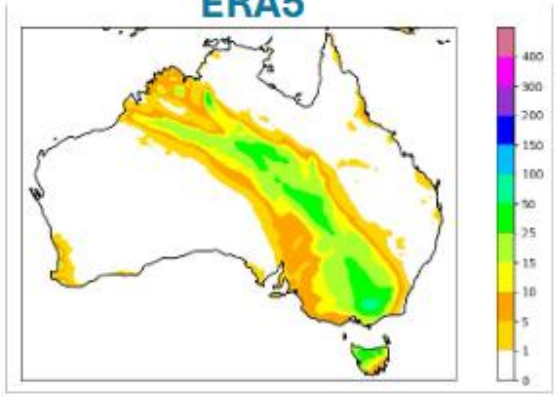
AWAP



GPM

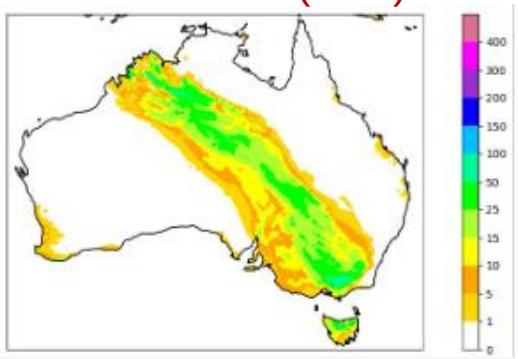


ERA5



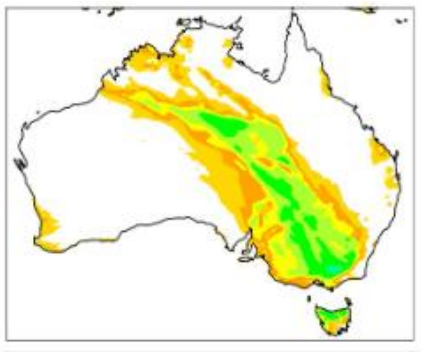
Forecasts for 30 May 2024 (48 hours lead time). Results interpolated to 0.25º x 0.25º

ACCESS-AE (em0)



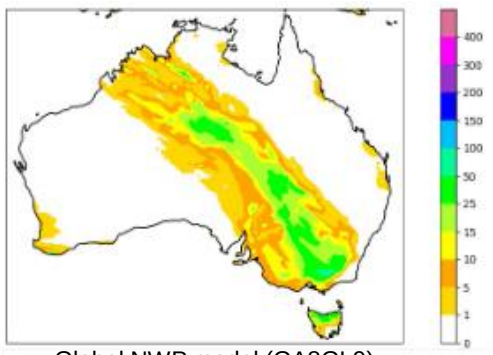
Pan-Australia regional NWP model (RAL3, central member shown)

ACCESS-G3



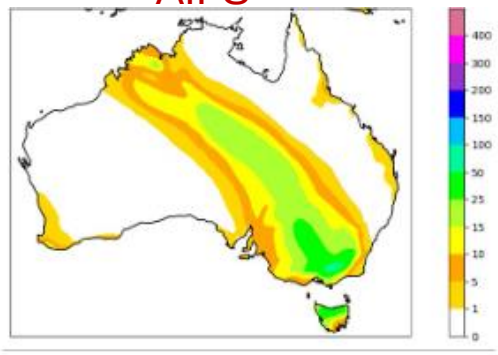
Global NWP model (GA6GL6)

ACCESS-G4



Global NWP model (GA8GL9)

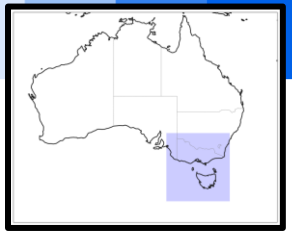
AIFS



From: Chen Li

OFFICIAL

Fraction Skill Score (FSS) results: JJA 2024 (4 starts per day: n=368)



Rainfall for VIC and Tasmania region

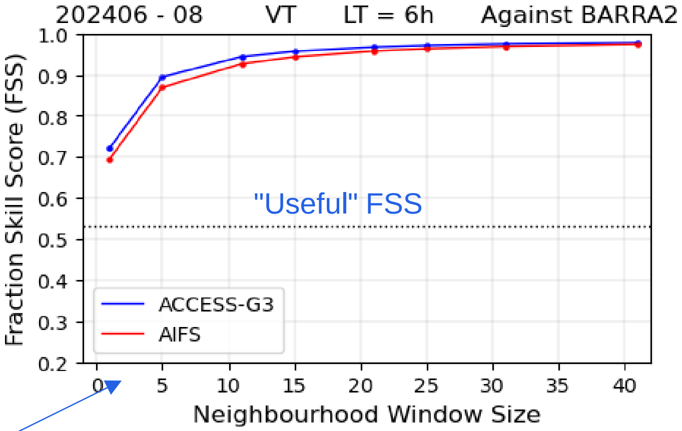
Threshold: **0.5mm/hr** | Threshold: **1mm/hr** | Threshold: **2mm/hr**

Lead Time: 6hr

Good

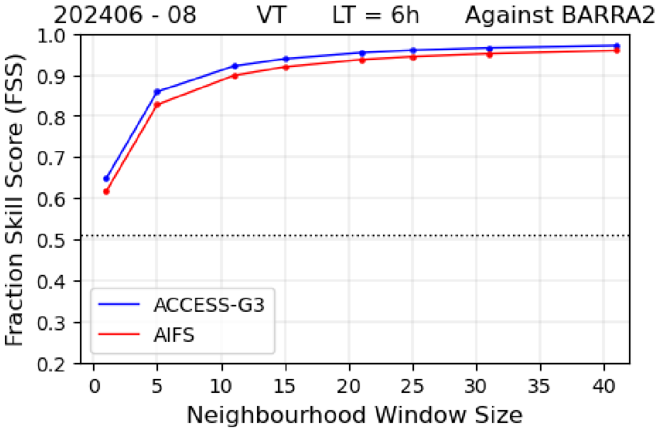
FSS

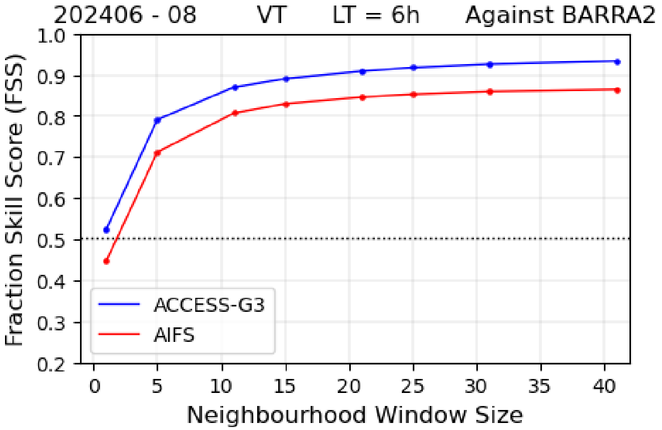
Poor



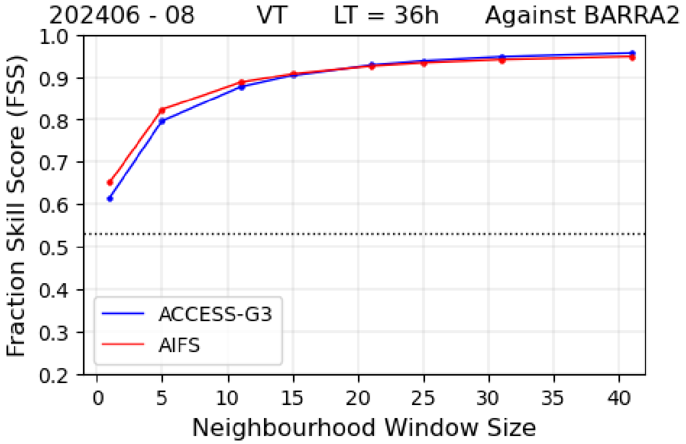
"1" window size is ~ 25km

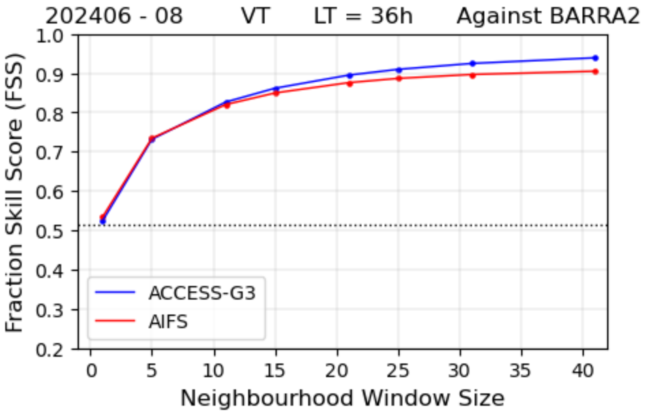
Spatial scale (width of squares)

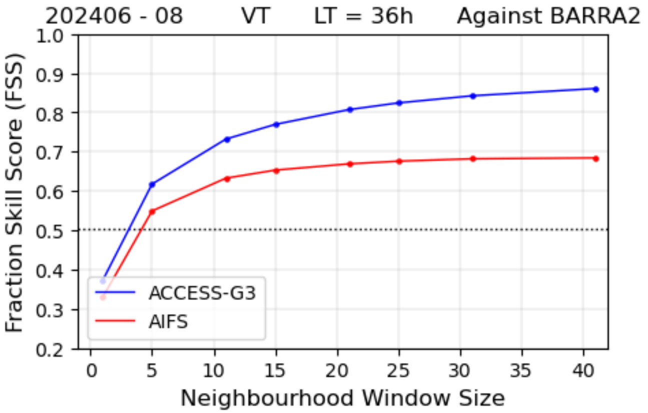




Lead Time: 36hr





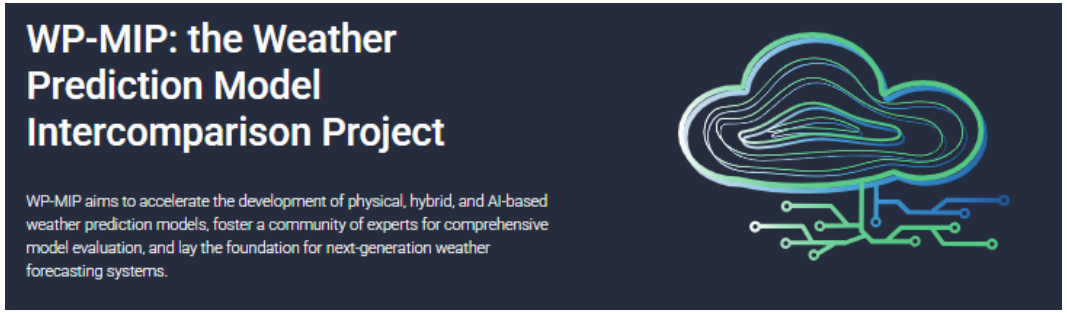


From: Chen Li

Future plans

Specific immediate term goals:

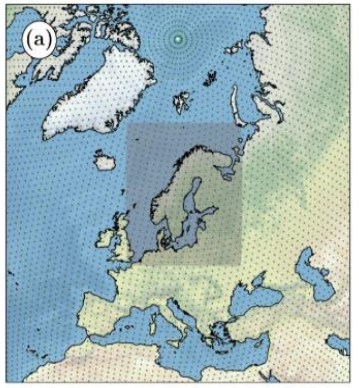
- More exploration of AIFS
- Contribute to WP-MIP (AIFS initialized with ACCESS-G3 ICs….)
- Implement AIFS-CRPS and evaluate for Australia (and to S2S)
- Implement Bris (Met Norway stretched grid model), an extension of AIFS
- Australian configuration of Bris with high resolution component over Australia, fine-tuned on BARRA reanalysis
- Integrating ML models into post-proc (IMPROVER) (Belinda's poster 2.2)
- Can an ML model drive ocean forecasts?

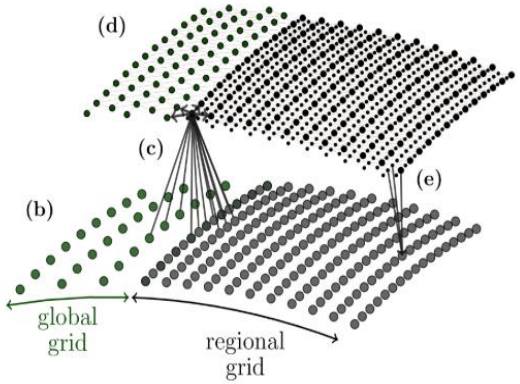


https://www.wcrp-esmo.org/activities/wp-mip

Regional ML weather model with a global stretched-grid (Bris model, Met Norway)

https://www.sigma2.no/research/bris-high-resolution-data-driven-weather-forecasting-model





https://arxiv.org/html/2409.02891v1

More generally:

- Building infrastructure and capability
- Building understanding of how to integrate global ML models into our forecast delivery pipeline
- Strengthening connections with key national and international players
- Starting an ongoing conversation with Bureau Services around global ML

The Bureau of Meteorology

Thank you

debbie.hudson@bom.gov.au