Developing Physics-oriented Diagnostic Tools for Model Evaluation and Improvement

PI: Zhuo Wang,
Department of Atmospheric Sciences
University of Illinois at Urbana-Champaign

Co-PIs: Stan Benjamin (NOAA ESRL)

Melinda Peng (NRL Monterey)

Ming Zhao (NOAA GFDL)

Model Evaluation

- Evaluation of model forecasts is an indispensible component of the model development effort.
- Standard performance-oriented metrics, such as ACC and RMSE, are routinely used in operational centers for forecast verification. NCAR DTC also developed model evaluation tools (MET).
- Physics-oriented diagnostics focus on the critical processes or phenomena and, through evaluation of key variables, shed light on the deficiencies of the physical parameterization or other errors in a model.
- In brevity, physics-oriented evaluation not only provides information on how well a model performs but also on why a model may fail in a certain aspect.

Over-arching Objective

- **Objective**: develop physics-oriented diagnostic tools to assist the development of the NOAA's Next Generation Global Prediction System (NGGPS) under the R2O Initiative
- **Expected outcome**: a suite of diagnostic tools with general applicability across models
- In-depth evaluation of model forecasts using physics-oriented metrics helps to "*improve the physical parameterizations to allow for efficient, accurate and more complete representations of physical processes and their interactions across scales*" and helps to "*achieve a world class global predication system*", which is highly relevant to the priority of the R2O initiative.

Two Foci

1. **Prominent Motion Systems**
   - Tropical Cyclones
   - MJO
   - Blocking
   - Teleconnections

   - **Level 1:** Performance-oriented evaluation
   - **Level 2:** Physics-oriented evaluation
   - **Level 3:** Evaluation of predictability

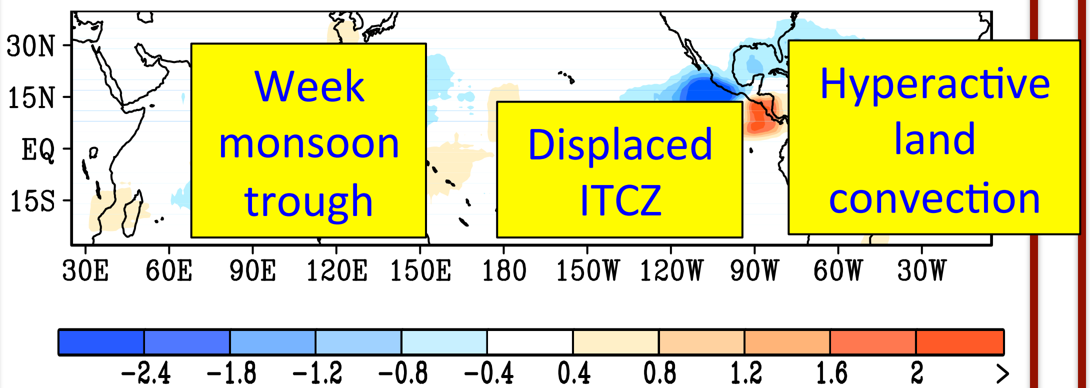
2. **Specific Physical Processes**
   - Moist convection: cumulus parameterization, CWV-precip relationship, Q1/Q2 diagnoses
   - Model representation of different cloud regimes: evaluation of the model clouds using satellite simulators

Outline of the Report

- Major Accomplishments in FY16
  - Evaluation of Tropical Cyclones and the MJO in GEFS (a brief summary)
  - Predictability of Tropical Cyclones at Different Time Scales
- Priority Focus for FY17
- Transition to EMC

Physics-oriented Evaluation to Link Tropical Cyclogenesis Biases to the Model Physics: An Example

Biases of Tropical Cyclogenesis in GEFS

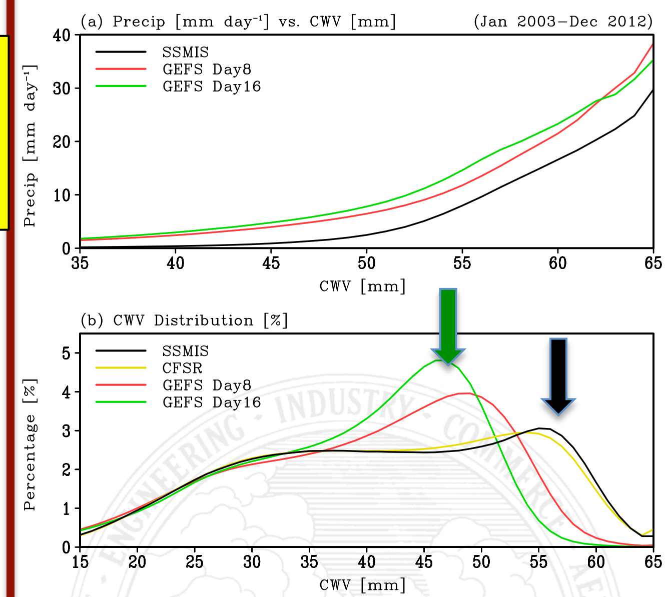


Large biases in tropical cyclogenesis exist on the regional scale.

TC genesis biases are related to biases in diff. large-scale circulations over diff. basins

- a weaker-than-observed monsoon trough over the West Pacific
- a southward displaced ITCZ over the East Pacific
- hyperactive Africa easterly waves over the East Atlantic

Evaluation Against Satellite Data



a dry bias in column water vapor (CWV) and **precipitation is initiated too early with respect to CWV accumulation** in the GEFS → **Deficiencies in the cumulus scheme**.

Why are we interested in predictability?

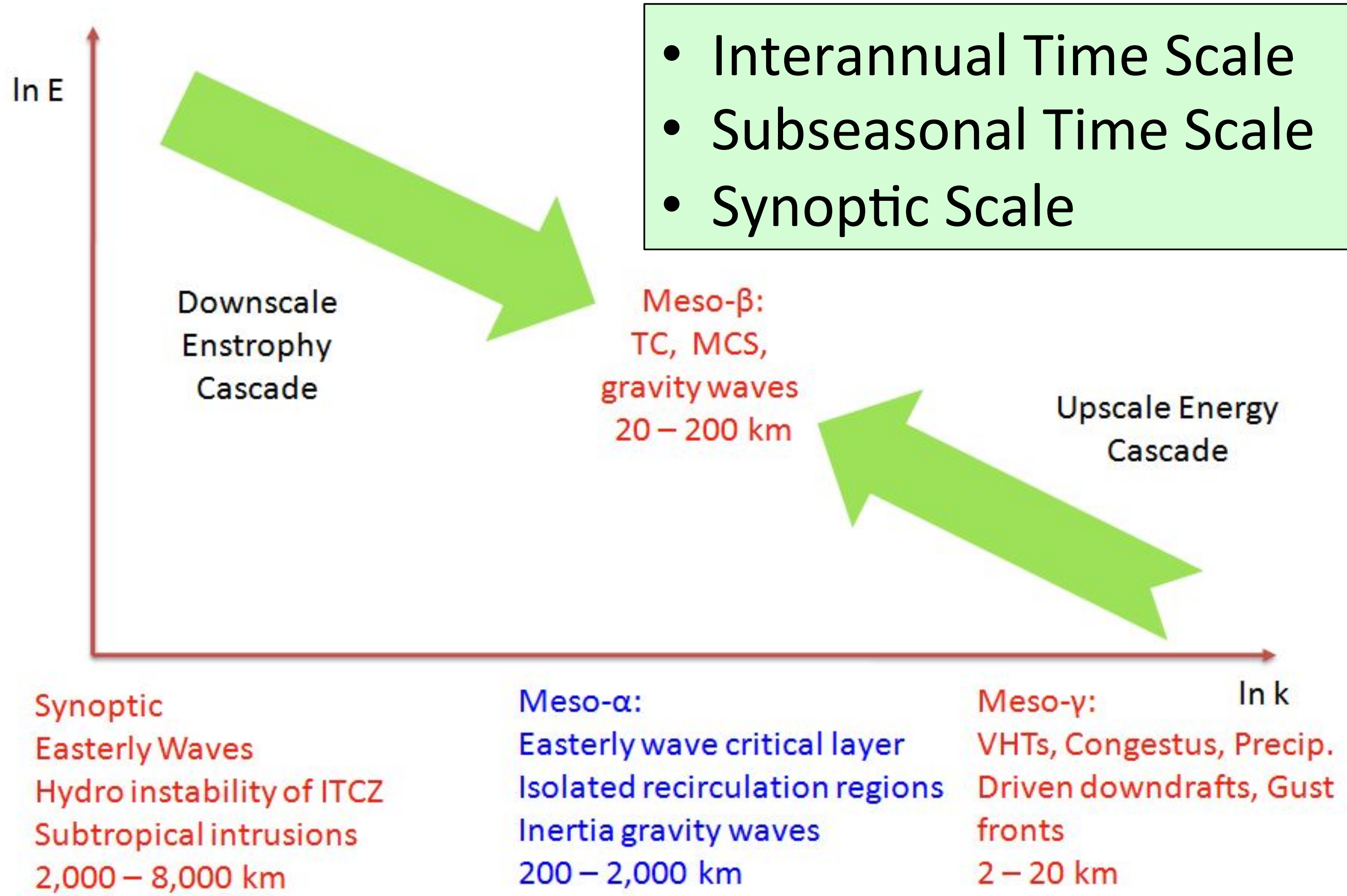
- “*Future advances in the quality of ISI predictions are closely tied to exploiting new sources of predictability or improving the representation of known sources of predictability in current forecast systems*” (NRC report)
  - Identify the sources of predictability
  - Evaluate the representation the sources
  - What are the key physical processes for skillful prediction of TCs? How are they represented?
  - Help us better interpret the prediction skill

Tropical Cyclogenesis

- Socio-economic impacts of TCs
- For extended range forecasts, skillful prediction of TC genesis location and time is critical for the skillful track and intensity prediction
- Skillful prediction of TCG near coast is of particular importance for storm preparedness

TC Formation: Multi-scale Interaction

- Interannual Time Scale
- Subseasonal Time Scale
- Synoptic Scale



Montgomery and Smith 2010

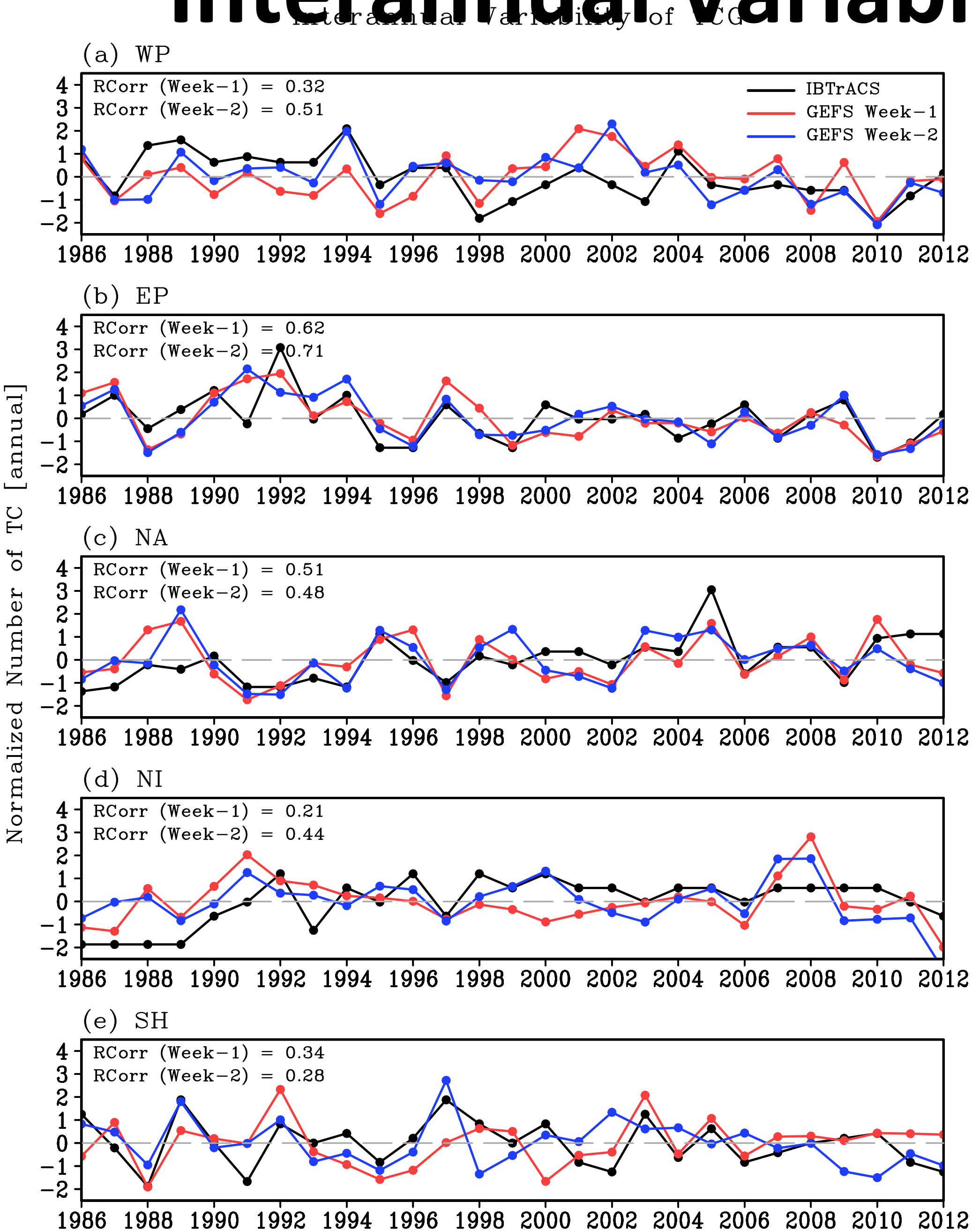
Data

- Global Ensemble Forecasting System Reforecast-2 (GEFS-R), with the forecast lead up to 16 days (1985-2012) (Hamill et al. 2013)
- GFDL Vortex Tracker (Marchok 2002)
- IBTrACS (Knapp et al. 2010)

TC indices

- **TC counts:** the total number of TCs within a certain period over a certain basin;
- **TC days:** the sum of the life time of all TCs measured in days;
- **Accumulated cyclone energy** (ACE): integrating the squares of the maximum sustained surface wind speed over the lifetime of a TC for all TCs, which is a function of TC numbers, lifetimes, and intensities.
- All indices are derived from GEFS week-1 or week-2 reforecasts – **potential prediction skill of GEFS on the subseasonal to seasonal time scale**.

Interannual Variability of TC Counts

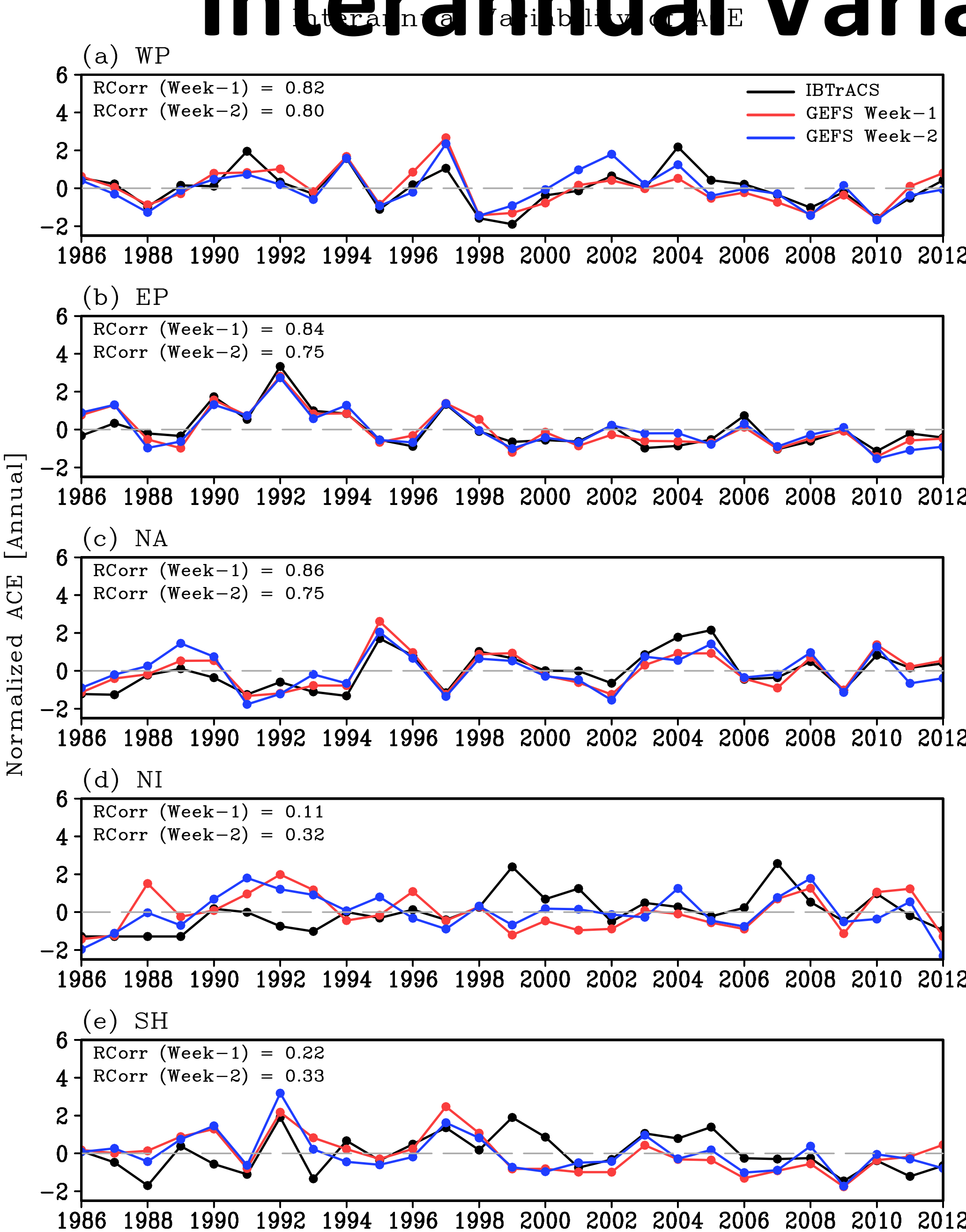


Rank Correlation with obs:

	Week-1	Week-2
WP	0.32	0.51
EP	0.62	0.71
NA	0.51	0.48
NIO	0.21	0.41
SH	0.34	0.28

Red: the significant coefficients with 95% confidence

Interannual Variability of ACE



Rank Correlation with obs:

	Week-1	Week-2
WP	0.82	0.80
EP	0.84	0.75
NA	0.86	0.75
NIO	0.11	0.32
SH	0.22	0.33

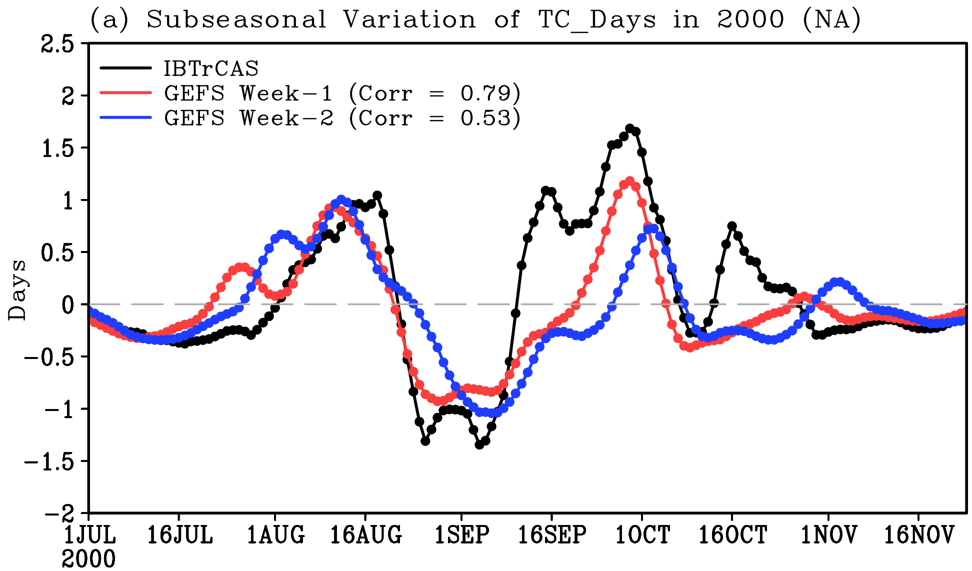
Red: the significant coefficients with 95% confidence

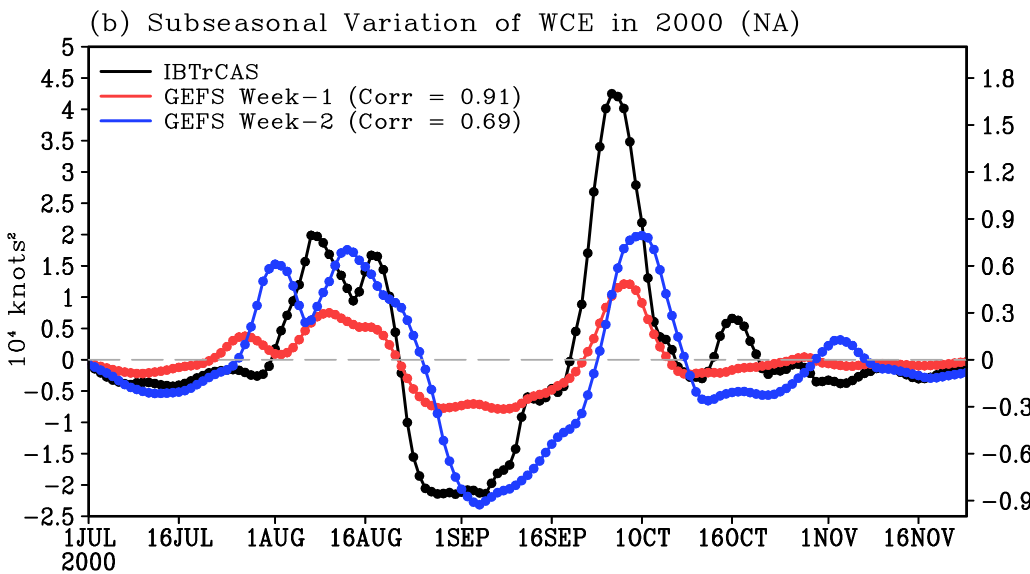
Pearson's correlations with Nino 3.4

Pearson's correlations with Nino 3.4 (Jun-Nov, 1985-2012)						
	TC frequency			*ACE*		
Basin	*Obs*	*Week-1*	*Week-2*	*Obs*	*Week-1*	*Week-2*
WP	*0.25*	0.36	0.35	0.72	0.73	0.76
EP	0.48	0.66	0.75	0.38	0.49	0.61
NA	-0.42	-0.66	-0.46	-0.48	-0.52	-0.55

All above the 95% confidence (0.32) except the one between Nino 3.4 and the observed TC counts over the WP.

Subseasonal Time Scale: Can GEFS capture the active and inactive periods of TC activity?





Black: 7-day running mean TC days or <u>weekly cyclone energy (WCE)</u> from IBTrACS, with the seasonal cycle removed.

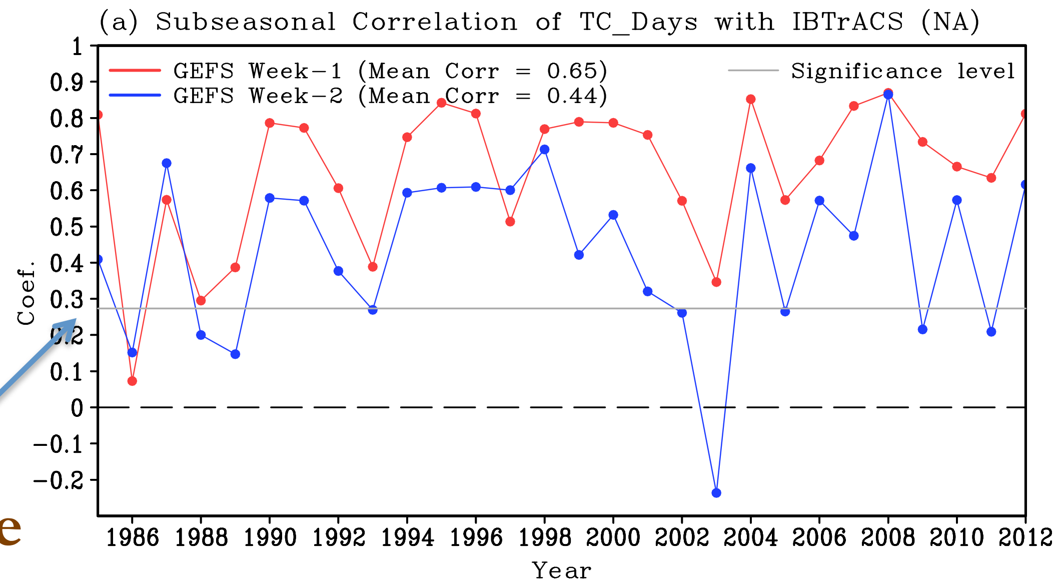
Red: Week-1 GEFS reforecasts

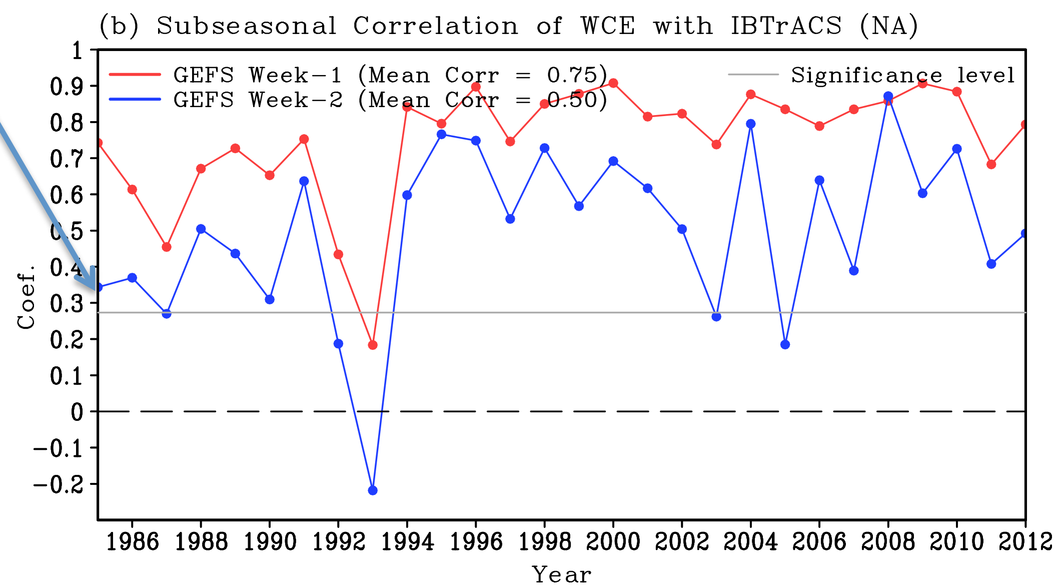
Blue: Week-2 GEFS reforecasts

Corr.	*Week-1*	*Week-2*
TC Days	0.79	0.53
WCE	0.91	0.69

GEFS captures the active and inactive periods for TC activity reasonably well in 2000… what about other years?

Time series of the Pearson's correlation between the observed and predicted TC days / WCE





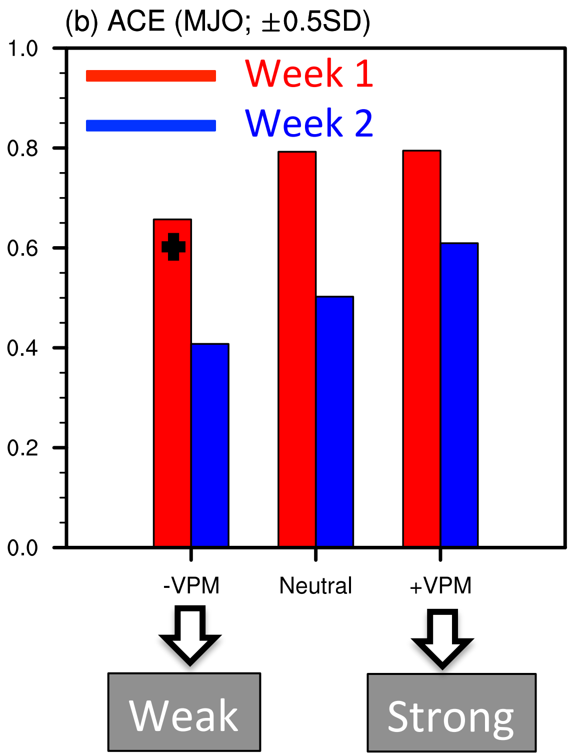
95% confidence

Reasonable skill in predicting the active and inactive TC periods (corr > 0.6) in most years with the lead time of 7-14 days

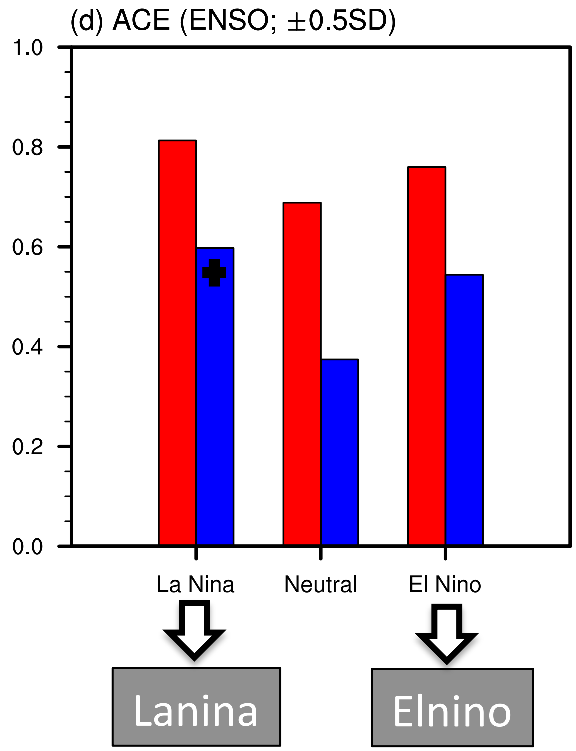
But still see large year-to-year variability of the correlation coefficients…?

What modulates the subseasonal prediction skill?

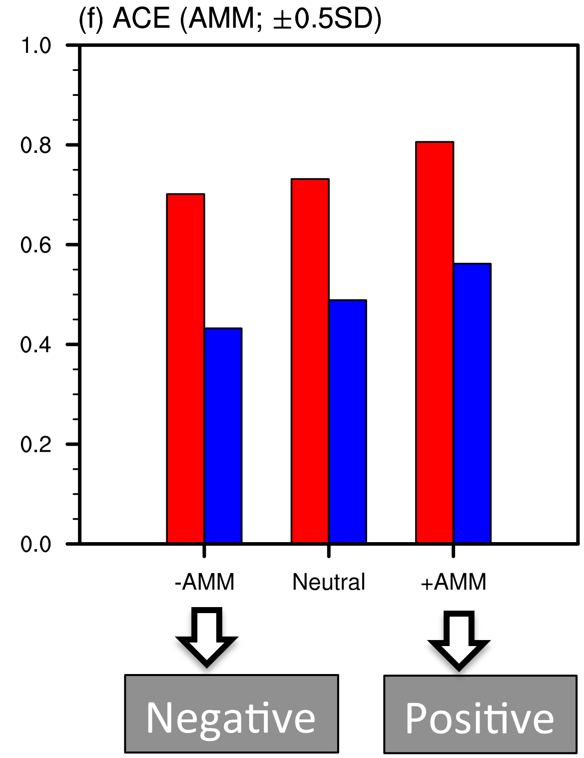
Impact of MJO



Impact of ENSO



Impact of AMM



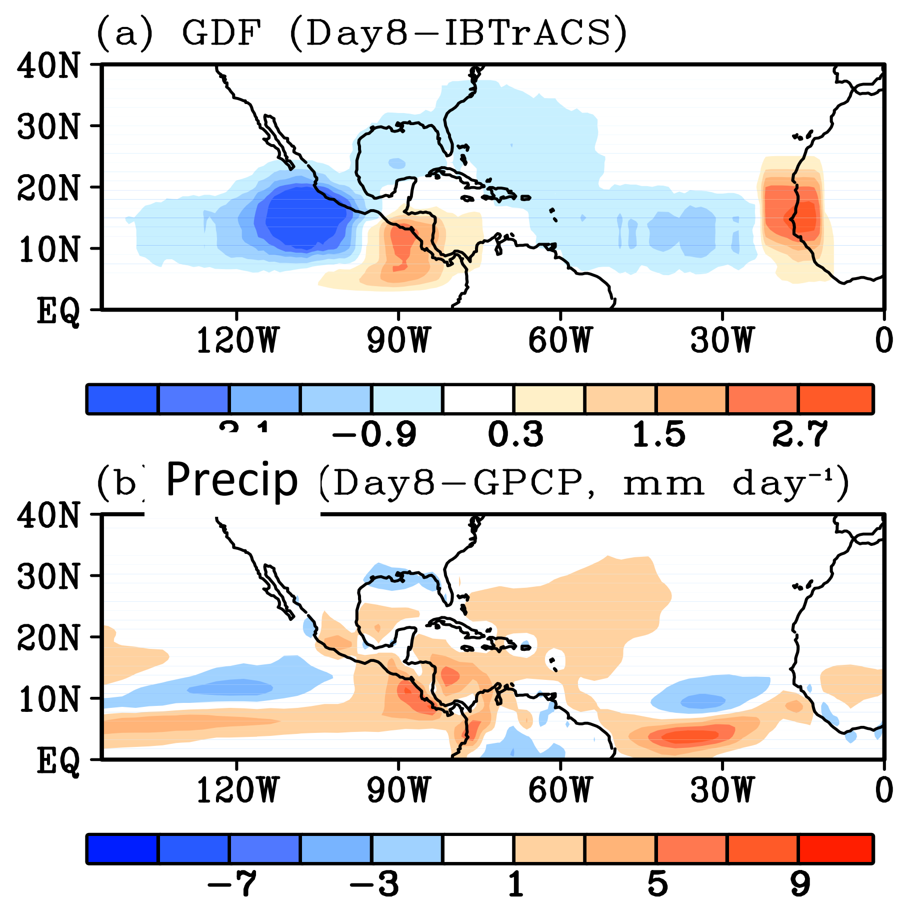
Prediction skill tends to be higher (weaker) in years of strong (weak) MJO activity.

Higher skill during Elnino or Lanina years than in neutral years

No significant impacts of AMM were found ← the ITCZ biases in the GEFS

Atlantic Meridional Mode (AMM) and TC Counts

Indices	Obv.	Week-1	Week-2
Nino3.4	-0.42	-0.66	-0.66
AMM	0.74	0.48	0.52



The AMM-TC correlation in the GEFS is weaker than the observation

Likely due to biases in both the Atlantic TCG and the ITCZ

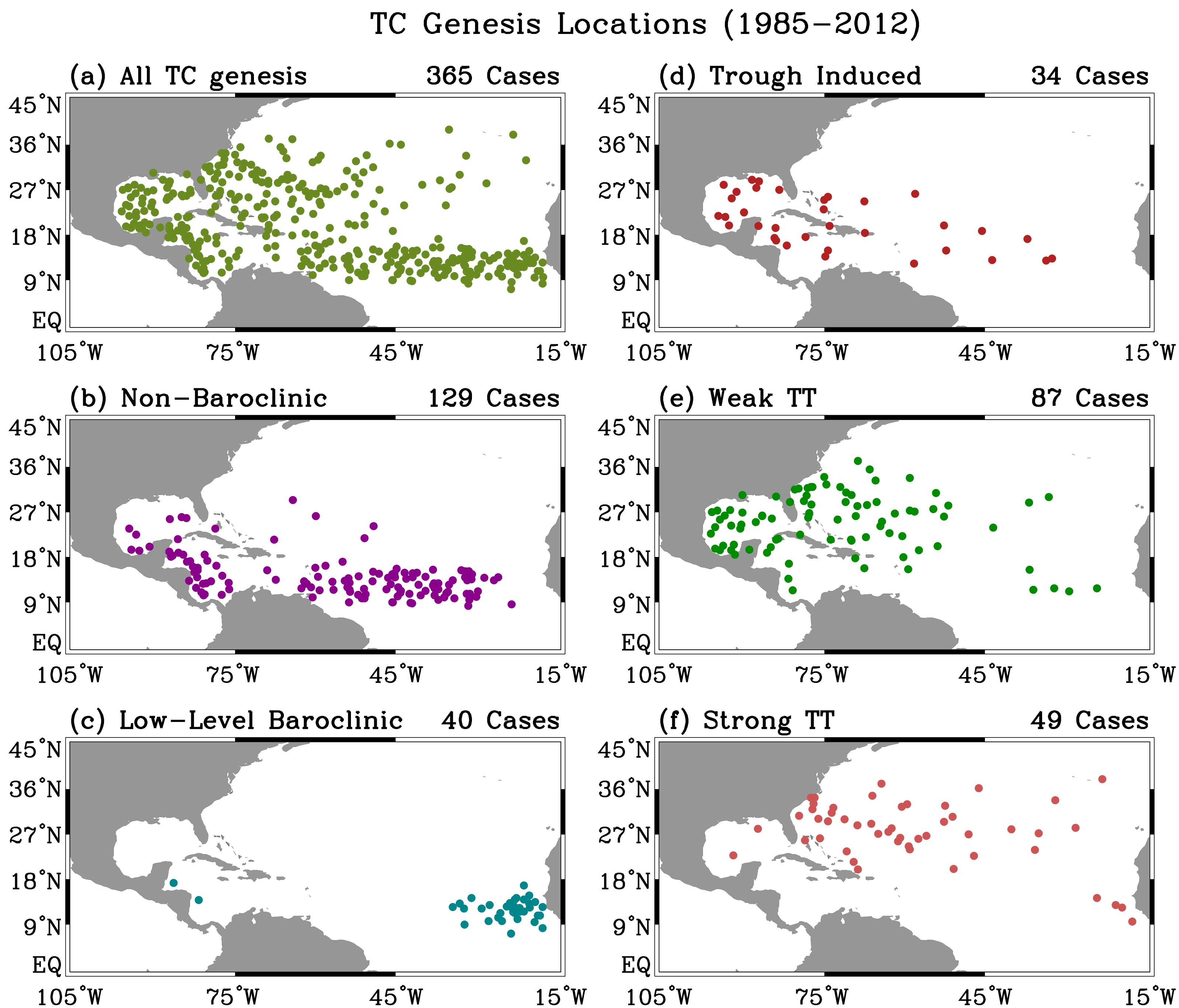
AMM: Untapped predictability

Synoptic Time Scale: TC genesis pathways

- Precursors provide the local environment for TCG.
- McTaggart-Cowan et al.*(2013)* defined **5 categories** of TC genesis based on the synoptic-scale environment;
- They are categorized by two metrics:

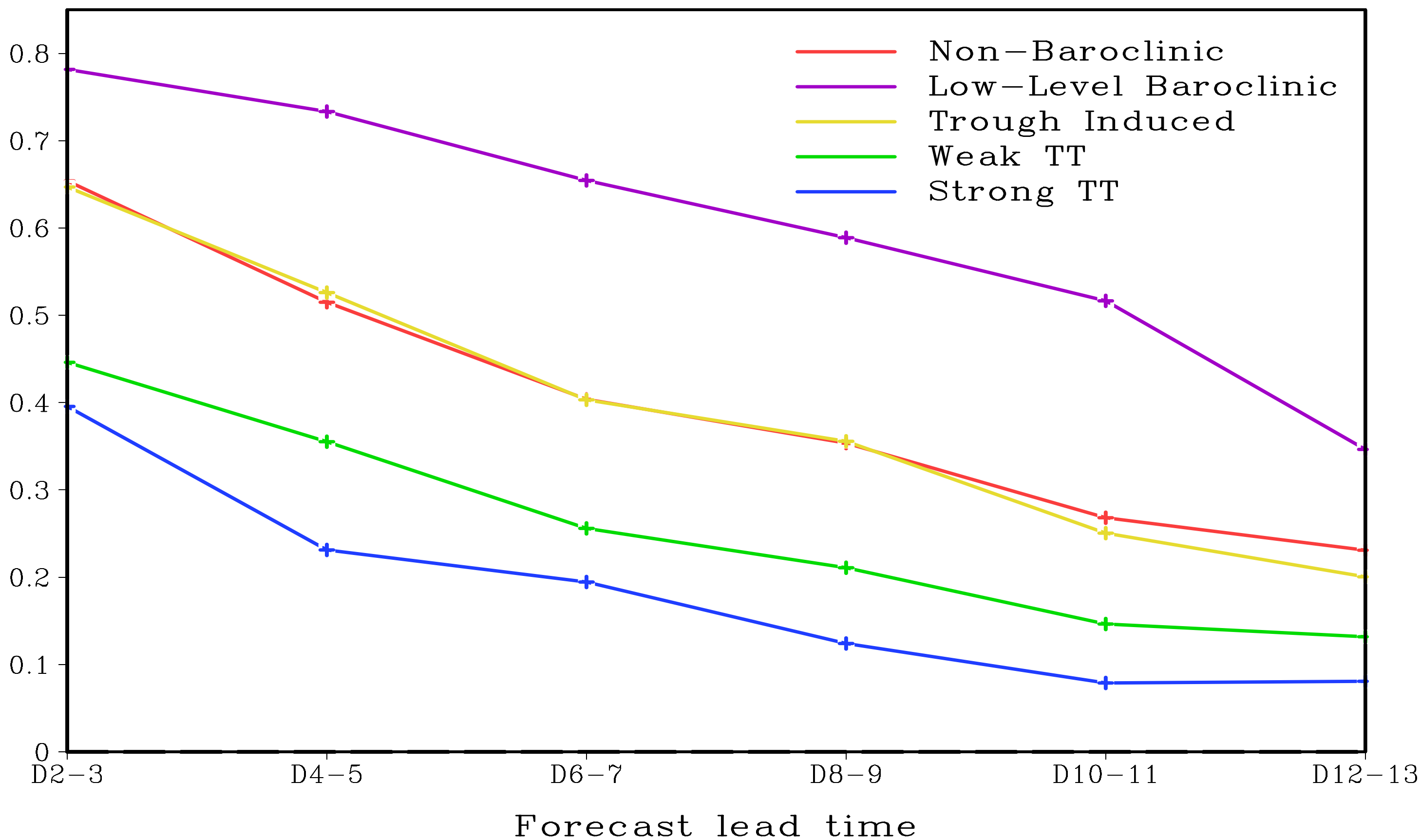
  1. **Upper-level forcing**
  2. **Low-level baroclinity**

	Non-baroclinic	*Low-level baroclinic*	*Trough induced*	*Weak TT*	*Strong TT*
Upper-level forcing	Low	Low	High	High	High
Low-level Baroclinity	Low	High	Low	Medium	High


TC Genesis Locations (1985–2012)
(a) All TC genesis
365 Cases
(b) Non–Baroclinic
129 Cases
(c) Low–Level Baroclinic
40 Cases
(d) Trough Induced
34 Cases
(e) Weak TT
87 Cases
(f) Strong TT
49 Cases
45°N
36°N
27°N
18°N
9°N
EQ
105°W
75°W
45°W
15°W

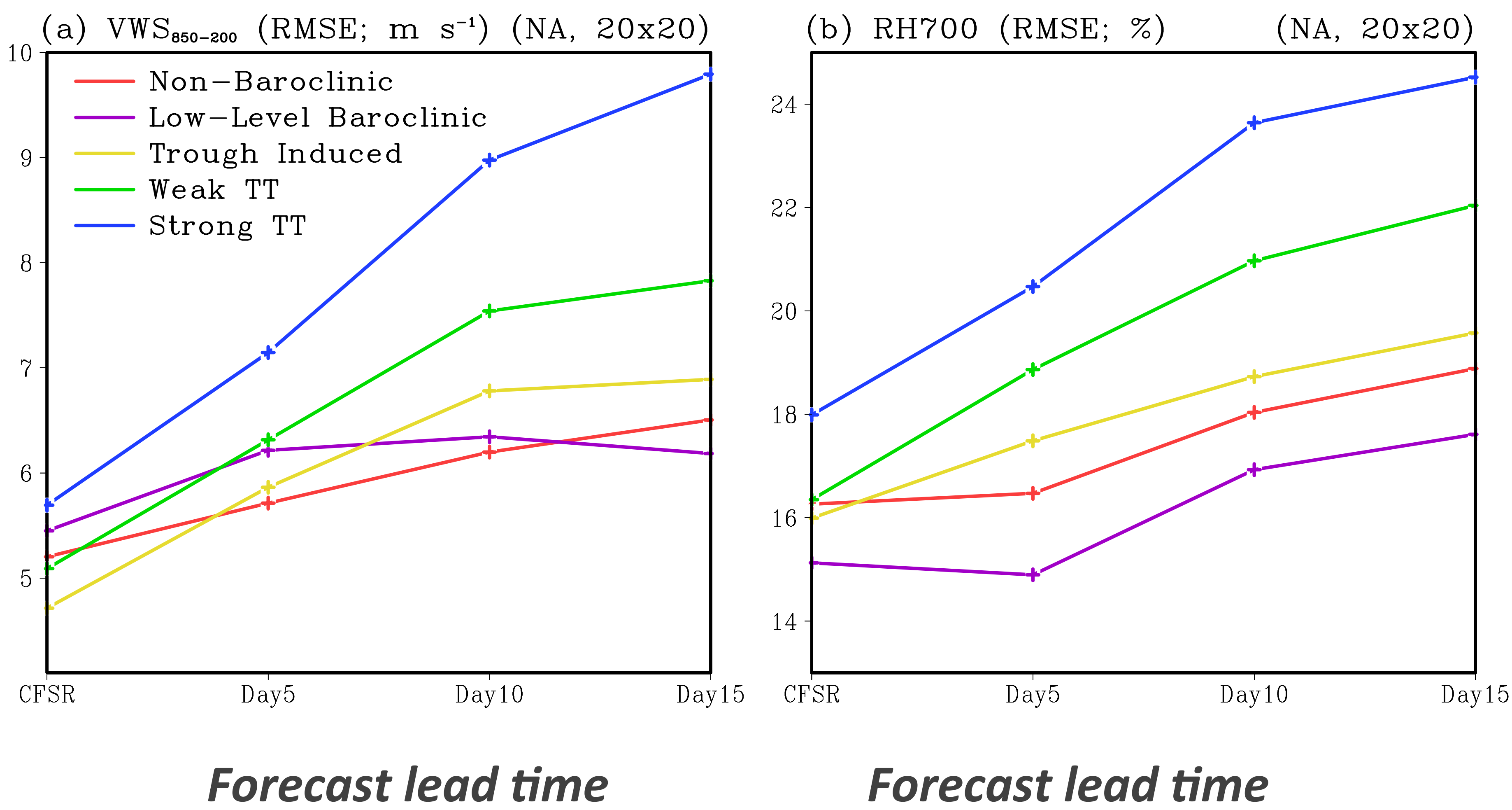
Synoptic Time Scale: Symmetric Extremal Dependence Index (SEDI)

SEDI: No degeneracy for rare binary events. **higher score-> higher skill**



Forecast skill: Low-level baroclinic > Non-baroclinic or Trough induced > Weak TT > Strong TT

Synoptic Time Scale:
RMSE in the Vicinity of Genesis



Root-mean-square error (RMSE) of vertical shear and RH700 over 20º x 20º box around the observed genesis center

Summary

- GEFS skillfully captures the interannual variability of TC counts and ACE over the North Pacific and the Atlantic, which can be partly attributed to the modulation of TCs by the ENSO.
- GEFS shows promising skill in predicting the active and inactive periods of TC activity over the Atlantic on the subseasonal timescales, but the skill has large year-to-year fluctuations. Our analysis shows possible impacts of the ENSO and the MJO on the subseasonal TC predictability.
- The impacts of the AMM were not well-represented in the model, which can be attributed to biases in the TCG distribution and the ITCZ location - an untapped source of predictability in the model.
- Local synoptic-scale environment affects TC predictability over the Atlantic: stronger upper-level forcing (weak TT and strong TT) is associated with weaker predictability..

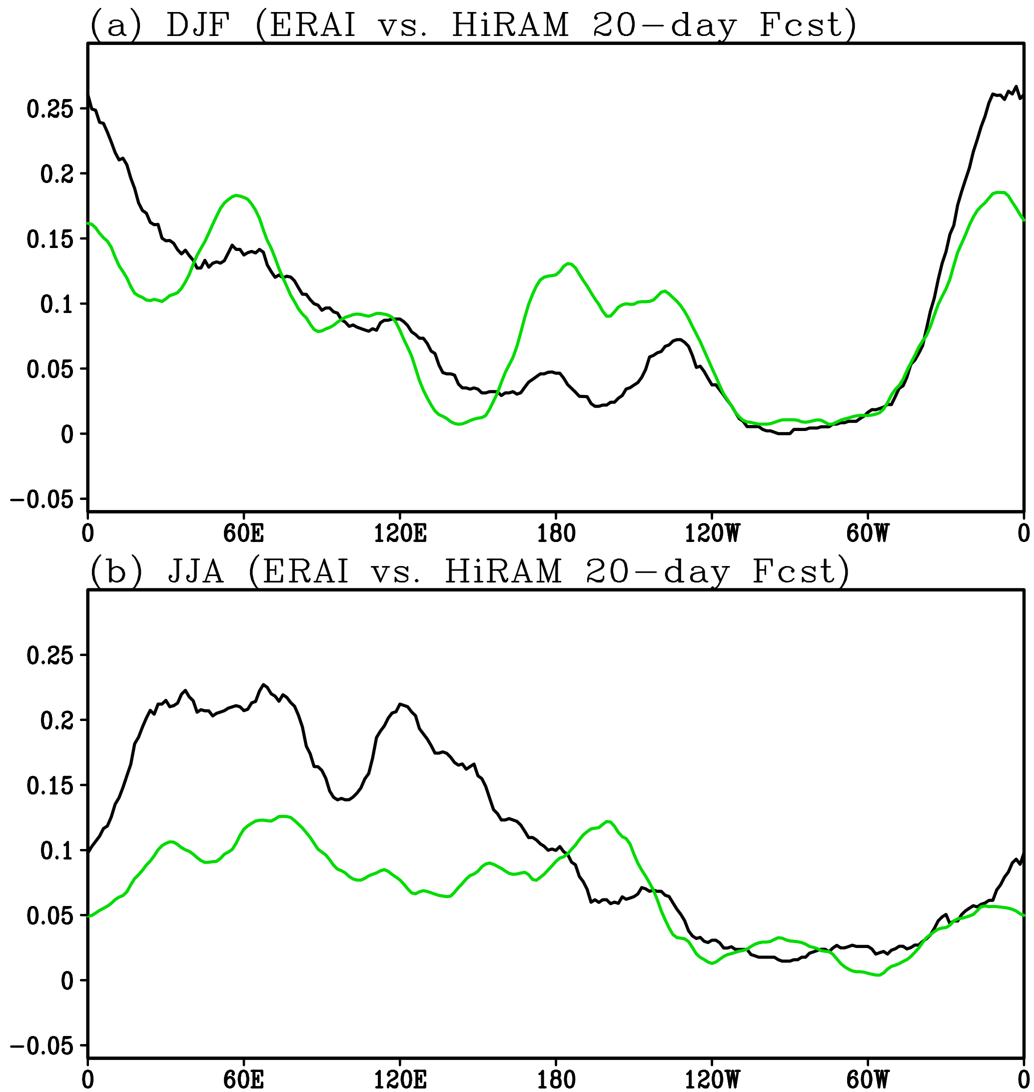
Priority Focus for FY17

1. Prominent Motion Systems
   - Tropical Cyclones
   - MJO
   - **Blocking**
   - **Teleconnections**

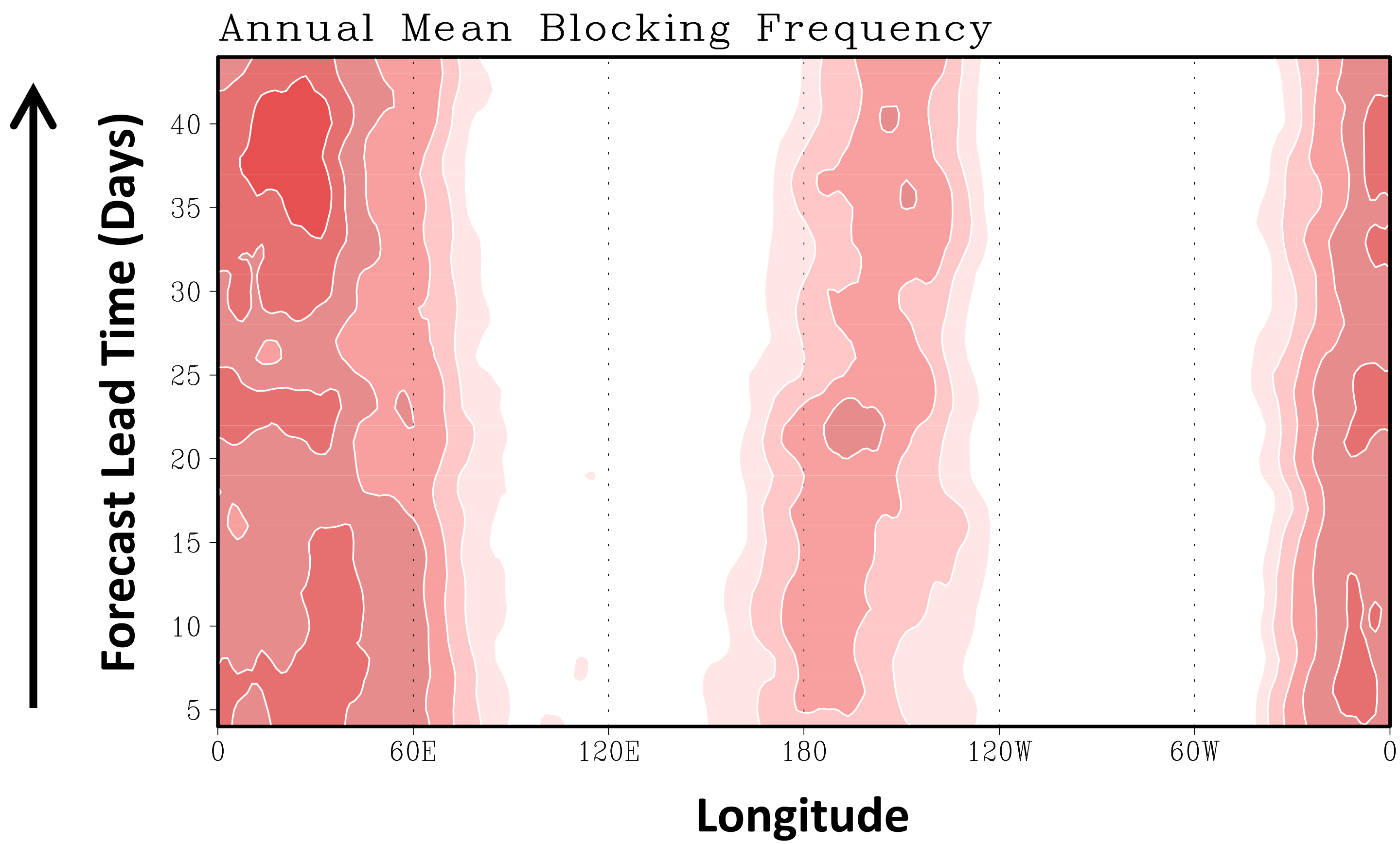
- **Level 1:** Performance-oriented evaluation
- **Level 2:** Physics-oriented evaluation
- **Level 3:** Evaluation of predictability

2. Specific Physical Processes
   - Moist convection: cumulus parameterization, CWV-precip relationship, Q1/Q2 diagnoses
   - Model representation of different cloud regimes: evaluation of the model clouds using satellite simulators

Blocking (DJF): ERA-Interim vs. GFDL Coupled HiRAM Hindcasts (2003-2013)



Annual Mean Blocking Frequency



- Blocking frequency does not decrease with forecast lead time
- An eastward shift of the blocking center over the Pacific

Transition of the Diagnostic Tools

- Diagnostic tools have been tested with different models of different spatial/temporal resolutions
  - Developed with GEFS; testing with FIM and CM4 hindcasts
  - A bug in the TC tracker was found in the testing with FIM hindcasts and reported to HWRF team.
- Transition to DTC global model Testbed with testing data and documents (working with Ligia Bernardet and Louisa Nance)
- Bi-weekly global model verification meeting organized by Fanling Yang

QUESTIONS?