An Assessment of Predictability and Prediction of NCEP GEFS for Subseasonal Forecast

Yuejian Zhu
Wei Li, Eric Sinsky, Hong Guan, Xiaqiong Zhou and Bing Fu
Environmental Modeling Center
NCEP/NWS/NOAA

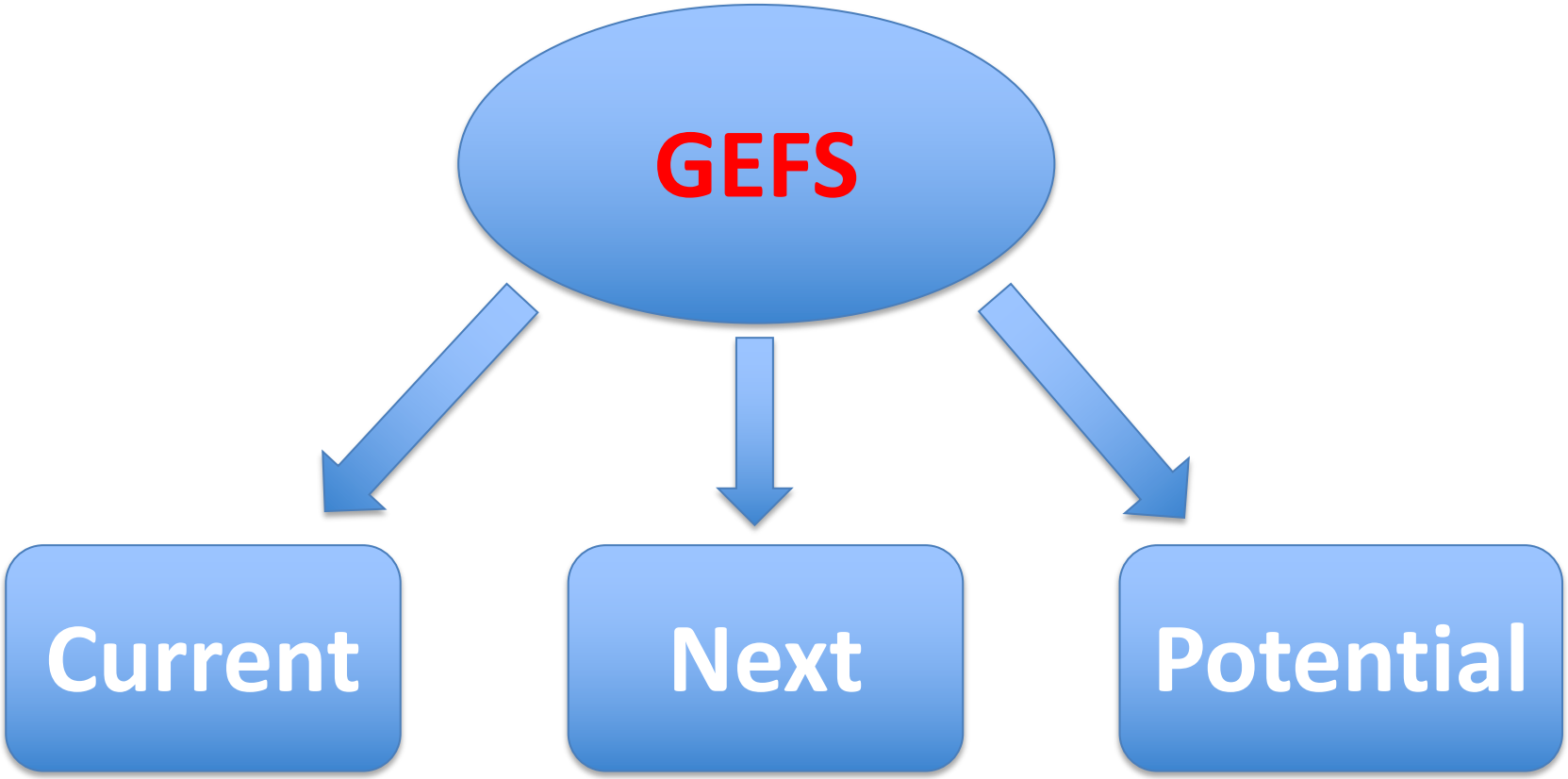
Present for 43rd Climate Diagnostic and Prediction Workshop
October 23-25 2018
Santa Barbara, California

Introduction

- Recently, the need for numerical guidance covering the **weeks 3&4** period has been increasing, driven primarily by economic requirements to support decision-makers and for preparedness to changes in climate. The NOAA is accelerating its efforts to improve numerical guidance and prediction capability for the extended range–S2S: the weeks 3&4 period that bridges the gap between weather and climate. Covering the extended-range period will enable NOAA to provide **seamless numerical guidance** to the public, protecting life and property. Thus, a better understanding of predictability and numerical model capabilities are necessary to enhance our capabilities of prediction beyond week-2.

- The **NCEP GEFS** has been very successful, providing reliable weather and week-2 probabilistic forecast guidance that has translated into valuable information for the general public. But the S2S **prediction capabilities** of the GEFS have only recently been evaluated. Specifically, these capabilities were evaluated as part of the **NOAA SubX** with a 18-year reforecast is used as a reference system. This study involves a comparison of SubX results with those from the newly developed **FV3-based GEFS**, which includes a different dynamical core, different horizontal resolution, different micro physics, etc. The upper limits of prediction skill will be investigated through these experiments with various evaluation metrics, which include extratropical circulations, MJO and other phenomena.

What do we present?



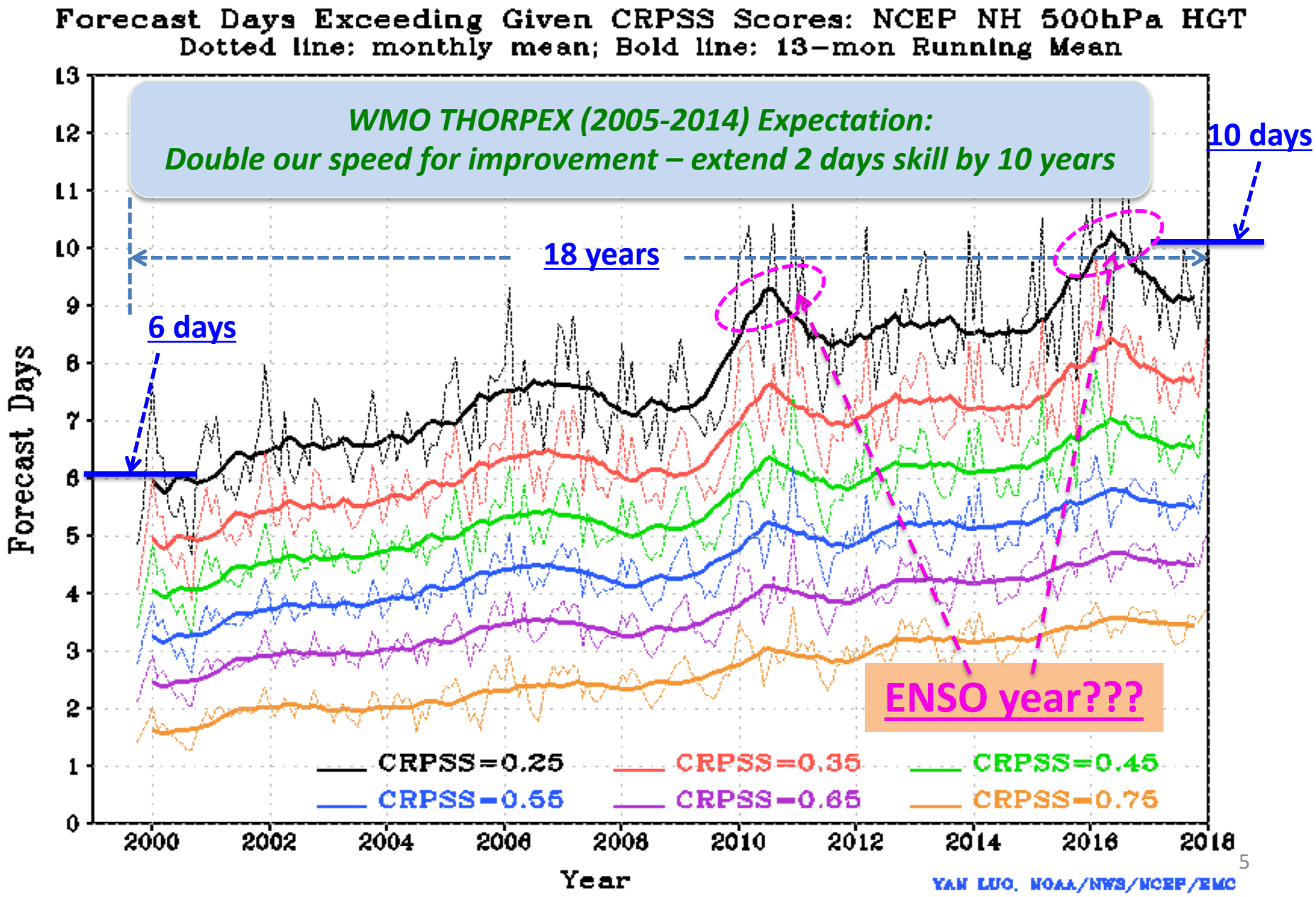
Current: Prediction skills of operational GEFS (35d forecast to support SubX)

Configuration of GEFS v11 and v11+

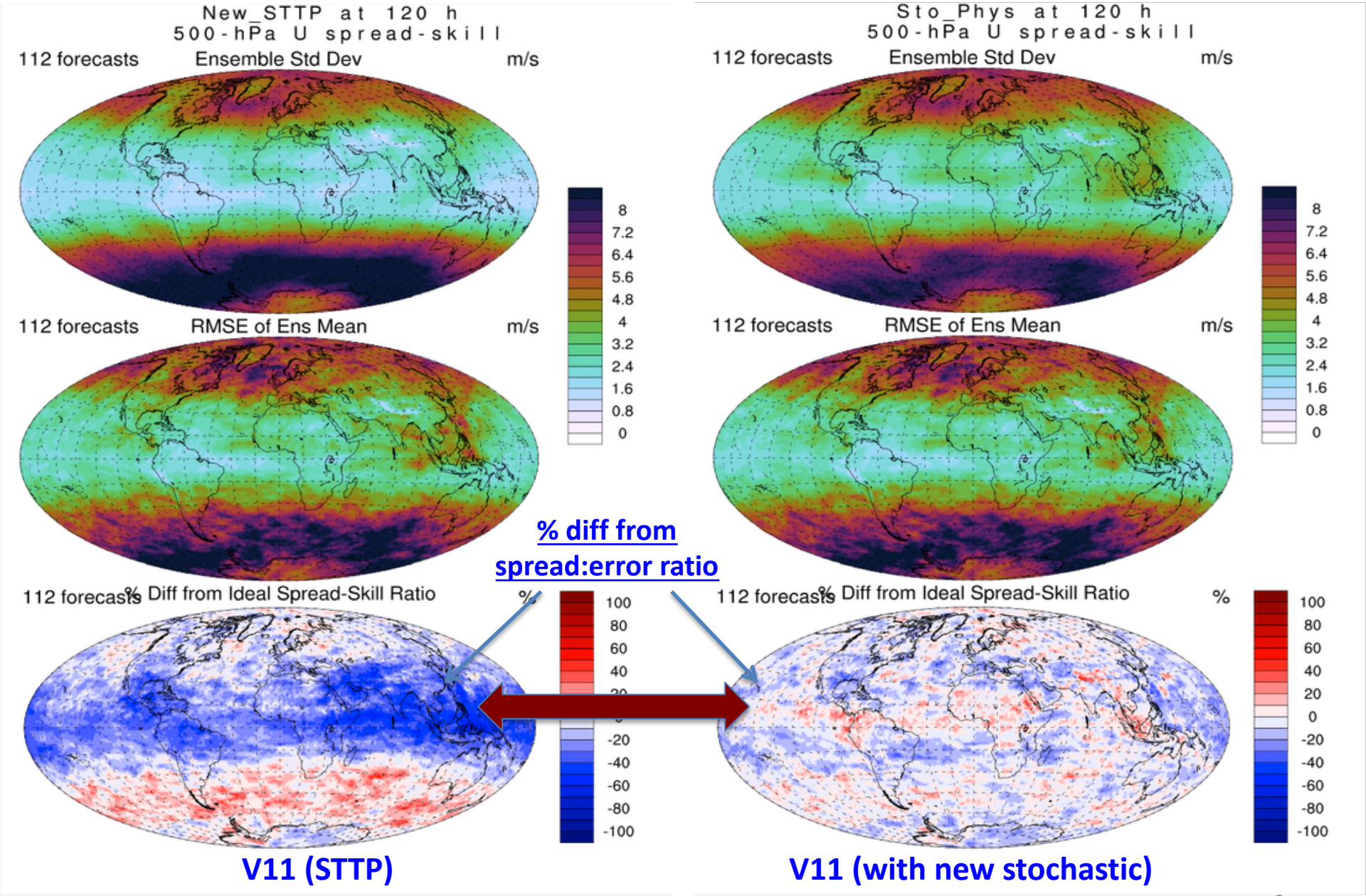
- Model: GSMv14 (spectrum model with semi-Lagrange time integration)
- Initial perturbation: F06 of EnKF analysis
- Model perturbation: STTP (stochastic total tendency perturbation)
- Resolutions: TL574L64 (0-8 days); TL384L64 (8-16 days)
- Forecast leads: out to 16 days (and 35 days)
- Members: 20 perturbed + control forecast
- Frequency: 4 times per day (00; 06; 12; 18UTC)
- Output data: 0.5d resolution globally

- ***GEFS v11 + to support SubX in real-time***:
  - ✓ SPPT+SHUM+SKEB (**SPs**) with control version of SST;
  - ✓ SPs with bias corrected CFSv2 forecast SST (**SPs+CFSBC**);
  - ✓ SPs with bias corrected CFSv2 forecast SST and scale aware convection scheme (**SPs+CFSBC+CNV**) ;

CRPSS for NH 500hPa geopotential height



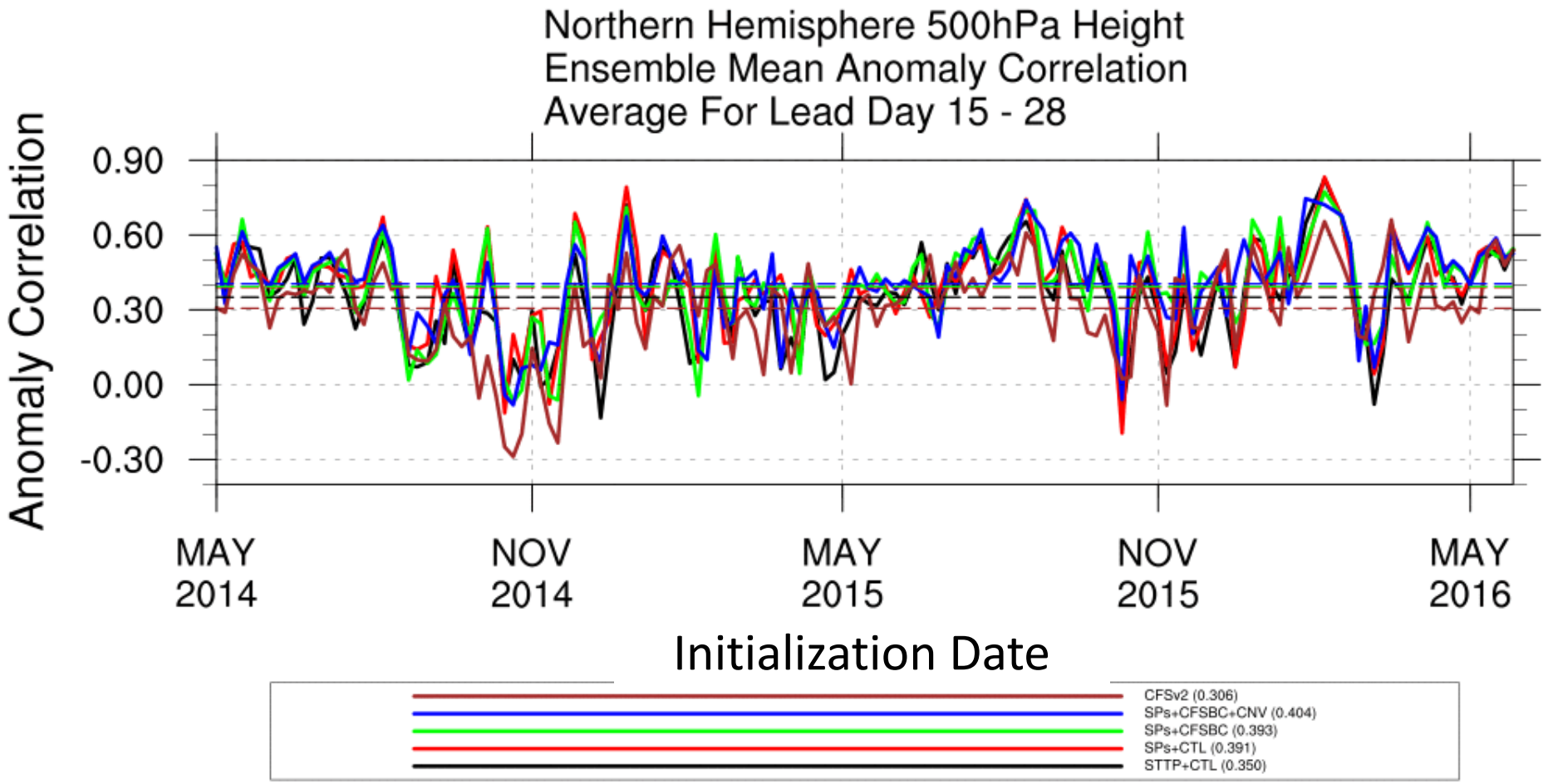
Change of ensemble spread from introducing new stochastic physics



Courtesy of Dr. Walter Kolczynski

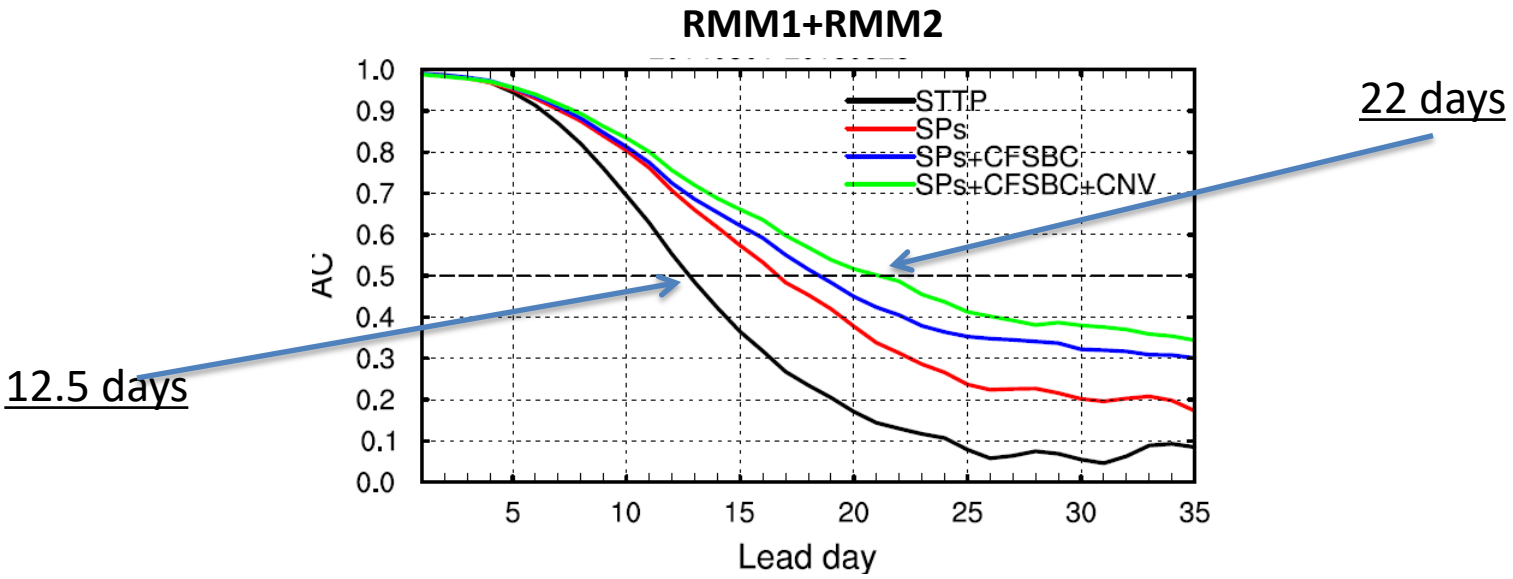
Skills of weeks 3&4 average of NH 500hPa height

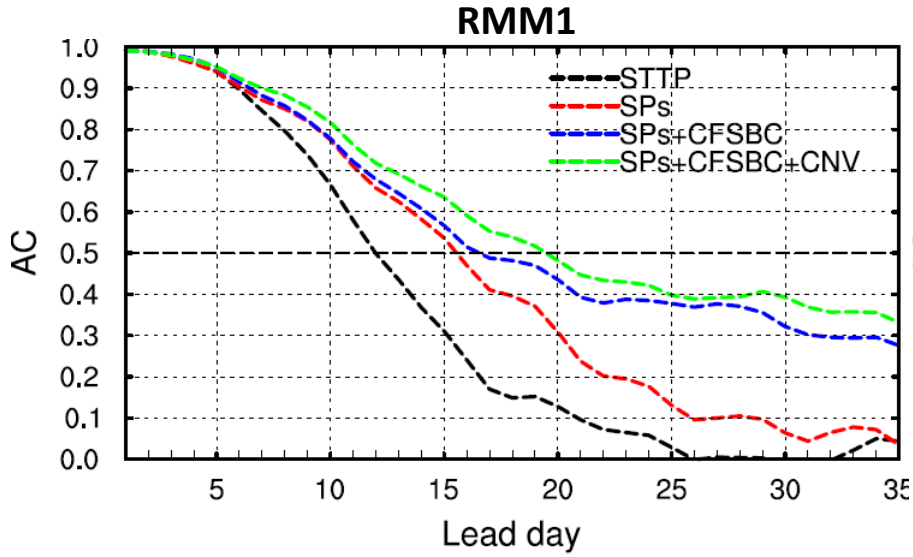
Period: April 2014 – April 2016

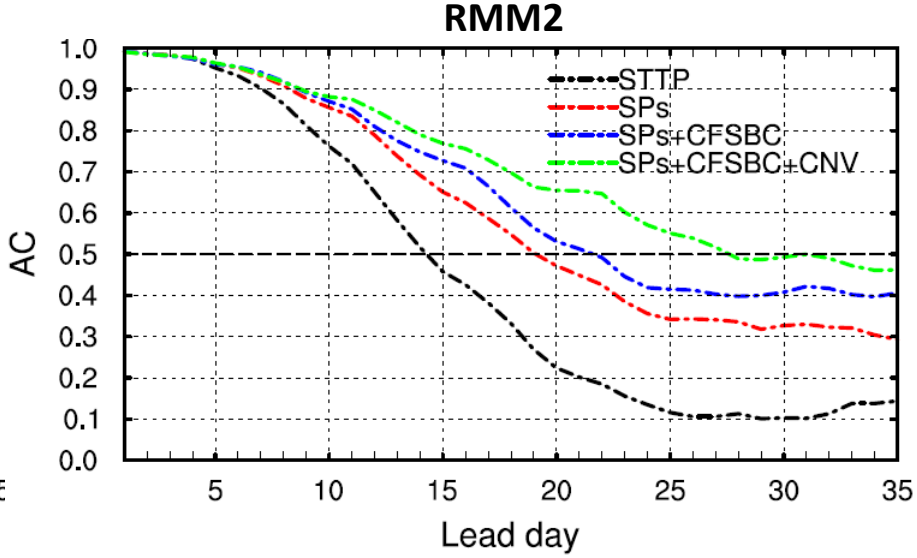


SPs+SST_bc+SA-CV (0.404) **CFSv2 (0.306)**

WH-MJO Forecast Skills for 2-yr Experiments







NCEP GEFS related publication on S2S (since 2016)

- Zhu, Y., X. Zhou, M. Pena, W. Li, C. Melhauser and D. Hou, 2017: *"Impact of Sea Surface Temperature Forcing on Weeks 3 & 4 Forecast Skill in the NCEP Global Ensemble Forecasting System"* Weather and Forecasting, Vol. 32, 2159-2173
- Zhu, Y., W. Li, E. Sinsky, H. Guan, X. Zhou and D. Hou, 2018: *"An Assessment of Subseasonal Forecast Using Extended Global Ensemble Forecast System (GEFS)"* STI Climate Bulletin, P150-153, doi:10.7289/V5/CDPW-NWS-42nd-2 018
- Zhu, Y., X. Zhou, W. Li, and et al., 2018:*"Towards the Improvement of Sub-Seasonal Prediction in the NCEP Global Ensemble Forecast System (GEFS)"* Journal of Geophysical Research, 6732-6745
- 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
- Guan, H., Y. Zhu, E. Sinsky, W. Li, X. Zhou, D. Hou, C. Melhauser and R. Wobus, 2018: *"Systematic Error Analysis and Calibration of 2-m Temperature for the NCEP GEFS Reforecast of SubX Project"* Weather and Forecasting (in final process)
- Liu, P., Y. Zhu, and et al., 2017 *"Climatology of Tracked Persistent Maxima of 500-hPa Geopotential Height"*, Climate Dynamics, 701-717
- Liu, P., Q. Zhang, C. Zhang, Y. Zhu, and et al., 2016: *"A Revised Real-Time Multivariate MJO Index"* Monthly Weather Review, Vol. 144, 627-642
- He, B., P. Liu, Y. Zhu, W. Hu 2017: *"Prediction and Predictability of Northern Hemisphere Persistent 2 Maxima of 500-hPa Geopotential Height Eddies in GEFS"* Climate Dynamics (final online version)
- Fu, J-X., W. Wang, Y. Zhu, and et al. 2018: *"Impacts of Different Cumulus Schemes on the Pathways Through Which SST Feedbacks to the Madden-Julian Oscillation"* Journal of Climate
- Pegion and co-authors, 2018: *"The Subseasonal Experiment (SubX): A multi-model subseasonal prediction experiment"*, Submit to BAMS (in review process)

Next: FV3 based GEFS (v12)

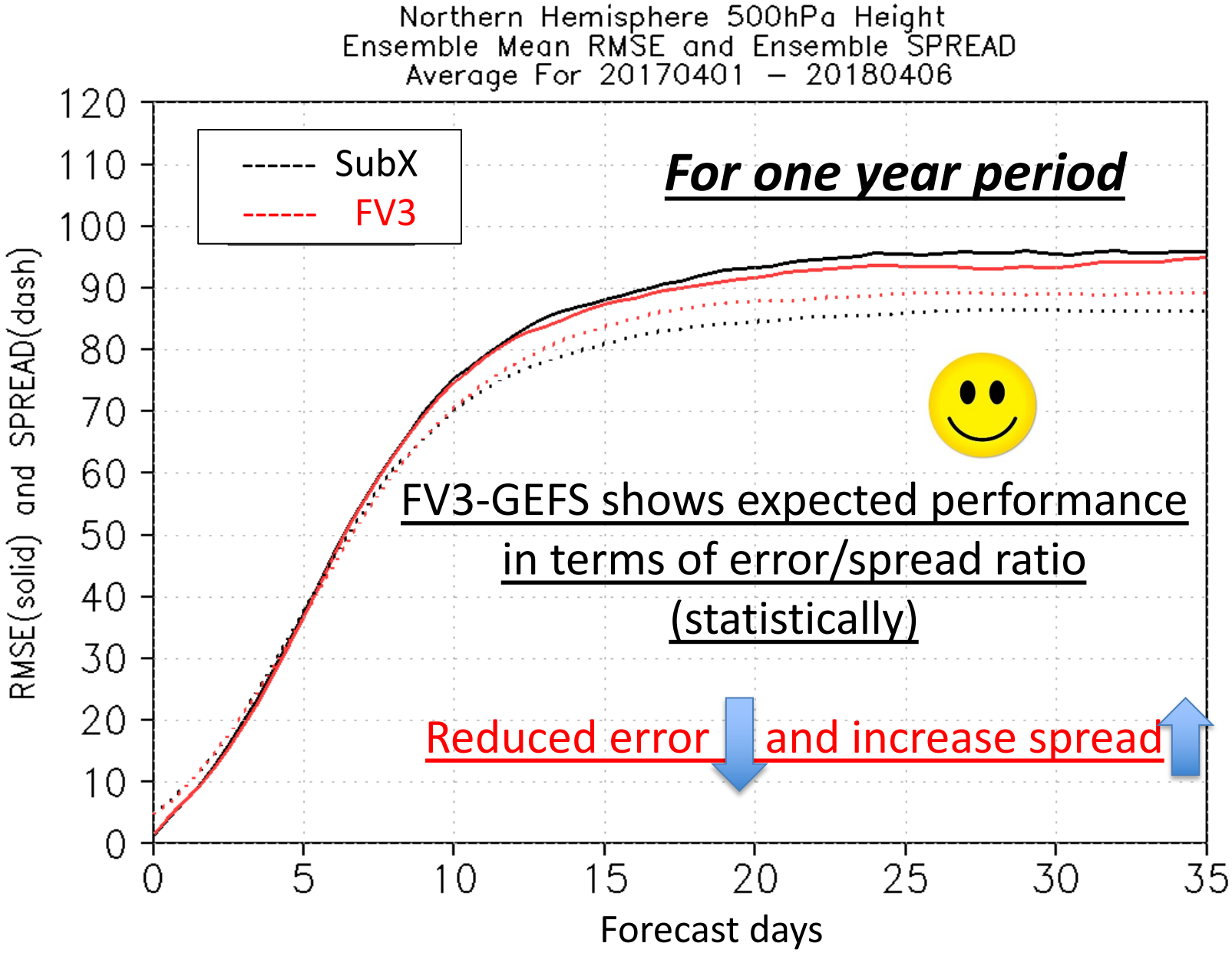
Configuration of GEFS v12 (plan)

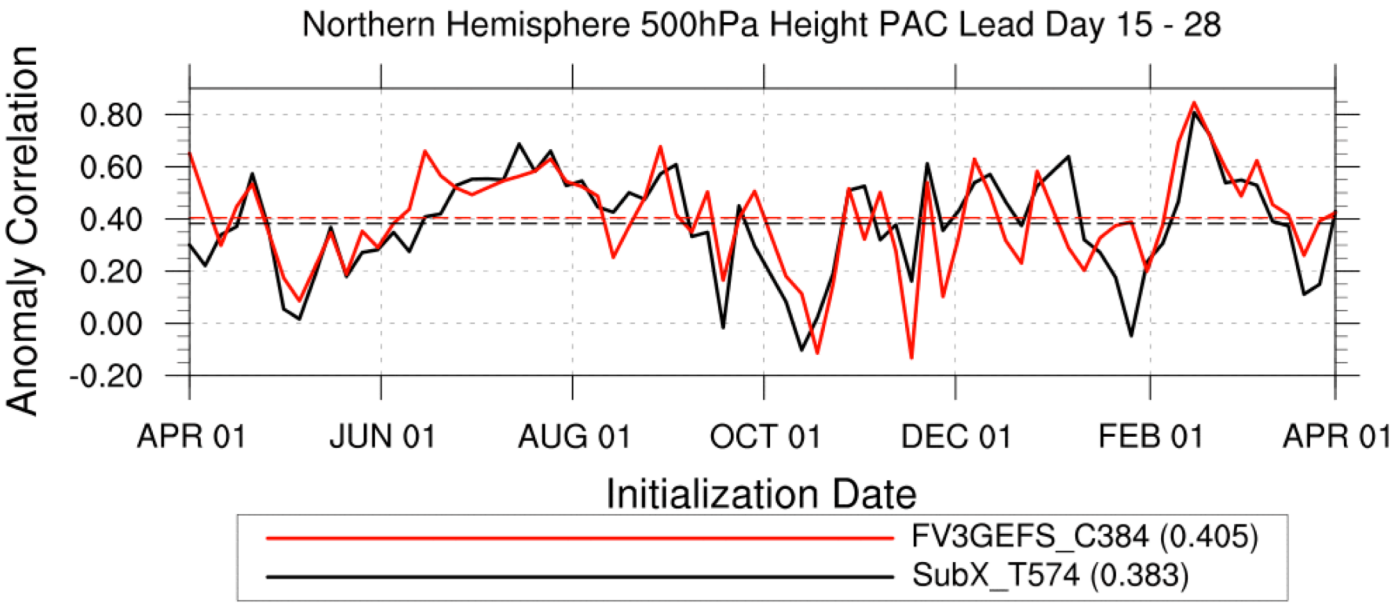
- What are the major difference from GEFSv11 (or V11+ SubX version)

  - ✓ Model dynamics – **FV3 (Finite-Volume Cubed-Sphere Dynamical Core)**
  - ✓ Horizontal resolution – C384 ~ 25km
  - ✓ Microphysics – GFDL MP
  - ✓ Tuned Stochastic Physics
  - ✓ 31 ensemble members (skills we have demonstrated are from 21 members)

  - ✓ Computation cost – factor of 4

RMSE and Ensemble Spread of NH 500hPa height





Weeks 3&4 average

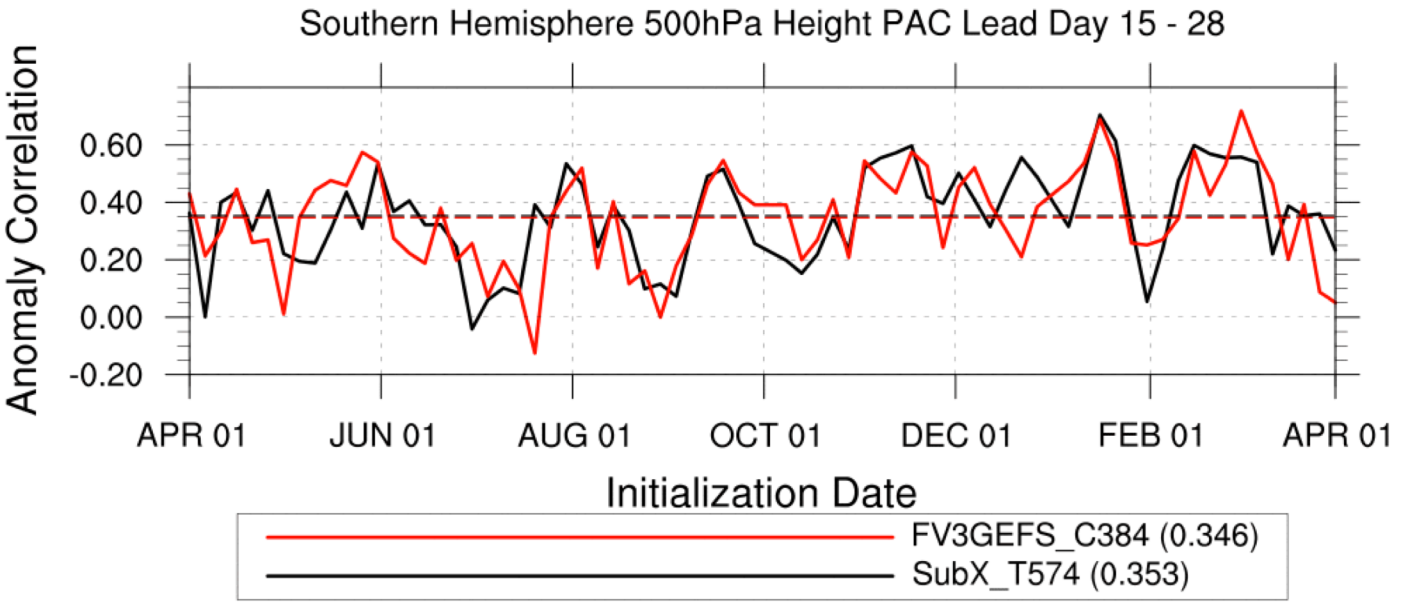
SubX --- 0.383

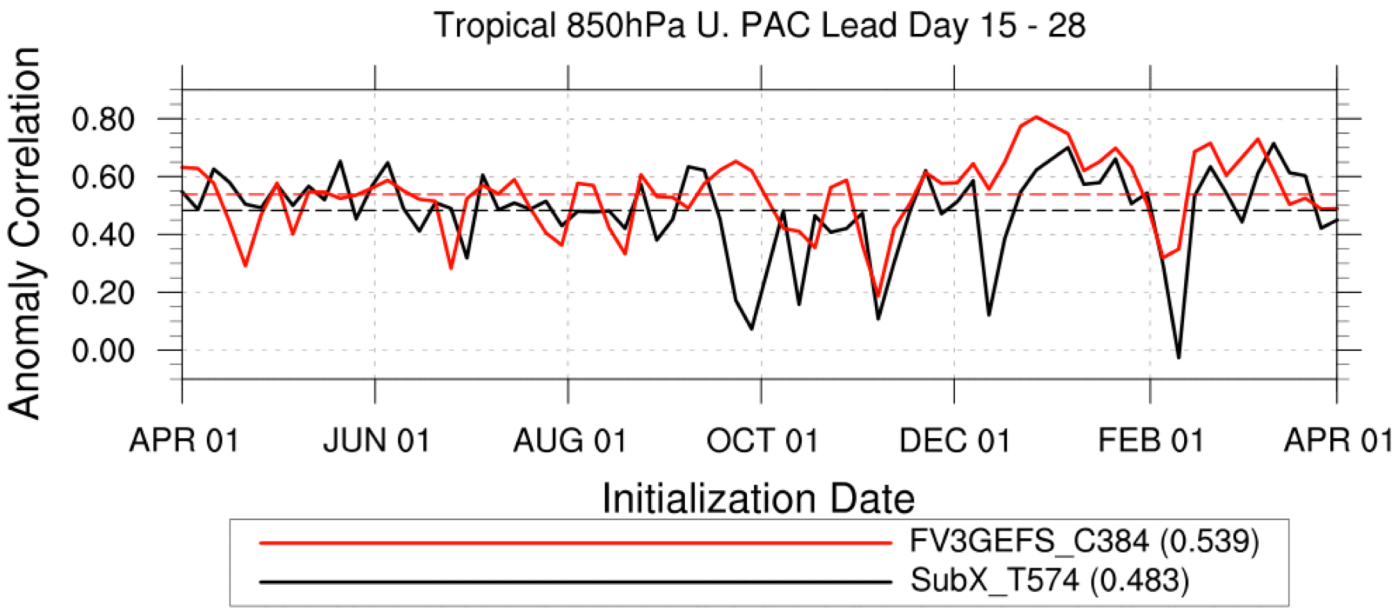
FV3 --- 0.405

Weeks 3&4 average

SubX --- 0.353

FV3 --- 0.346





Weeks 3&4 average

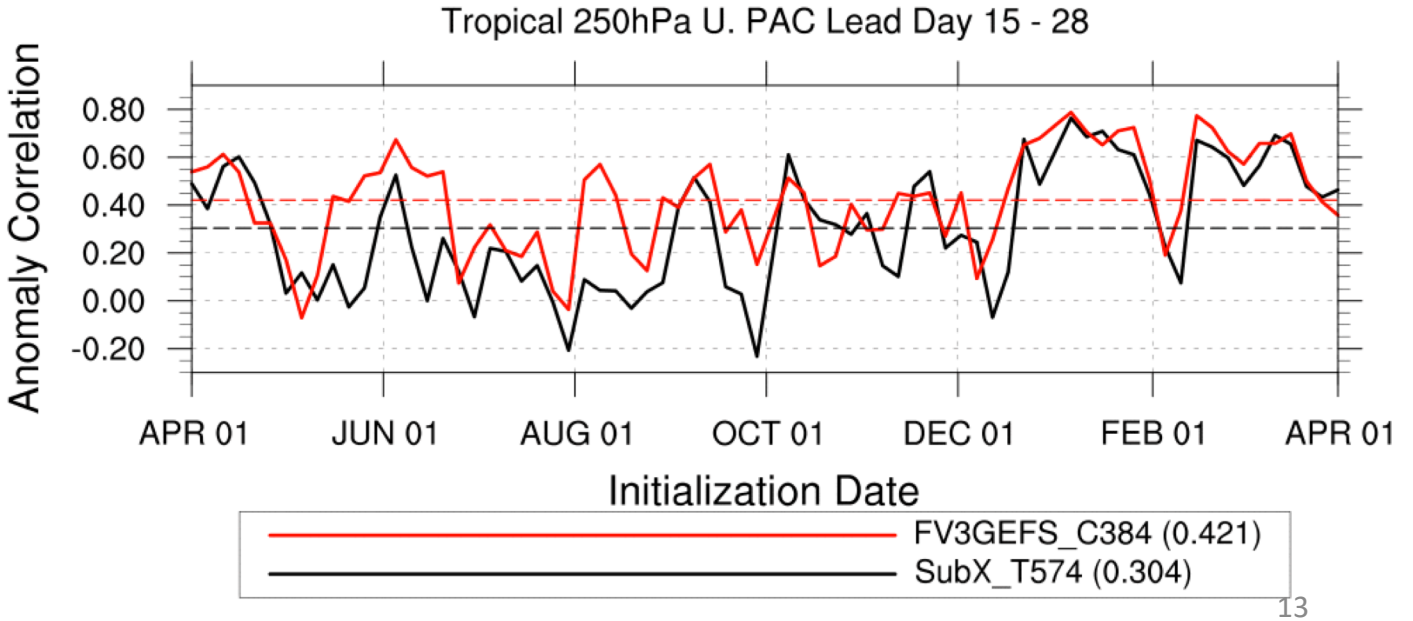
SubX --- 0.483

FV3 --- 0.539

Weeks 3&4 average

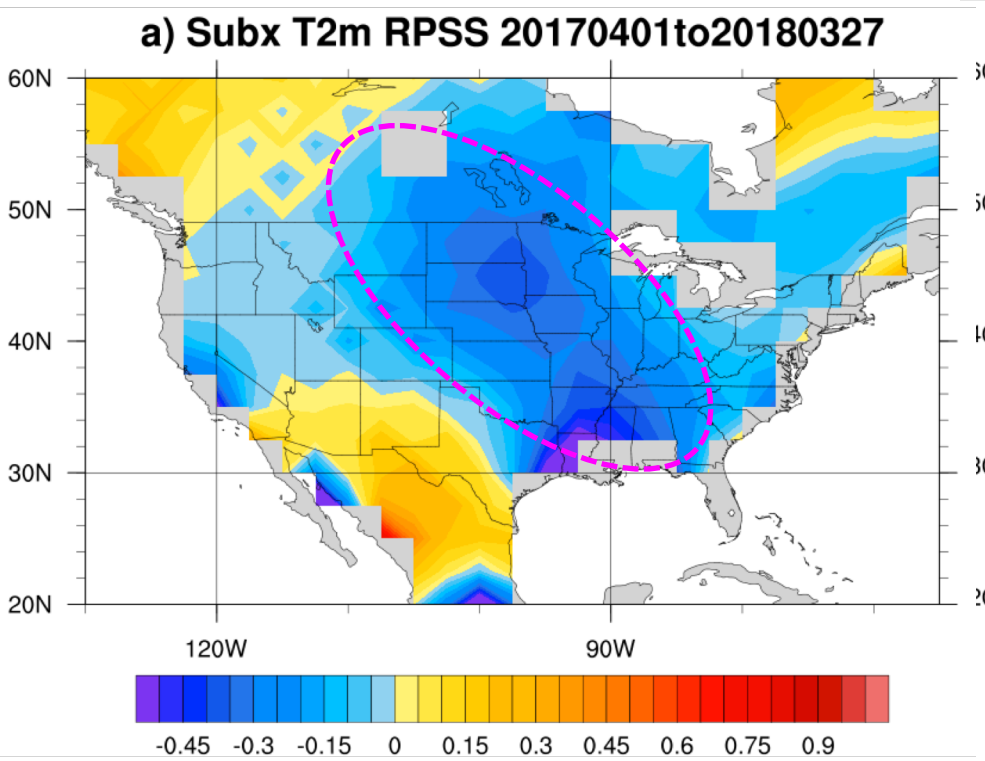
SubX --- 0.304

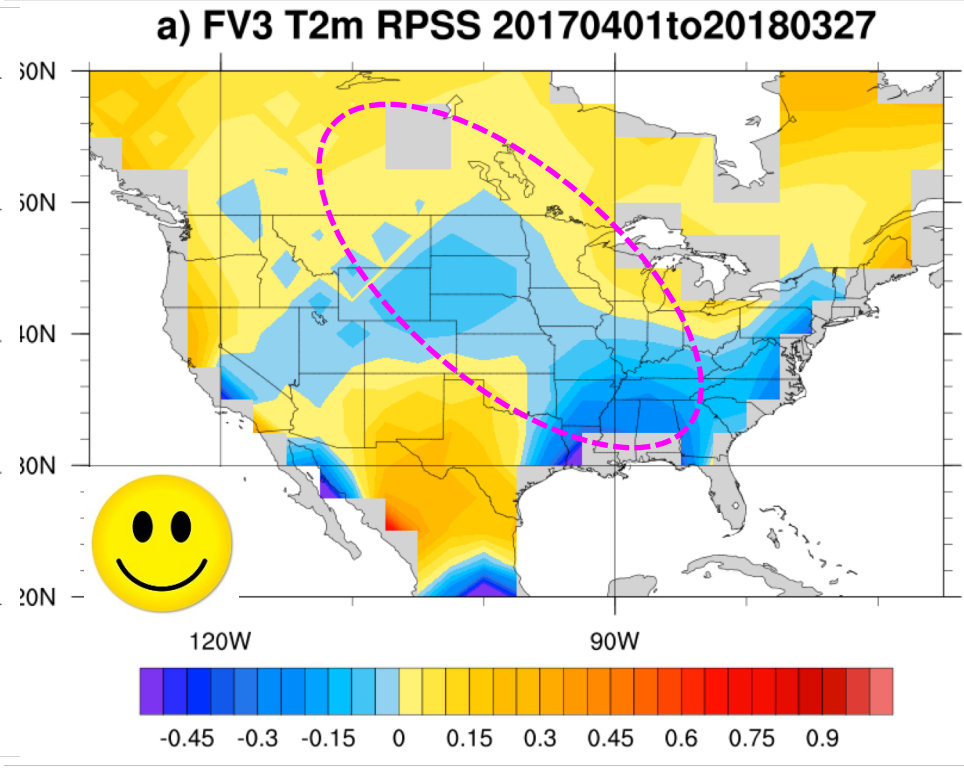
FV3 --- 0.421



RPSS scores for one years 35 days forecasts

Weeks 3&4 average



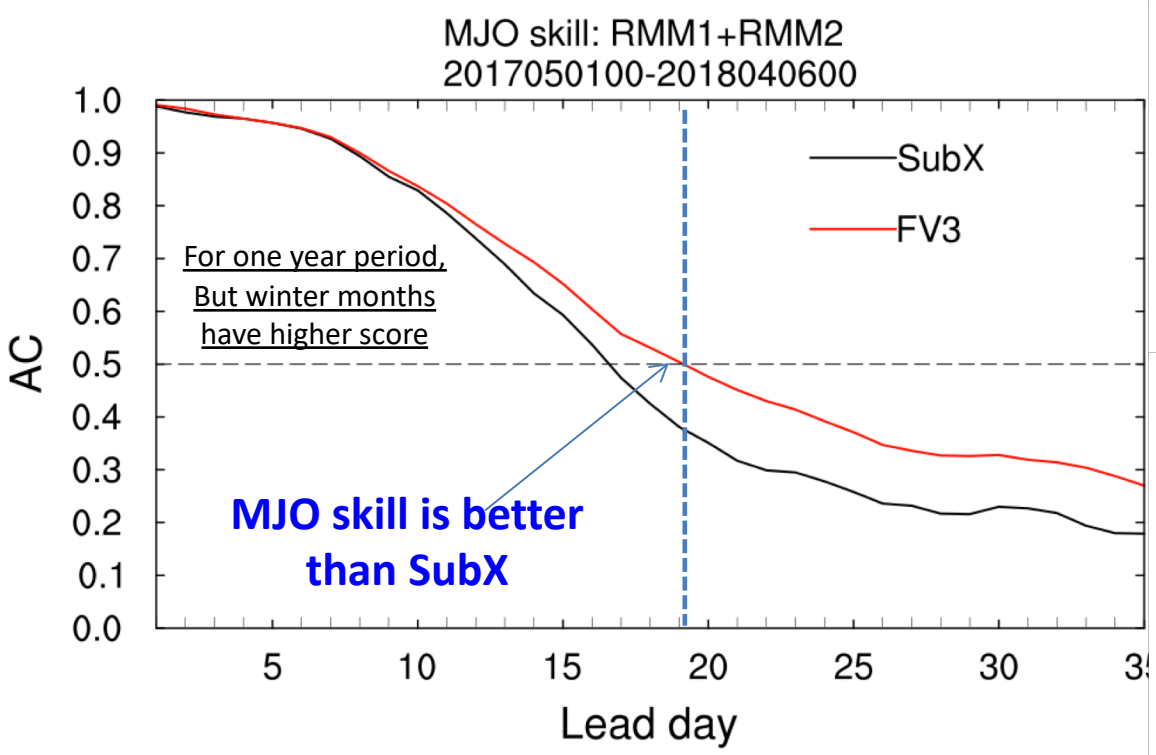


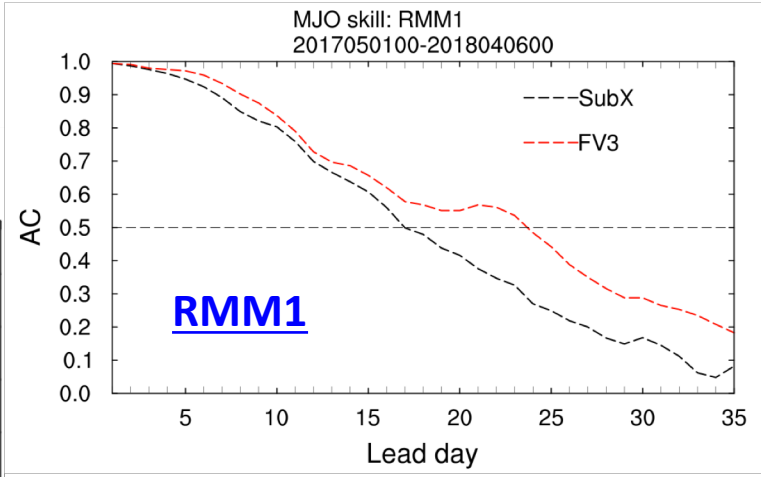
FV3-GEFS indicates an big improvement of T2m for CONUS

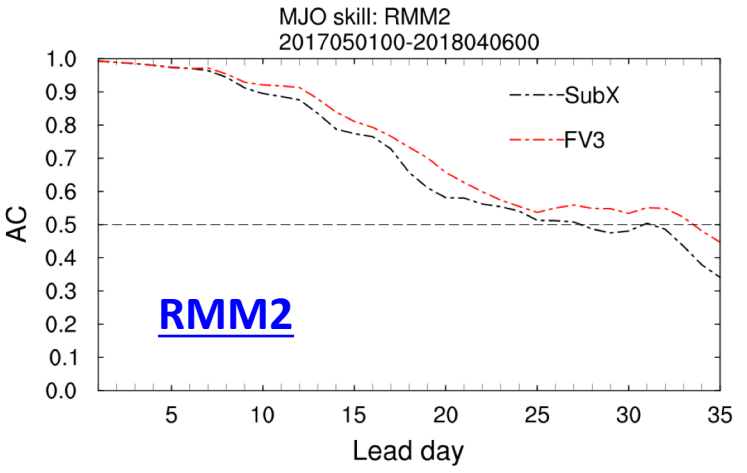
For raw ensemble forecast (no calibration)

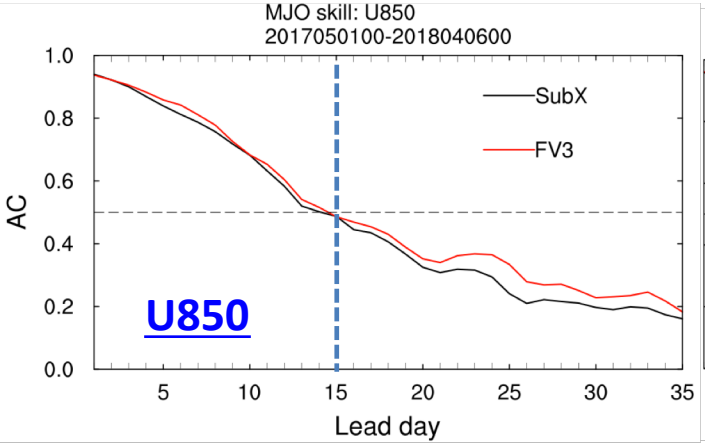
Truth: own analysis or f00 at 2.5d resolution

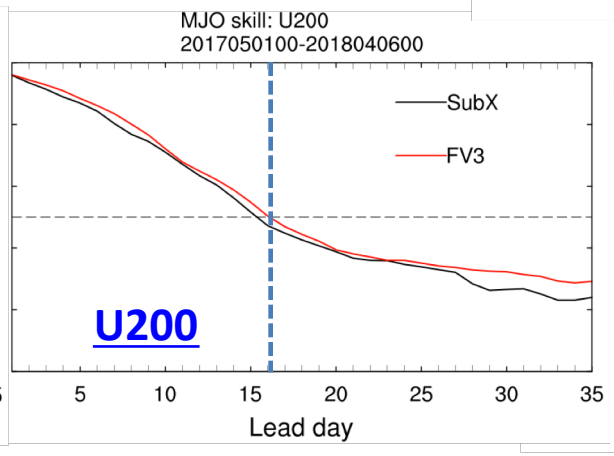
Tropical Prediction Skills

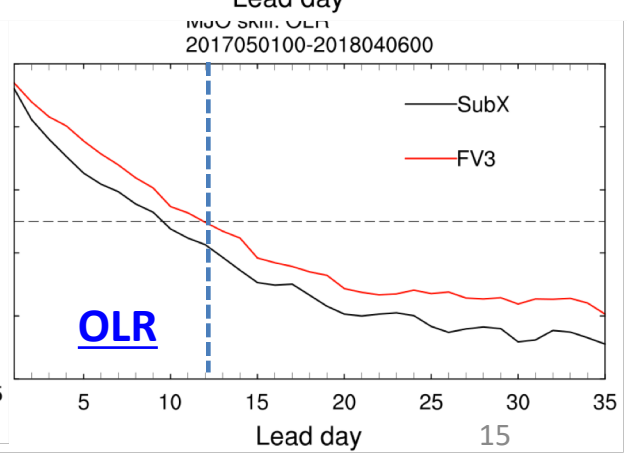












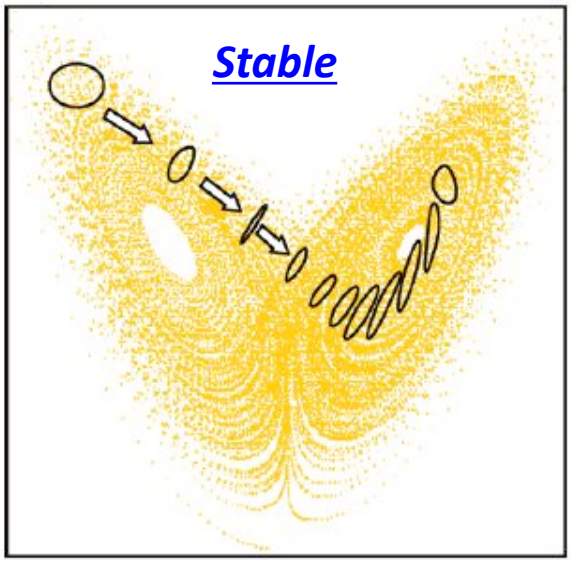
Potential - Predictability

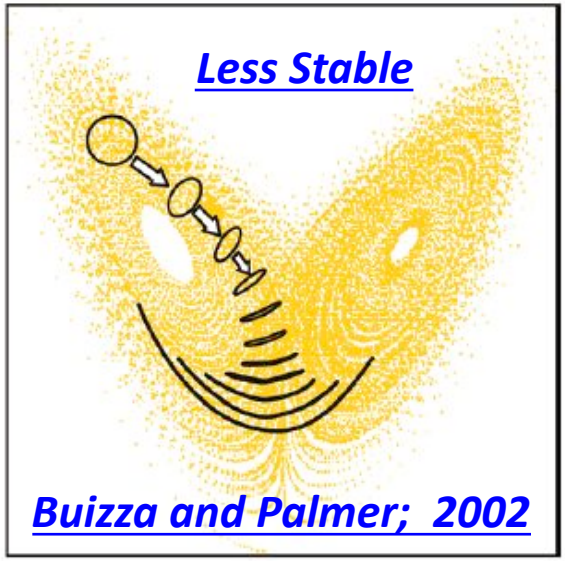
Our assumptions:

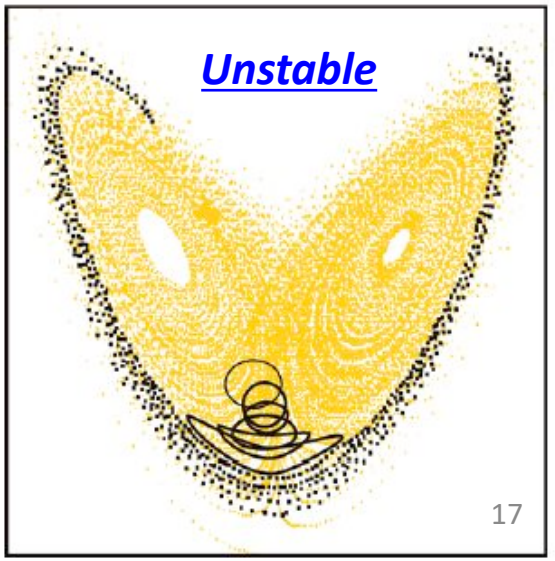
1. Model is perfect
2. Ensemble system is perfect
3. Ensemble mean will have best average solution, but smooth out small scales
4. Errors come from observation uncertainties and chaotic system

Background

- ✓ **Predictability** is the degree to which a correct prediction or forecast of a system's state can be made either qualitatively or quantitatively
- ✓ Charney (1951) indicated that forecast skill would break down, but he attributed it to **model errors** and errors in the **initial conditions**
- ✓ Lorenz (1963) discovered that even with a perfect model and almost perfect initial conditions the forecast loses all skill in a finite time interval **because chaotic system**
- ✓ Now, we are getting closer to the **2 week limit of predictability**, and we have to extract the maximum information

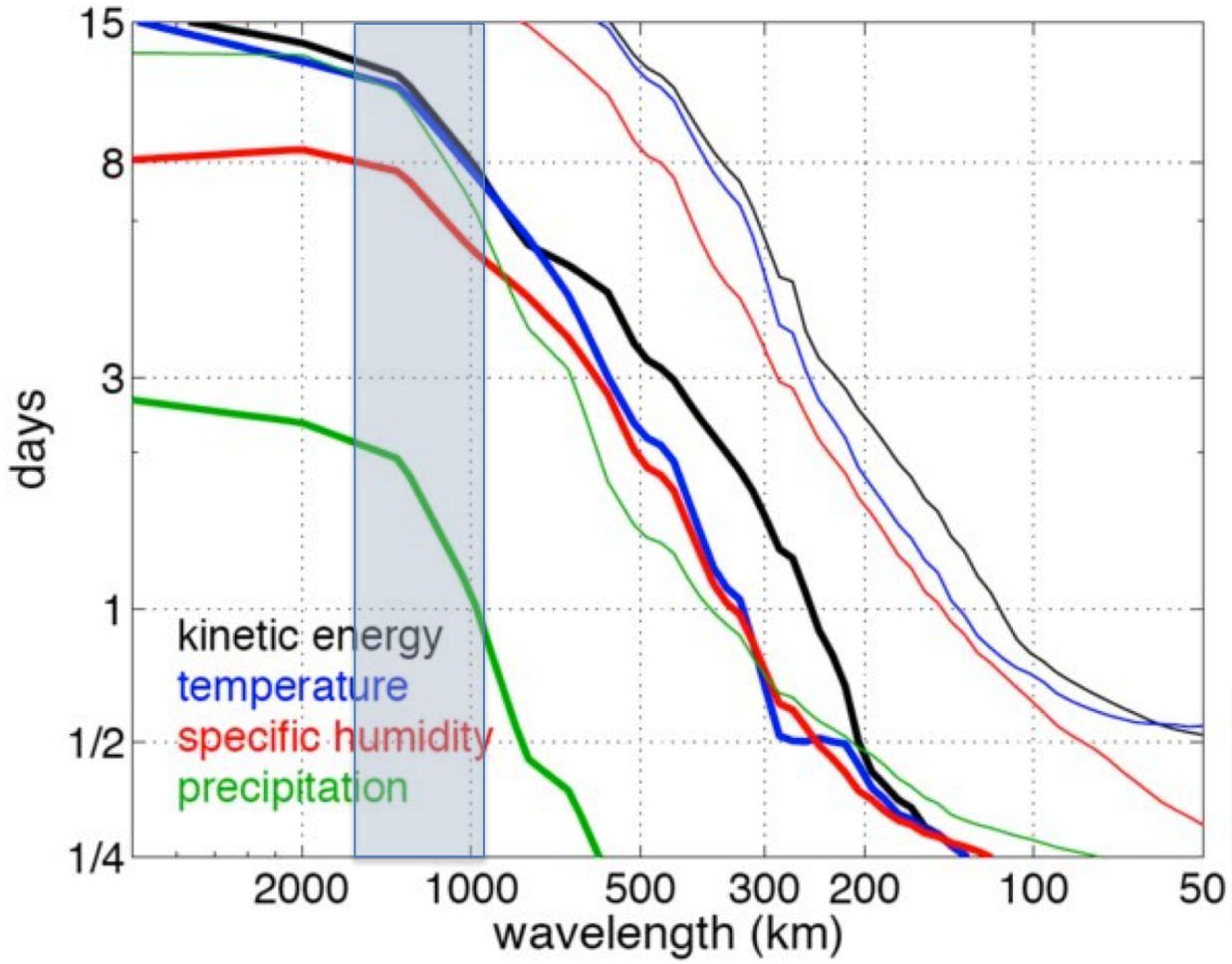






One example of many interesting studies

Ying and Zhang, 2017; JAS - Practical and Intrinsic Predictability of Multiscale Weather and Convectively Coupled Equatorial Waves during the Active Phase of an MJO



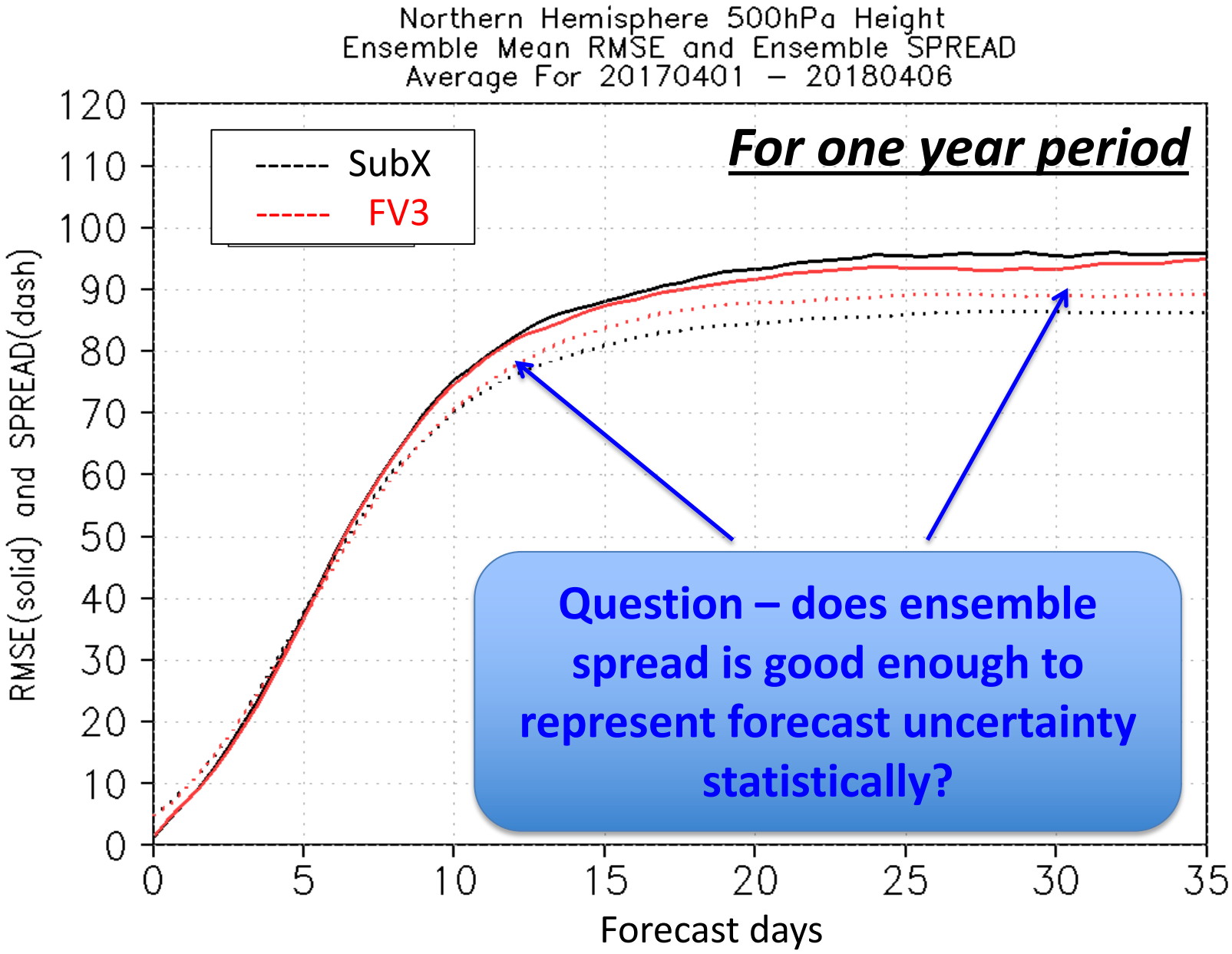
Predictable timescale (days) for kinetic energy, temperature, humidity, and precipitation as a function of horizontal wavenumber (labeled as corresponding wavelength in km). Intrinsic predictability limits are shown in thin lines, and practical predictability limits in thick lines.

Based on all these referred studies – we could explore "predictability" to useful prediction kills

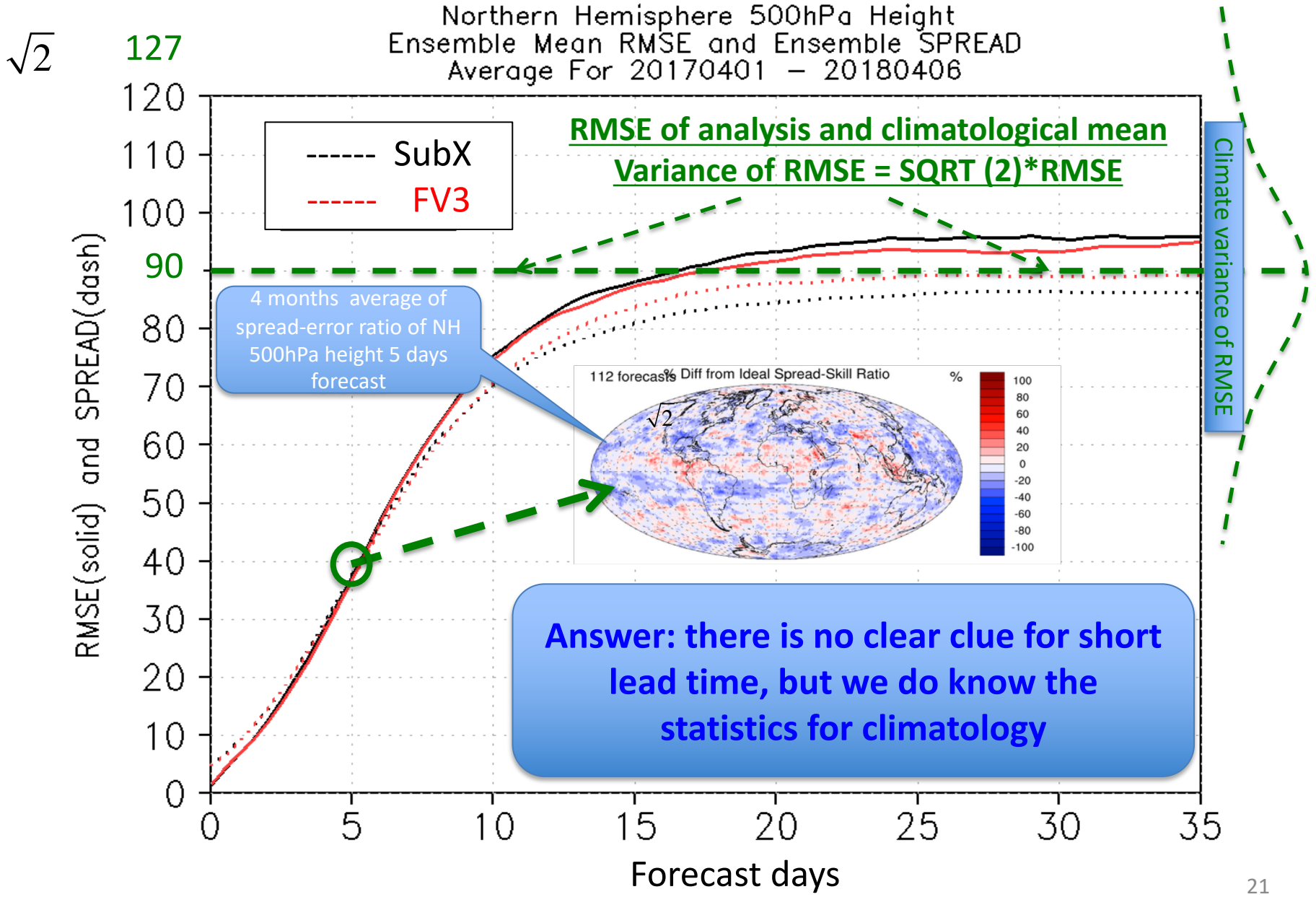
This investigation will focus on

- State-of-art global ensemble forecast system (GEFS)
  - Present initial uncertainty (EnKF) and model uncertainty (SPs)
- Principal assumptions (hypotheses) are
  - Ensemble system is perfect
  - Ensemble spread really represents true forecast uncertainty
  - All individual perturbed forecast could be proxy truth (and equal)
  - Ensemble mean will be best forecast solution for large scale forecast
- Large scale systems (or events) in terms of
  - Spatial resolution
  - Temporal resolution
- Calculation of anomaly correlation in terms of
  - Pattern
  - Time series of domain average
- Prediction skills are based on
  - NH 500hPa geopotential height - PAC
  - Tropical MJO RMM1+RMM2 (850hPa and 200hPa zonal wind and MJO)
- Prediction skills are presented for
  - Useful and true skills for current system
  - Potential useful skills – kind of predictability

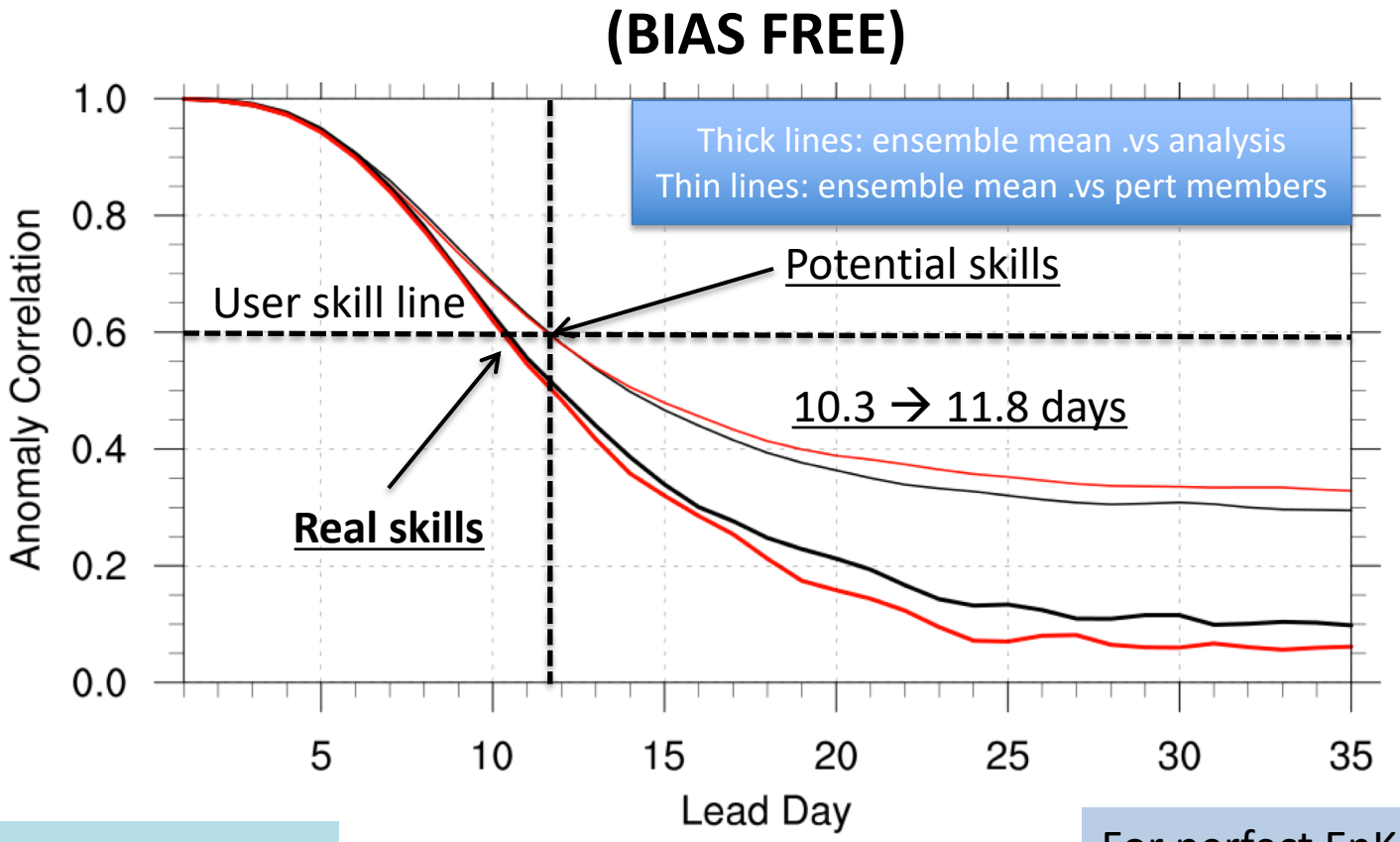
RMSE and Ensemble Spread of NH 500hPa height



RMSE and Ensemble Spread of NH 500hPa height



Over-all prediction and potential prediction skills for NH 500hPa height extra-tropics (day-to-day)



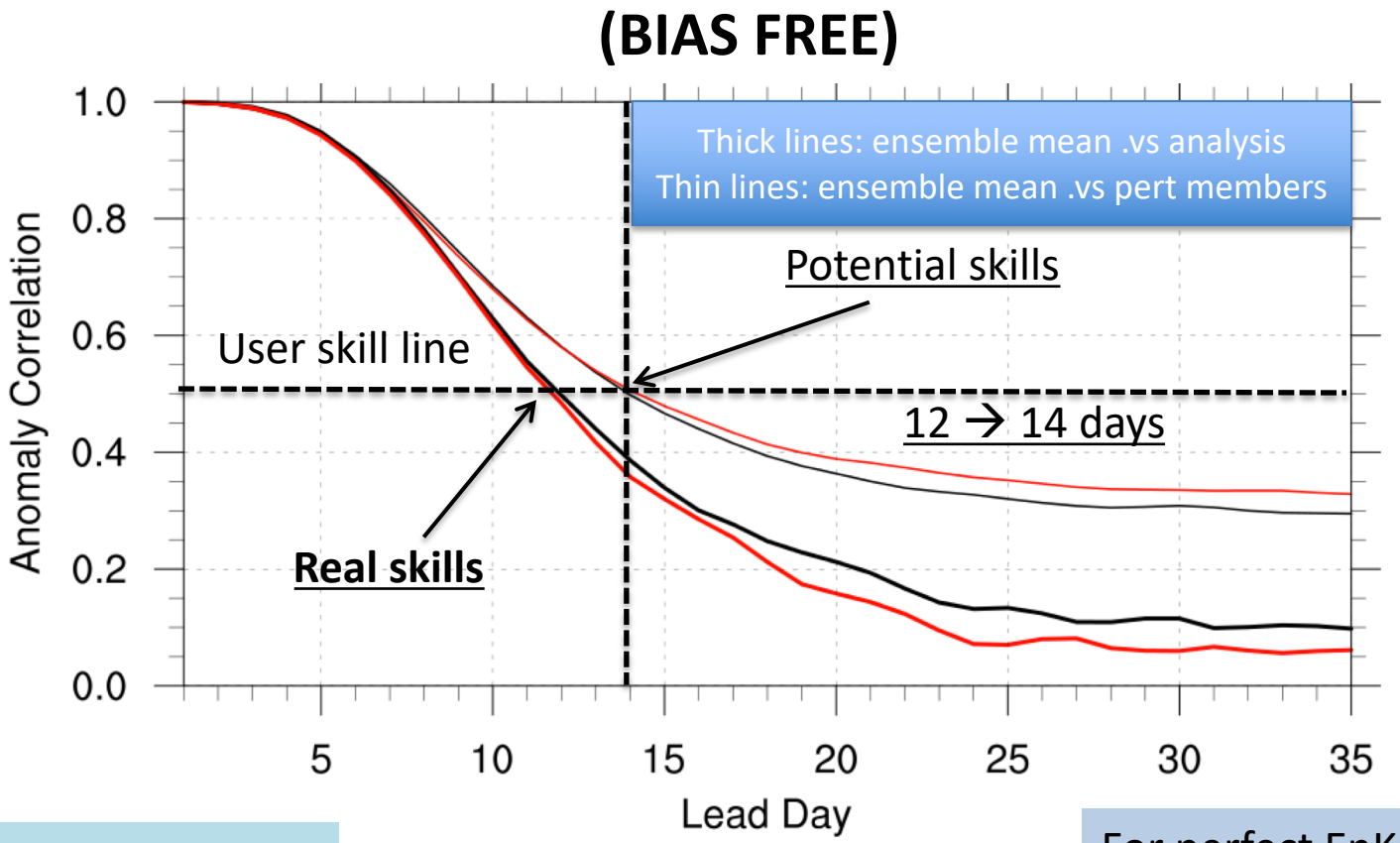
Please note that GEFS has limited ensemble size (21)

SubX

FV3GEFS

For perfect EnKF system, all initial analyses are equal, all forecast should be trues if model is perfect

Over-all prediction and potential prediction skills for NH 500hPa height extra-tropics (day-to-day)



Please note that GEFS has limited ensemble size (21)

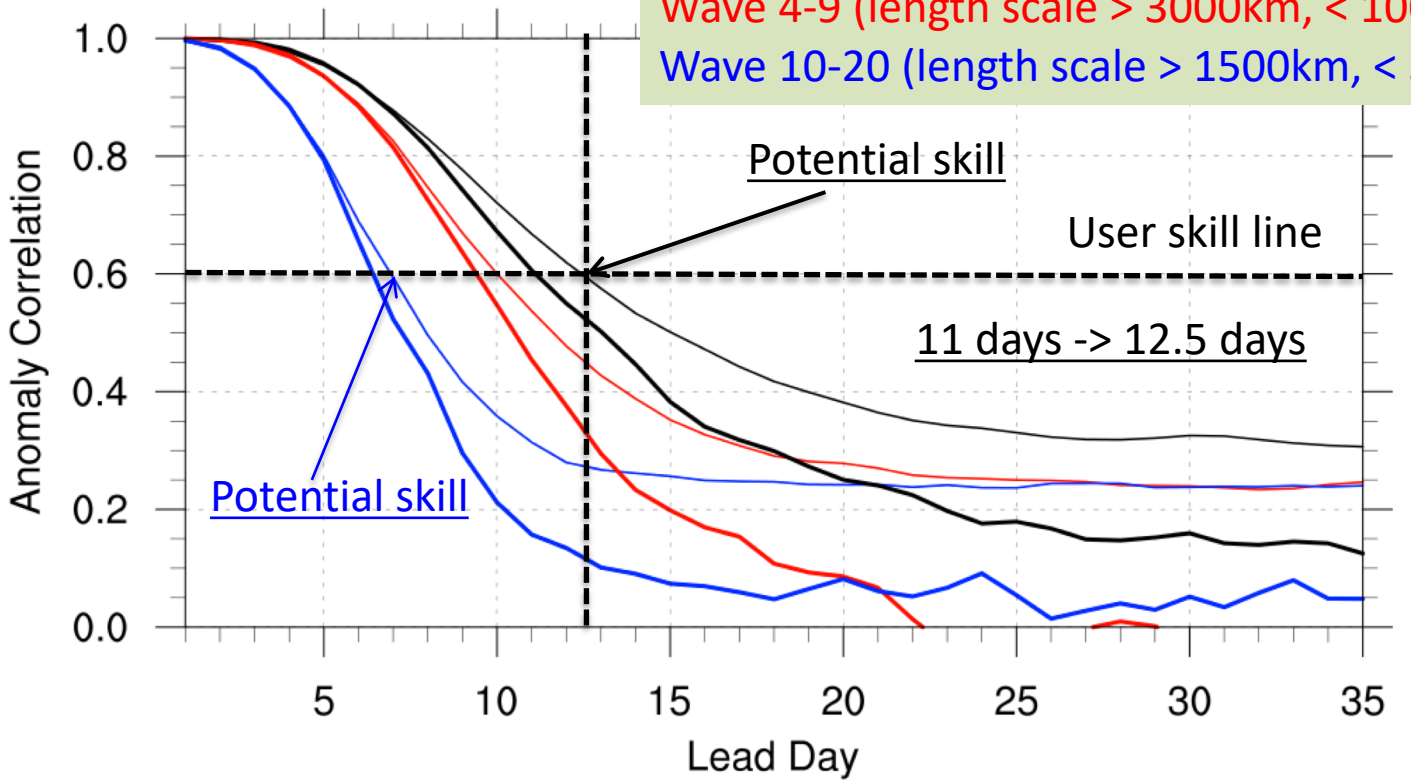
SubX

FV3GEFS

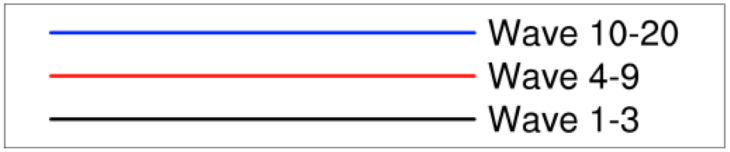
For perfect EnKF system, all initial analyses are equal, all forecast should be trues if model is perfect

Prediction and predictability for NH 500hPa height extra-tropics (diff. scales)

Wave 1-3 (length scale > 10000 km)
Wave 4-9 (length scale > 3000km, < 10000km)
Wave 10-20 (length scale > 1500km, < 3000km)



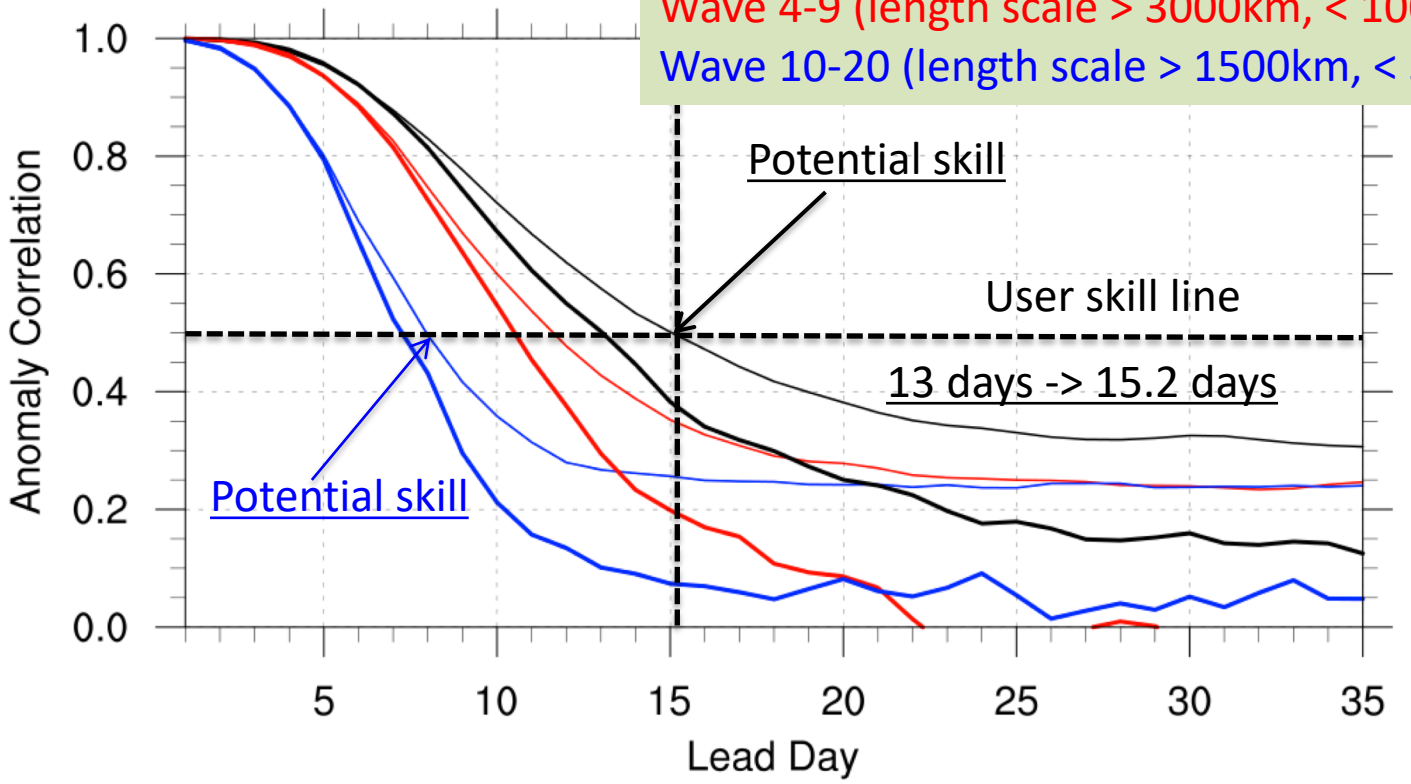
Skills are based on FV3-GEFS version



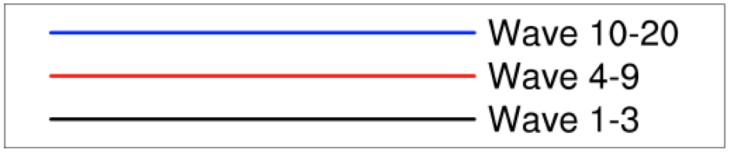
GEFS SubX version has similar skill

Prediction and predictability for NH 500hPa height extra-tropics (diff. scales)

Wave 1-3 (length scale > 10000 km)
Wave 4-9 (length scale > 3000km, < 10000km)
Wave 10-20 (length scale > 1500km, < 3000km)

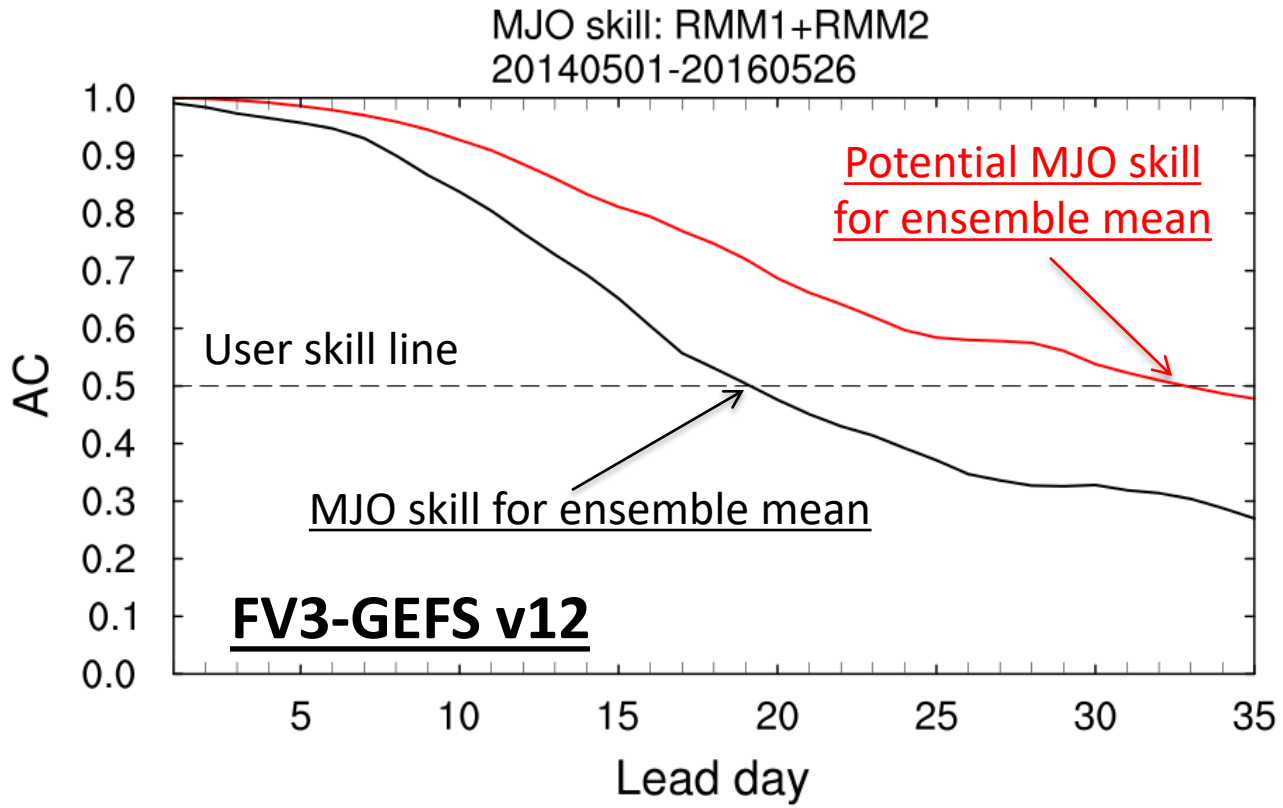


Skills are based on FV3-GEFS version



GEFS SubX version has similar skill

Prediction and predictability of MJO



Discussion: Black line shows the MJO skills from current 1-year FV3-GEFS (v12) experiment (ensemble mean .vs analysis); Red line shows the potential MJO skills from the same 1-year experiment, but uses ensemble mean against ensemble control. We have assumed 1). Ensemble system is perfect; 2). Ensemble mean has best performance of large scale solution (TRUE); 3). Ensemble control forecast is perfect if model is perfect; 4). Ensemble control is independent of ensemble mean (and/or each perturbed forecast). **Q: does this indicate that there is large room for us to improve MJO prediction?**

Summary

Prediction of Current and next GEFS

- Weeks 3&4 average
  - SubX version is better than CFSv2
  - FV3 version is similar to SubX
- Weeks 3&4 bias
  - FV3 version shows much better bias than SubX (not shown)
  - FV3 version has better RPSS skill than SubX for CONUS T2m
- Tropical area
  - SubX version has better MJO skills than CFSv2
  - FV3 version shows better MJO skills than SubX

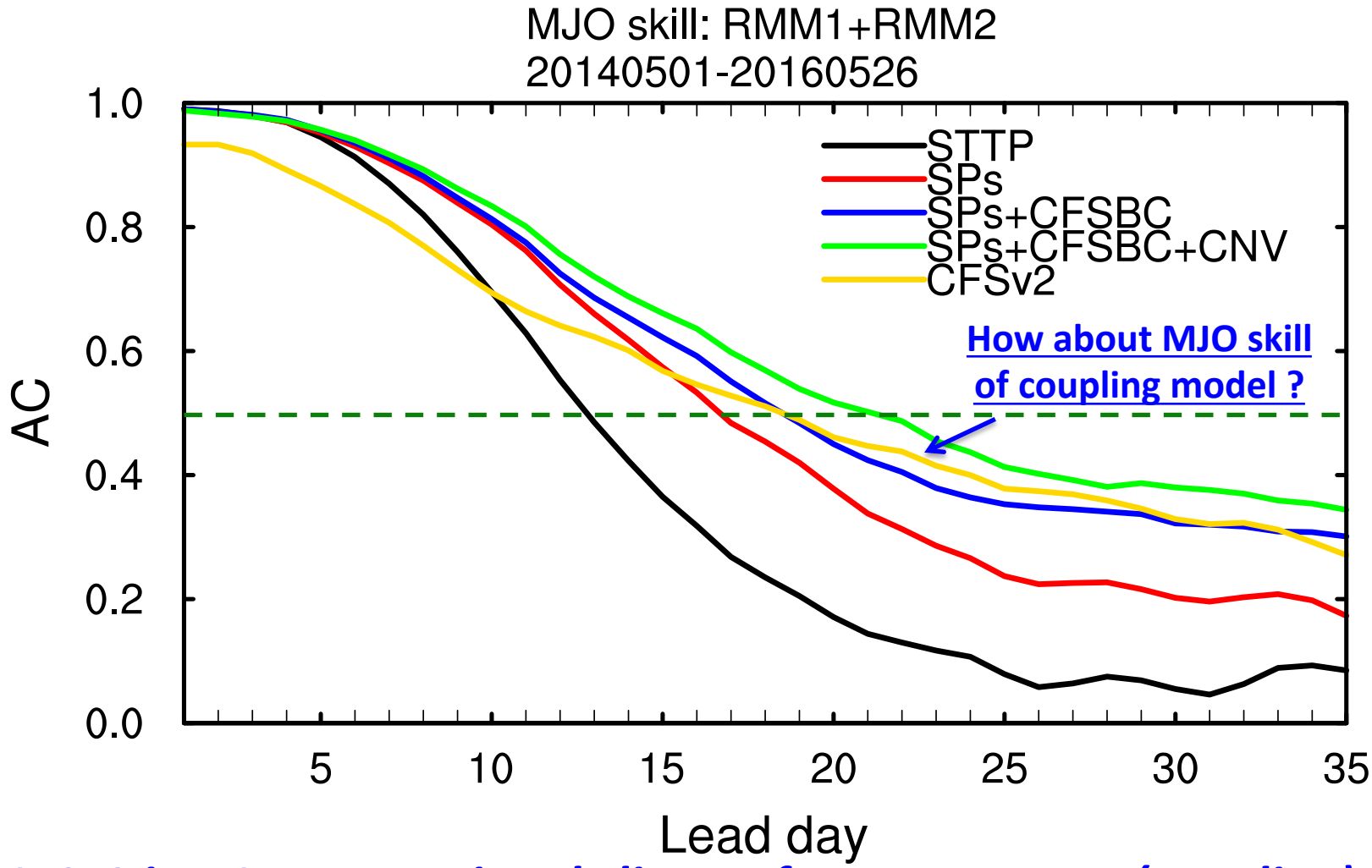
Potential prediction skills

- Overall
  - Potential 2+ days skillful forecast could be added from current prediction through improving our model
- Different scales
  - Planetary scales (>1000km) could be 15+ days skill
  - Large scales (wave 4-9) could be 10+ days
  - Synoptic scales (wave 10-20) could be 8+ days
- Tropical area
  - Large potential to enhance tropical prediction, such as MJO

Will continue to investigate probabilistic forecast skills!!!

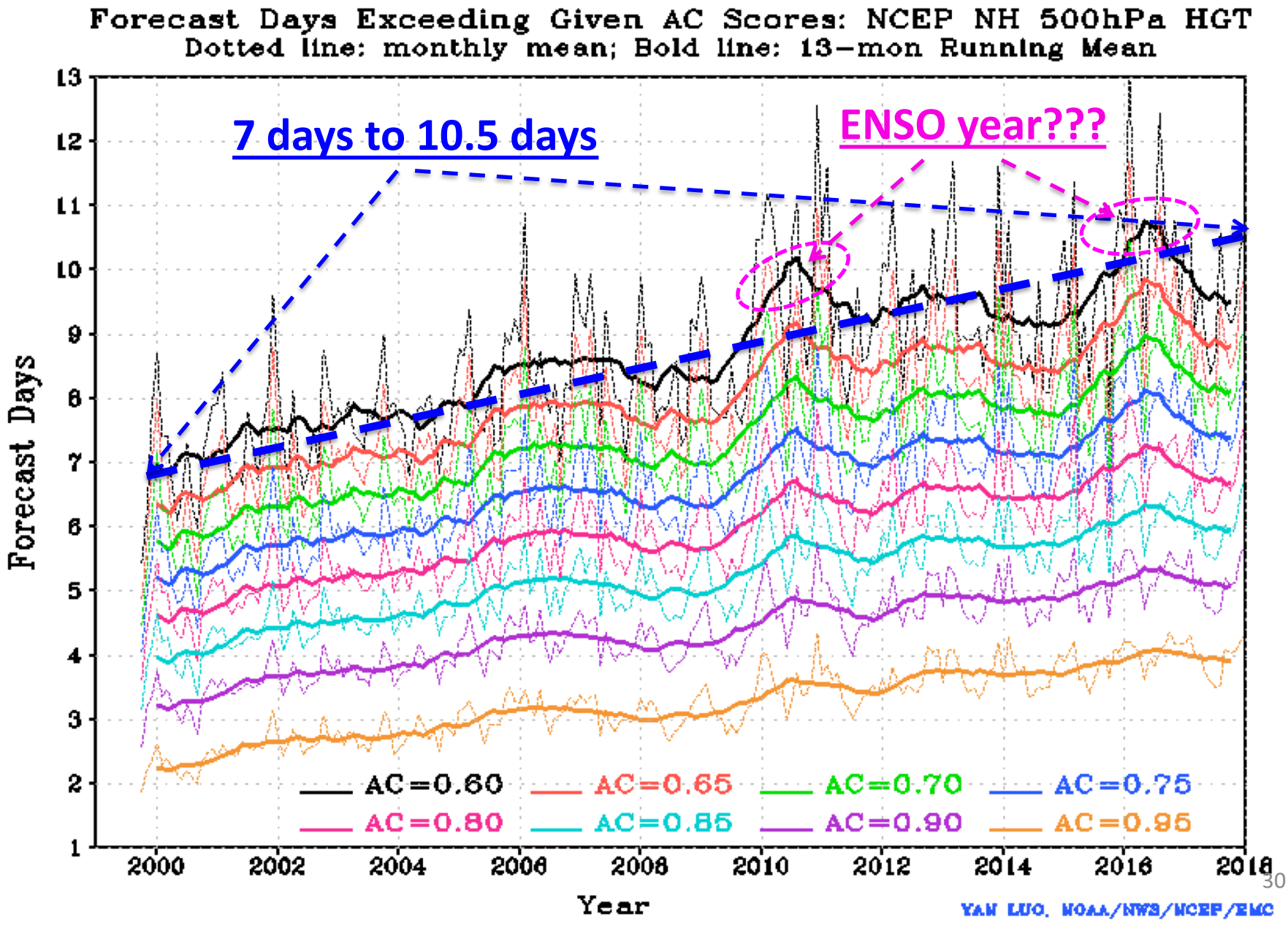
Thanks for your attention!!!

GEFS week 3&4 forecasts (May 2014-May 2016)



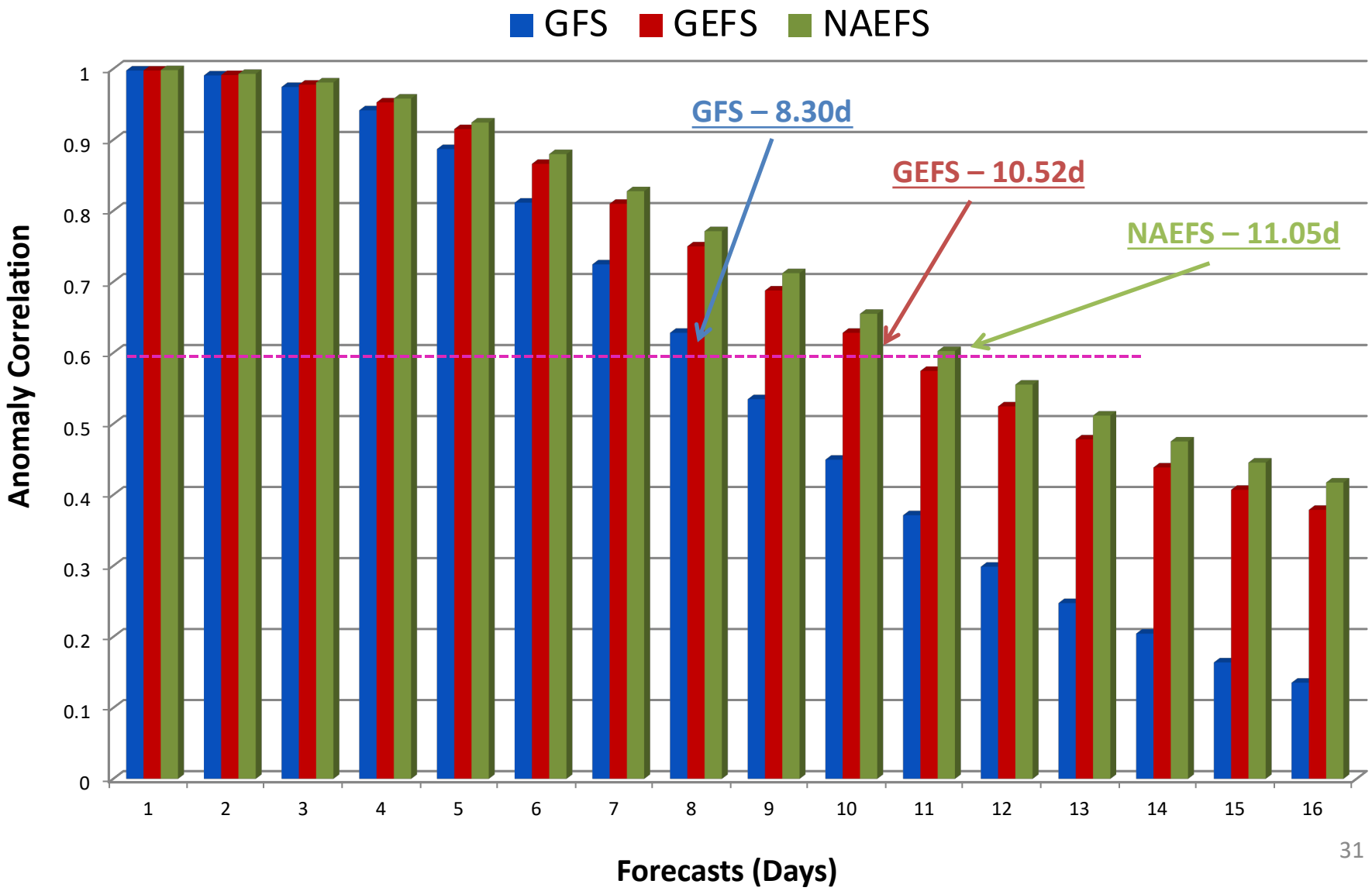
CFSv2 is NCEP operational climate forecast system (coupling) implemented on 2011 – 16 members leg (24 hours) ensemble

AC for NH 500hPa geopotential height



NH Anomaly Correlation for 500hPa Height

Period: January 1st – December 31st 2016

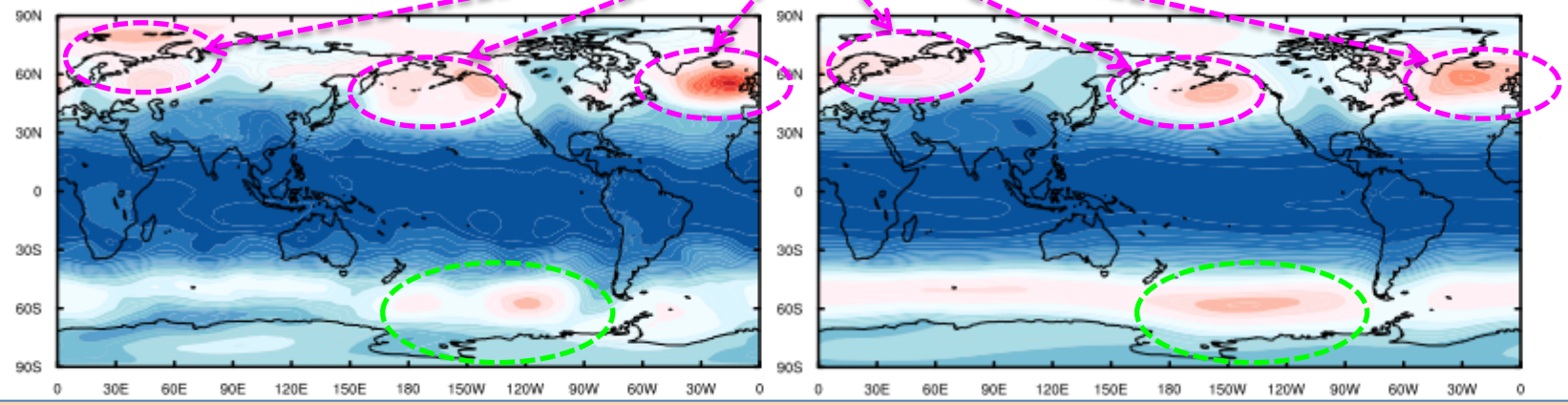


RMS error and spread distributions for 500hPa height of GEFSv11

20130901 – 20140228 (6 months) 5-day running mean
For different lead times

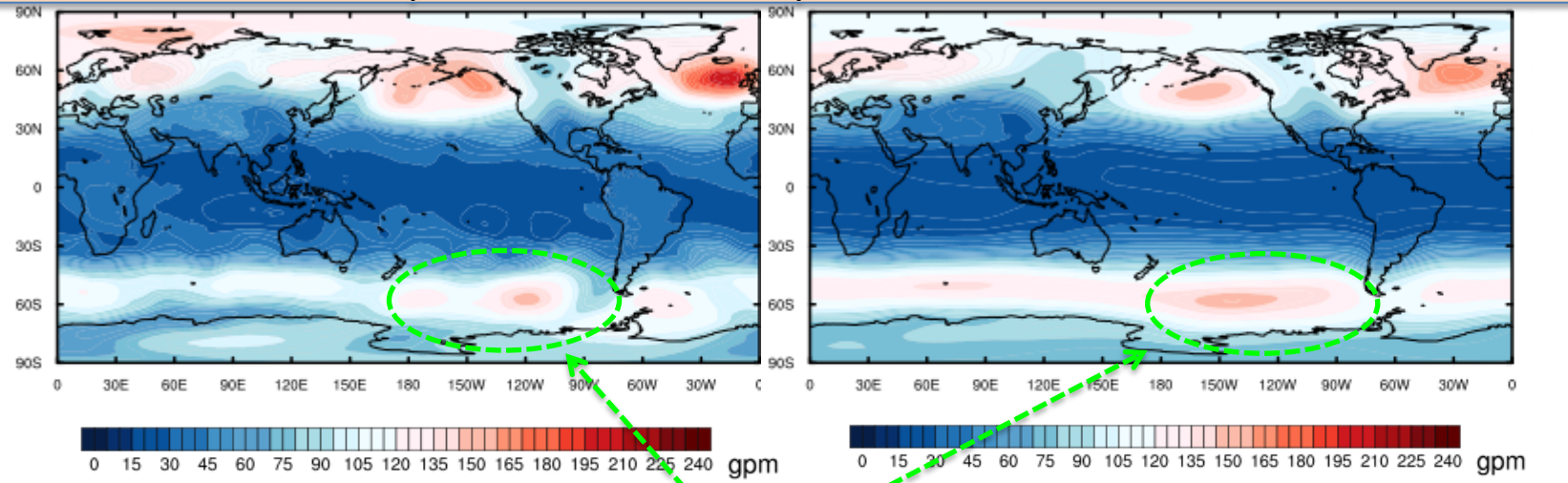
RMSE Uncertainty is Well represented **SPREAD**

Lead day = 19



How about tropical area??? We may need to look at wind field!!!

Lead day = 23



Over dispersion ?