Designing an optimal strategy for GMAO S2S ensemble forecasts

Andrea Molod
NASA/GSFC GMAO
Work of: Anna Borovikov, Siegfried Schubert
And GMAO's Seasonal Prediction Development Group and Software Infrastructure Team

GMAO Seasonal Prediction Group

GMAO Seasonal Prediction group uses coupled Earth-System models and analyses, in conjunction with satellite and *in situ* observations, to study and predict phenomena that evolve on sub/seasonal to decadal timescales. A central motivation for GMAO is the innovative use of NASA satellite data to improve forecast skill

- **Atmosphere/Ocean Coupled Model Development**
- **Ocean Analysis Development**
- **Development of Initialization/Perturbation Strategy for ensembles of Sub/Seasonal Forecasts**
- **Coupled Assimilation Strategy Development**

- **Production of Coupled Data Assimilation (Re)Analysis**
- **Production/Dissemination of Sub/Seasonal Forecasts**

- **Evaluation/Assessment of Forecast Fidelity**
- **Evaluation/Assessment of Assimilated Ocean State**

- **Predictability Studies**

GEOS-S2S-2 was released in November 2017 (Molod et al., 2020)
GEOS-S2S-3 "one-way" weakly coupled reanalysis ("GiOcean") and retrospective forecast suite are ~50% completed, due for release and near-real time use in early 2025

GODDARD EARTH SCIENCES

GEOS-S2S-3 System Characteristics

Model
- AGCM: Recent GMAO NWP (including aerosol model) + two-moment cloud microphysics
- OGCM: **MOM5, ~0.25 deg, 50 levels,**; **Ice Sheet runoff to proper location**
- **New "atmosphere-ocean interface layer" - diurnal warming and cool layer**
- Sea Ice: CICE-4.0
- Forecasts: **initialized from "GiOcean" NRT assimilation, new perturbation/ensemble strategy;**
- Hindcasts: **initialized from "GiOcean" "late look" reanalysis, new perturbation/ensemble strategy;**

Coupled Ocean Data Assimilation System – Coupled Reanalysis "GiOCEAN", "GiOcean-NRT"
- atmosphere is "replayed" to **"GEOS_IT"**; precipitation correction over land, **modified "replay"**
- **Aerosol is "replayed" to analyzed aerosol optical depth**
- Penny et al. (2013) LETKF code/system, set here using **(updated)** static background error statistics;

Observations
- nudging of SST and sea ice fraction from GEOS-IT boundary conditions, **new technique for sea ice;**
- assimilation of *in situ* Tz and Sz including Argo, XBT, CTD, tropical moorings;
- assimilation of satellite along-track ADT (Jason, Saral, ERS, GEOSAT, HY-2A, CryoSat-2);
- sea ice concentration from the National Snow and Ice Data Center (NSIDC).
- **assimilation of SMAP, Aquarius sea surface salinity**

GODDARD EARTH SCIENCES

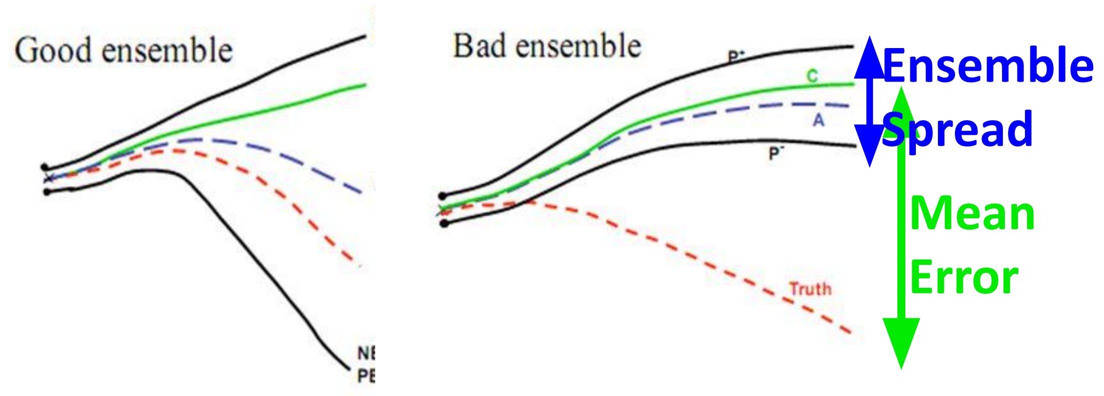
GEOS-S2S-3 Forecast Ensemble Strategy

Motivation for Change in Ensemble Perturbation Strategy:

- GEOS-S2S Tropical Pacific SST was found to be under-dispersive early in the forecast and over-dispersive later (Molod et al., 2020). This prompted the change in the ensemble perturbation strategy.

Evaluate "confidence" by comparing:

- **Ensemble spread (distance among members)**
- **Mean Error (mean of error of individual ensemble members)**



Kalnay, 2003

Following *Barnston et al* (*2015)* we compare the mean intra-ensemble standard deviation (σ) with the standard error of the estimate *(SEE)*:

$$SEE = SDy\,\sqrt{1 - cor_{xy}^2},$$

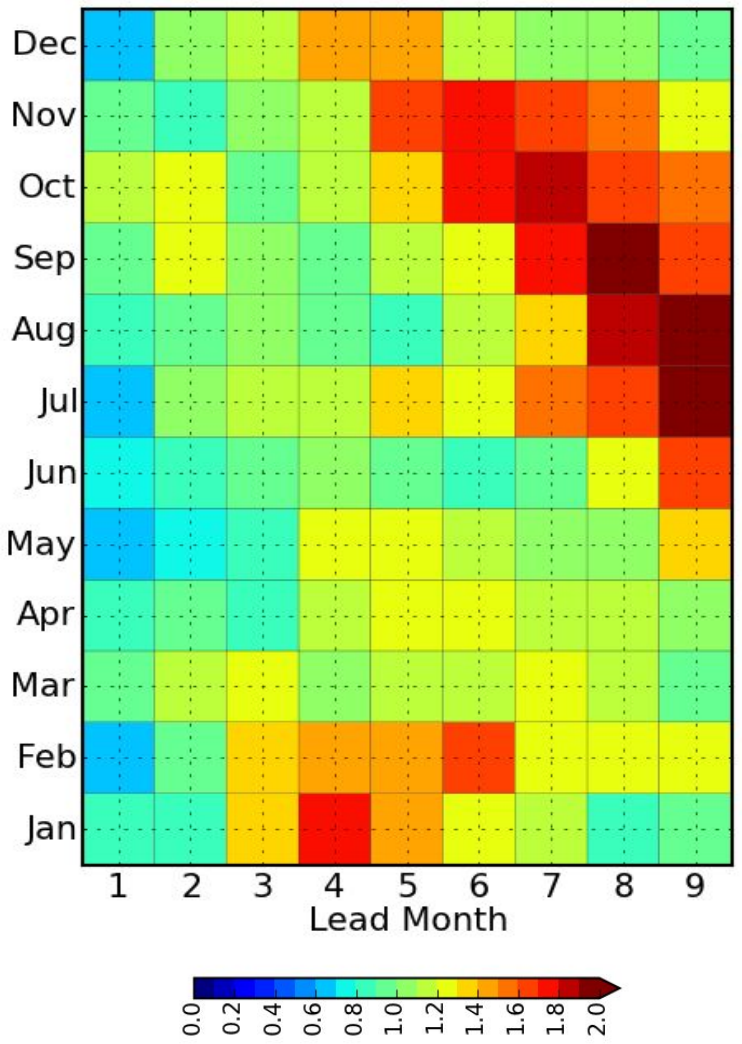
where *SDy* is the standard deviation of the observation *(y)*, and cor_{xy}^{2} is the squared correlation between the ensemble mean forecast *(x)* and the observation.

σ: standard deviation of the intra-ensemble spread

R = σ/SEE, which should be close to 1 for a perfect model

if *R < 1* the model is under dispersive
if *R > 1* model is over dispersive

R = σ/SEE (GEOS-S2S-2)

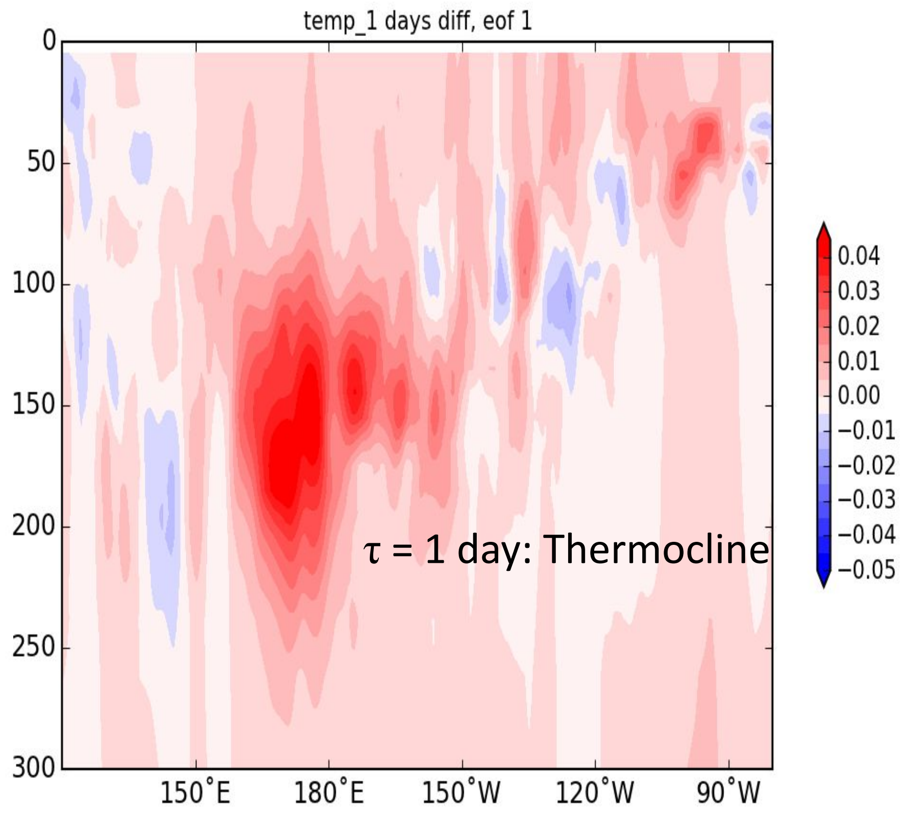


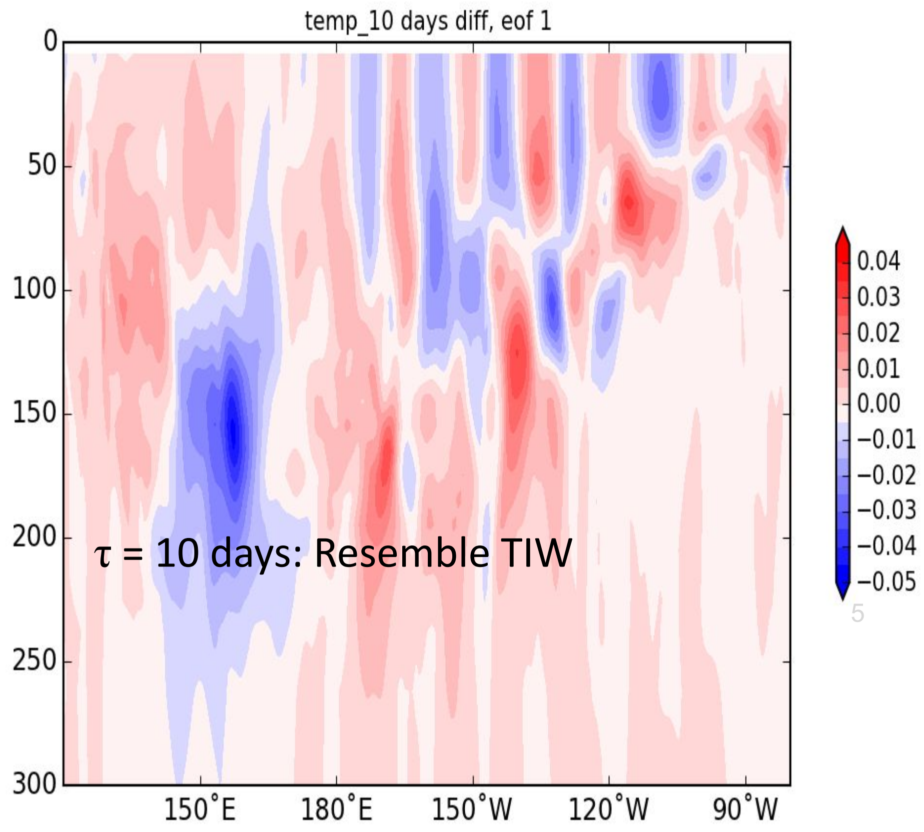
Ensemble Forecasts – Perturbation Strategy

GEOS-S2S-2: Perturbations are scaled differences in AODAS states, 1-day differences for subseasonal forecasts, 5-day differences for seasonal.

GEOS-S2S-3: "Synchronized Multiple Time-lagged (SMT)" - Perturbations for combined forecasts are randomly selected from 1-day through 10-day differences in AODAS states. These spatial structures are closely related to the optimal perturbations that would be obtained from a singular value decomposition of the linear propagator **A** ($\Delta\vec{X}_\tau(t) \equiv \vec{X}(t+\tau) - \vec{X}(t) \approx A_\tau \vec{X}(t)$,) and presumably be sampling preferentially those perturbations with the largest growth rates.

Typical structure of SON ocean temperature perturbations, shown as the leading EOF of the Pacific equatorial x-z cross section of temperature averaged between 2°S-2°N.

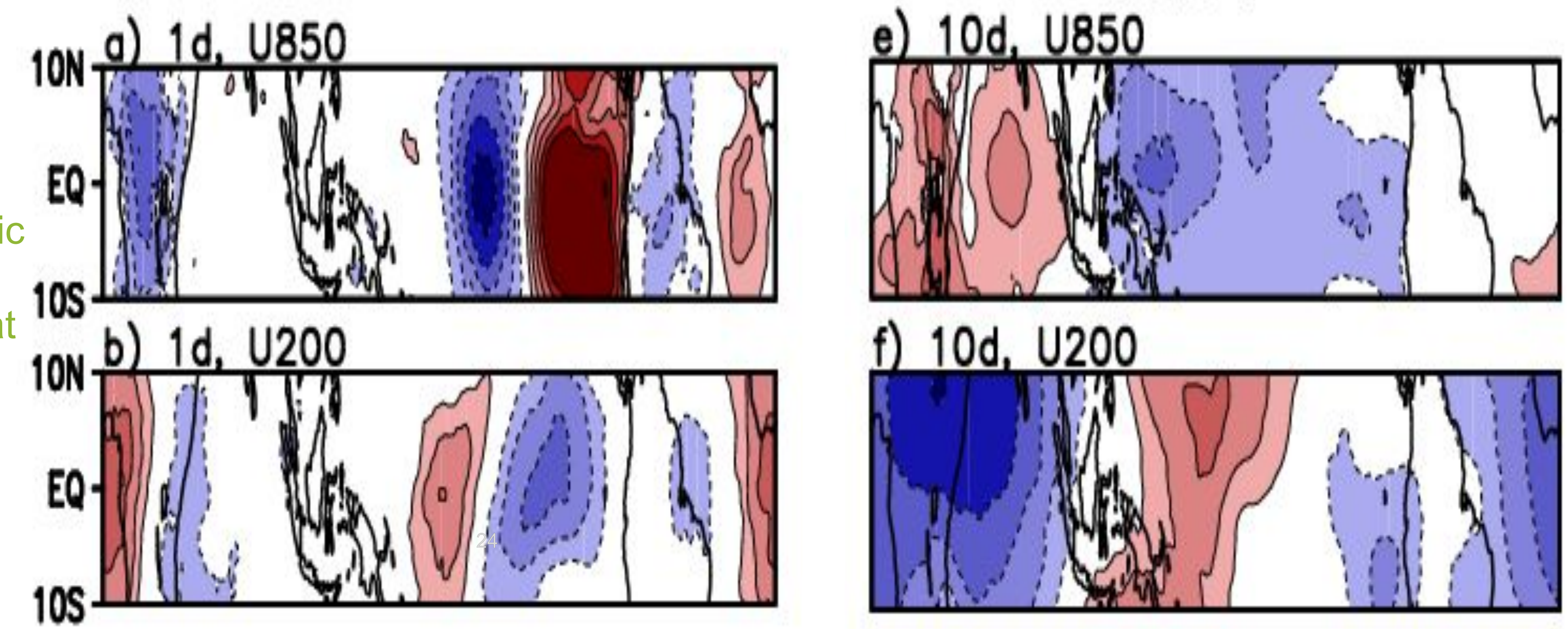




From: Schubert et al., 2019

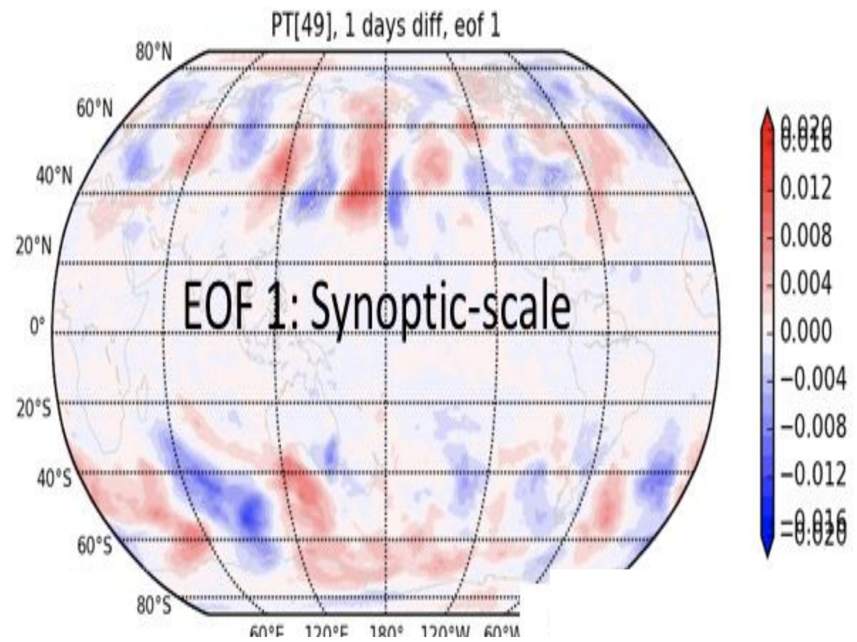
Ensemble Forecasts – Perturbation Strategy

Typical (leading EOF) structure of atmospheric perturbations in the tropics for zonal wind at 850mb and 200mb during DJF.

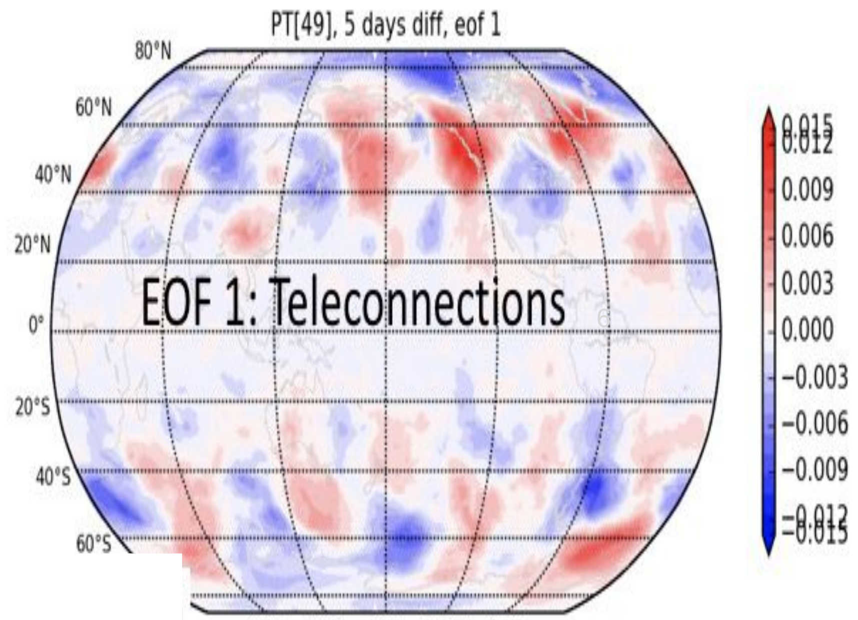


Typical (leading EOF) structures of SON 450 mb Potential Temperature perturbations.

τ = 1 day



τ = 5 day



By varying the separation time between nearby analysis states we are able to generate a wide array of different types of atmospheric and oceanic perturbations that represent physically realistic and important modes of variability.

From: Schubert et al., 2019

GODDARD EARTH SCIENCES

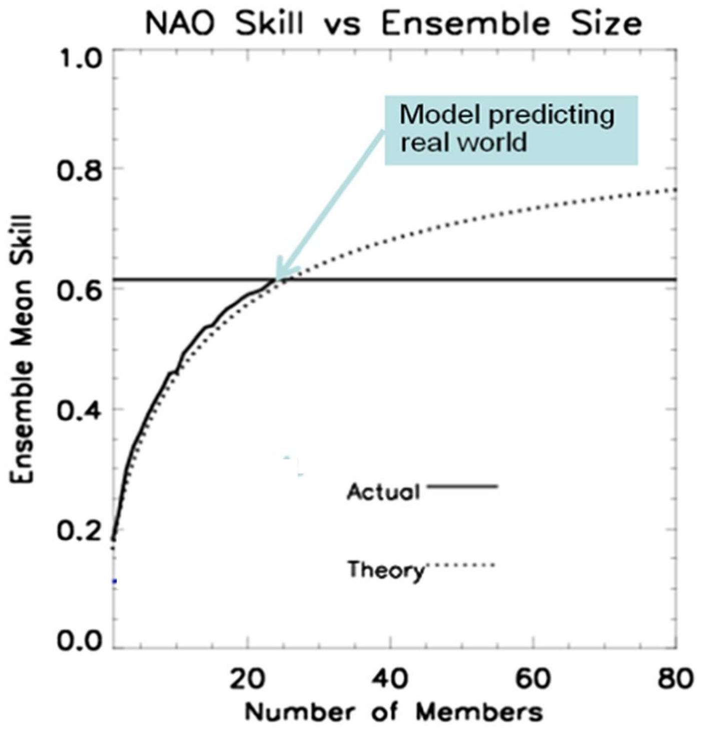
Ensemble Forecasts – Ensemble Size

Motivation for Change in Ensemble Size:

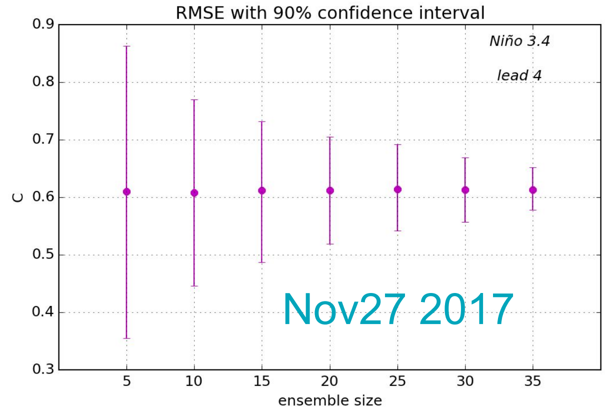
- Extratropical skill was lower than the best state-of-the-art systems because of the small ensemble size (eg., Scaife et al., 2018). This prompted the change in ensemble size and the new approach to the number of ensembles.

- Little evidence of additional skill from ensemble size beyond a few months. This prompted a sub-sampling strategy for extending selected ensemble members

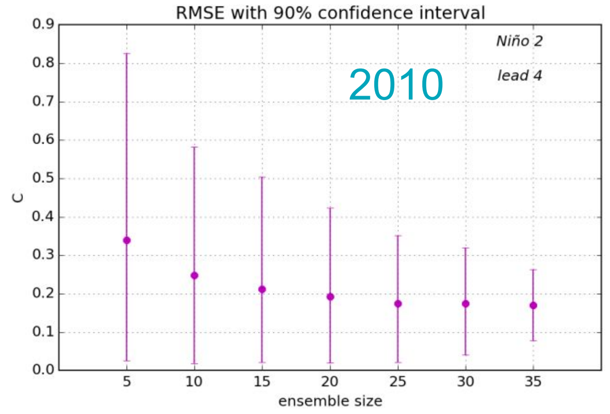
NAO at 1 month lead
(Scaife & Smith, 2018)



ENSO at 4 months lead



Most often – error remains steady with larger ensemble size



Rare – error continues to drop with larger ensemble size

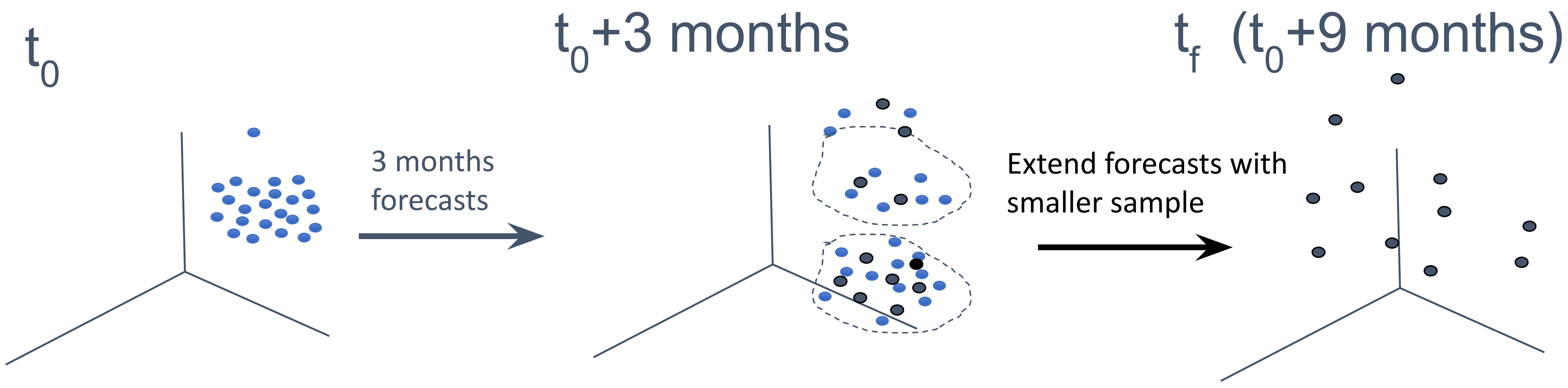
We need lots of ensemble members for short lead, but not for long leads.

Analysis performed by: Anna Borovikov

GODDARD EARTH SCIENCES

Ensemble Forecasts – Ensemble Size

Forecast Ensemble Strategy – Sub-sampling



Initial ensemble size should be as large as possible for carrying out three month forecasts

After the first 3 months most atmospheric teleconnections and land impacts have little skill tied to initial conditions.

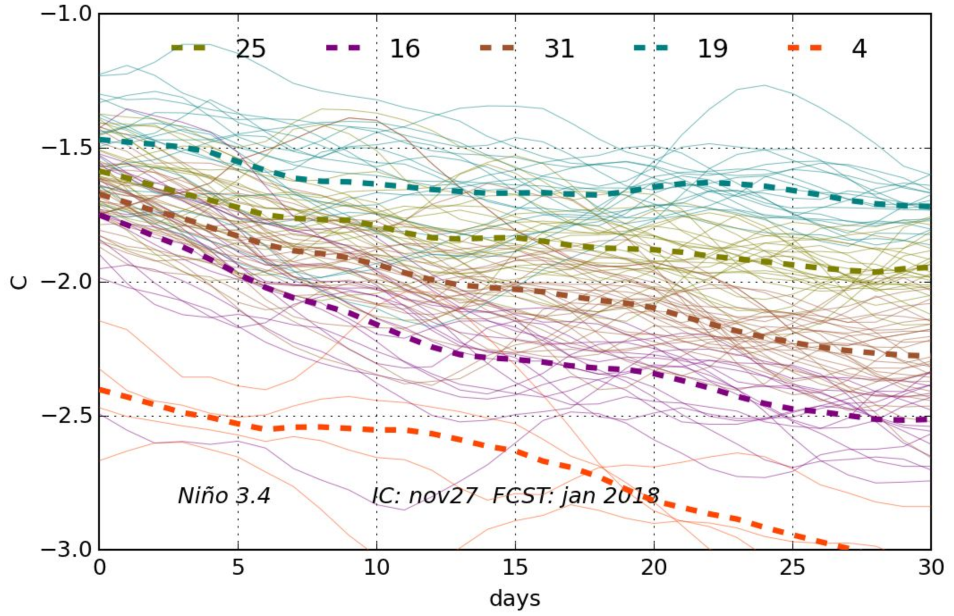
Stratify based on emerging directions of error growth using K-means clustering, and allocate proportionally within each stratum

This is done based on the Nino3.4 index since it is largely ENSO that matters after the first 3 months.

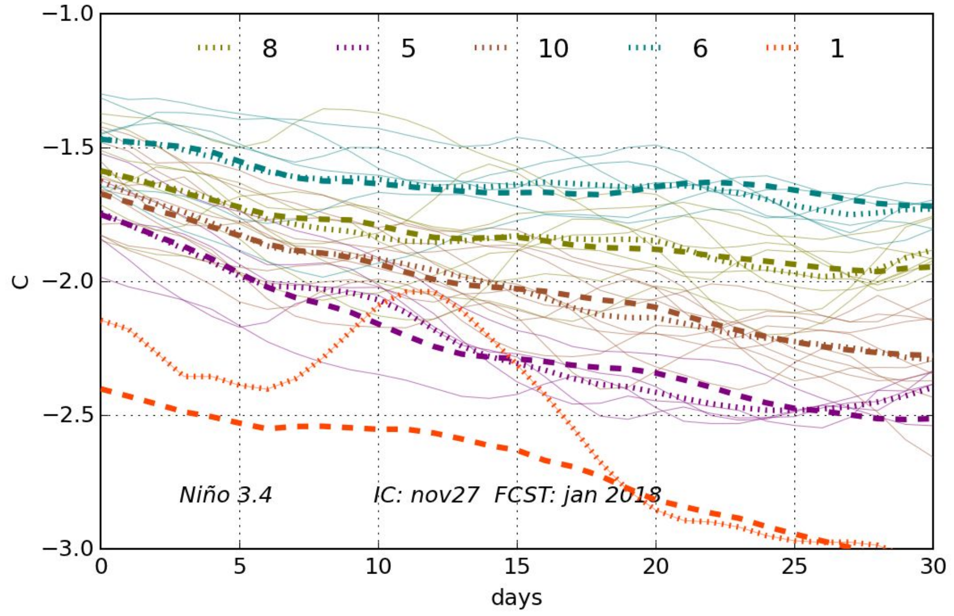
Final long-lead forecasts with the small subset of the initial larger ensemble

From: Schubert et al., 2019

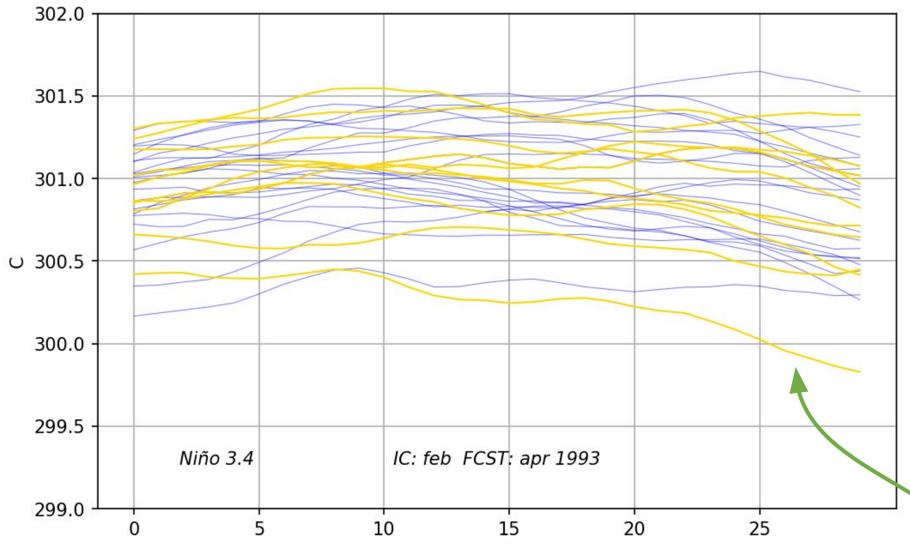
Ensemble Forecasts – Ensemble Size

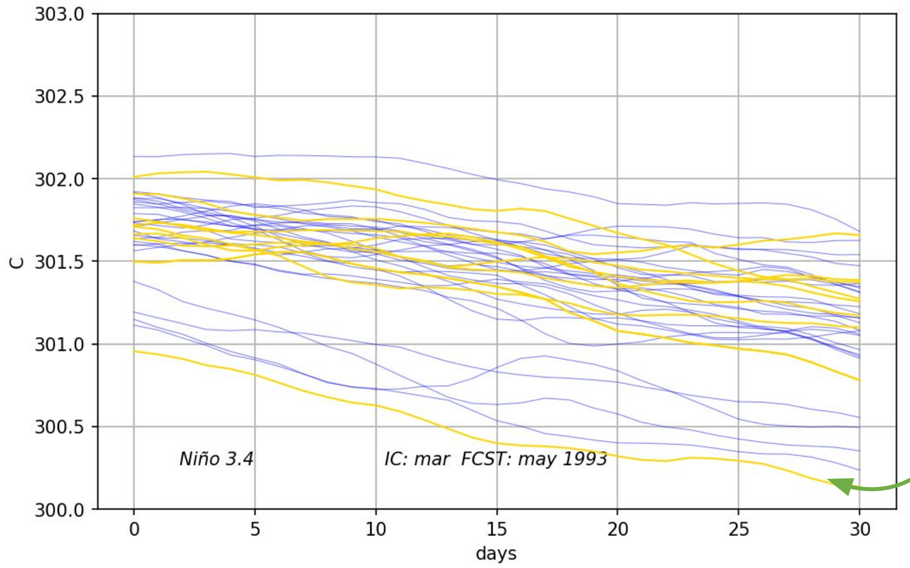


EXAMPLE:
original
clusters
and means



EXAMPLE:
sub-sampled
clusters
and means
(dotted lines)
original means
(dashed lines)

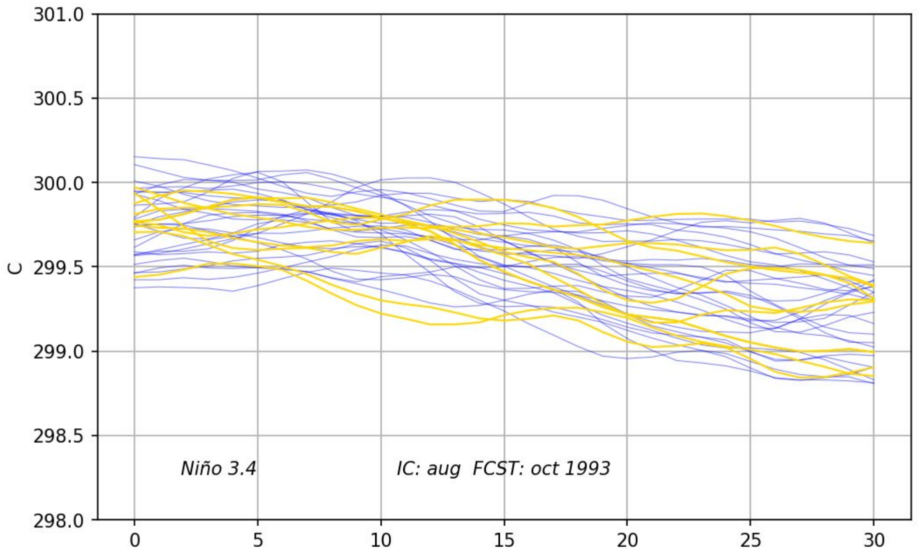


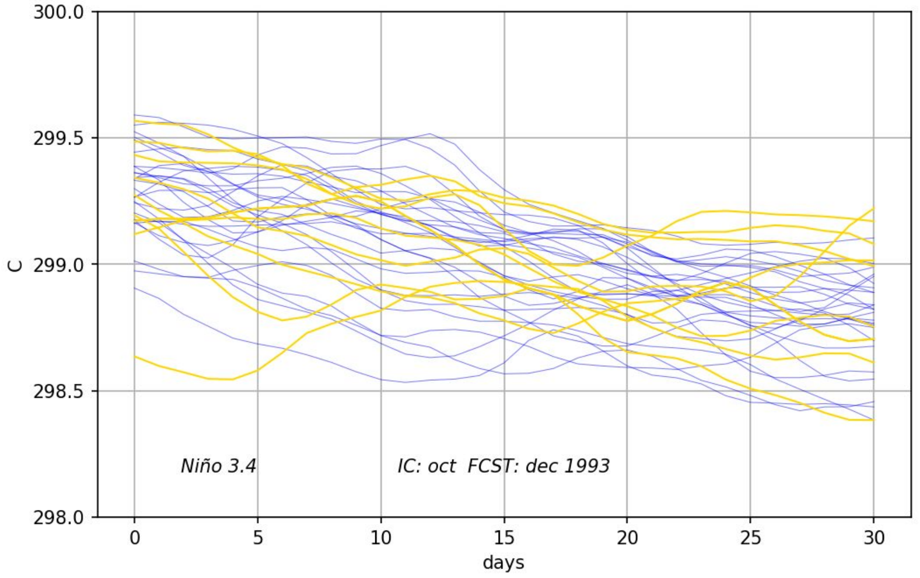


Examples of the routine diagnostic plots.

"outlier" members picked up

tight or wide spread reflected well

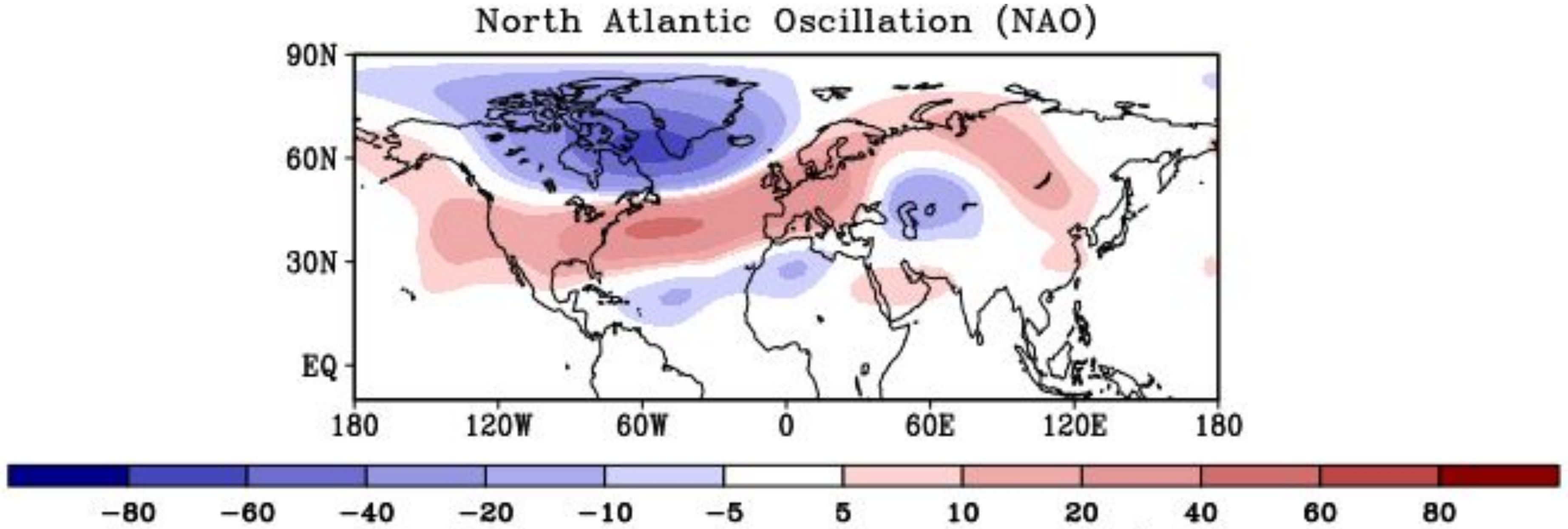




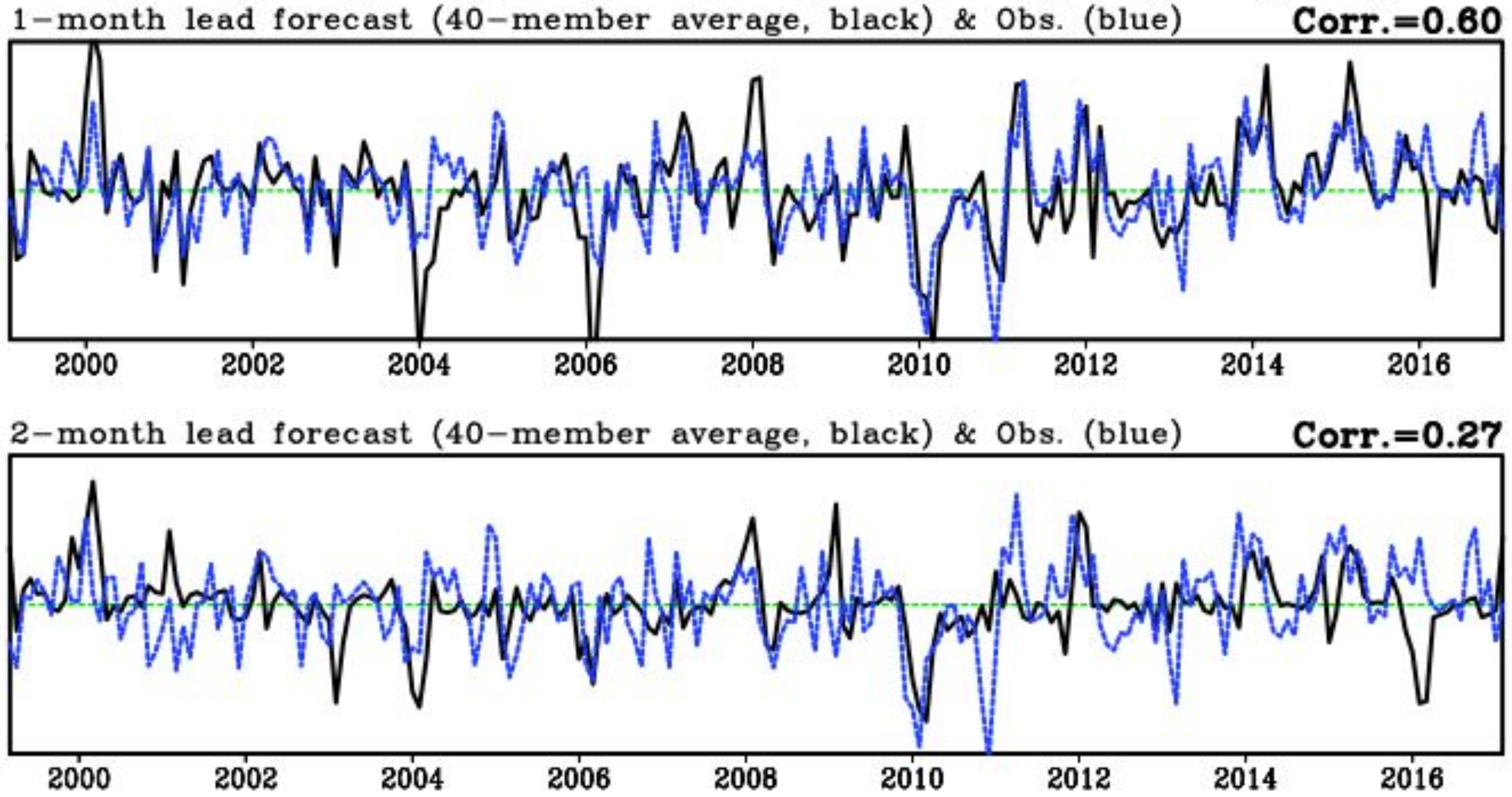
GEOS-S2S-3 Near Real-Time Sub/Seasonal Prediction Suite

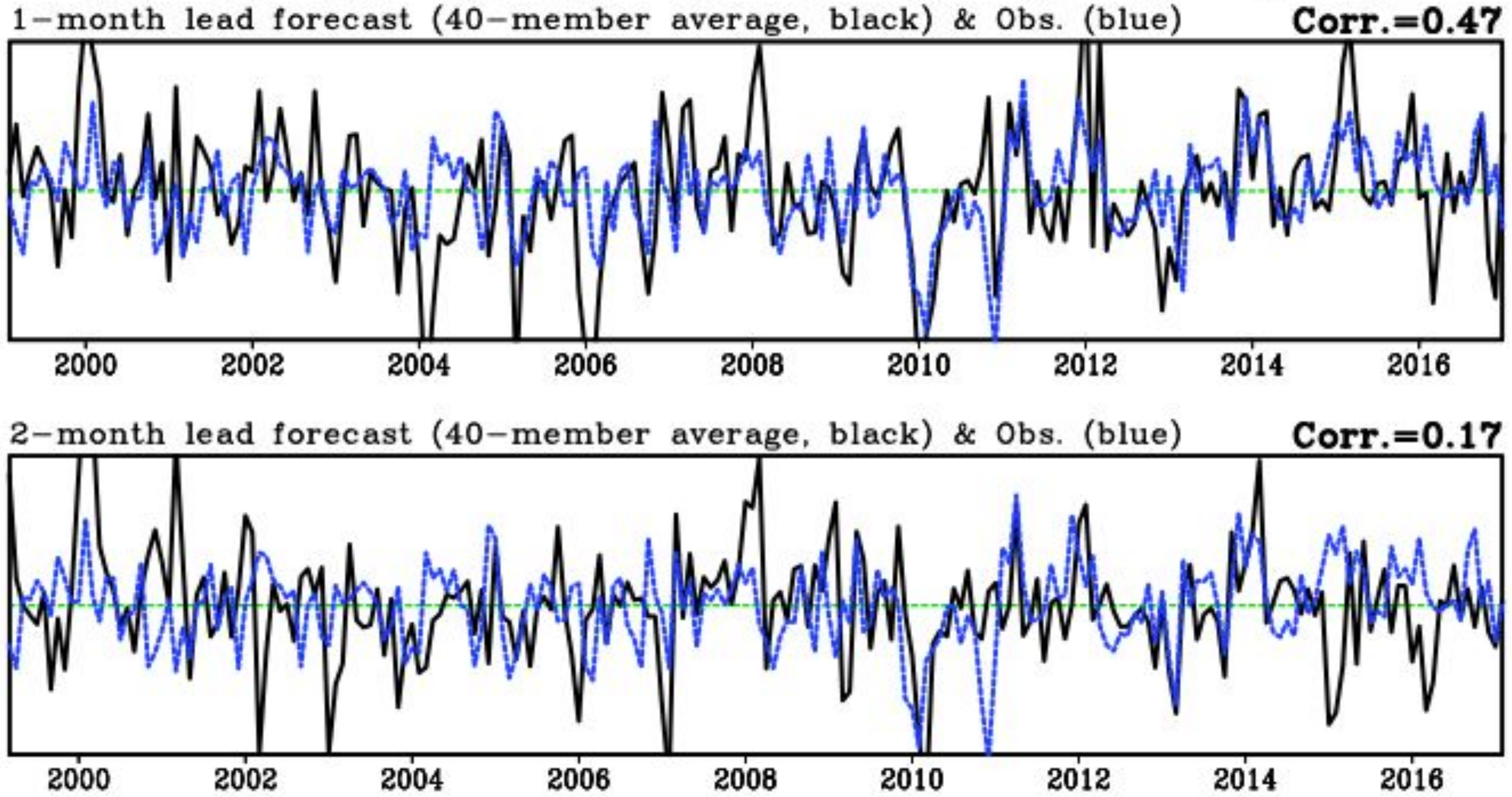
	Sub/Seasonal
Length of Forecast	9 months
Frequency of forecasts	Every 5 days
Number of Ensembles	40 member lag/burst for first three months, selection of 10 members for remaining 6 months
Frequency of submission	Once per week OR once per month (as needed)
Retrospective Initial Conditions from	“GiOcean – Late Look” GEOS-S2S-3 AODAS
Retrospective Forecasts	1982-2022
Near-real time Initial Conditions from	“GiOcean – First Look” GEOS-S2S-3 AODAS

GEOS-S2S-3: Forecast Evaluation - NAO



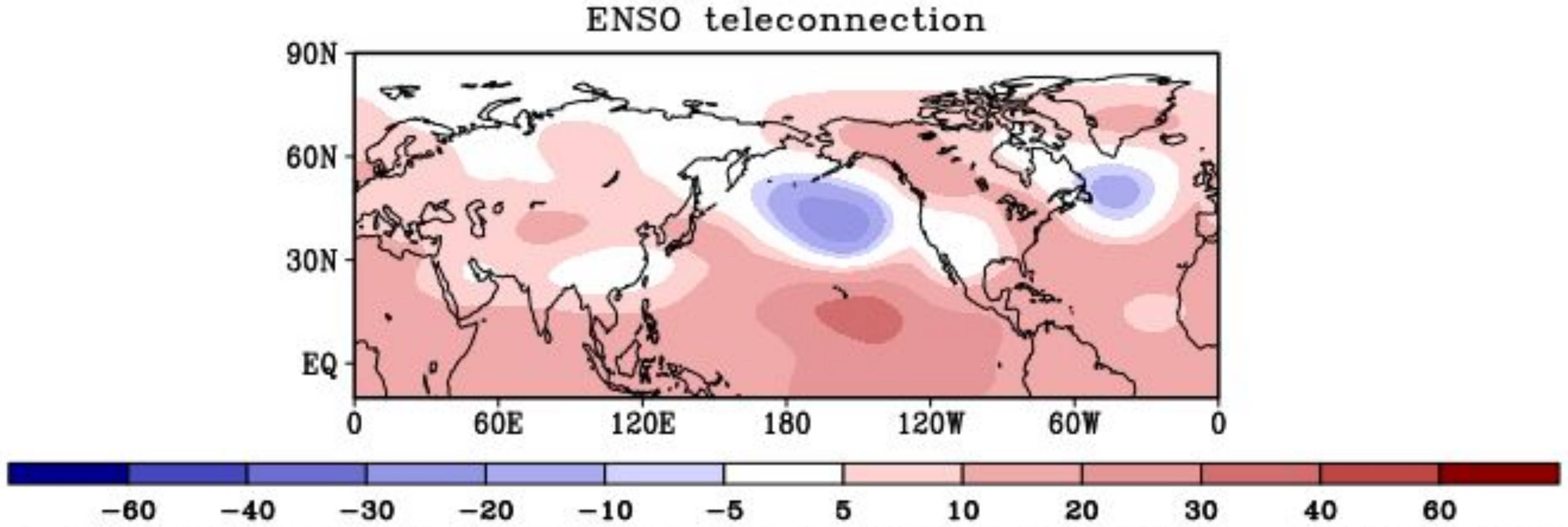
Analysis of Young-Kwon Lim



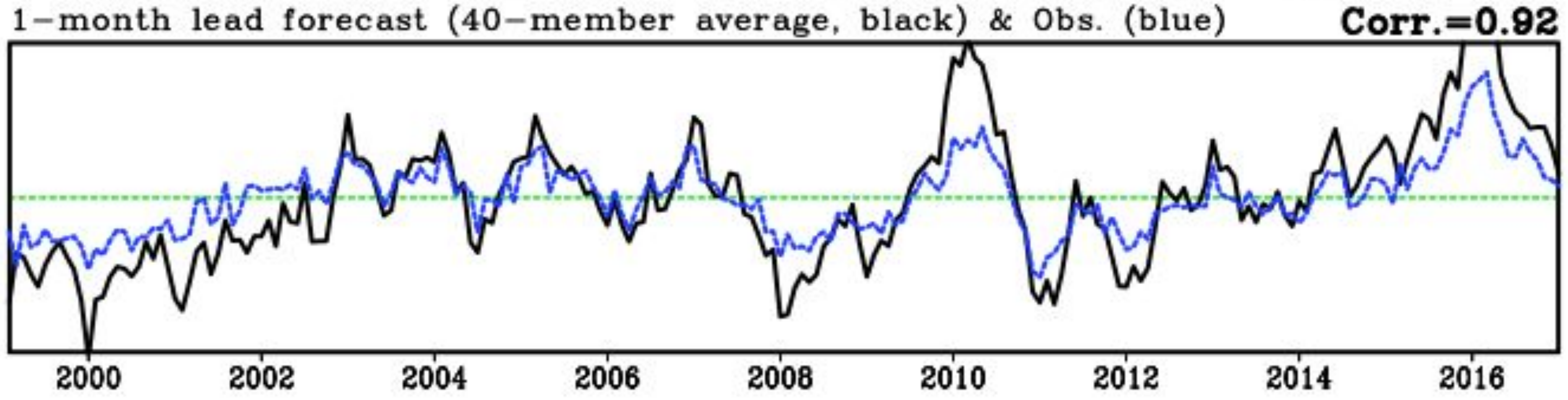


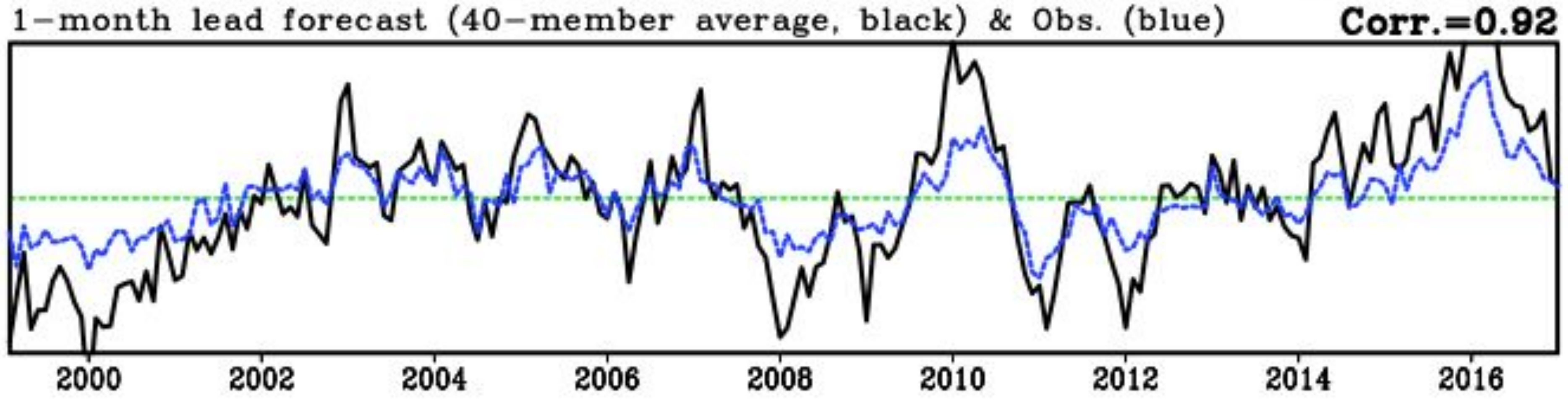
GODDARD EARTH SCIENCES

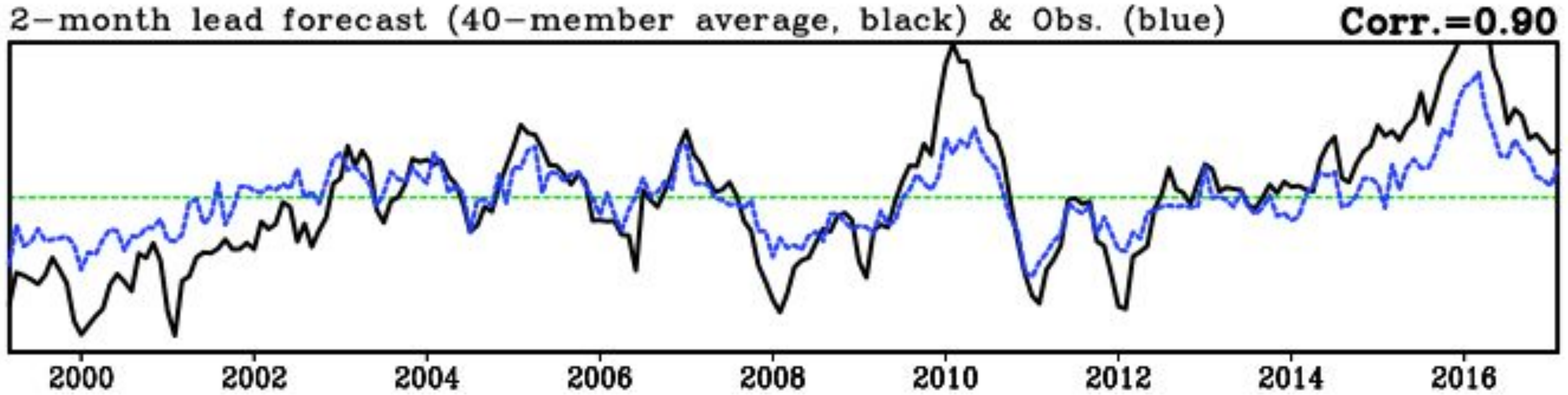
GEOS-S2S-3: Forecast Evaluation – ENSO Teleconnection

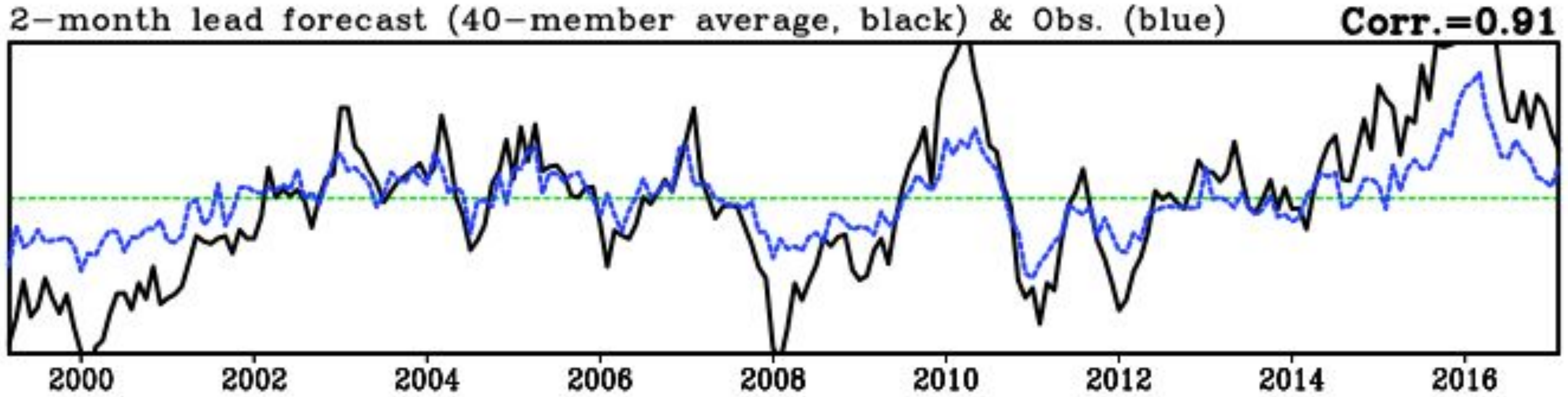


Analysis of Young-Kwon Lim









GODDARD EARTH SCIENCES

Summary

- GEOS-S2S-3 Ensemble Perturbation Strategy now approximates perturbations of the most important and unstable structures – this strategy has improved forecast confidence

- GEOS-S2S-3 Ensemble Size has expanded to 40 ensemble members for the first 3 months of lead time and extends 10 of those to 9 months. This strategy has improved NAO forecast skill as well as other atmospheric modes.

Schubert Siegfried, Anna Borovikov, Young-Kwon Lim, and Andrea Molod, 2019. **Ensemble Generation Strategies Employed in the GMAO GEOS-S2S Forecast System**. *NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2019-104606*, Vol. **53**, 75 pp.