The Development of UFS Coupled GEFS for Subseasonal and Seasonal Forecasts

Yuejian Zhu,
Bing Fu, Bo Yang, Hong Guan, Eric Sinsky,
Wei Li, Jiayi Peng, Xianwu Xue and Dingchen Hou
EMC/NCEP/NWS/NOAA

Phillip Pegion
(PSL/OAR/NOAA)

Xin-Zhong Liang and Sanghoon Shin
University of Maryland

Acknowledgements:
EMC Coupling and Model groups; CPCs and WPO/OSTI/CTB

Present for 47th CDPW (Virtual)
10/26/2022

Highlights:

- Introduction
- Background
  - GEFS SubX system
  - GEFSv12 and reforecasts
- Experiments - UFS coupled GEFS
  - Configurations - UFS P5/GFSv15/GEFSv12
  - Tested the different horizontal resolutions (C192/C384)
  - Tuned stochastic parameters for atmosphere
- Summary

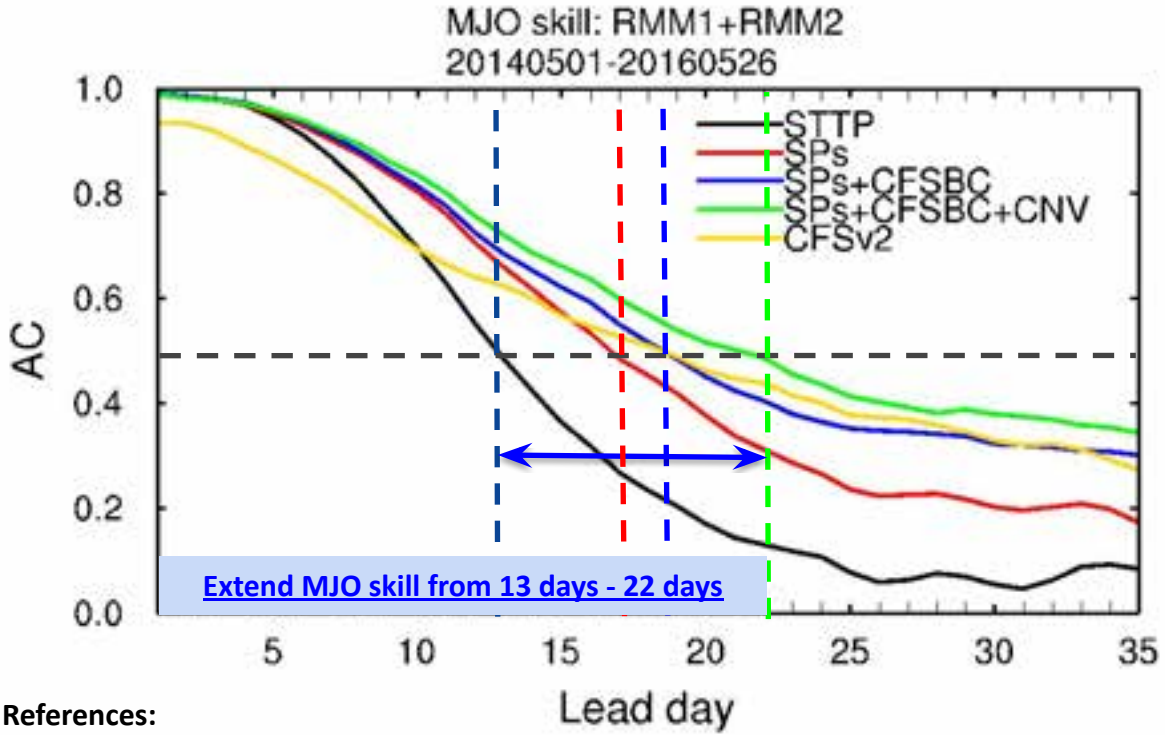
Introduction

- In collaboration with **Unified Forecast System (UFS) Research to Operation/Subseasonal to Seasonal (R2O/S2S) and Medium Range Weather (MRW) applications**, focus of this mainly CTB funded project is to improve NCEP Global Ensemble Forecast System (GEFS) for subseasonal prediction from current operational GEFS.

- The **GEFS v12** was implemented in NCEP operation in September 2020, which is the 1st UFS application with coupling to Wave Watch 3 (WW3) ensembles with 31 members and about 25 km horizontal resolution, out to 35 days to cover subseasonal prediction. The operational GEFS has demonstrated forecast capability and excellent performance for subseasonal time scales including MJO predictions, surface temperature and precipitation through 31 years reforecasts before it was implemented.

- A **quantify forecast uncertainty** is the main work through the integration of **a fully coupled UFS** and various stochastic physical perturbation schemes to initialize **Ensemble Prototype** (EP) experiments toward **GEFSv13 implementation**. The EP experiments are closely configured to UFS prototype experiments. Currently, a UFS prototype 5 (P5) based EP-1 has been completed with optimum atmospheric model perturbations which run a 2-year period (initialized once per week) with 10 perturbed and 1 unperturbed members, out to 35 days. A full evaluation of the EP1 will be presented in terms of various evaluation metrics.

Background (1) - Improvement of MJO skills through NOAA SubX project

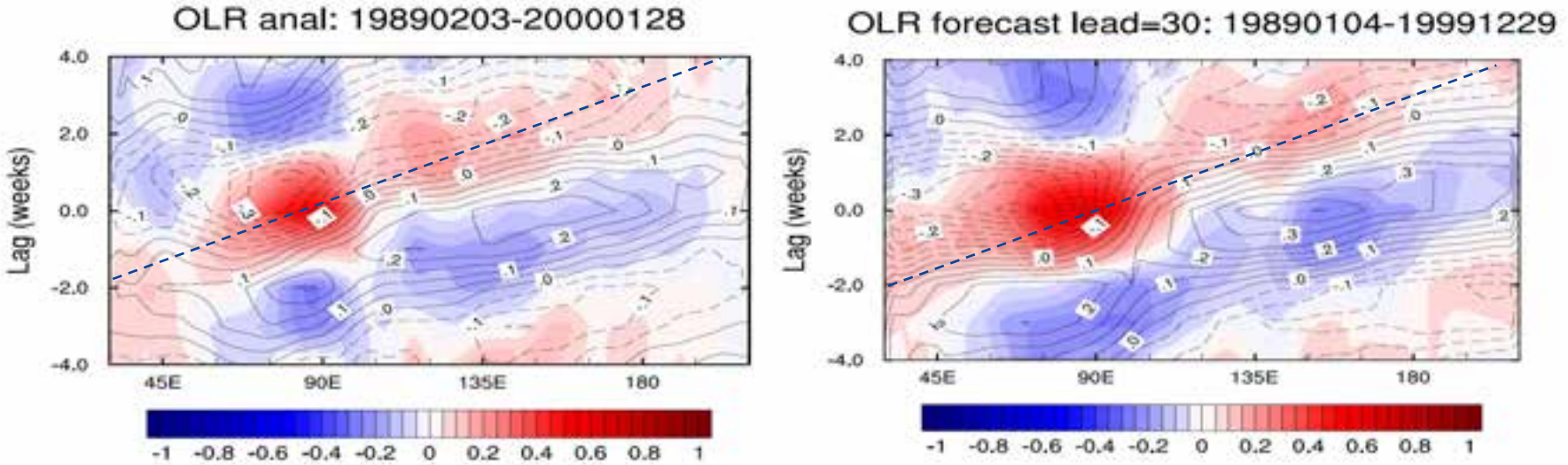


- GEFS SubX system was supporting NOAA CTB subseasonal experiment (SubX)
- Reference: *Pegion, K., and co-authors, 2019: The Subseasonal Experiment (SubX): A multi-model subseasonal prediction experiment, Bull. Amer. Meteor. Soc. 100 2043-2060*
- 20 years reforecast; 11 members, once per week, out to 35 days
- Real-time forecast started from July 2017
- MJO scores period:
  - May 2014-May 2016)

References:

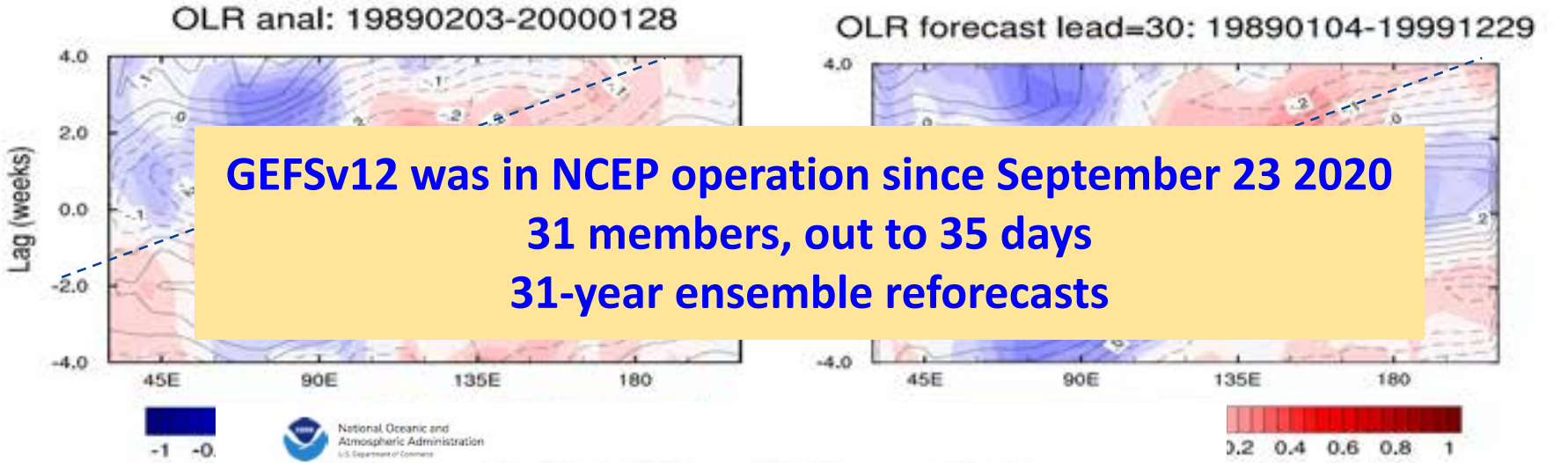
1. *Zhu, Y., X. Zhou, W. Li, D. Hou, C. Melhauser, E. Sinsky, M. Pena, B. Fu, H. Guan, W. Kolczynski, R. Wobus and V. Tallapragada, 2018: Towards the Improvement of Sub-Seasonal Prediction in the NCEP Global Ensemble Forecast System (GEFS). JGR, p6732-6745*
2. *Li, W., Y. Zhu, X. Zhou, D. Hou, E. Sinsky, C. Melhauser, M. Pena, H. Guan and R. Wobus, 2018: Evaluating the MJO Forecast Skill from Different Configurations of NCEP GEFS Extended Forecast. Climate Dynamics.*

Background (2) - Propagation of MJO in GEFSv12 (1989-1999)



1. *Spatial and time correlation (anomaly) in the **Central Indian Ocean** / time-lag of 11 years analysis (CFSR; left) and 30-day forecast (GEFSv12 ensemble mean; right). The correlation coefficient of OLR is in shaded and 850 zonal wind is in contours. The statistics indicate that there is a very good eastward propagation of signal (or MJO) from Indian Ocean. However, it is challenging to capture northward propagation of Intra-Seasonal Oscillations.* Courtesy of Dr. Wei Li
2. ***GEFSv12 and reforecast References:***
   a. *Zhou, X., Y. Zhu, D. Hou, B. Fu, W. Li, H. Guan, E. Sinsky, W. Kolczynski, X. Xue,Y. Luo, J. Peng, B. Yang, V. Tallapragada, P. Pegion, 2022: The Development of the NCEP Global Ensemble Forecast System Version 12 (Wea. Forecasting)*
   b. *Guan, H., Y. Zhu, E. Sinsky, B. Fu, W. Li, X. Zhou, X. Xue, D. Hou, J. Peng, M. Nageswararao, V. Tallapragada, T. M. Hamill, J. S. Whitaker, G. T. Bates, P. Pegion, S. Frederick, M. Rosencrans, A. Kumar, 2022: (GEFSv12 reforecast dataset for supporting subseasonal and hydrometeorological applications (MWR)*

Background (2) - Propagation of MJO in GEFSv12 (1989-1999)



GEFSv12 was in NCEP operation since September 23 2020
31 members, out to 35 days
31-year ensemble reforecasts

1. *Spatial and*
   (GEFSv12 e
   indicate tha
   northward
2. ***GEFSv12 an***
   a. *The*
      Sins
   b. *GEF*

FSR; left) and 30-day forecast
s in contours. The statistics
ver, it is challenging to capture

D. Hou, B. Fu, W. Li, H. Guan, E.
Forecast in process)
- H. Guan, Y. Zhu, E. Sinsky, B.
Fu, W. Li, X. Wei, A. Kade, B. Hou, J. Peng, M. Pagowski, V. Tallapragada, Y. Ma, ..., J. S. Whitaker, G. T. Bates, P. Pegion, S. Frederick, M. Rosencrans, A. Kumar (MWR, accepted)

National Oceanic and Atmospheric Administration

NOAA upgrades Global Ensemble Forecast System

Foundation for weather forecast models gets a big boost from the FV3

Focus areas: Weather Topics: modeling

September 23, 2020 — NOAA's Global Ensemble Forecast System (GEFS) is undergoing significant upgrades to expand its capabilities and improve weather forecasting. Meteorologists at the National Weather Service use GEFS to produce medium-range weather forecasts and to issue watches and warnings during high-impact weather events, including hurricanes, blizzards, and extreme heat and cold.

The configurations of UFS coupled GEFS

Experiments period: October 2017 - September 2019

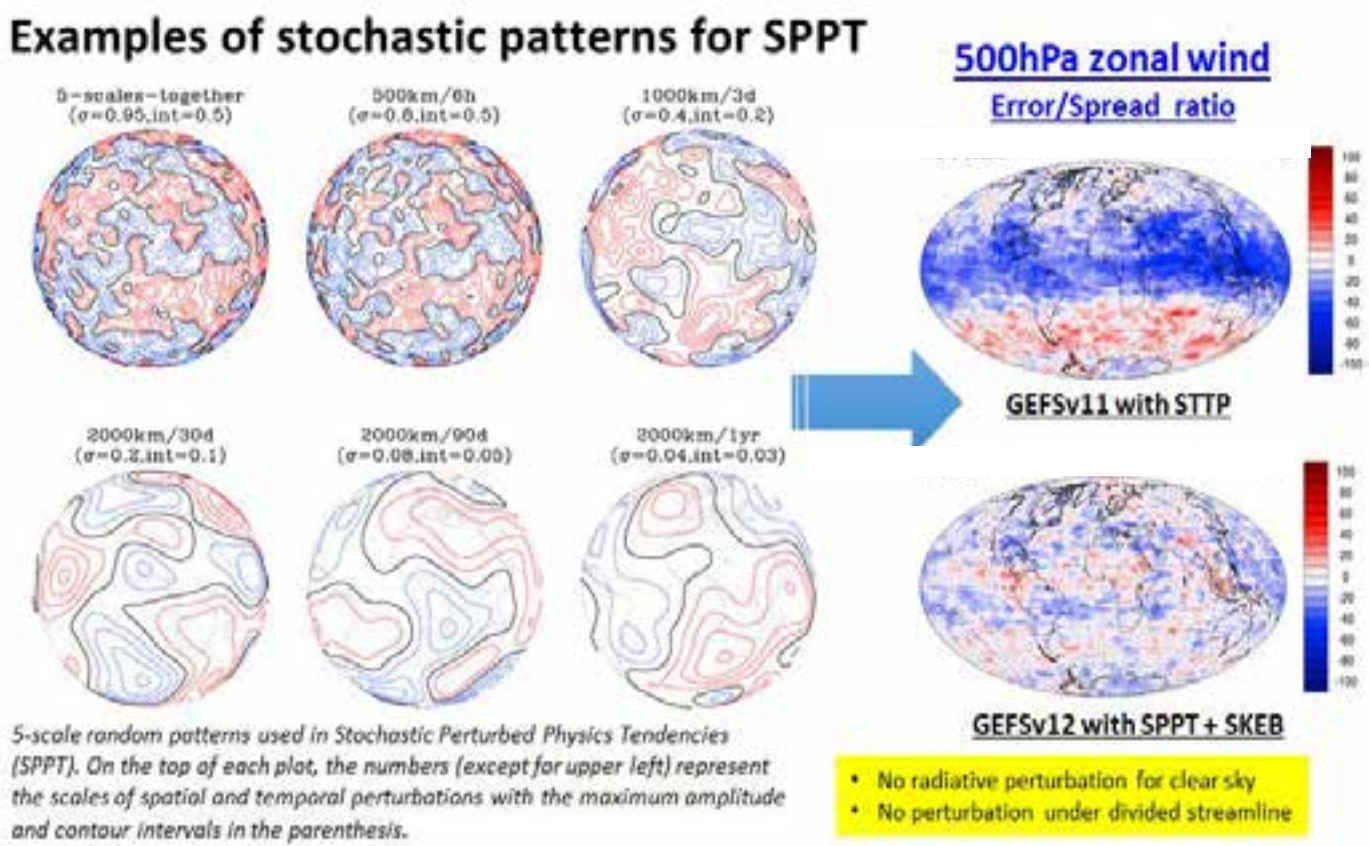
- GEFSv12 Reforecast - benchmark or reference
  - C384L64 **uncoupled** - GFSv15/GEFSv12 configurations
  - Initial analysis and perturbations
    - GEFSv12 reanalysis (and 3D-IAU replay)
  - Stochastics
    - SKEB (0.6); SPPT-5 scales (0.8,0.4,0.2,0.08,0.04)

Ensemble Prototype 1 (EP1)

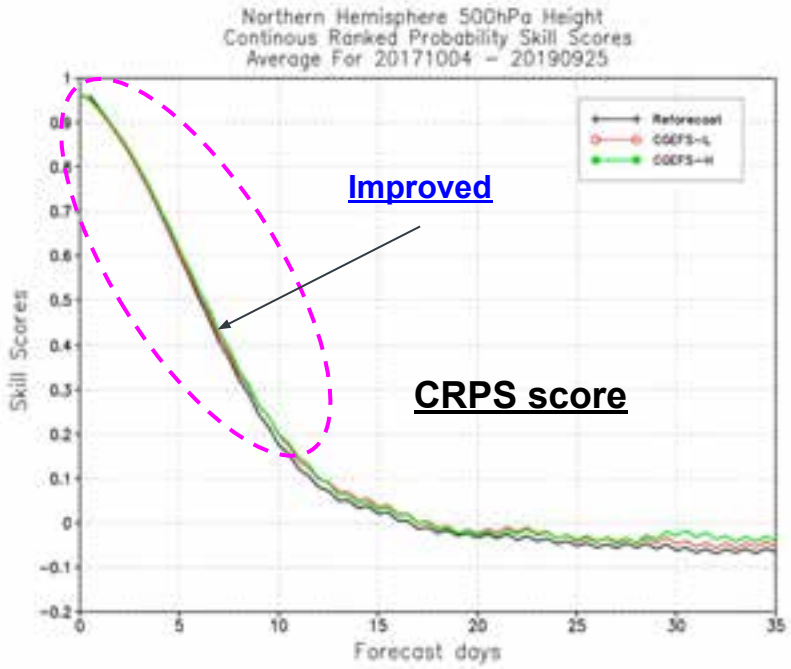
- CGEFS-L - coupled GEFS at low (C192) resolution
  - C192L64 (P5 coupling) + MOM6(0.5)+ CICE6 (0.5) + WW3 ---> early version of P5
    - ATM IC: Initial analysis and perturbations for atmosphere - GFSv15 retrospective cycling data
    - OCN IC: CFSv2 analysis interpolated to 0.5 (SST and salinity only)
    - ICE IC: CPC's seaice analysis
    - WAV IC: no wave IC
  - Stochastics
    - **SKEB (0.7) + SPPT (30%off: 5 scale (0.56, 0.28, 0.14, 0.056, 0.028))**
    - Stochastic function are the same as reforecast except for the amplitude
- CGEFS-H - coupled GEFS at high (C384) resolution
  - C384L64 (P5 coupling) + MOM6(0.25) + CICE6(0.25) + WW3 ---> final version of P5
  - ICs: Same as EXP-L (C192) except adding wave IC
- Challenges -
  - Initial conditions of ocean - no cycling, no dynamical field (current); No initial perturbations for ocean and land

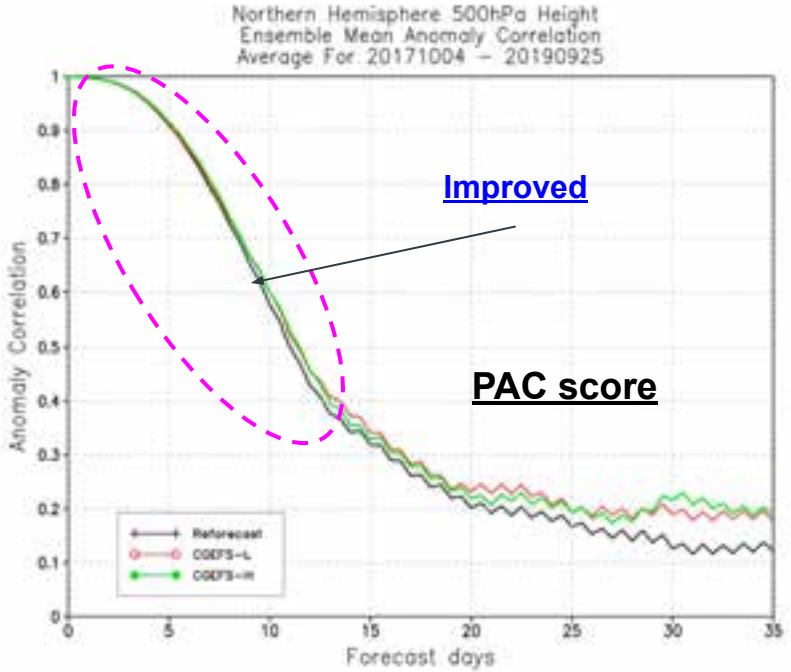
Model uncertainty of GEFSv12: SPPT and SKEB

- **SKEB**: Estimate energy lost each time step and inject this energy in the resolved scales. a.k.a stochastic energy backscatter (SKEB; Berner et al. 2009)
- **SPPT**: perturb the results from the physical parameterizations, and boundary layer humidity (Palmer et al. 2009), and inspired by Tompkins and Berner 2008, we call it SPPT
- ***Replace STTP for GEFSv12 with SPPT and modified SKEB (amplitude reduced to 0.5 from 1.0)***



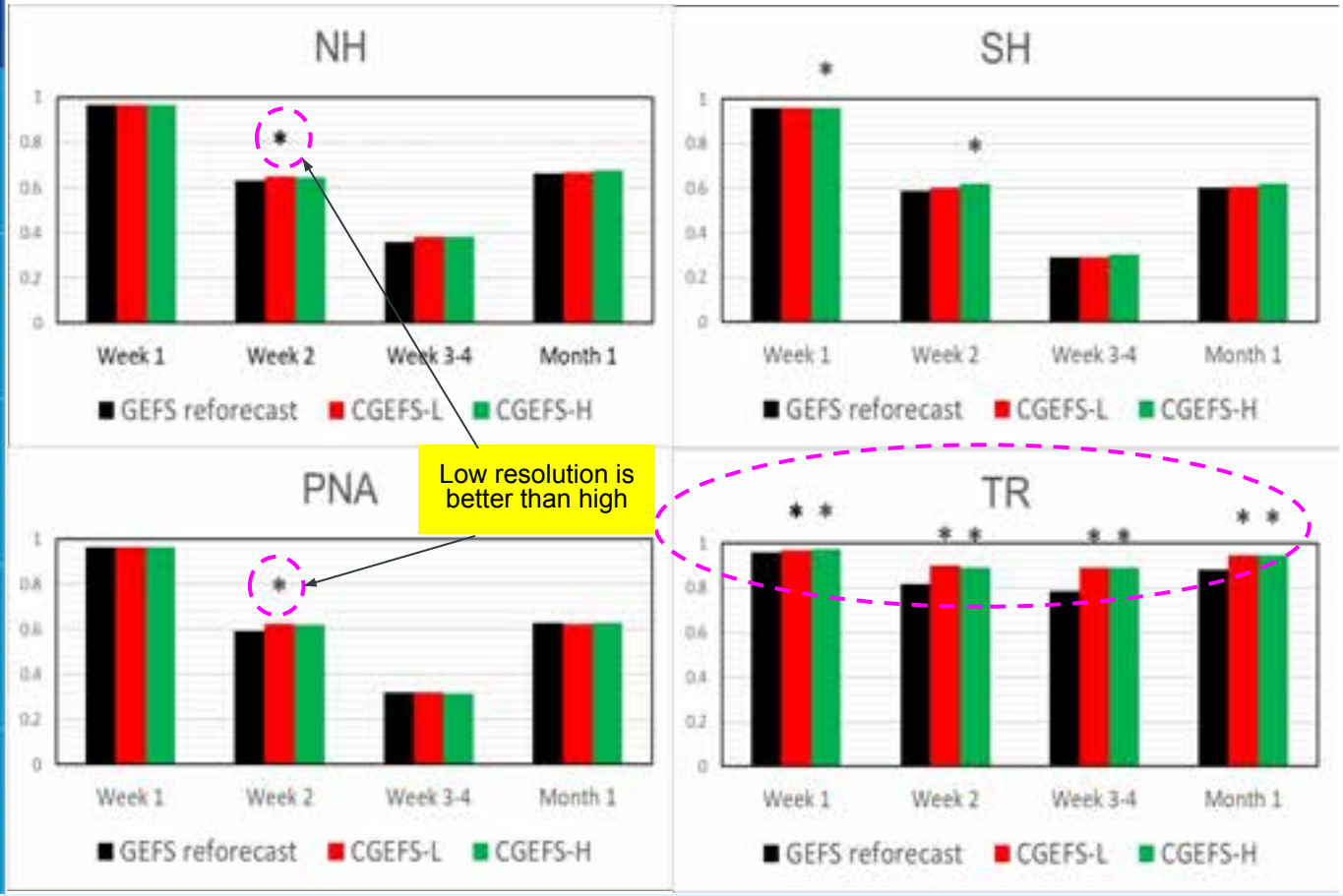
The skills of North Hemisphere 500hPa geopotential height





- *Day-to-day verification of Northern Hemisphere 500hPa geopotential height for the GEFSv12 reforecast (black), CGEFS-L (red) and CGEFS-H (green). a) CRPS scores (Left); b) Pattern anomaly correlations (right).*
  - *CGEFS-H and GEFS-L are better than GEFS reforecast for all lead-time, "significant" better for week-1 and week-2 (see next slide for all different regions and average of lead-time.*

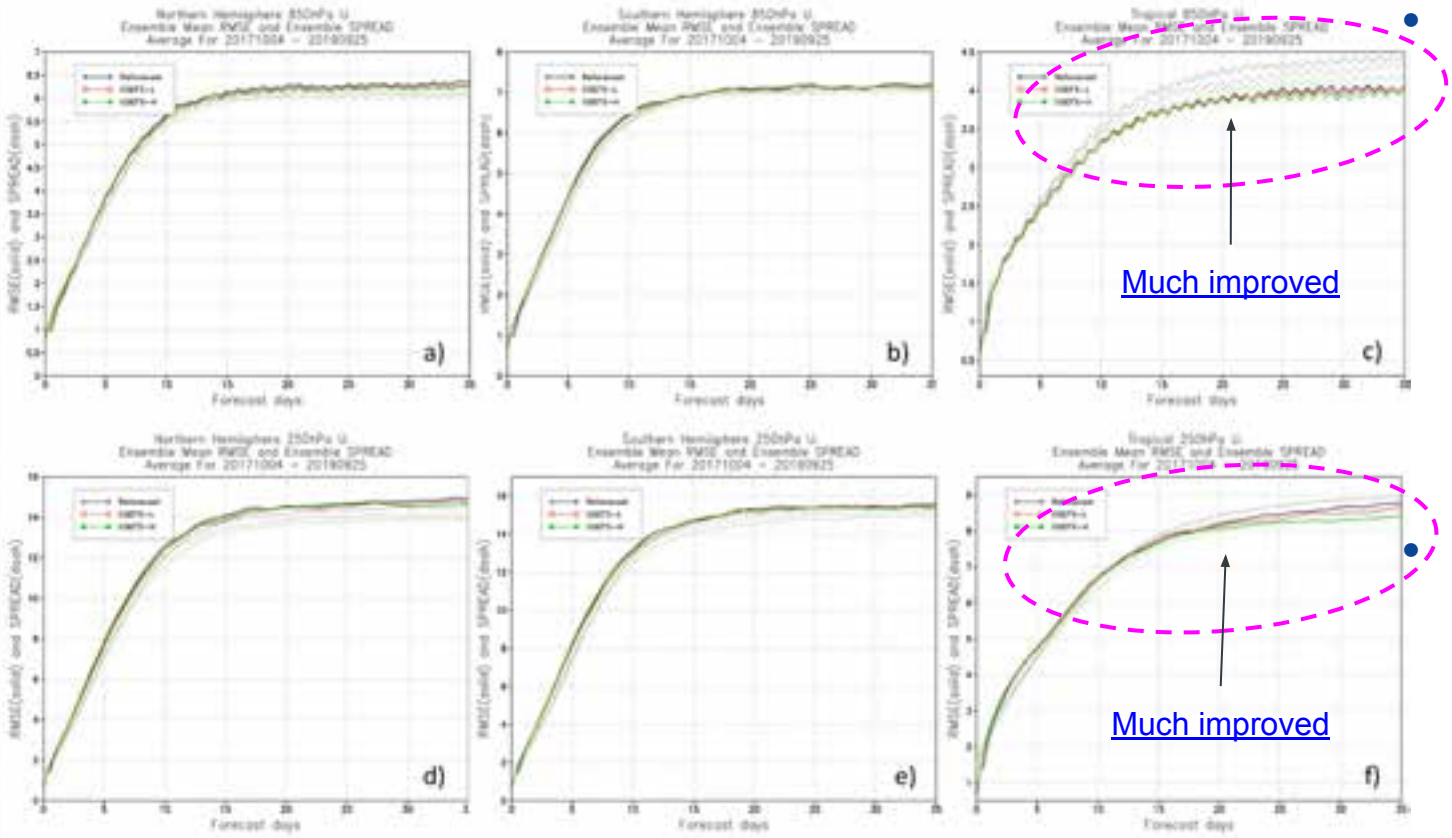
The skills of 500hPa geopotential height (week-1; week-2; weeks 3&4 and month)



The case-averaged 500hPa geopotential height AC for the GEFSv12 reforecast (black), CGEFS-L (red) and CGEFS-H (green). The asterisks (*) signify that the difference between CGEFS and the reforecast average AC is statistically significant at 95%. The AC scores are for week-1, week-2, weeks 3&4, monthly, NH (top left), SH (top right), TR (bottom right) and PNA (bottom left) domains.

Tropical improvement is "significantly" for all lead-time.

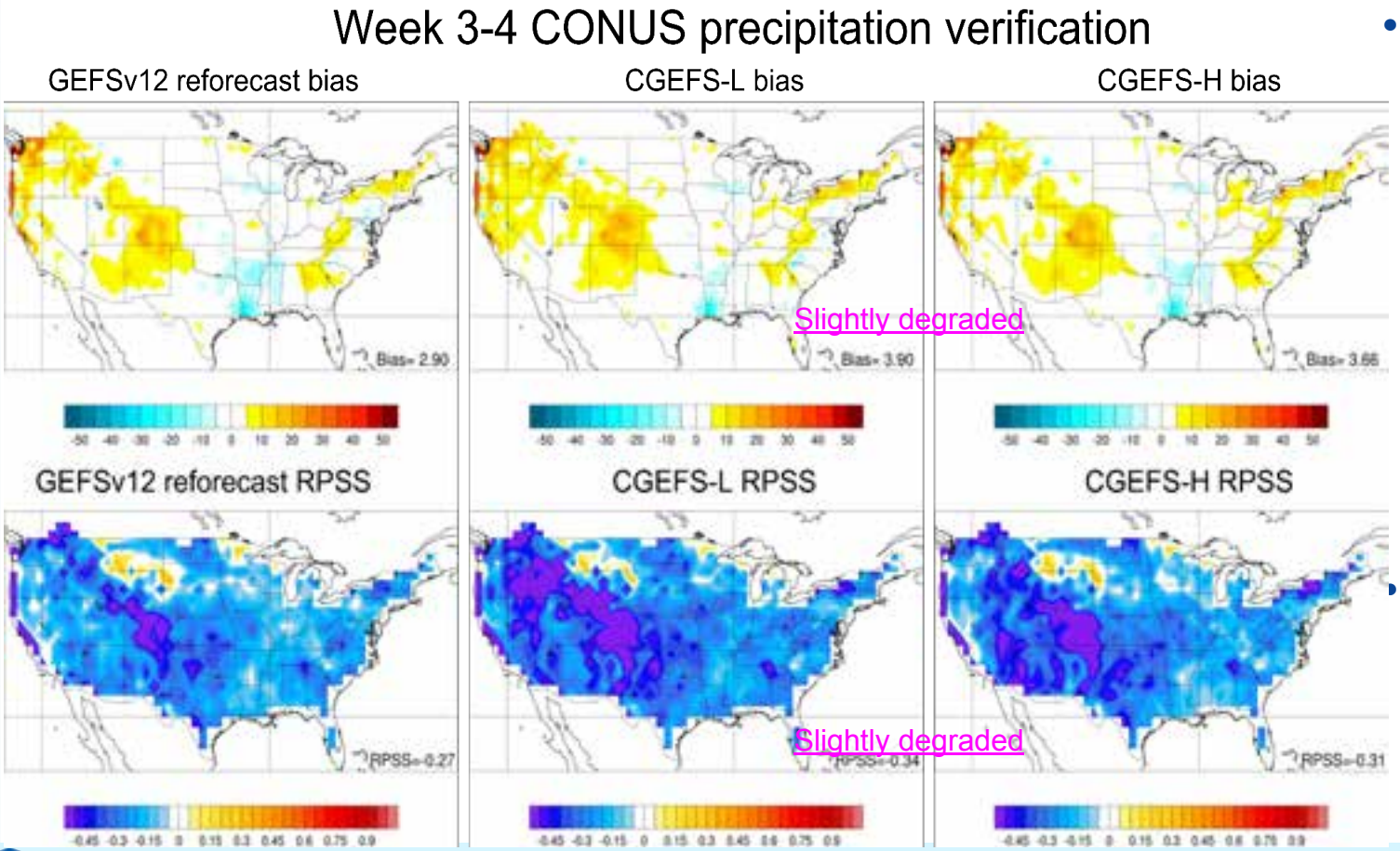
RMS error / spread of 850hPa and 250hPa zonal wind



The RMS errors of the ensemble mean (solid lines) and ensemble spread (dashed lines) of 850hPa zonal wind for NH (a), SH (b) and TR (c). The similar plots are for 250hPa zonal wind of NH (d), SH (e) and TR (f). There are three statistics (reforecast - black; CGEFS-L - red; CGEFS-H - green) to comparison

A biggest improvement is tropical region - reduce ensemble spread and RMS error, in particular for 850hPa tropical zonal winds.

The Bias and RPSS of the CONUS precipitation

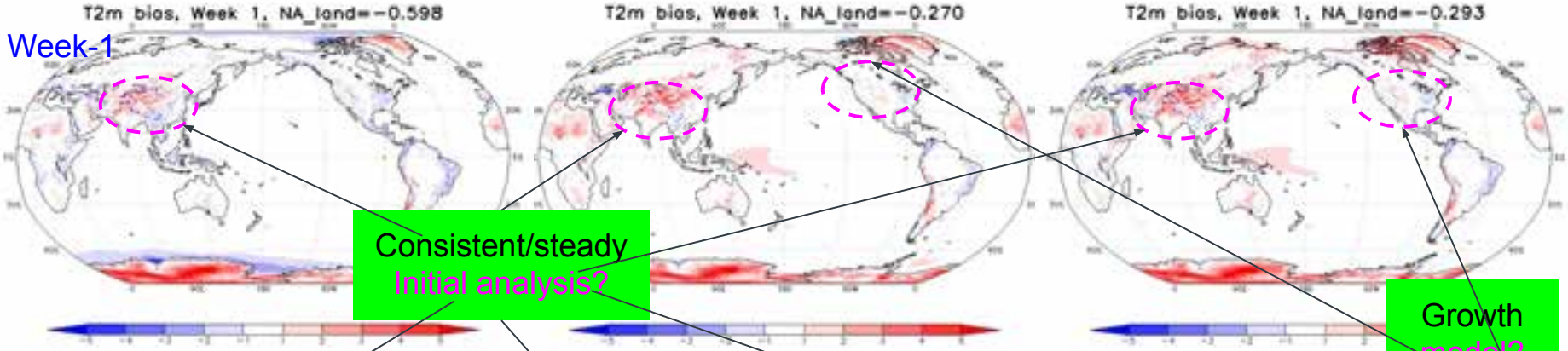


- *The bias (unit: mm; top row) and RPSS scores (dimension less; bottom row) of the CONUS precipitation for the GEFSv12 reforecast (left column), CGEFS-L (center column) and CGEFS-H (right column) of weeks 3&4 forecast. The numbers on the bottom right are the CONUS domain average scores.*
- *Both CGEFS-L and CGEFS-H are slightly increased bias and degraded RPSS scores, but it is not significant.*

Bias estimation of 2m temperature against ERA5

GEFSv12 refcst | **CGEFS-L** | **CGEFS-H**

Week-1

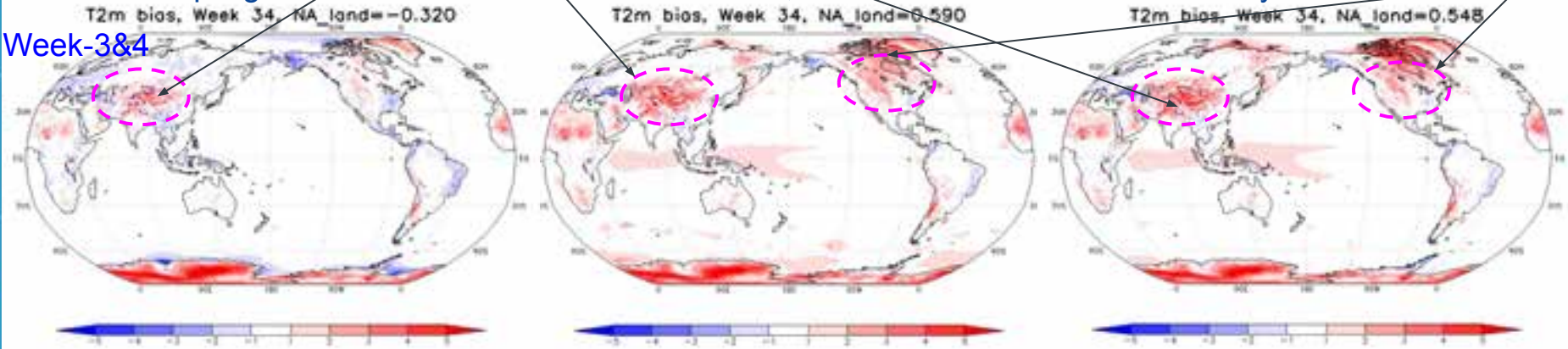


Consistent/steady
Initial analysis?

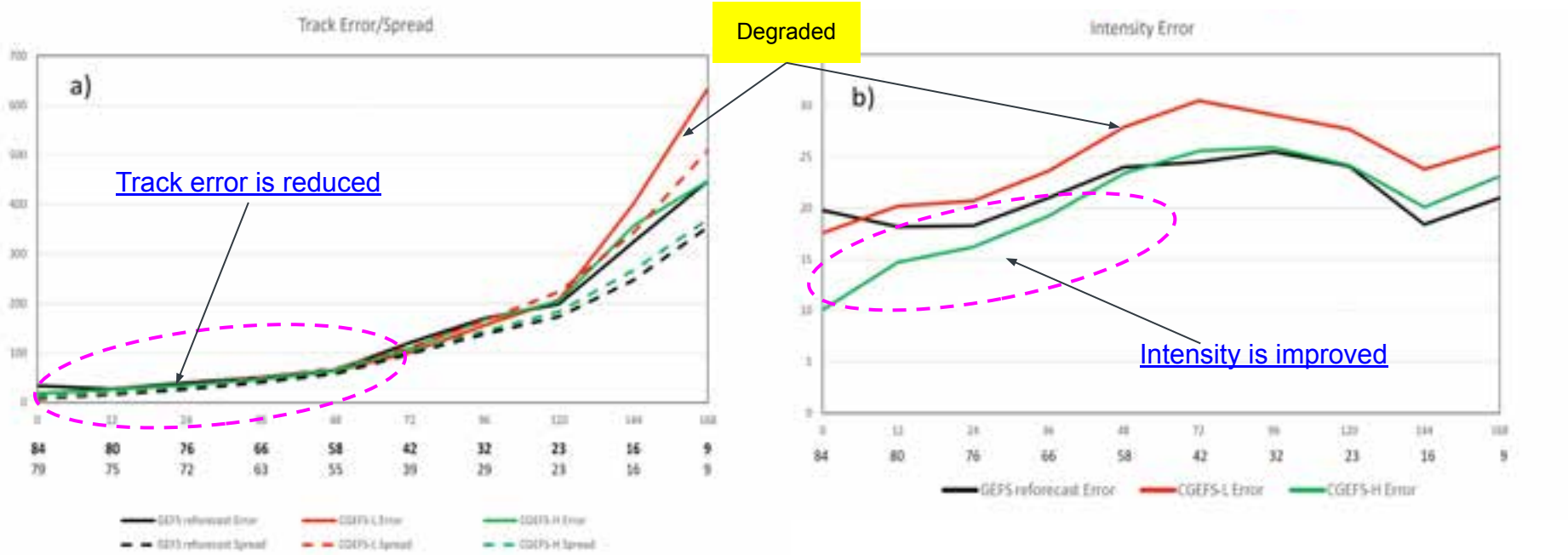
Growth
model?

Coupling forecasts are warmer than reforecast, CGEFS-H and CGEFS-L are very similar

Week-3&4

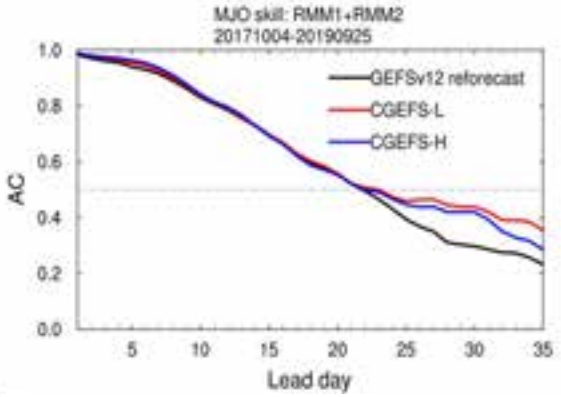


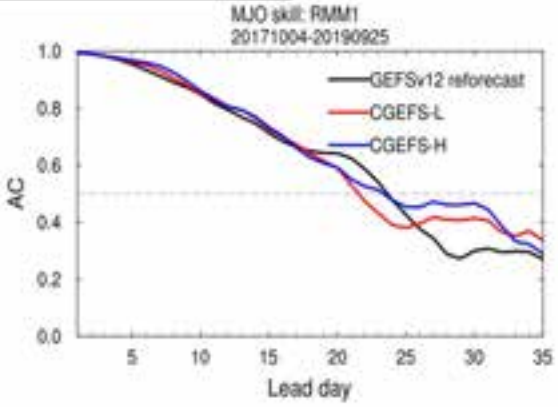
Tropical Storm Predictions (RMSE/Spread; Intensity)

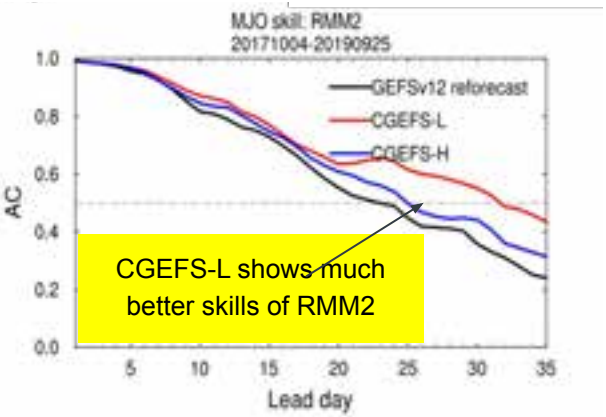


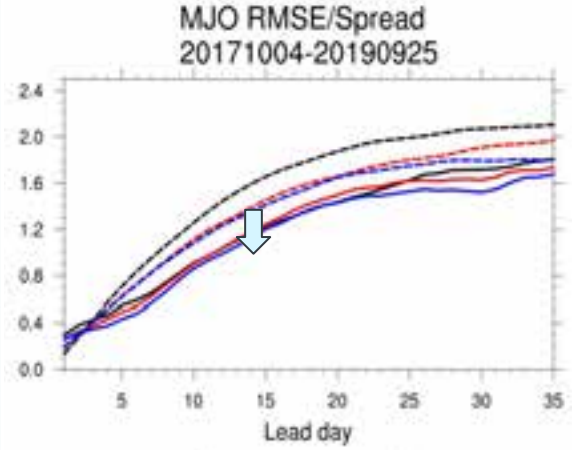
a). *Tropical storm tracks (solid line) and spread (dash line) for all three basins of three experiments (GEFSv12 reforecast (Black), CGEFS-L (red) and CGEFS-H (green)). The first, second and third rows of x-axis labels are respectively forecast lead hours, numbers of cases used to calculate track error (bolded; 2017-2019) and number of cases used to calculate spread (2018-2019). The y-axis is track error (unit: Nautical Mile (NM)).* ***b)***, *Tropical storm intensity error comparison of GEFS reforecast (black), CGEFS-L (red) and CGEFS-H (green) for the average of all domains. The first and second rows of x-axis are respectively forecast lead hours and numbers of cases used to calculate the intensity error. The y-axis is intensity error (unit: knot=Nautical Mile per Hour).*

MJO skills (RMM1+RMM2; RMM1; RMM2; RMS error/Spread)



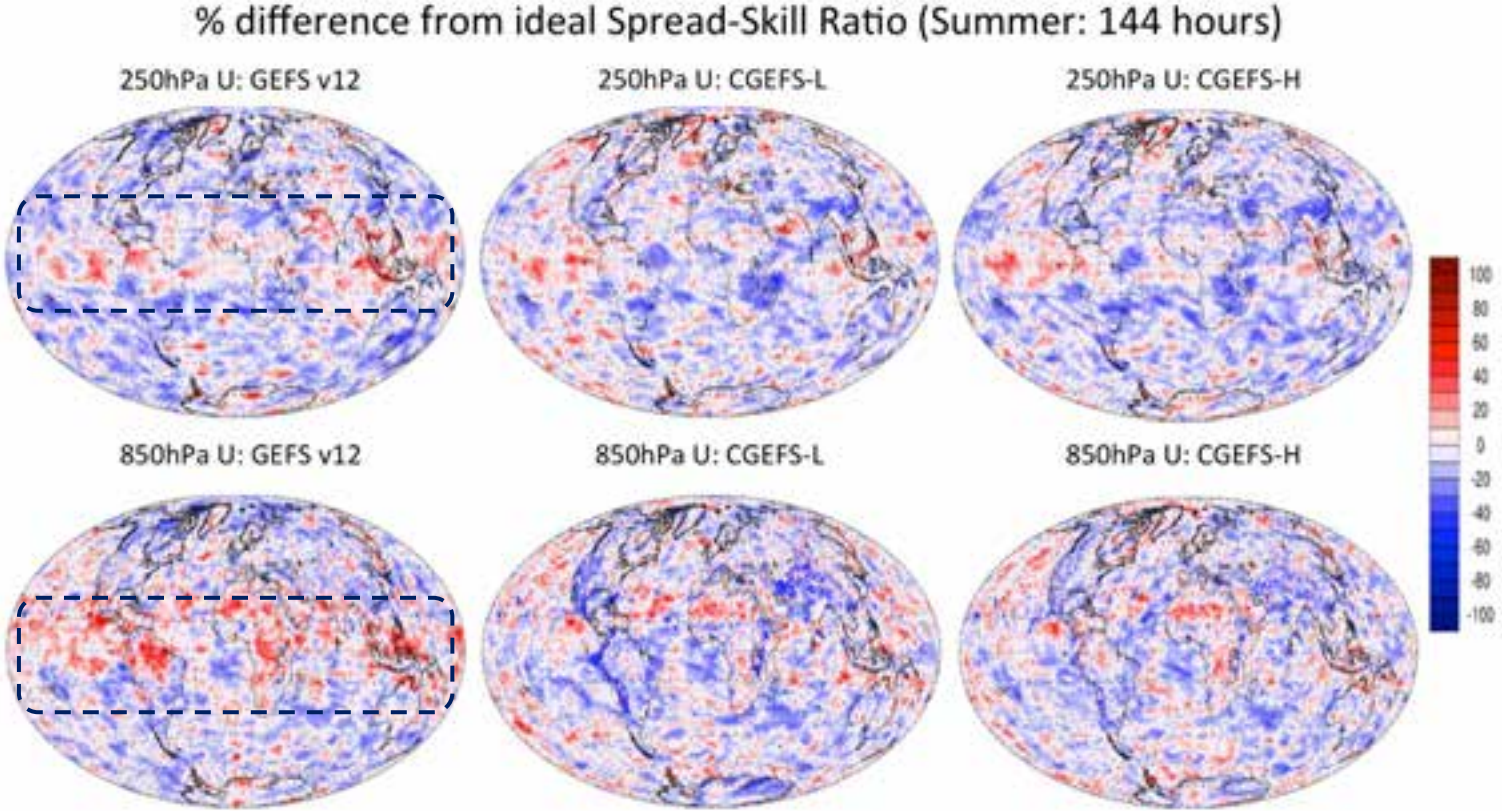






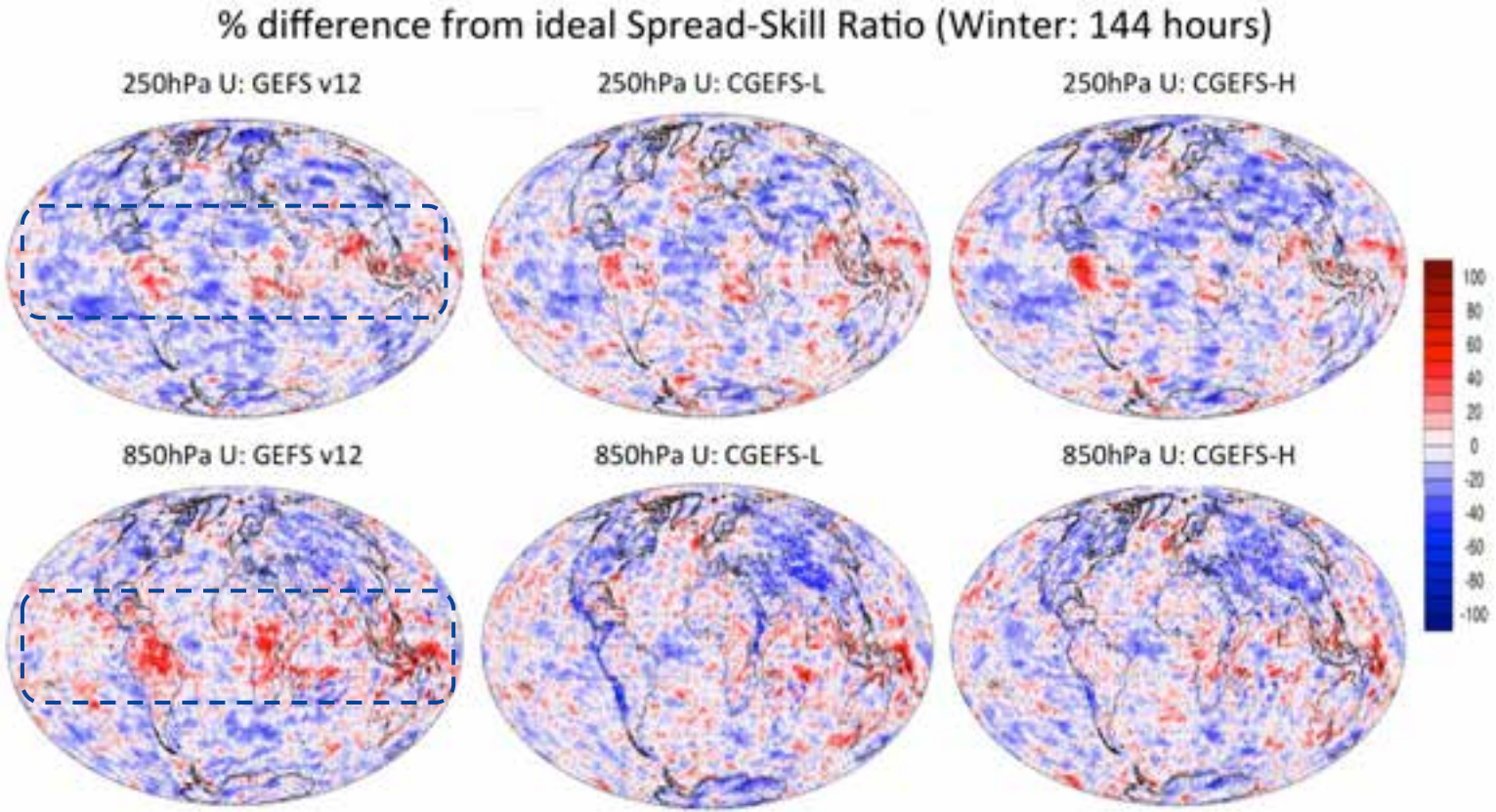
- *Top - MJO skills for the GEFSv12 reforecast (black), CGEFS-L (red) and CGEFS-H (blue). The combined MJO index (RMM1+RMM2) is on the top left, the RMM1 index is on the top middle, and the RMM2 index is on the top right.*
  - *Both CGEFS-H and CGEFS-L have better MJO skills than GEFSv12 reforecast, especially for RMM2.*
- *Bottom left - The RMS errors of ensemble mean (solid lines) and spread (dashed lines) of MJO for GEFSv12 reforecast (black), CGEFS-L (red) and CGEFS-H (blue). The unit of RMS error is based on the RMM index of MJO events (Unit: nondimension)*
  - *CGEFS-H has smaller RMS error for all lead-time*
  - *CGEFS-H and CGEFS-L reduce the MJO spread (a lot), but still over-dispersion*

Spread-Error Ratio (April - September, 52 cases, 144-hr)



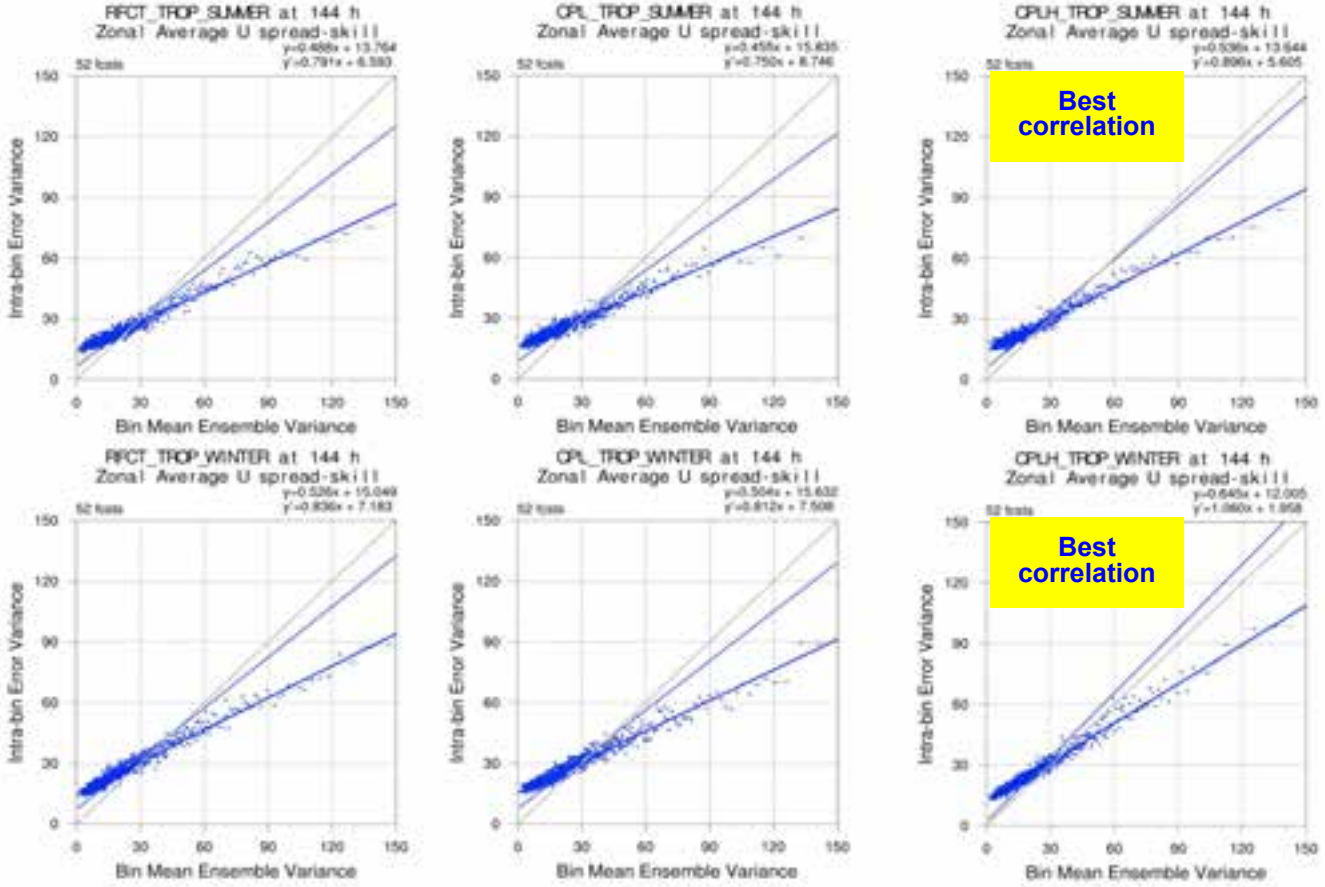
The boreal summer six months (April - September) error-spread ratio map of 850hPa zonal winds (upper row) and 250hPa zonal winds (lower row) at 144 hours (6 days) forecast for the GEFSv12 reforecast (left column), CGEFS-L (middle column) and CGEFS-H (right column).

Spread-Error Ratio (October - March, 52 cases, 144-hr)



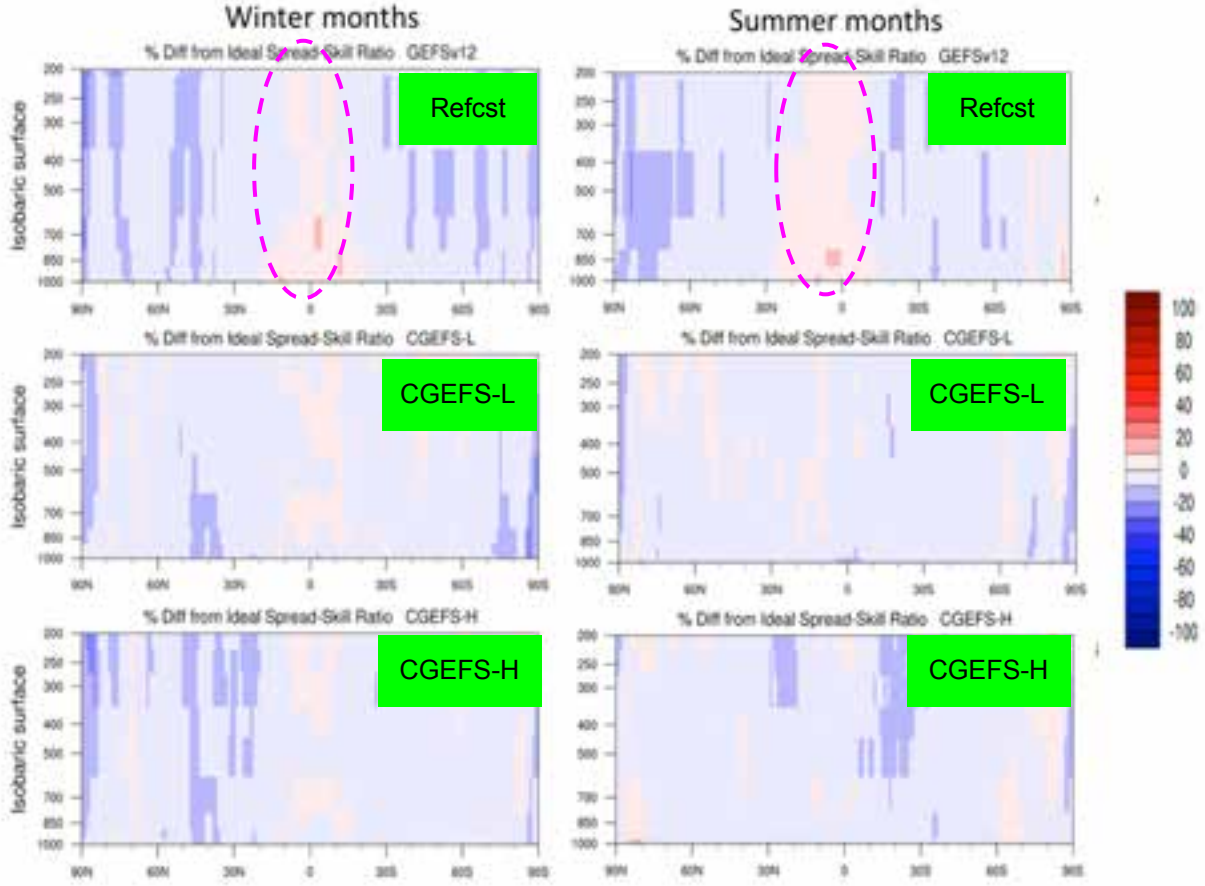
The boreal winter six months (October - March) error-spread ratio map of 850hPa zonal winds (upper row) and 250hPa zonal winds (lower row) at 144 hours (6 days) forecast for the GEFSv12 reforecast (left column), CGEFS-L (middle column) and CGEFS-H (right column).

The Scatter Plot - Linear Variance Method for Diagnostics



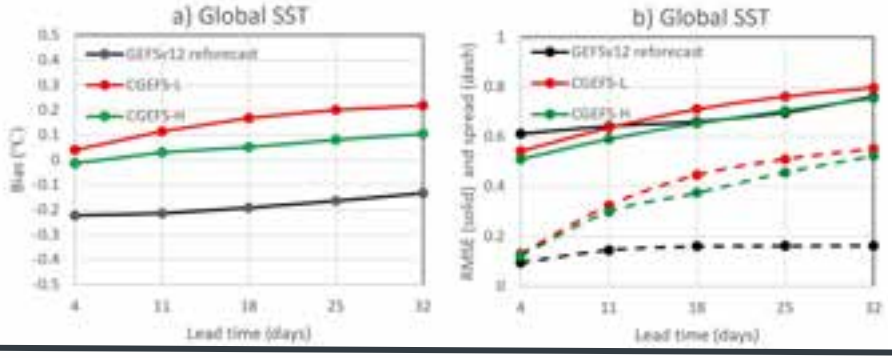
- *Linear Variance Method for Diagnostics (Kolczynski, 2011)*
- *The scatter plot of ensemble bin variance and mean error variance for boreal summer (6 months; upper row) and boreal winter (6 months; lower row) of tropical zonal winds on the 250hPa and forecast lead-time at 144 hours (6 days). The GEFSv12 reforecast is on the left column, CGEFS-L is on the middle column and CGEFS-H is on the right column.*
- *Solid blue line is for regression line; dot blue line is for assimilated regression line if there are sufficient ensemble members.*

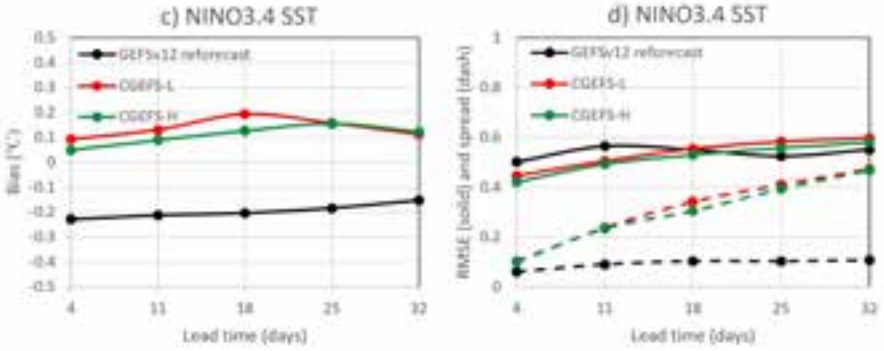
Vertical cross section of the RMS error - spread ratio for zonal winds



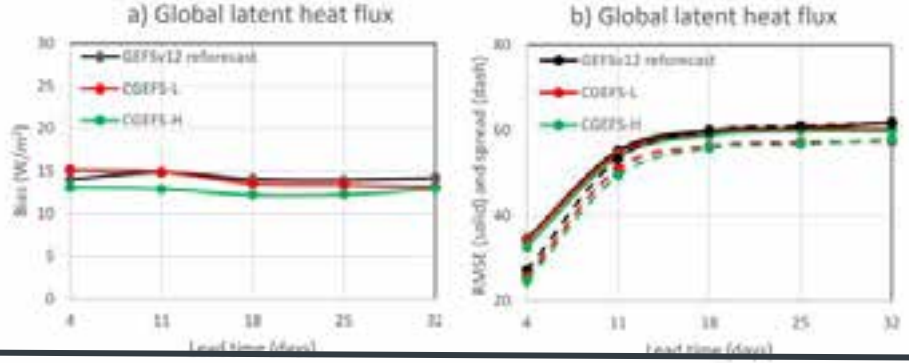
- *The vertical cross section of the RMS error and ensemble spread ratio for boreal winter six months (left column) and boreal summer six months (right column) of zonal wind from surface (1000hPa) to 200hPa in vertical, and for the GEFSv12 reforecast (top), GGEFS-L (middle) and CGEFS-H (bottom)*
- *Forecast lead-time - 144 hours (6 days)*
- *Tropical over-dispersion for GEFSv12 reforecast is our major concern*
- *CGEFS-L reduces tropical over-dispersion*
- *CGEFS-H reduces tropical over-dispersion*

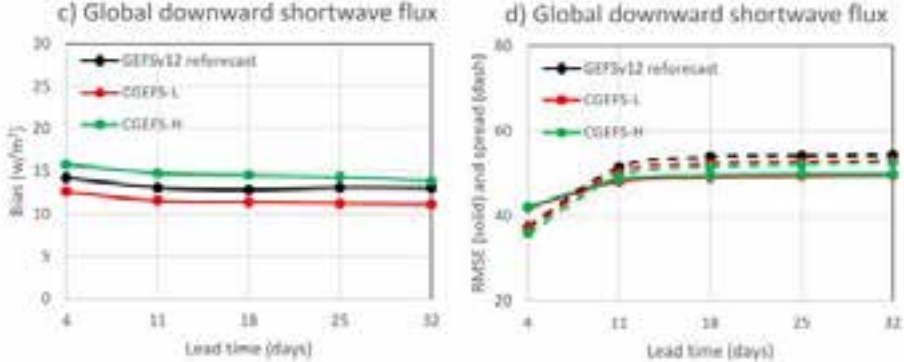
The biases and RMSE/SPRDs of SST, DSWRF, LHTFL for different domains





The SST bias, RMS error and spread of the global (50N-50S and ocean only) and Nino 3.4 domains for the GEFSv12 reforecast (black), CGEFS-L (red) and CGEFS-H (green or blue)





The DSWRF and LHTFL's bias, RMS error and spread of the global (50N-50S and ocean only; top) and Nino 3.4 domains (bottom) for the GEFSv12 reforecast (black), CGEFS-L (red) and CGEFS-H (green)

Summary of EP1 (CGEFS-H and CGEFS-L)

- Followings are preliminary conclusions based on the statistics and diagnostics:
  - Overall, coupled GEFS (low/high resolutions) extended prediction skill from GEFSv12 reforecast (or current operational GEFS).
  - CGEFS-H is better than CGEFS-L for weather prediction.
  - CGEFS-H has a closed skills of CGEFS-L for extended range (week-2; weeks 3&4) forecasts, but it indicates the higher model resolution does not take advantage for extended range forecasting.
  - The forecast uncertainties of coupled GEFS (low/high resolutions) are adjusted very well from updated SPPT and SKEB to enhance probabilistic forecast skill.
  - From a limited case, the TS track of CGEFS-H is slightly better for short lead-time, and intensity is much better than CGEFS-L and GEFSv12 reforecast
  - The MJO skill of coupled systems is better than reforecast for all lead-time, CGEFS-H reduces ensemble spread, but still over-dispersion
  - The CONUS precipitation of coupling is slightly degraded than reforecast but it is insignificant.
  - Coupling system introduced a slightly warm bias of surface temperature for CONUS which could be related to the land model and other physical processes.
  - The tropical forecast spread may still be larger than error which results the MJO uncertainty is over-dispersion (or under-confident).
- Challenges -
  - 11 members is not sufficient to represent full uncertainties
- Manuscript for publication -
  - Quantify the Coupled GEFS Forecast Uncertainty for the Weather and Subseasonal Prediction
    - In the review process

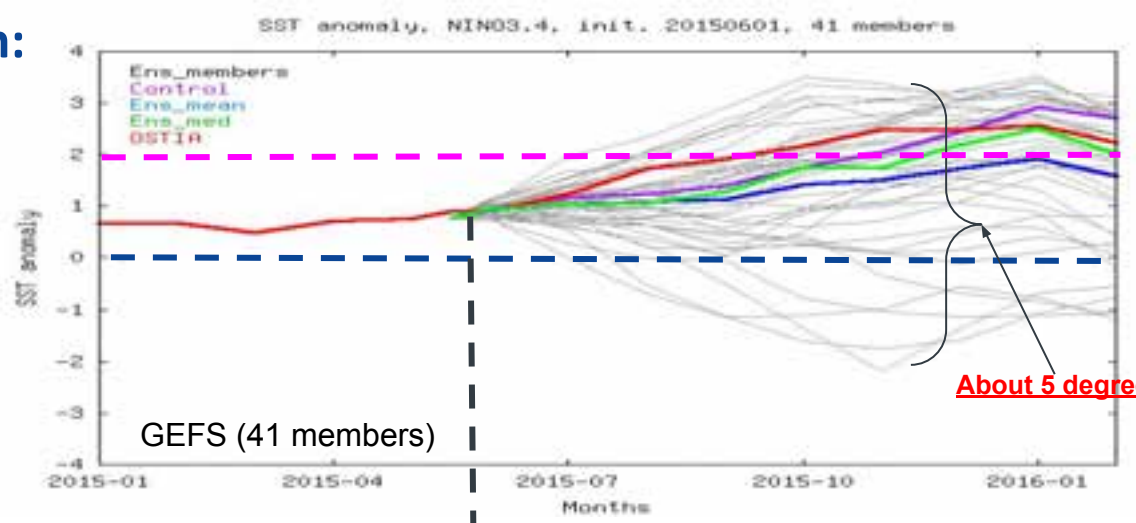
Seasonal Prediction - Configurations

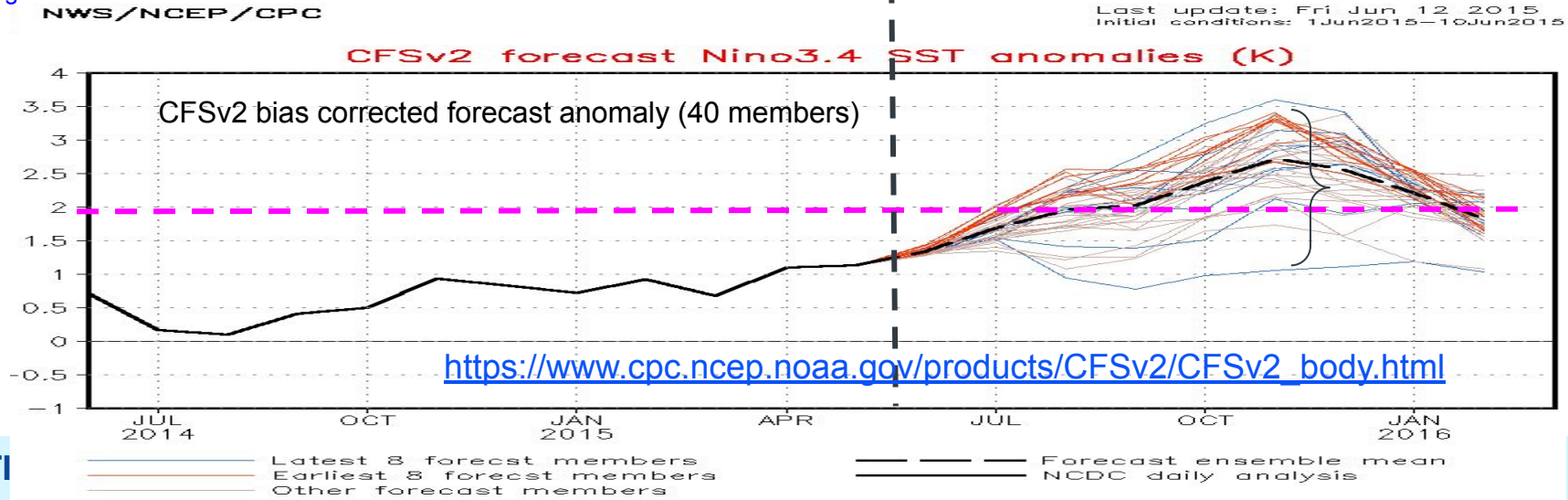
- Models
  - Atmosphere - **C192L64** (P5, GFSv15, GEFSv12)
    - Retain NSST to assimilate SST diurnal variability, expect to reduce cold SST bias
  - Ocean - **MOM6 (0.5d)**
  - Seaice - **CICE6 (0.5d)**
  - Wave - WW3 does not include in this experiment
- Stochastics
  - Atmosphere: 5-scale SPPT (30% off); SKEB (0.7)
    - They are the same as CGEFS-H and CGEFS-L.
  - Ocean:
    - ePBL perturbations (perturbed TKE generation and dissipation)
    - Perturbed SST, salinity and layer of thickness.
    - All 5 scales, [0.8,0.4,0.2,0.08,0.04]
- Initial conditions
  - June 1st 2012 (neutral case); June 1st 2015 (El Nino case); June 1st 2017 (La Nina case)
  - Atmosphere - GFSv15 retrospective analysis (and perturbations)
  - Ocean - GODAS interpolated to 0.5d including ocean current (full fields)
  - Ice - CPC's analysis
- Members and forecast length
  - 41 members; out to 9 months

Nino 3.4 indexes comparison: "El Nino Case"

Discussion:

- Right plum is coupled GEFS seasonal run, no bias correction
- Bottom plum is CFSv2 seasonal run, with bias correction
- Initial state of SST: ~1 degree for both systems
- CFSv2 demonstrated a good forecast in terms of trend and peak
- Coupled GEFS is very well either even without bias correction, but spread is too large.

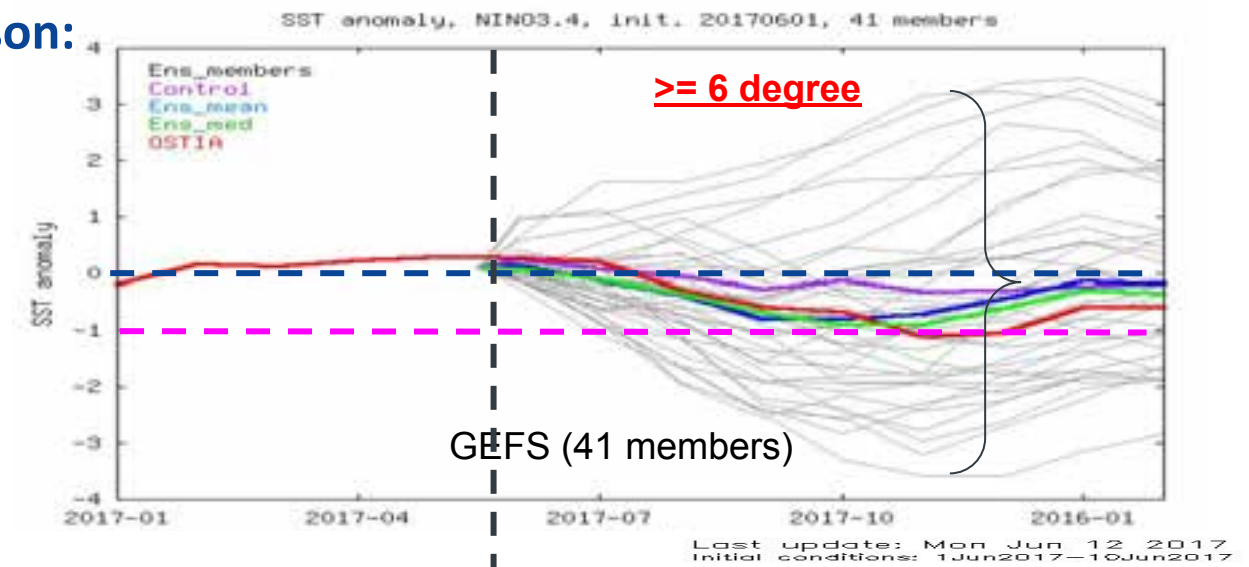




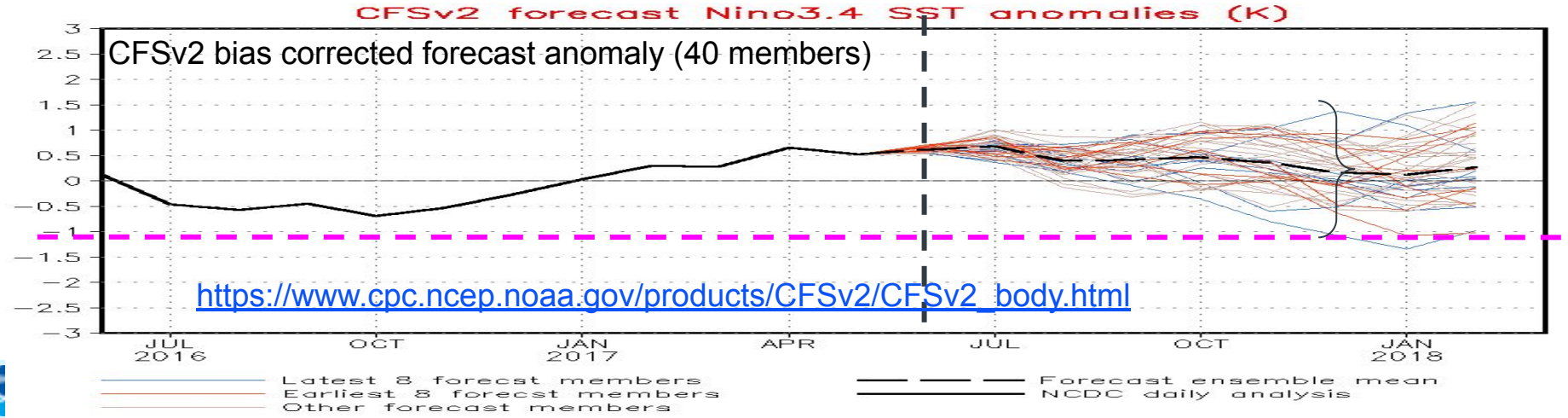
Nino 3.4 indexes comparison: "La Nina Case"

Discussion:

- Right plum is coupled GEFS seasonal run, no bias correction
- Bottom plum is CFSv2 seasonal run, with bias correction
- Initial state of SST: 0.2 (GEFS); 0.6 (CFSv2)
- CFSv2's prediction tends to very weak, no indication of La Nina event
- Coupled GEFS is very well capture the trend in the summer-fall time, but return to normal SST earlier, which could be a winter warm bias.



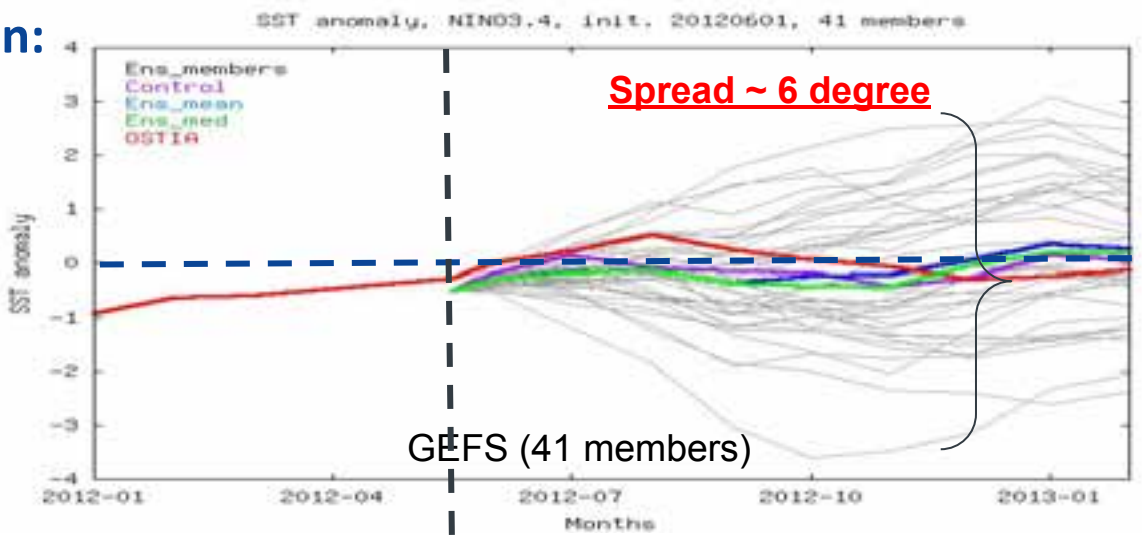
NWS/NCEP/CPC

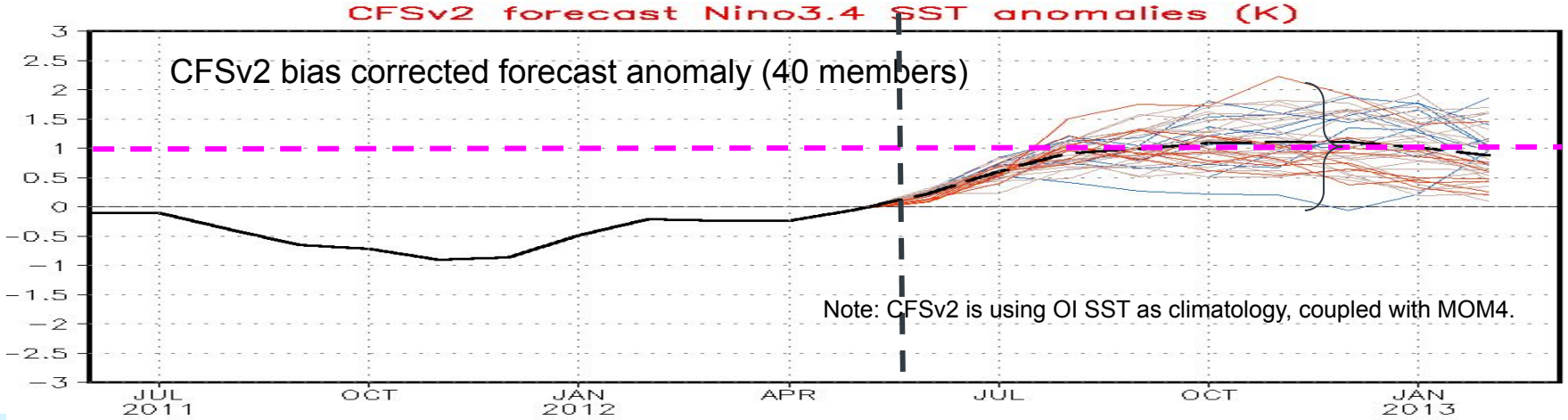


Nino 3.4 indexes comparison: "Neutral Case"

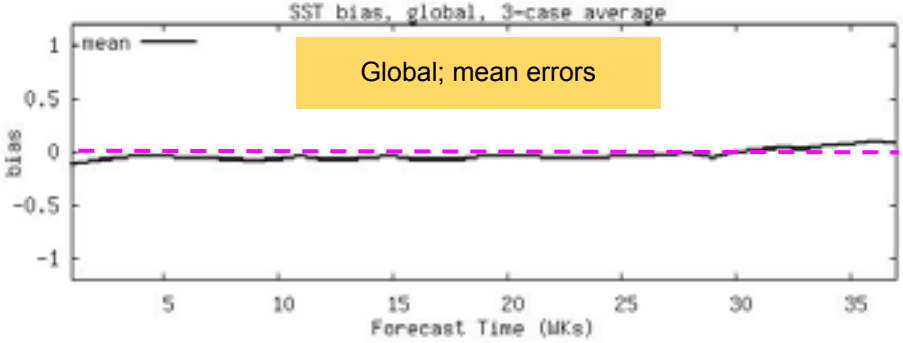
Discussion:

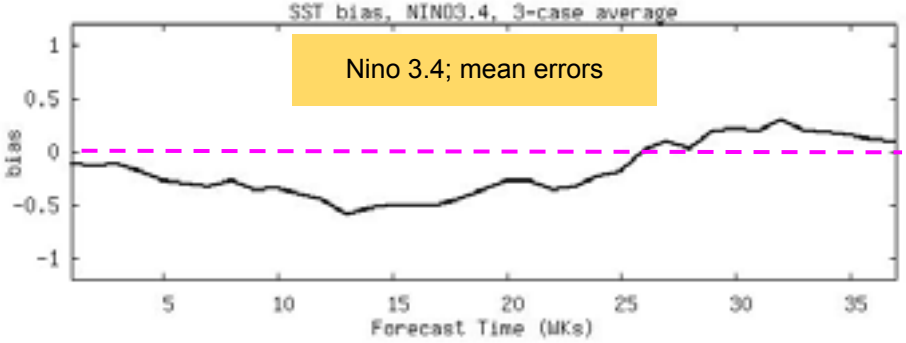
- Right plum is coupled GEFS seasonal run, no bias correction
- Bottom plum is CFSv2 seasonal run, with bias correction
- Initial state of SST: -0.3 (GEFS); 0.2 (CFSv2)
- CFSv2's prediction tends toward positive anomaly, reached 1d around fall (El Nino?).
- Coupled GEFS is a good prediction relatively for this neutral case, but the spread is still large.

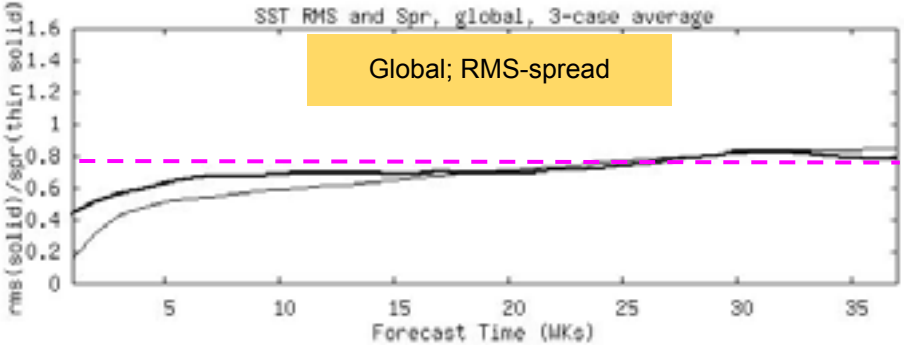


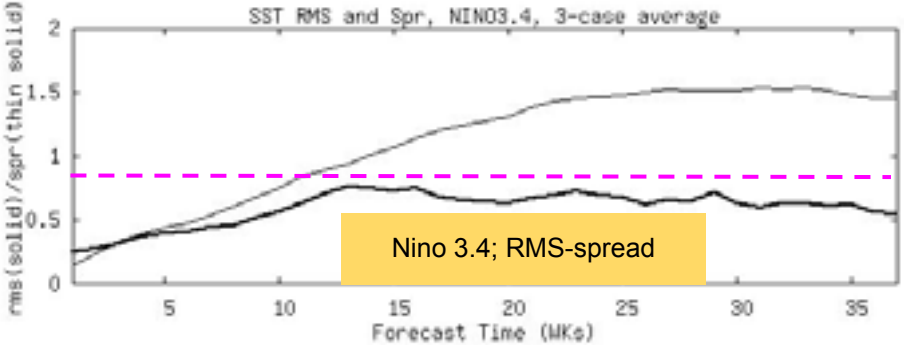


SST mean errors, RMS errors and spread (against OSTIA)



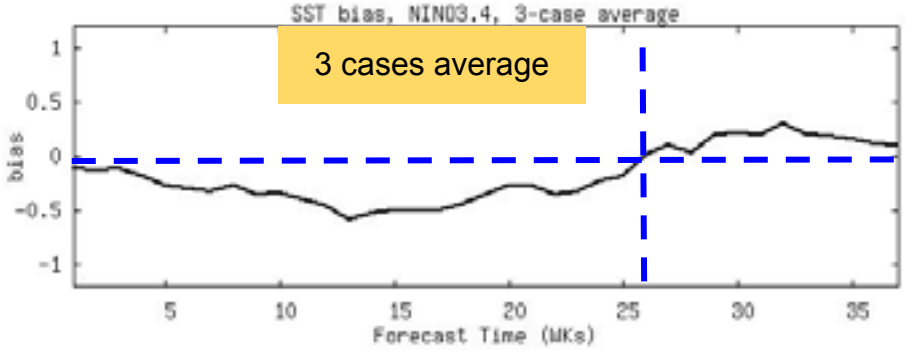


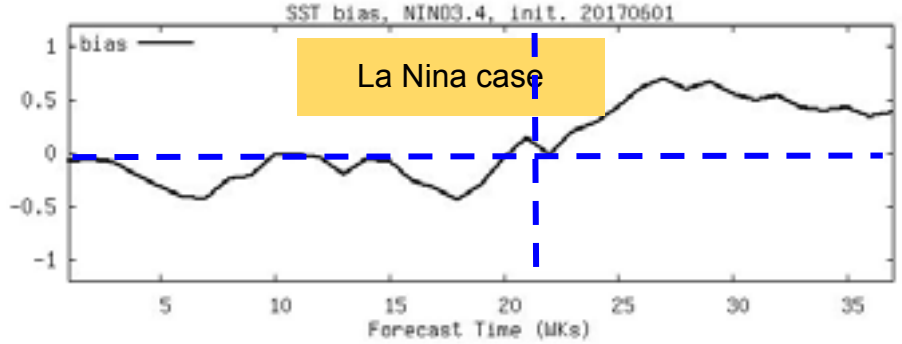


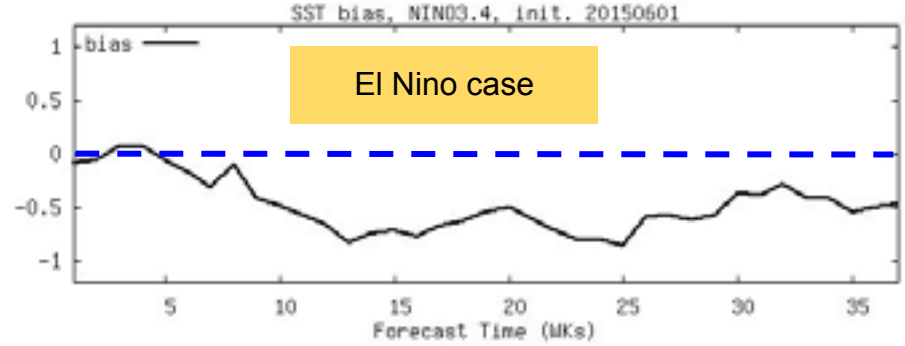


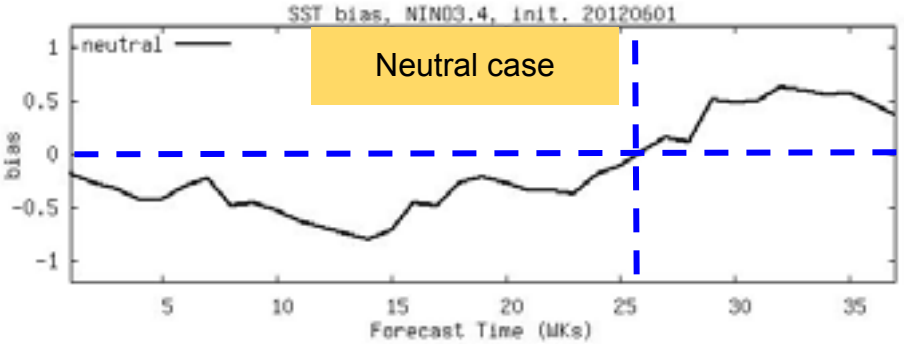
- Three cases (El Nino - 2015; La Nina - 2017; Neutral - 2012) average of SST bias, RMS error and ensemble spread
- Global average (weekly; left) and Nino 3.4 domain average (weekly; right); mean errors (top); and RMS error and spread (bottom)
- Nearly perfect of SST mean errors, ratio of RMS error and spread for global domain, but not Nino 3.4

SST bias of Nino 3.4 (against OSTIA)





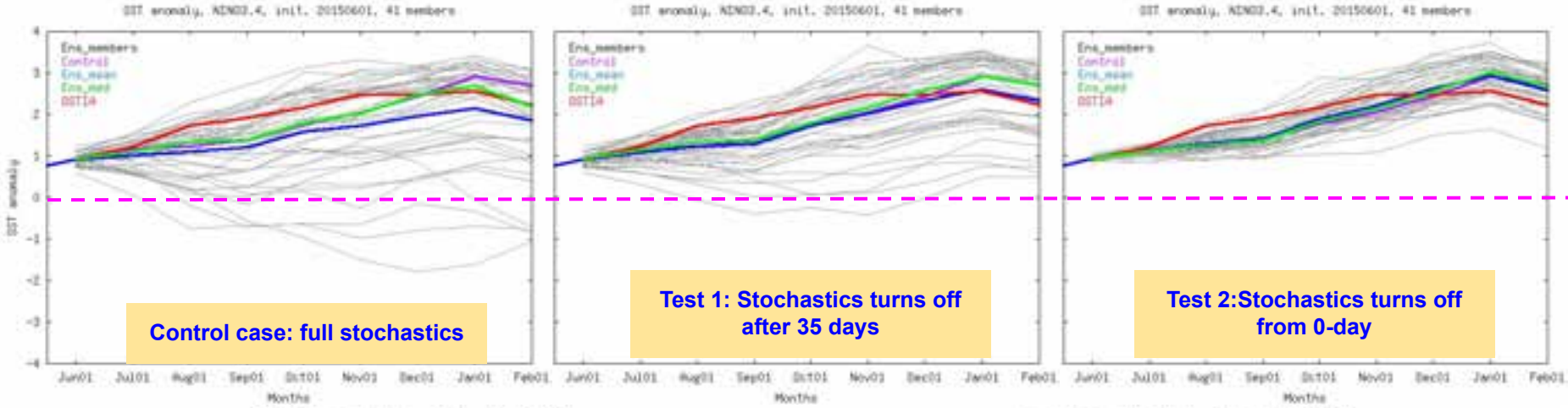


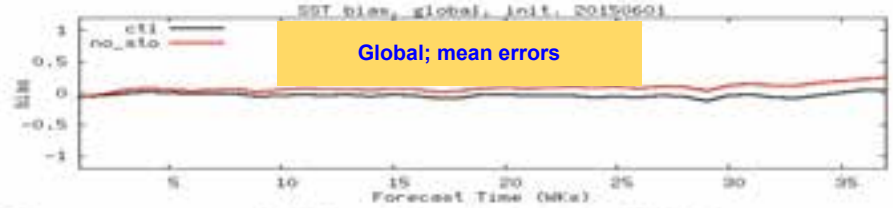


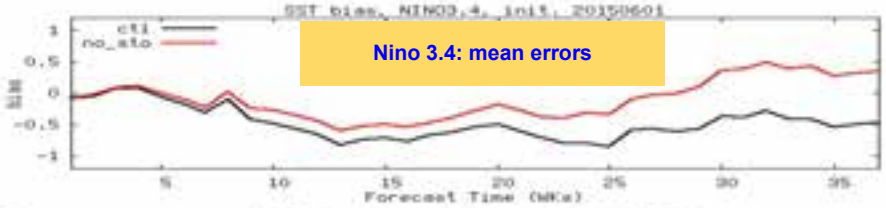
Discussion:

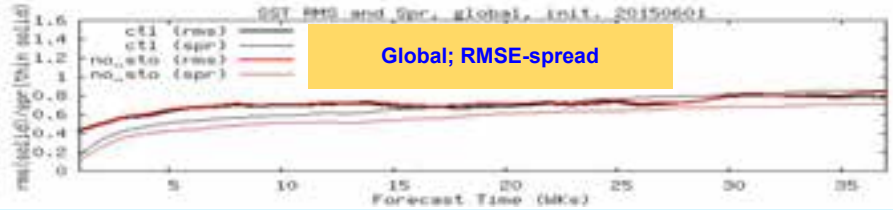
- 3 cases average time evolution of GEFS seasonal prediction - all cases are initialized on June 1st for 9 months
- Overall: from summer to fall, all three experiments have a cold bias, for winter time, they tend to warm
- On the average - summer is a cold bias, but winter is a warm bias

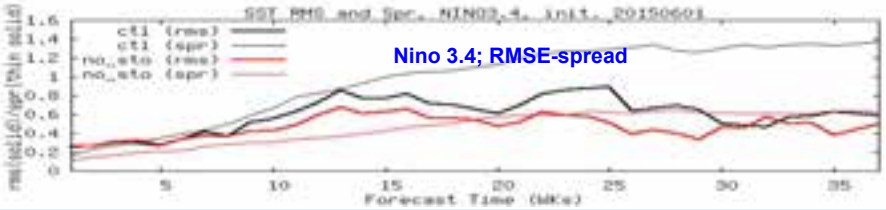
SST mean errors, RMS errors and spread (against OSTIA)



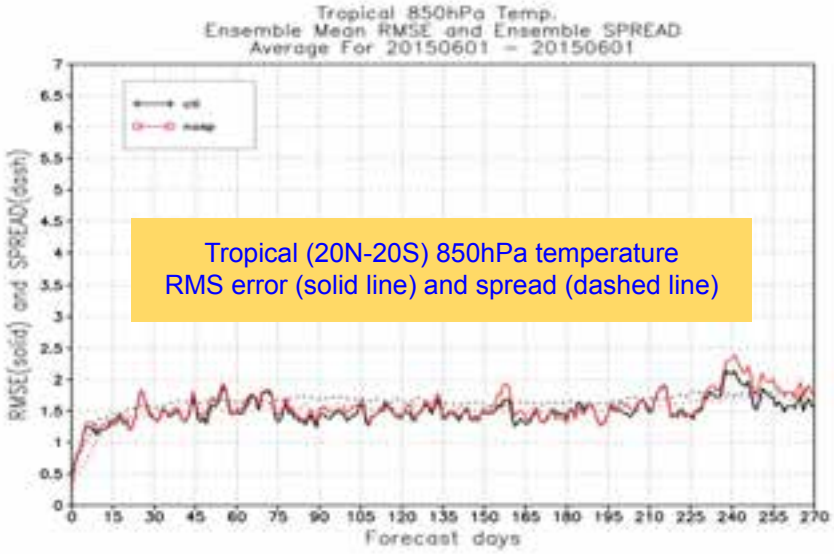


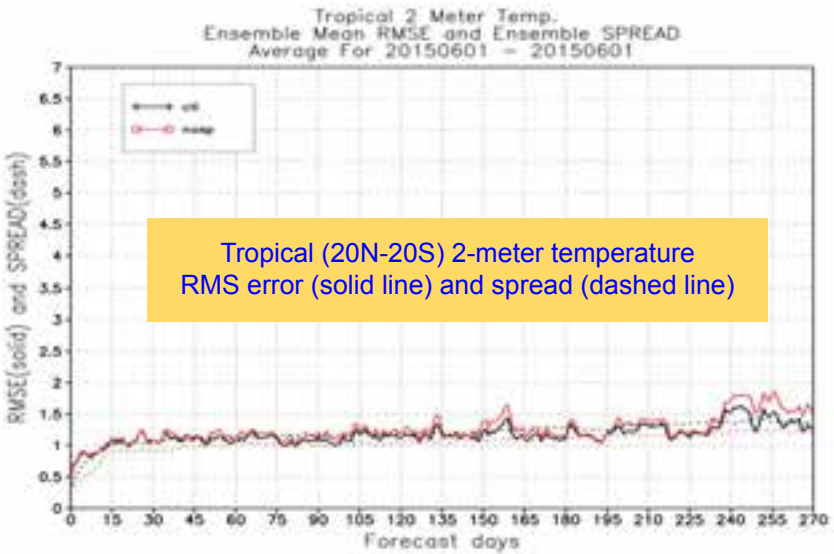






Tropical atmosphere temperature changes from stochastics





Discussion:

- This is for El Nino case only (initial time: 6/1/2015, and 9 months forecast)
- There is only a slightly difference for tropical atmosphere temperature from with/without stochastic schemes.
- However, it may indicate the large difference from 2-meter temperature to SST (previous slide). The more investigations are required to find out the reasons of large spread of ENSO prediction.

Short Summary of ENSO prediction

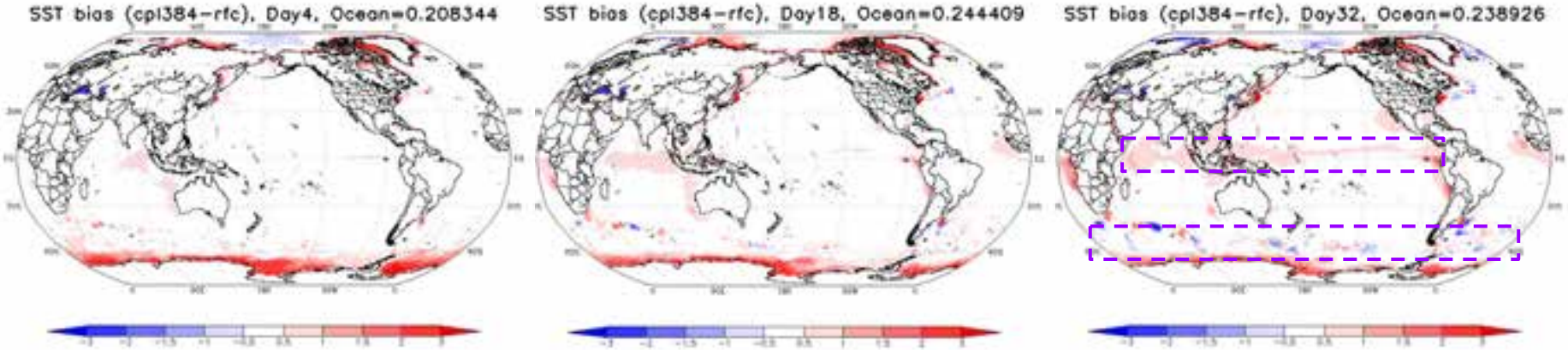
- The UFS coupled GEFS has demonstrated improved capability of the ENSO prediction, which could be preliminary UFS Seasonal Forecast system (SFS)
- Based on three special cases (El Nino; La Nina and Neutral) initialized on June 1 without bias correction, the preliminary UFS SFS shows even and/or better performance than CFSv2
- Like other seasonal prediction systems, there is systematic error around tropical and Nino 3.4 region. Consistently, it shows a cold bias for summer, and a warm bias for winter
- Ensemble spread of Nino 3.4 SST is much larger than the RMS error. It is larger than the other seasonal forecast systems, such as CFSv2 (NCEP), SPEAR (GFDL) and SEAS5 (ECMWF) - will continue to investigate this.
- Continue to work on the diagnostics of current experiments
- Plan to initialize the model for other seasons

Thanks for your attention!!!

Questions?

Bias Difference (CGEFS-H - Reforecast) of SST

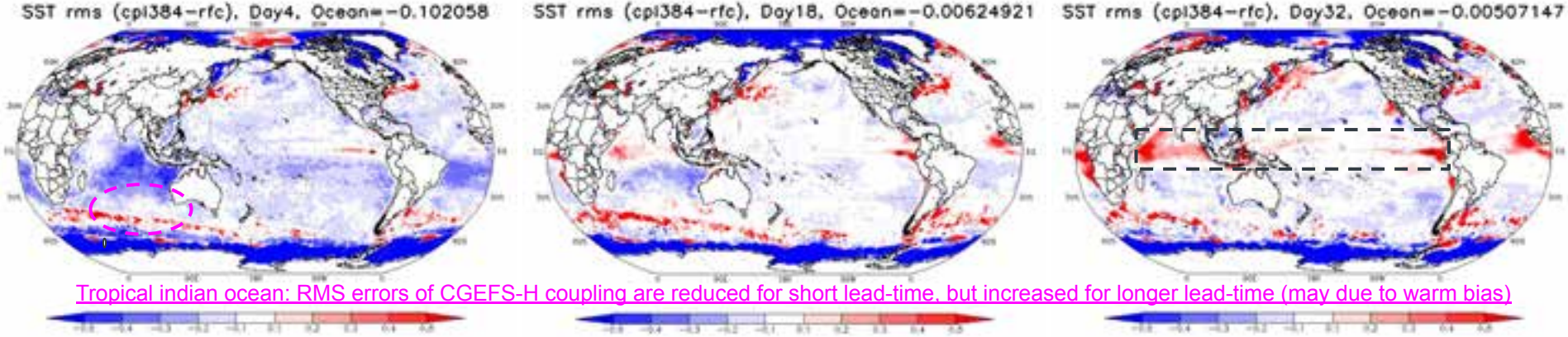
Day-4 **Day-18** **Day-32**



Discussion:

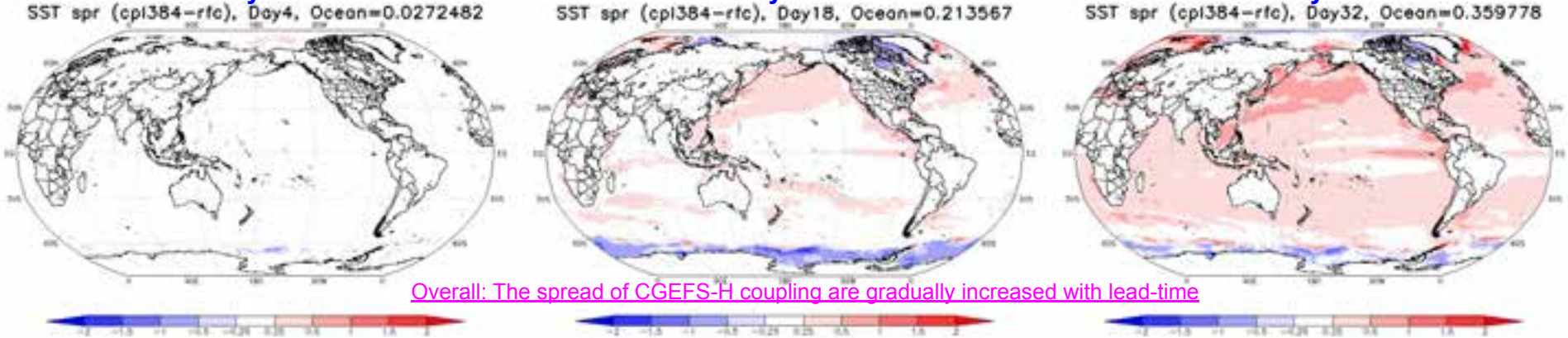
- All forecasts against OSTIA for bias calculation, plots show the difference of CGEFS-H and reforecast
- Main characters indicate the warm bias around tropical ocean, from Indian ocean to West Pacific ocean
- Southern ocean (around 50S-60S) - the difference (or bias difference) shows a lot of small scale pattern mainly due to the active Antarctic Circumpolar Current (ACC) - *see the similar characters for RMS errors next slide*
- Other large variability is from the Western Boundary Currents -- Gulf Stream, Kuroshio, Brazil, East Australian Current and Agulhas. All known and expected.
- Reference: http://web.gps.caltech.edu/~andrewt/research/acc.html

RMS/Spread Difference (CGEFS-H - Reforecast) of SST



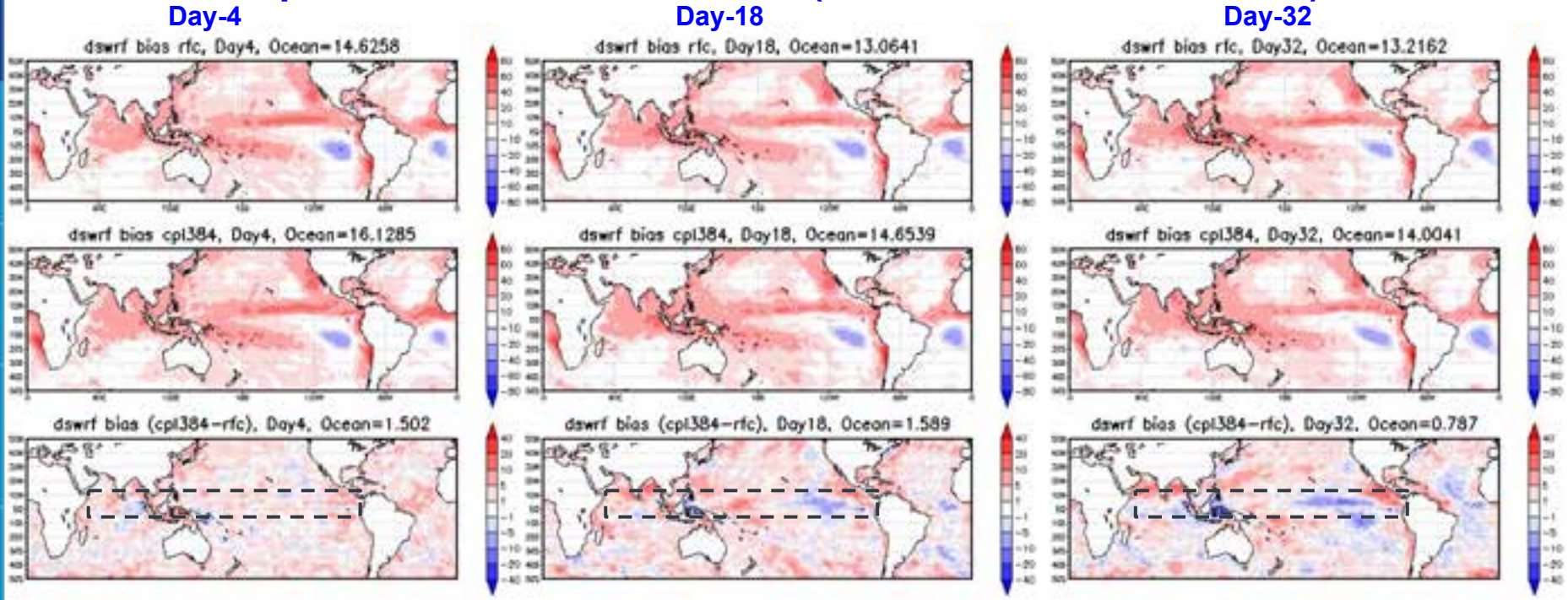
Tropical indian ocean: RMS errors of CGEFS-H coupling are reduced for short lead-time, but increased for longer lead-time (may due to warm bias)

Day-4 **Day-18** **Day-32**



Overall: The spread of CGEFS-H coupling are gradually increased with lead-time

Comparison of the DSWRF bias (CGEFS-H and Reforecast)

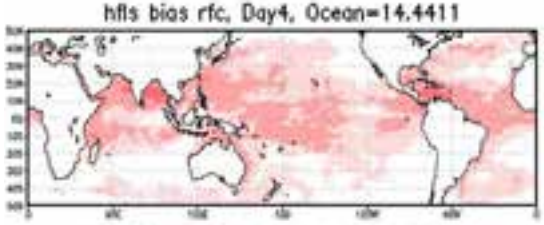


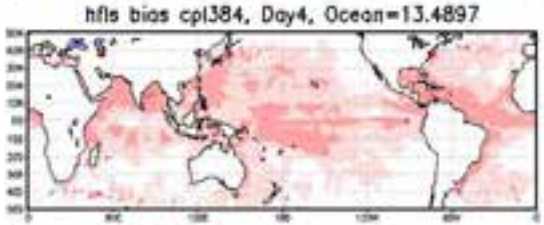
Summary and discussion:

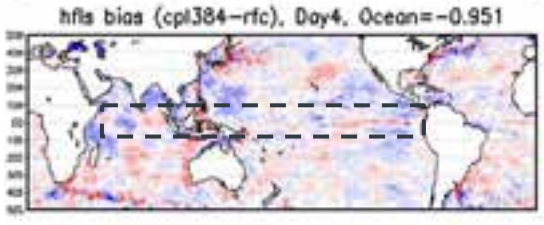
- Overall (globally), positive bias is dominant except for south-east Pacific Ocean and south Atlantic Ocean
- Coupling has slightly positive bias than reforecast
- CGEFS-H increases warm bias slightly of most ocean areas except for tropical around Indian ocean, maritime continent and central-east Pacific, which may indicate an increase in thick clouds and reducing DSWRF.

Comparison of the LHTFL bias (CGEFS-H and Reforecast)

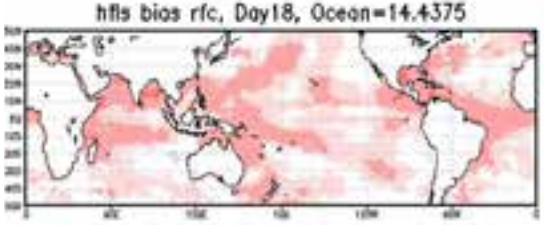
Day-4

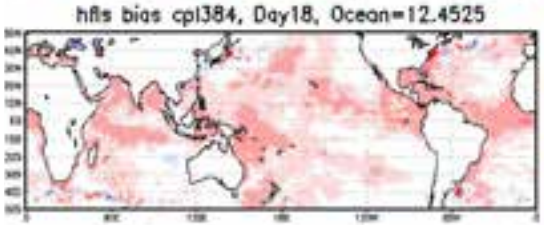


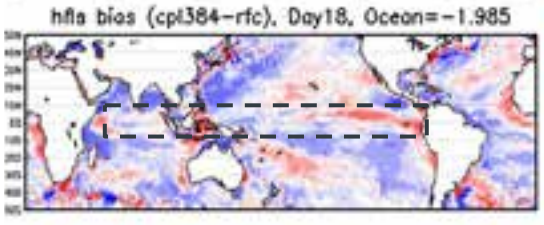




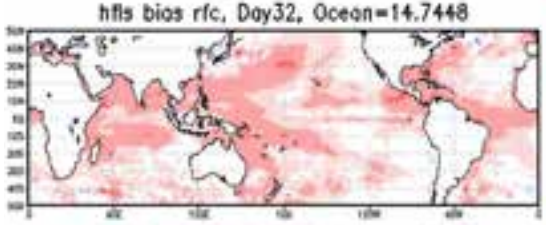
Day-18

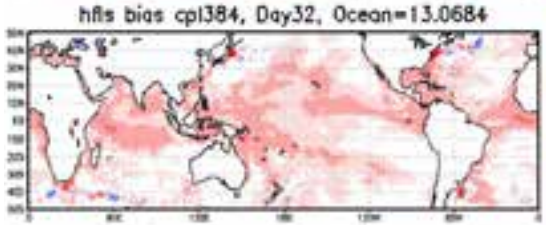


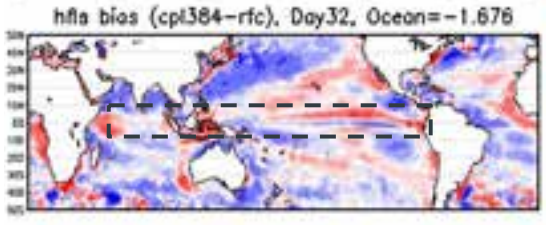




Day-32



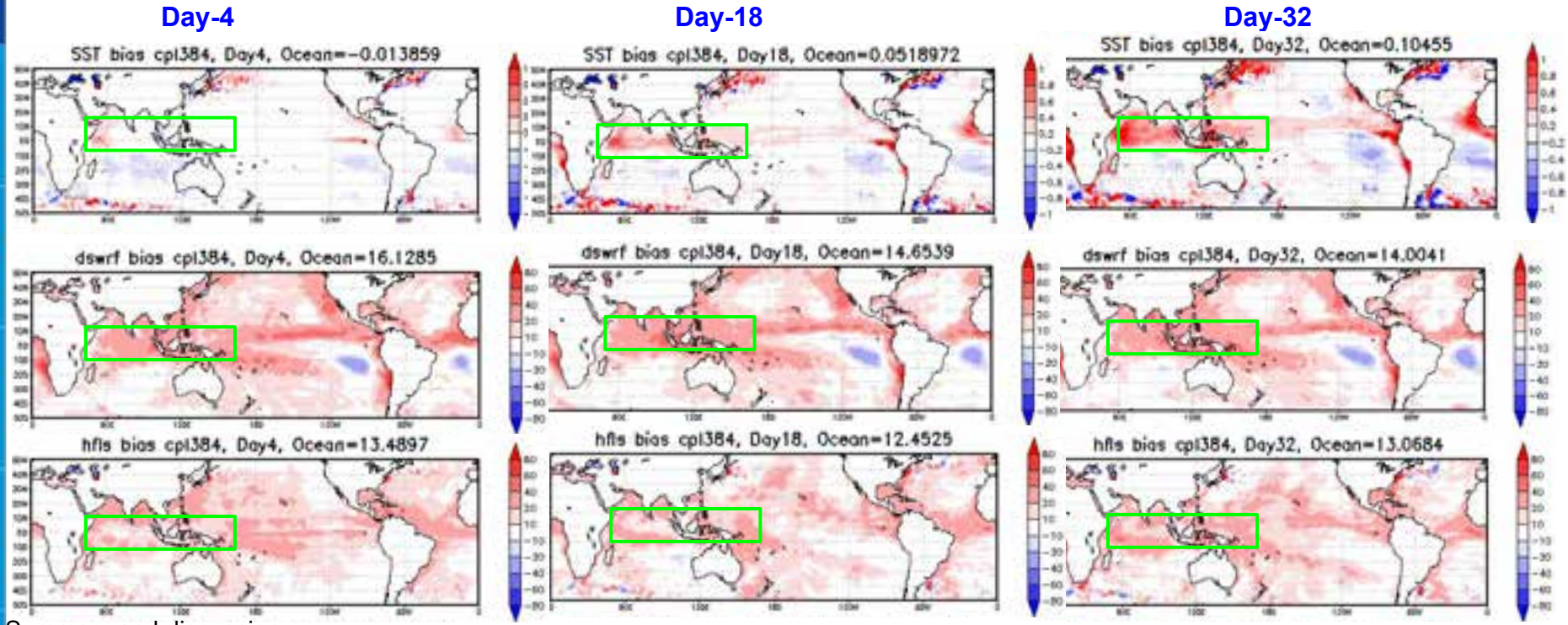




Summary and discussion:

- Positive bias are dominant for both runs
- Overall (globally), C384 coupling has less positive bias (May indicate that the rainfall is reduced slightly?)
- There is large bias reduced around west Pacific ocean
- There is the difference of bias around maritime area (which may be a good one?)

Contributions of the SST, DSWRF and LHTFL bias (Coupling C384)



Summary and discussion:

- SST: Warm bias increases with lead time over tropical (Indian Ocean and west Pacific Ocean)
- DSWRF: Larger positive bias persists for the same location
- LHTFL: Larger positive bias slightly increases with lead time for the same location
- Persistent warming effect from the DSWRF bias dominates cooling effect from the LHTFL bias. This leads a growing warm bias in SST