

Comparing Forecast Skills on Seasonal Climate Conditions: From Normal to Extreme

Cheng-Ta Chen and Shuo-Li Lin

National Taiwan Normal University, Dept. of Earth Sciences



APEC Climate Center 2008 Symposium, 2008 August 19-21

Motivation

- Extreme seasonal variation can have more impact on society and environment
- But are they more difficult to predict?



Multi-model ensemble system

- DEMETER system: 7 coupled global circulation models

Partner	Atmosphere	Ocean
ECMWF	IFS	HOPE
LODYC	IFS	OPA 8.3
CNRM	ARPEGE	OPA 8.1
CERFACS	ARPEGE	OPA 8.3
INGV	ECHAM-4	OPA 8.2
MPI	ECHAM-5	MPI-OM1
UKMO	HadCM3	HadCM3

4 start dates per year
 6 months hindcasts
 9 member ensembles
 3 ocean analyses, 4 \pm SST pert
 ERA-40: ocean forcing and
 atmospheric initial conditions

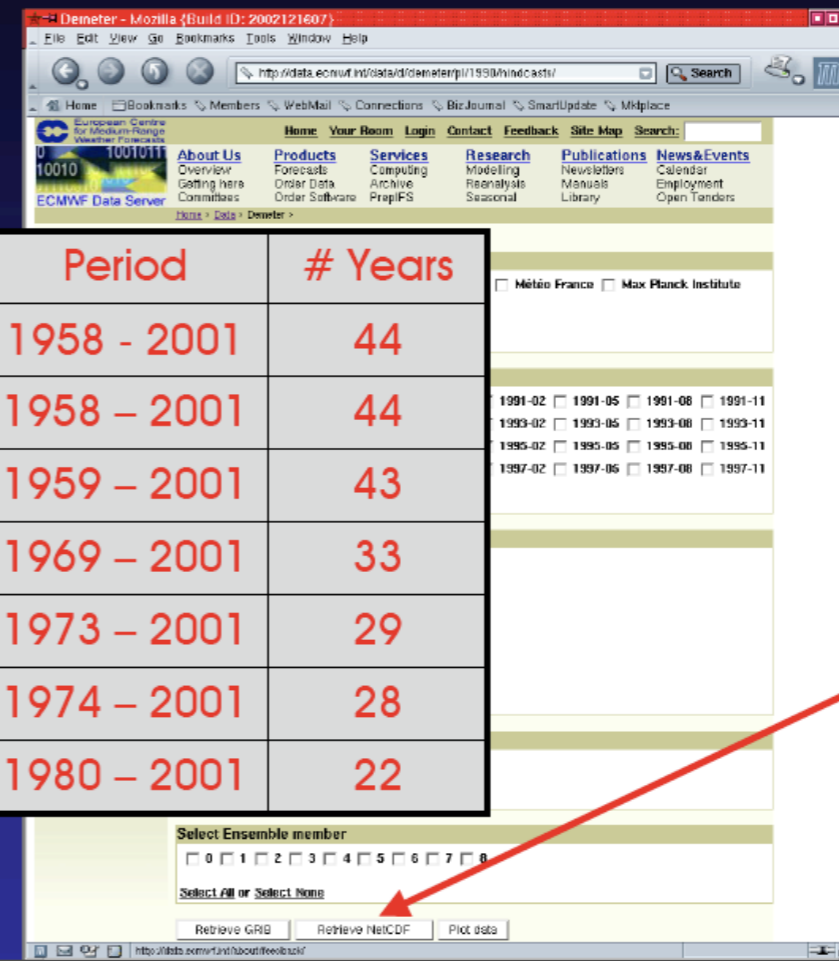
One-tier coupled
 forecast system

- Hindcast production for: 1980-2001

Start from Feb., May, Aug., Nov.
 for 6 month hindcast

<http://data.ecmwf.int/data>

	Period	# Years
ECMWF	1958 - 2001	44
CNRM	1958 - 2001	44
UKMO	1959 - 2001	43
MPI	1969 - 2001	33
INGV	1973 - 2001	29
LODYC	1974 - 2001	28
CERFACS	1980 - 2001	22



Retrieve
 NetCDF

Hindcast skill verification

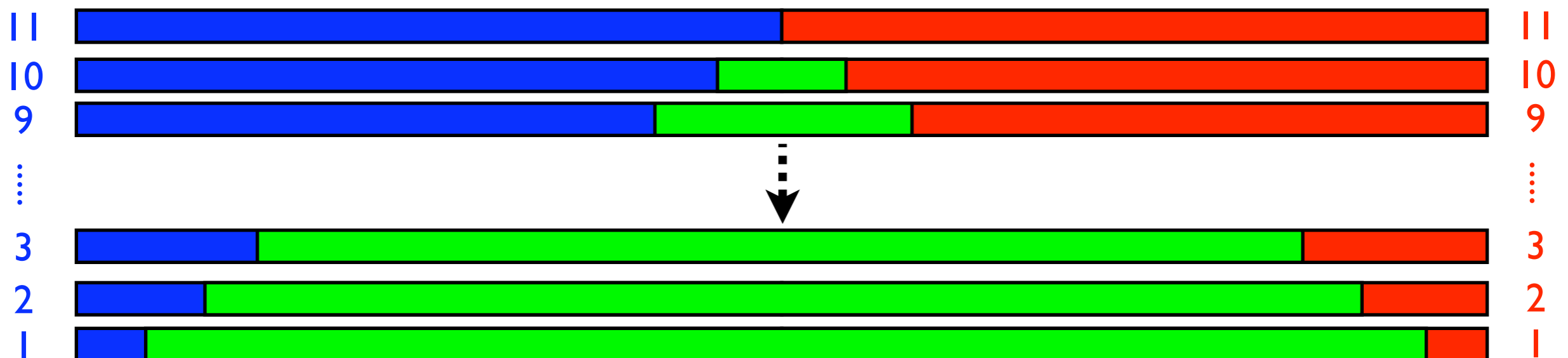
- Cross-validated hindcast data for 1980-2001 period

For each individual coupled model hindcasts from 1980-2001

189(21x9) cross-validated members to determine climatological PDF and threshold for 3-categories forecasts

determine the 3-categories probability of the forecast year based on 9 members

From 22-years of data, we can decide the rarity by gradually expand the normal part of the whole record



Hindcast skill verification

- 63(9)-members multi(single)-model ensemble forecast

-Categories

Either deterministic (which one more likely) or

probabilistic (probability for each categories)

11-0-11, 10-2-10, 9-4-9, 8-6-8, 7-8-7, 6-10-6,

5-12-5, 4-14-4, 3-16-3, 2-18-2, 1-20-1

- Equitable threat score (ETS) for deterministic forecast

$$ETS = \frac{hits - hits_{random}}{hits + misses + false\ alarms - hits_{random}}$$

How well did the forecast "yes" events correspond to the observed "yes" events (accounting for hits due to chance)?

- Grid-point and area aggregated ROC skill score

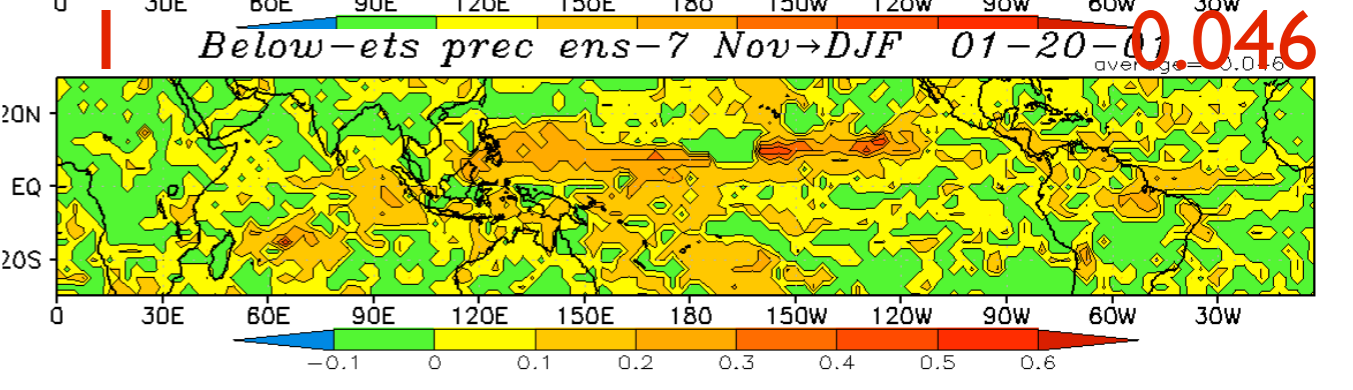
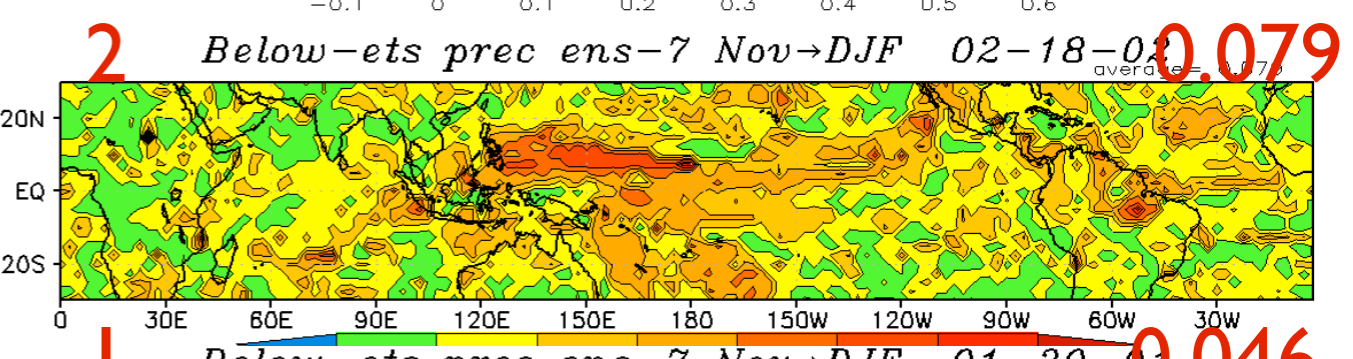
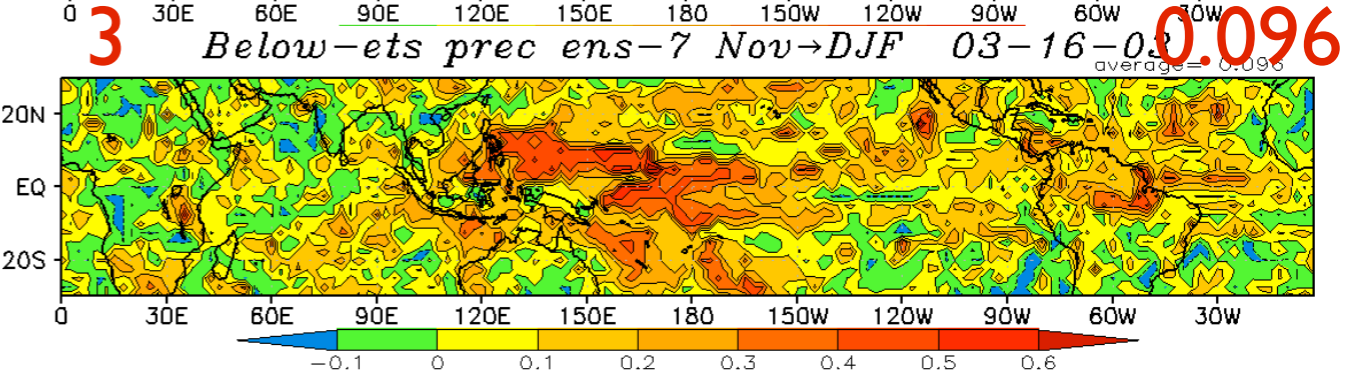
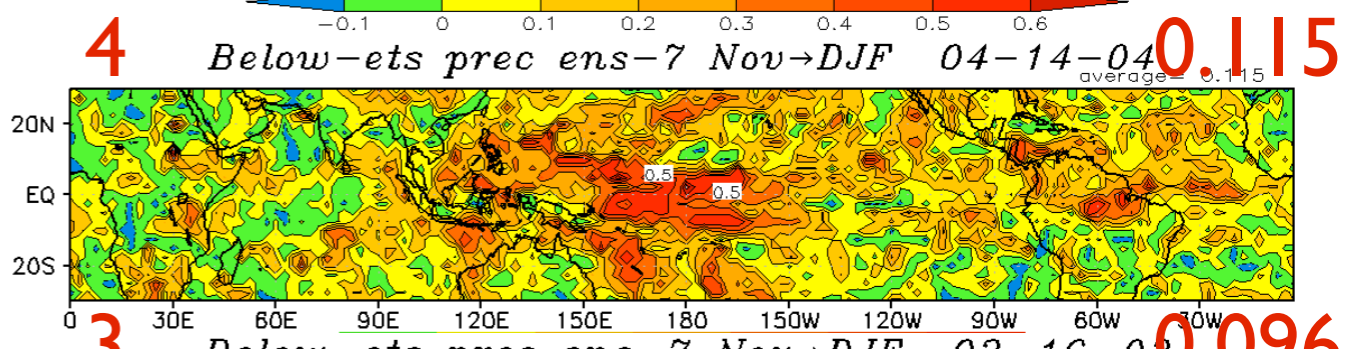
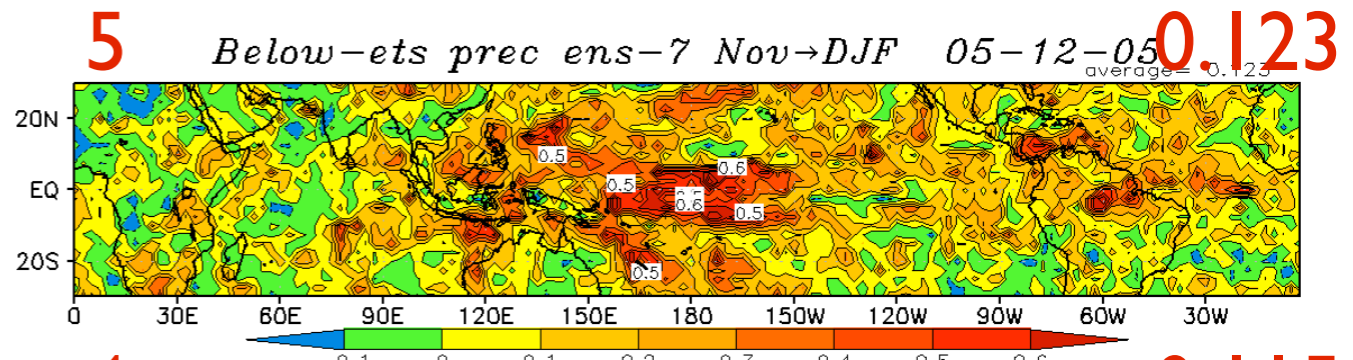
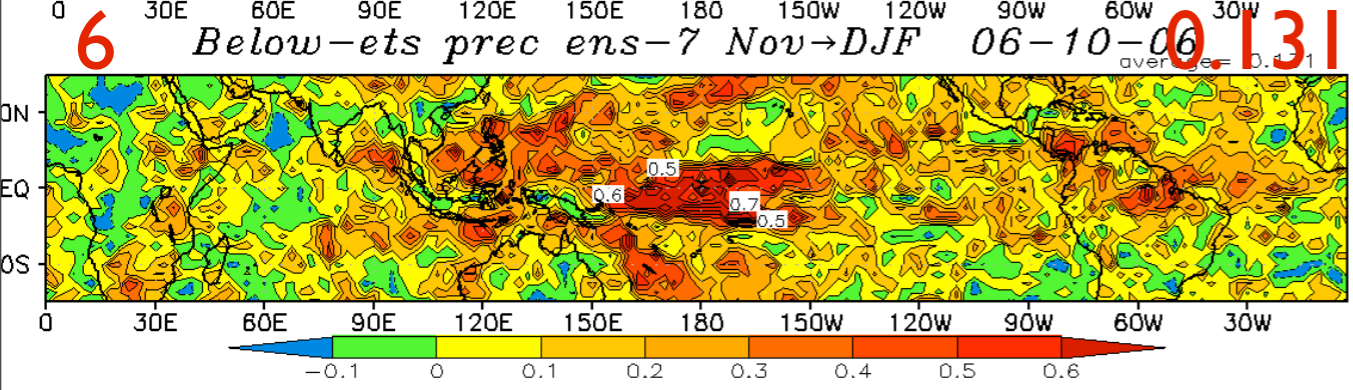
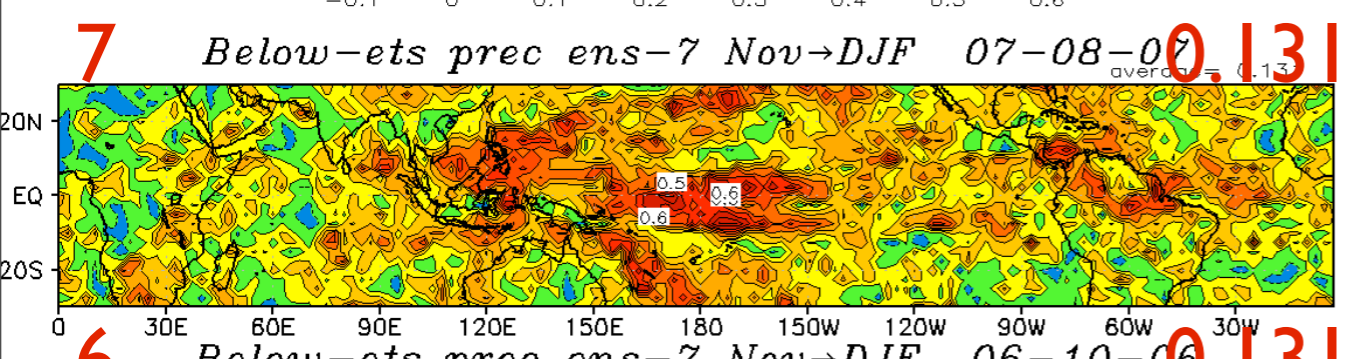
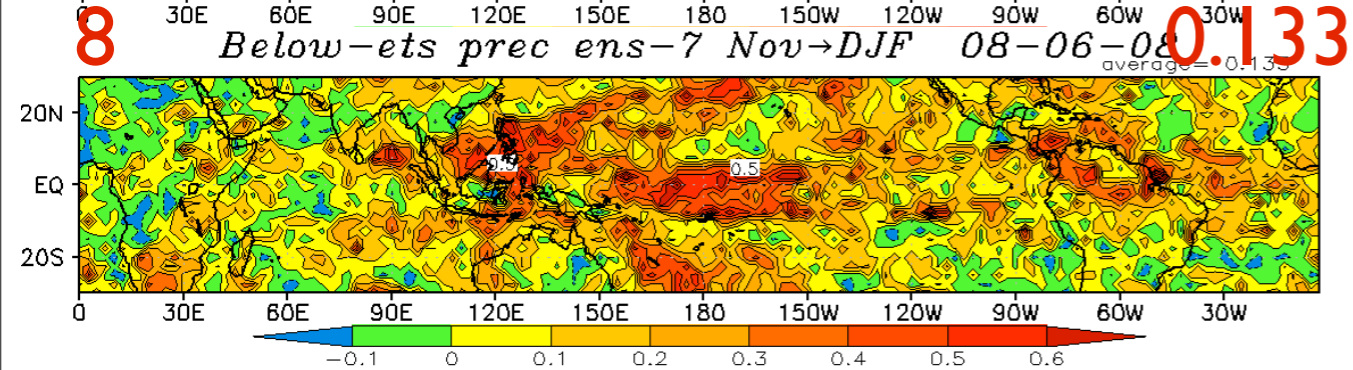
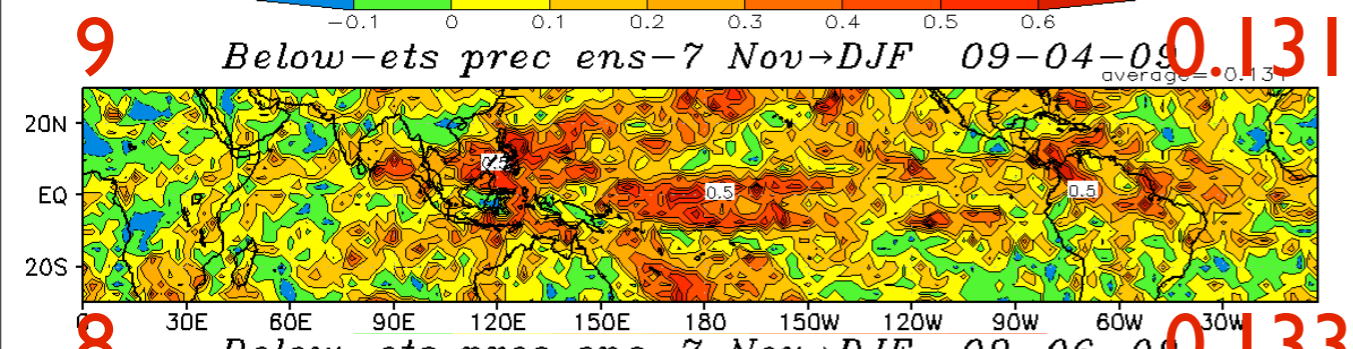
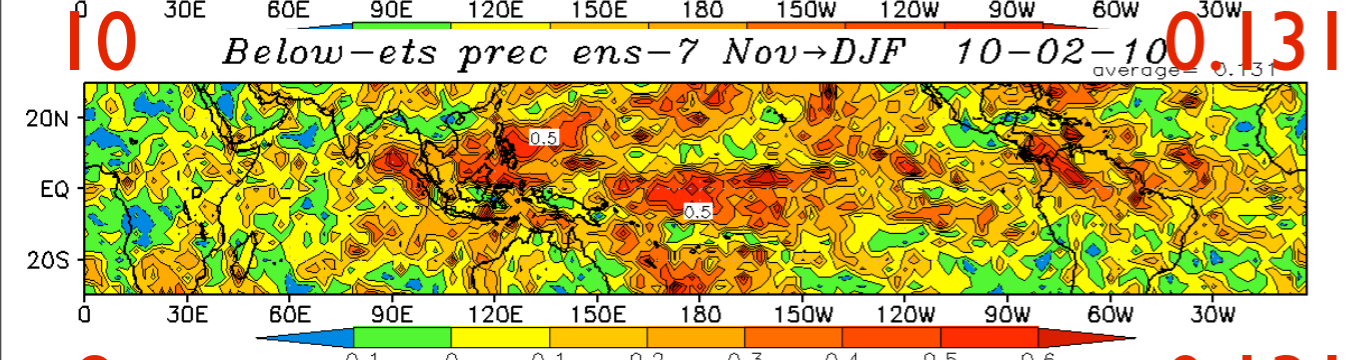
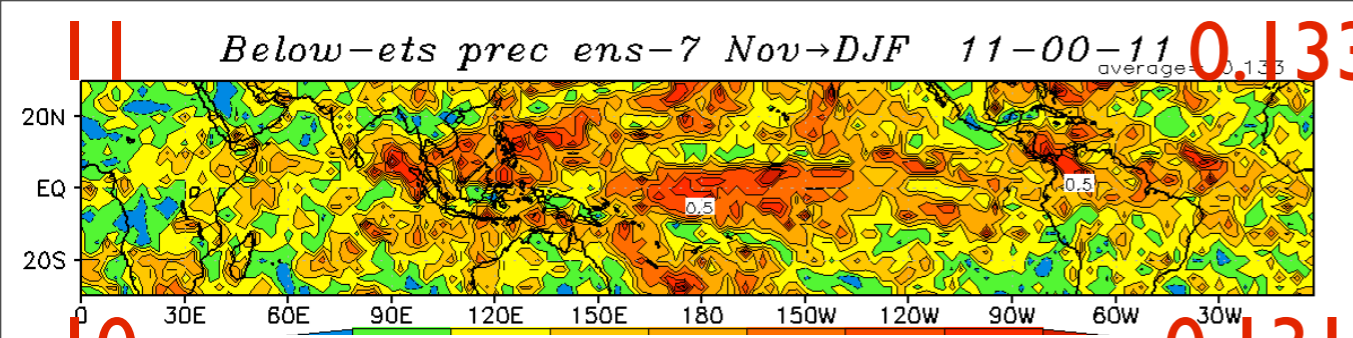
What is the ability of the forecast to discriminate between events and non-events?

- Area aggregated reliability diagram and brier skill score

What is the relative skill of the probabilistic forecast over that of climatology, in terms of predicting whether or not an event occurred?

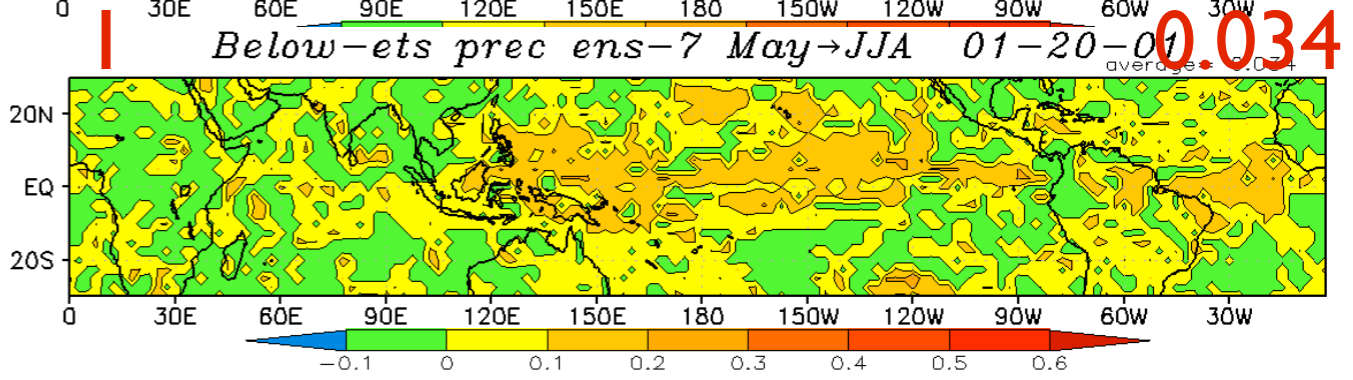
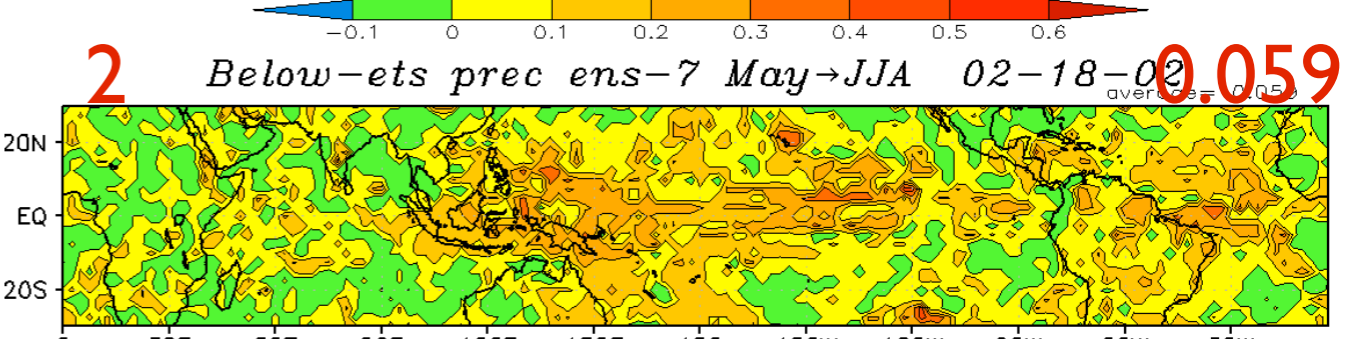
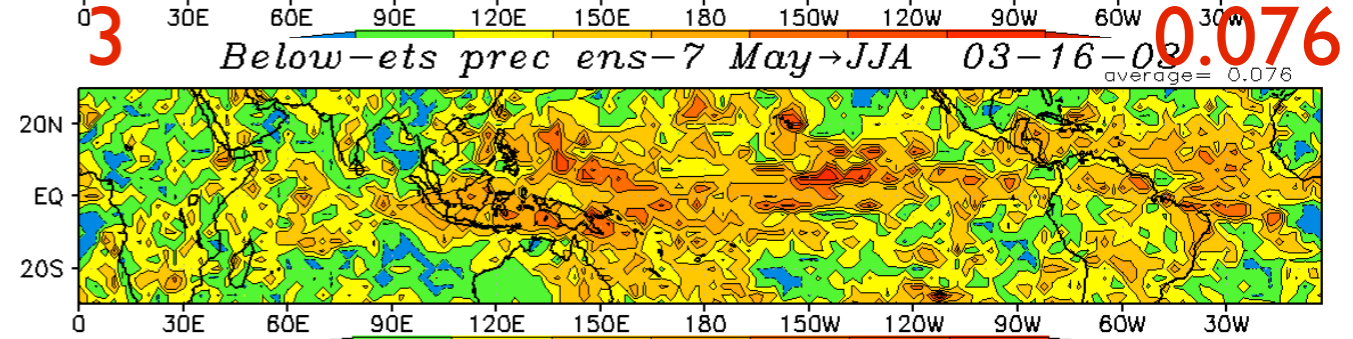
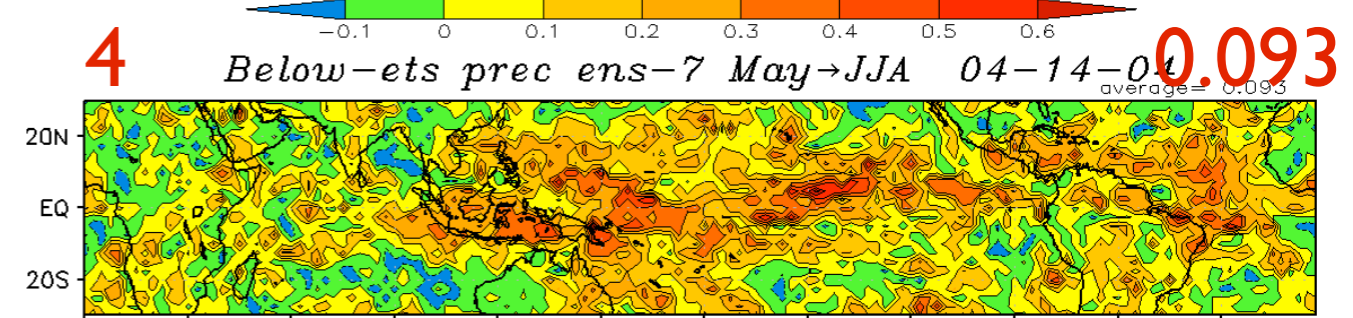
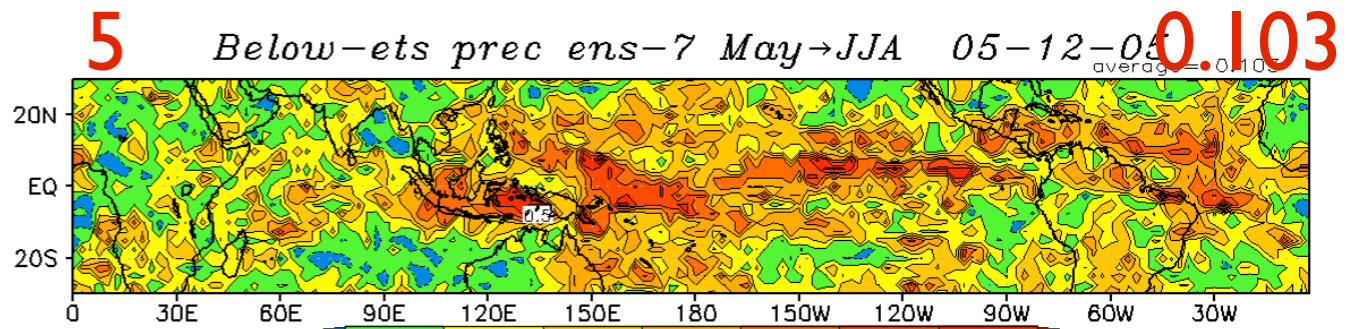
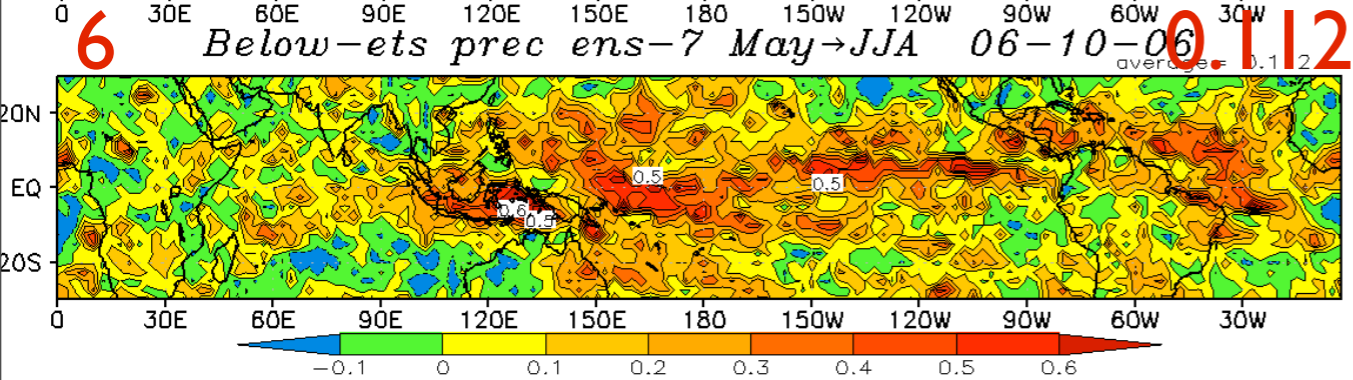
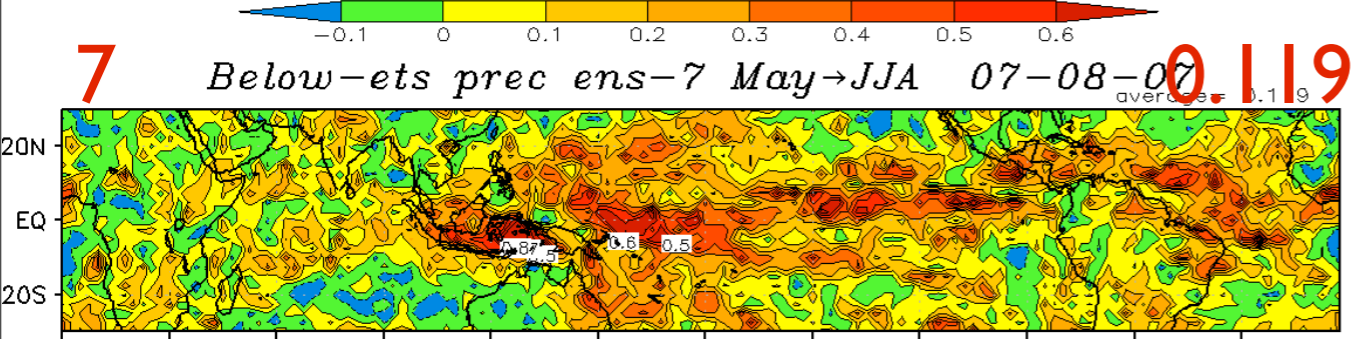
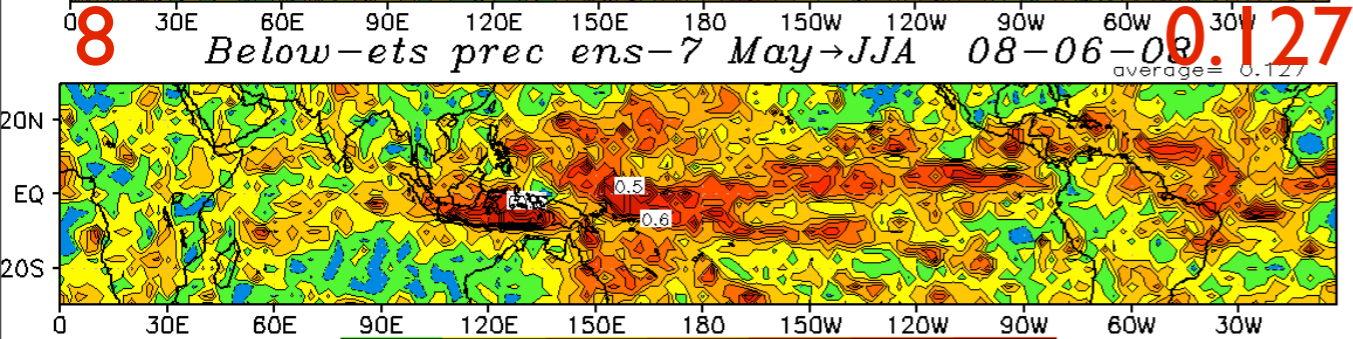
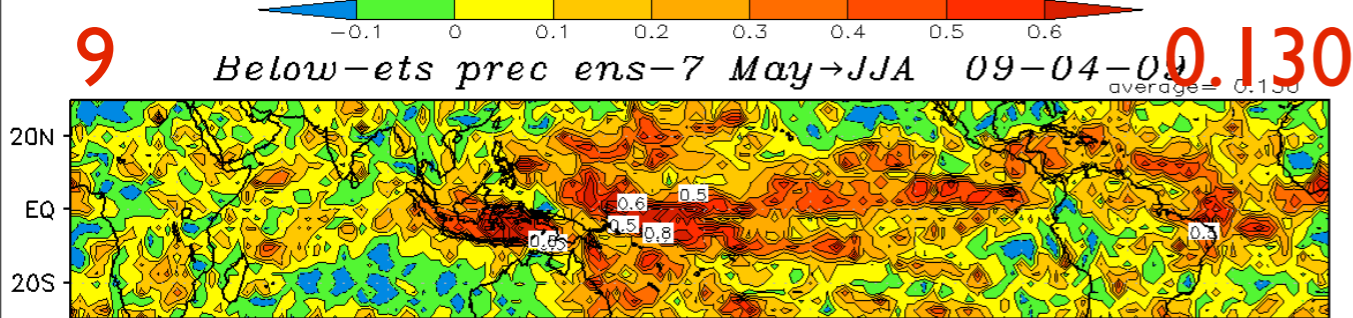
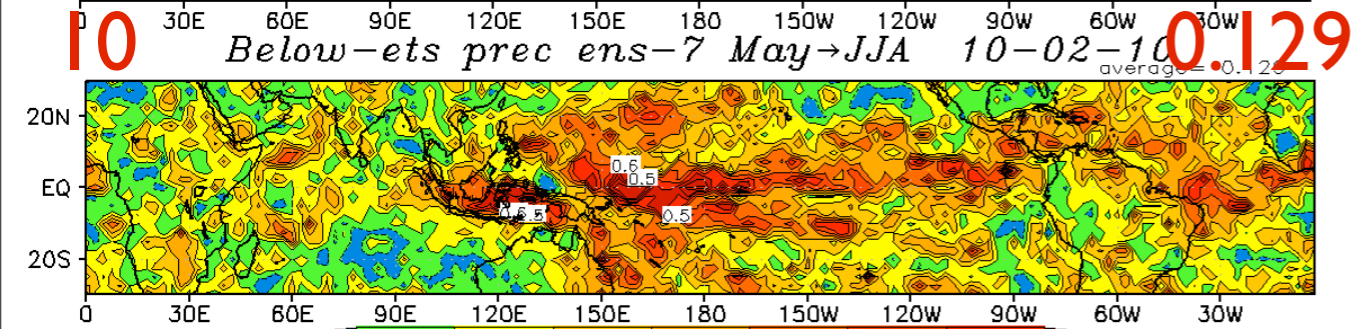
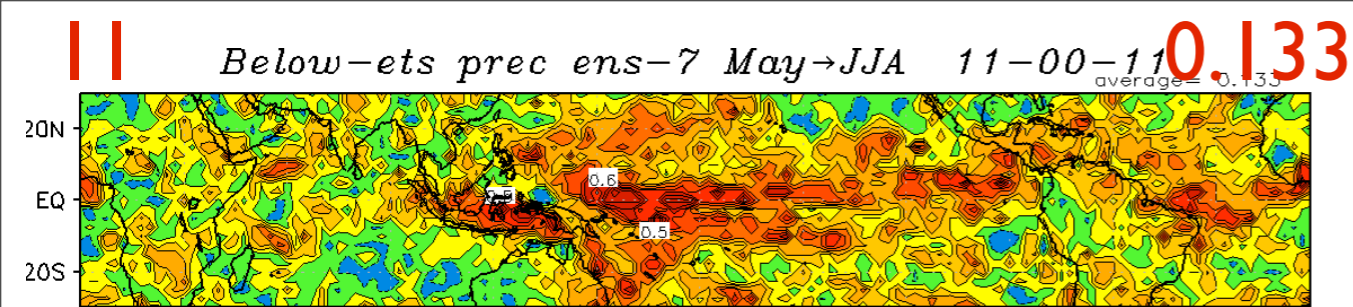
ETS score, Nov → DJF, Tropics

Below Normal → Well Below Normal



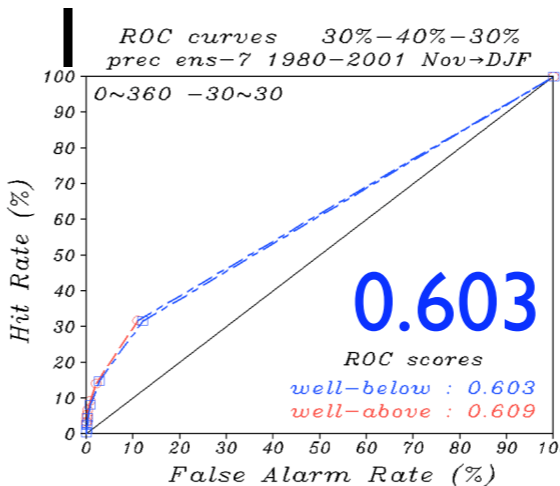
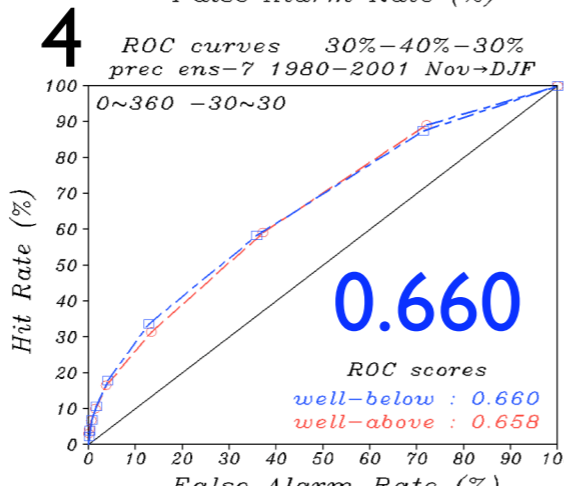
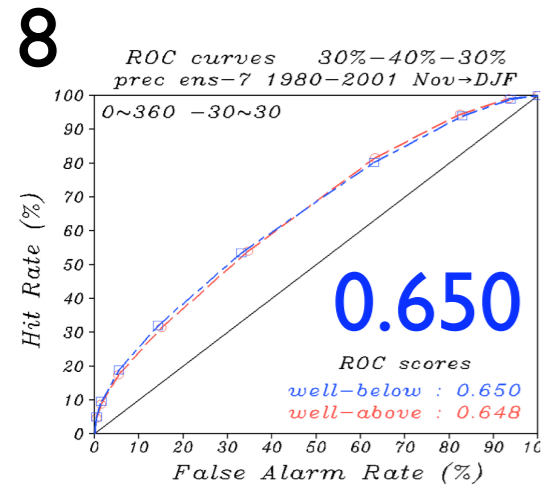
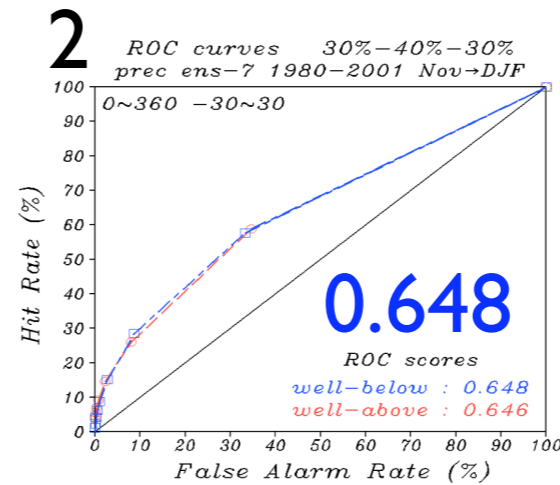
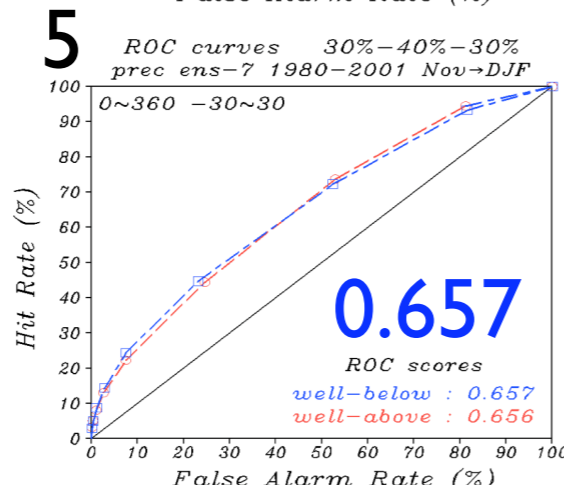
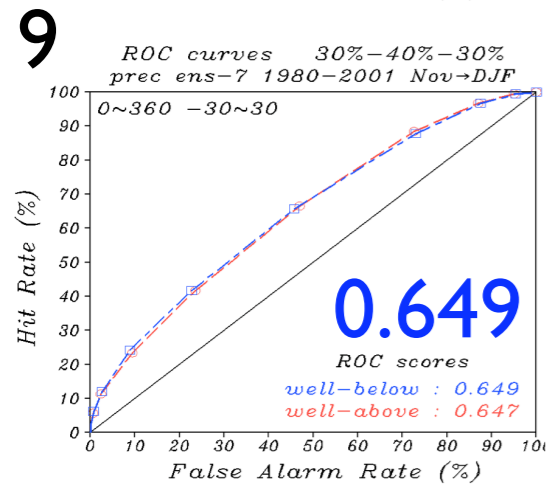
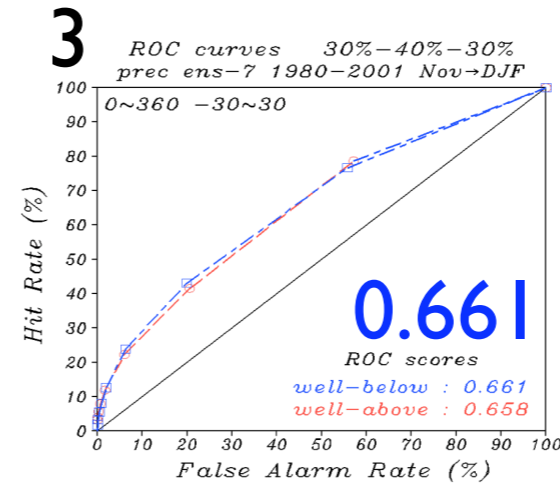
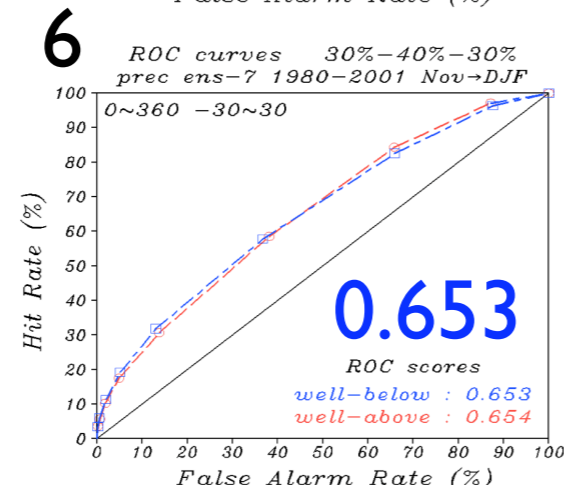
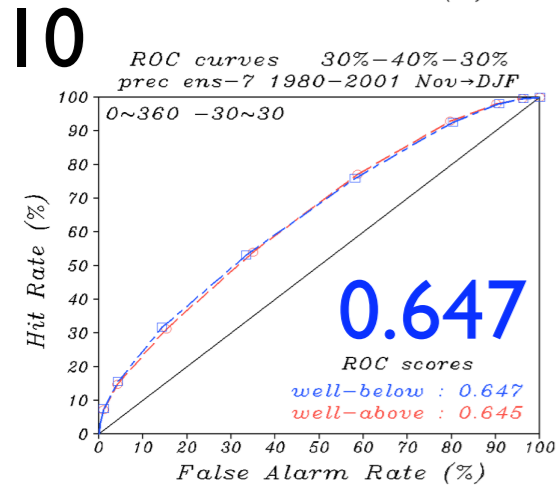
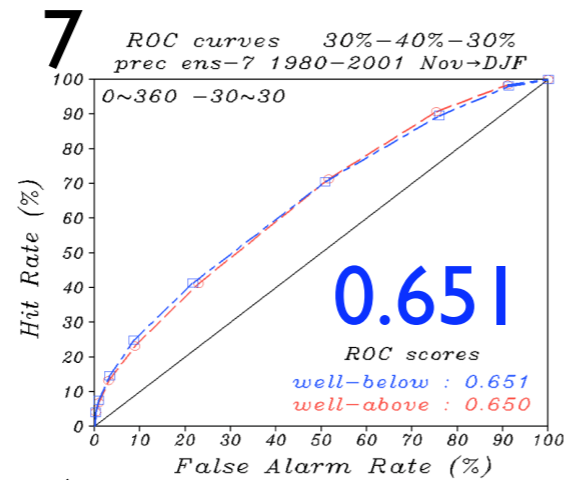
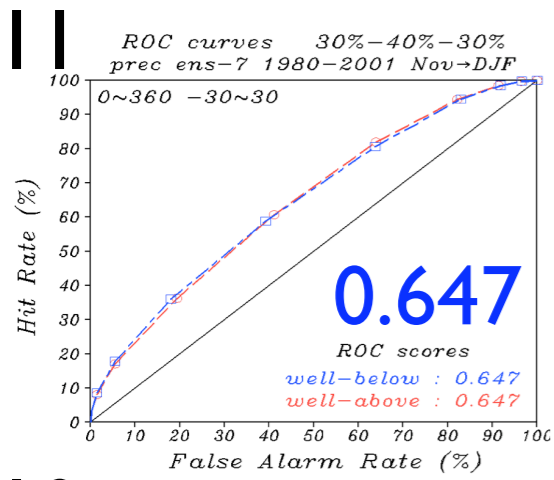
ETS score, May → JJA, Tropics

Below Normal → Well Below Normal



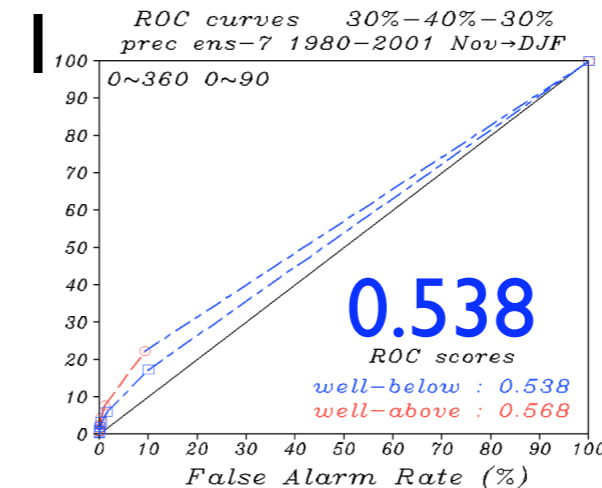
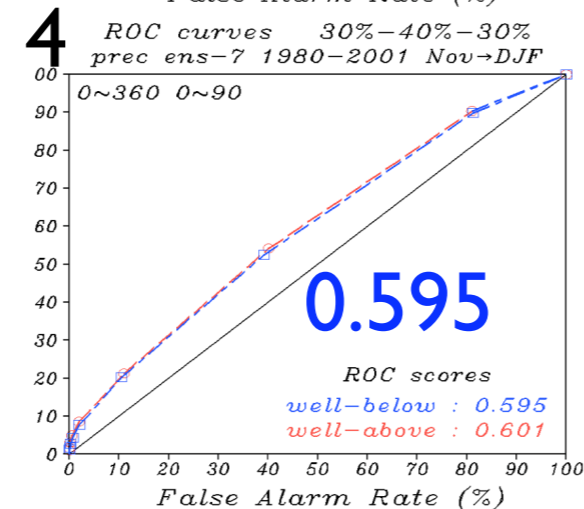
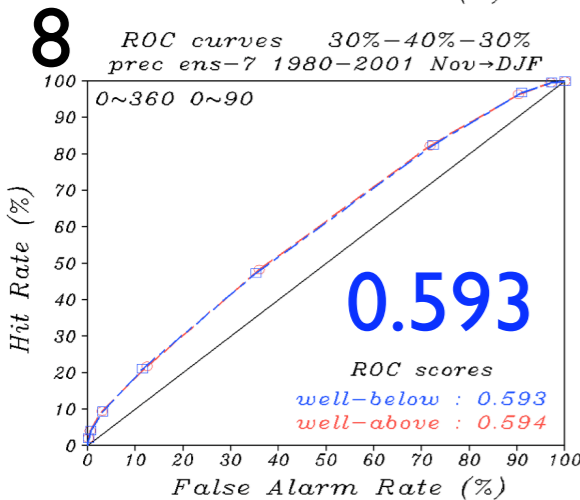
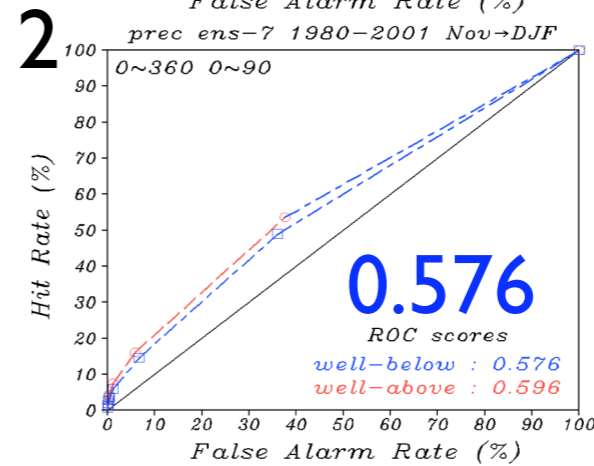
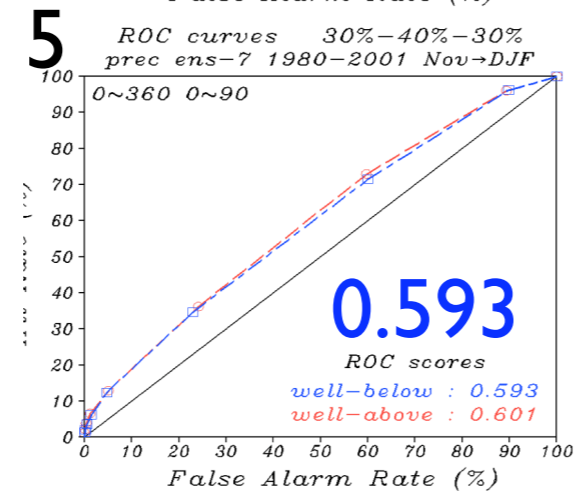
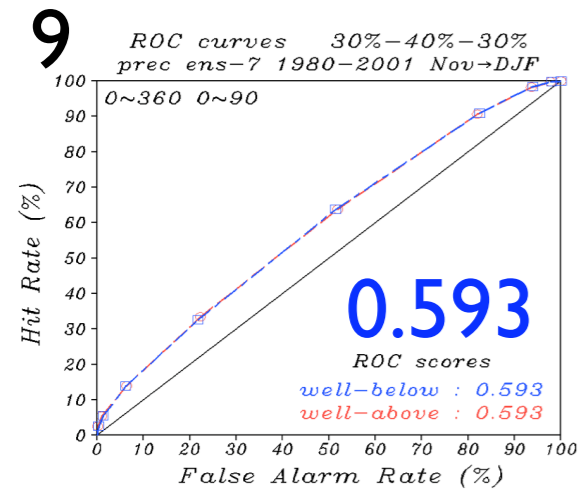
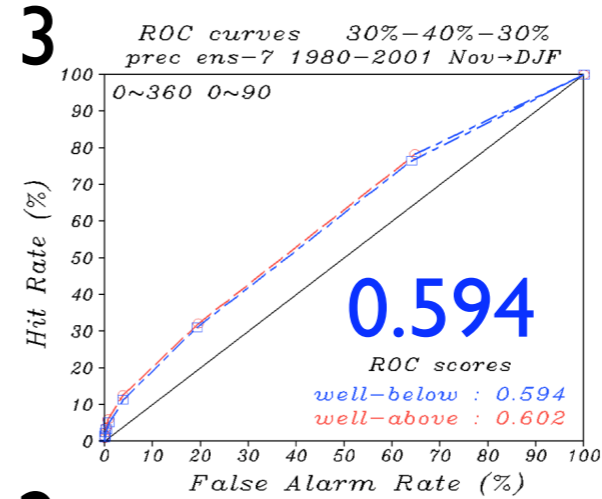
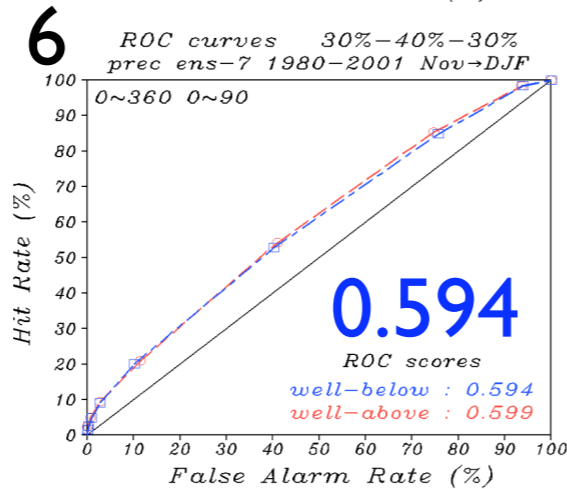
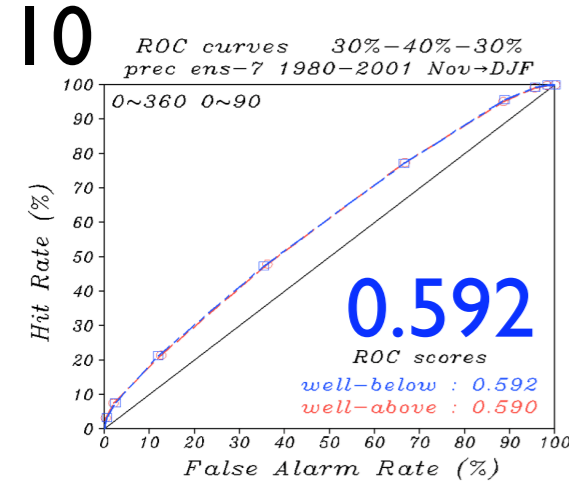
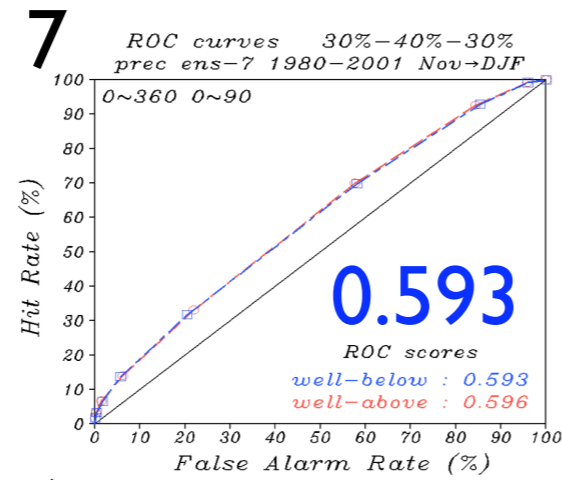
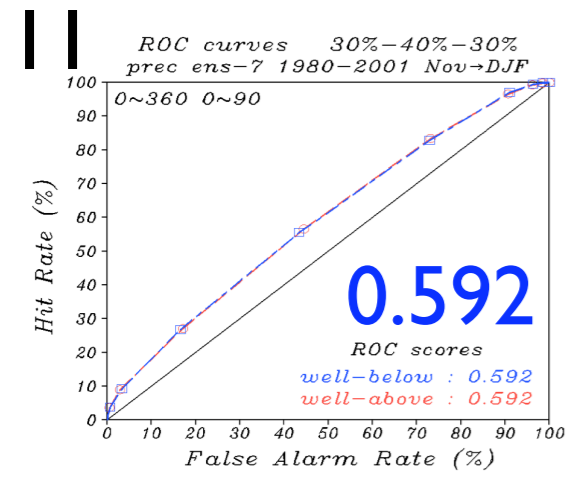
ROC score, Nov → DJF, Tropics

Below Normal → Well Below Normal



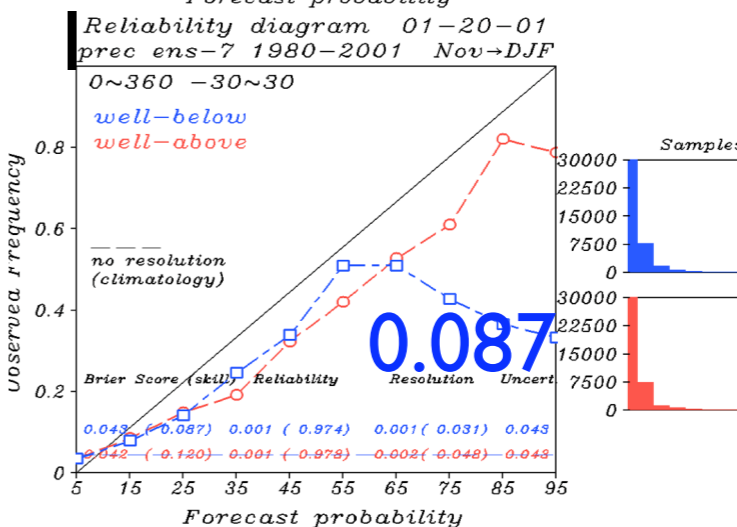
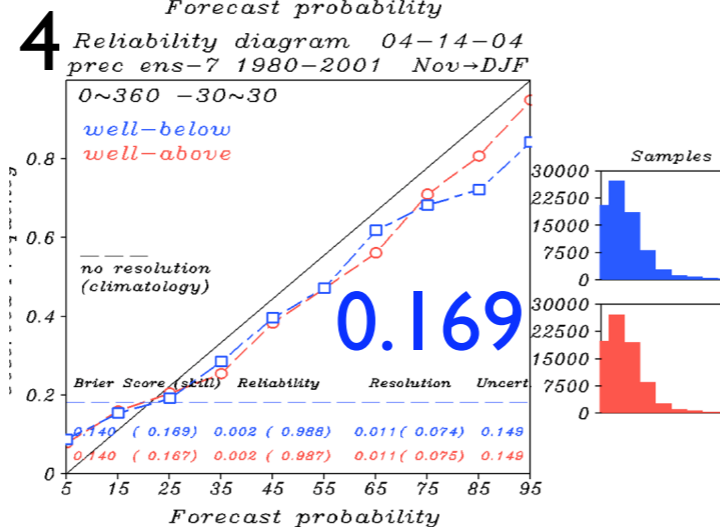
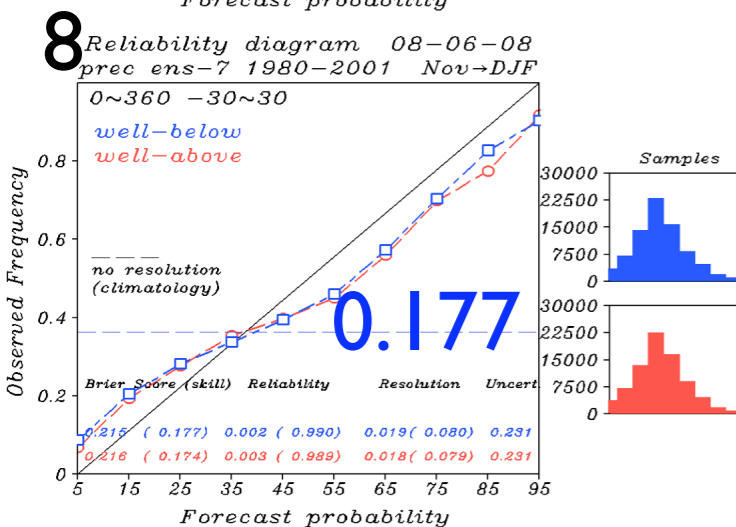
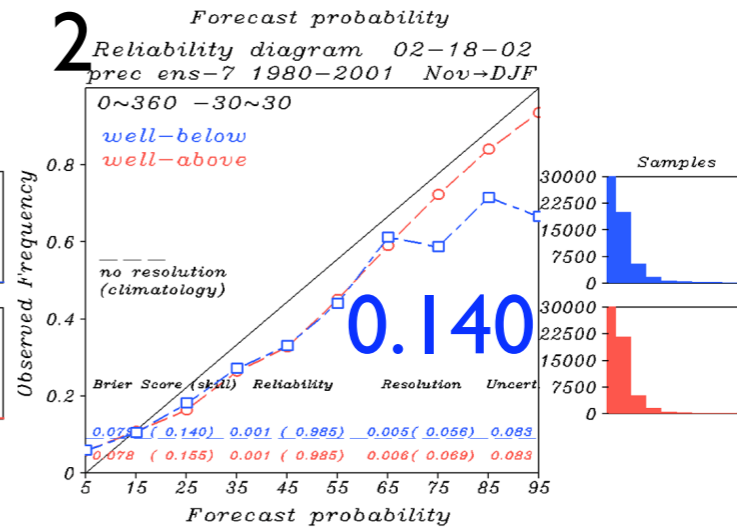
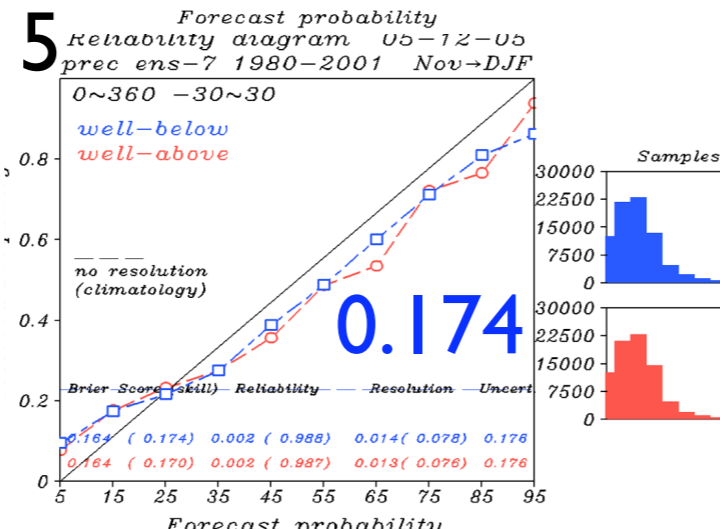
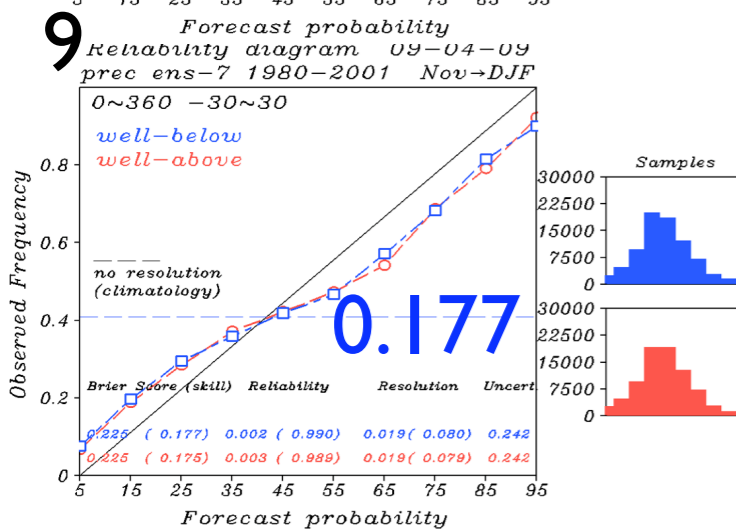
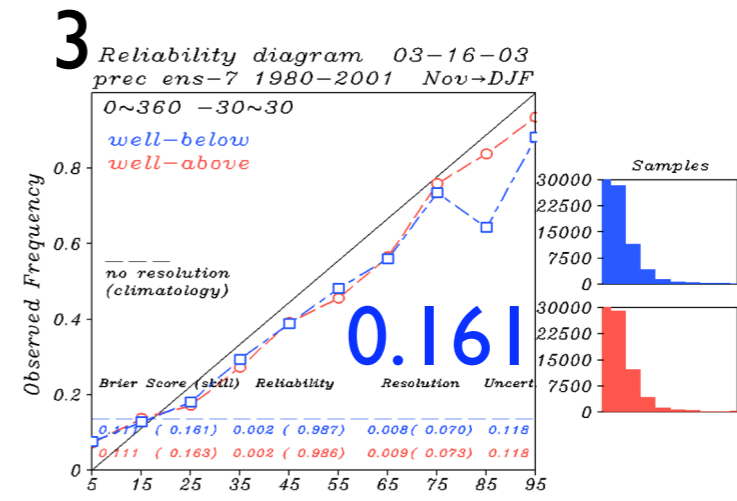
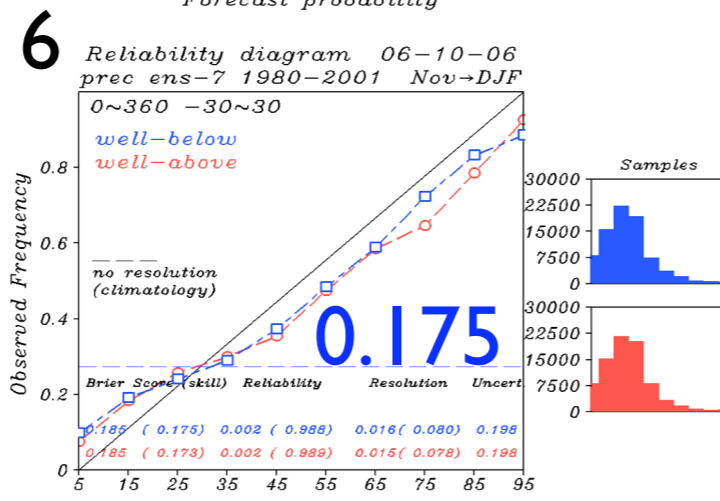
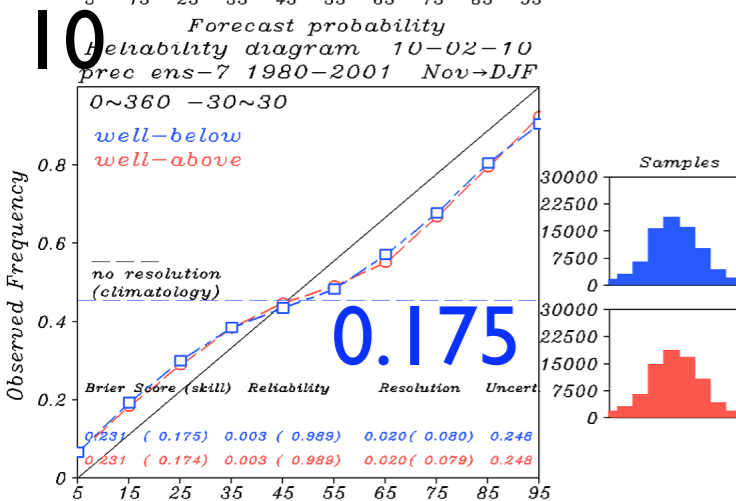
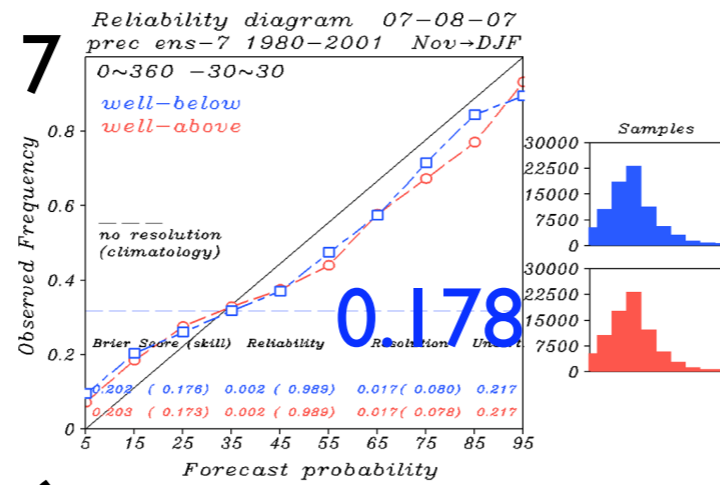
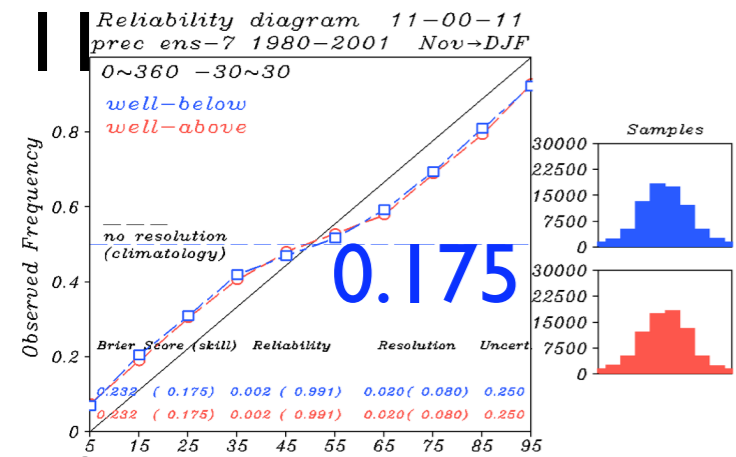
ROC score, Nov → DJF, Northern Hemisphere

Below Normal → Well Below Normal



Reliability diagram & freq. distribution

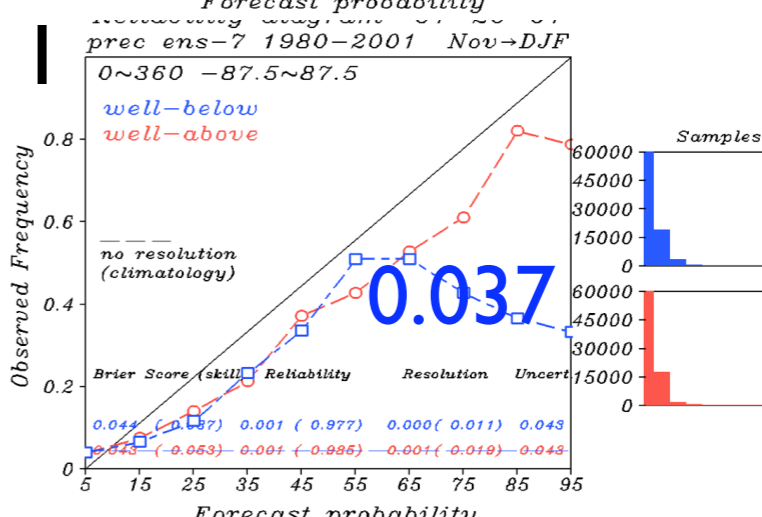
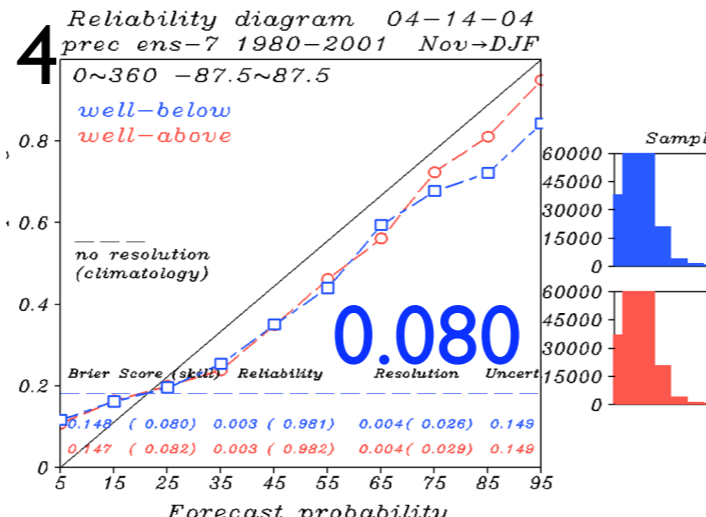
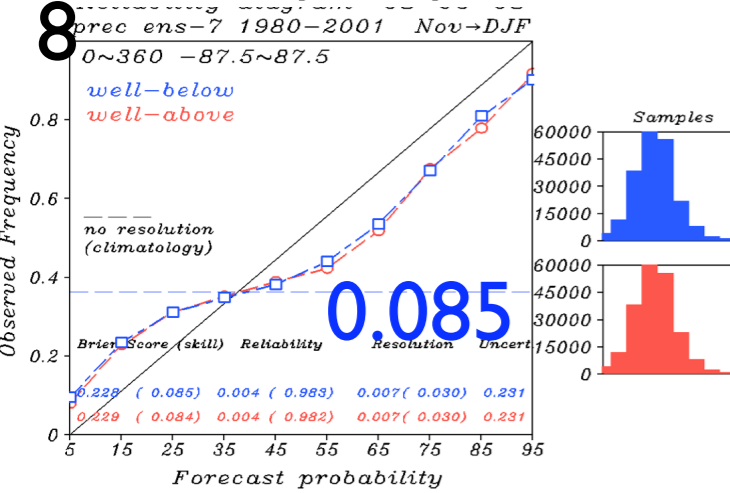
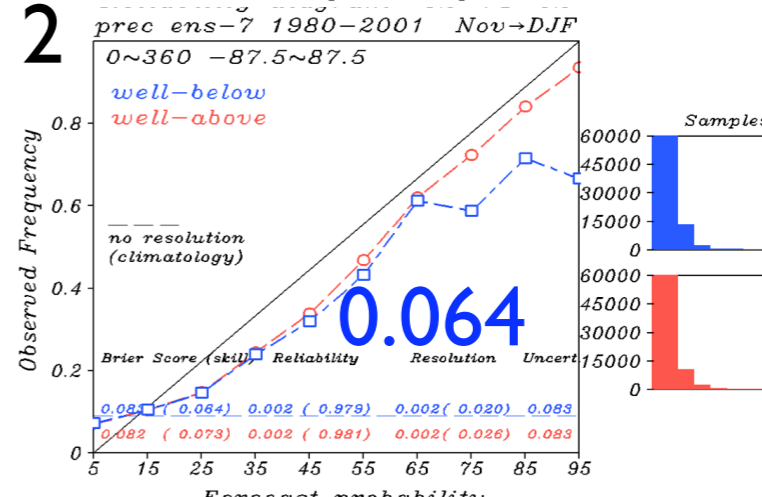
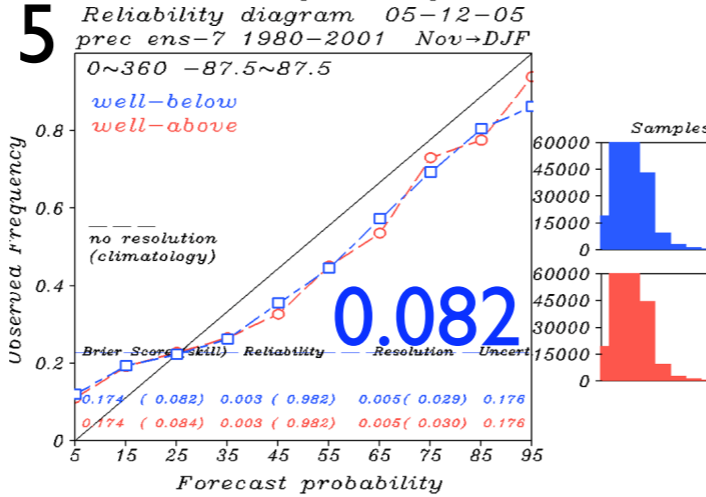
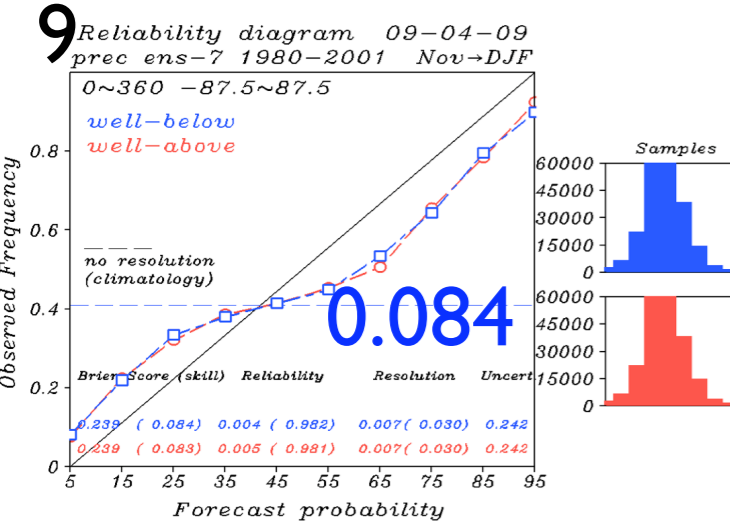
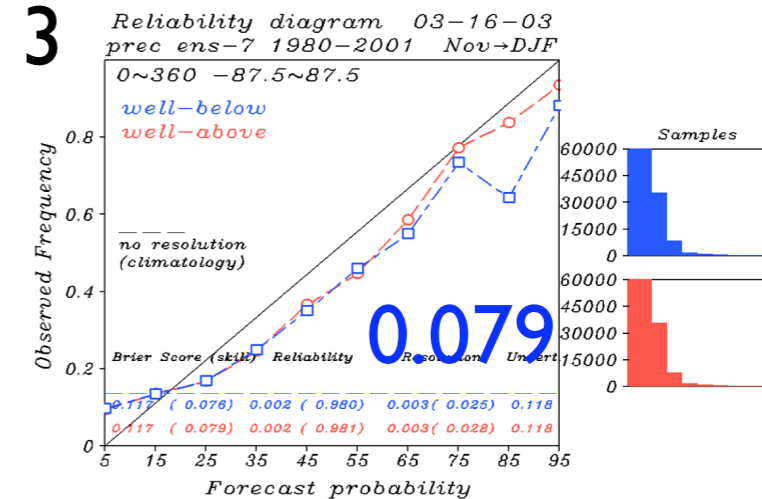
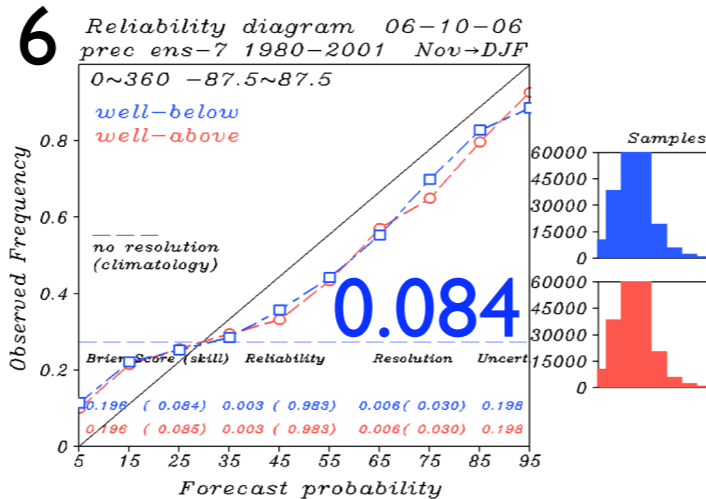
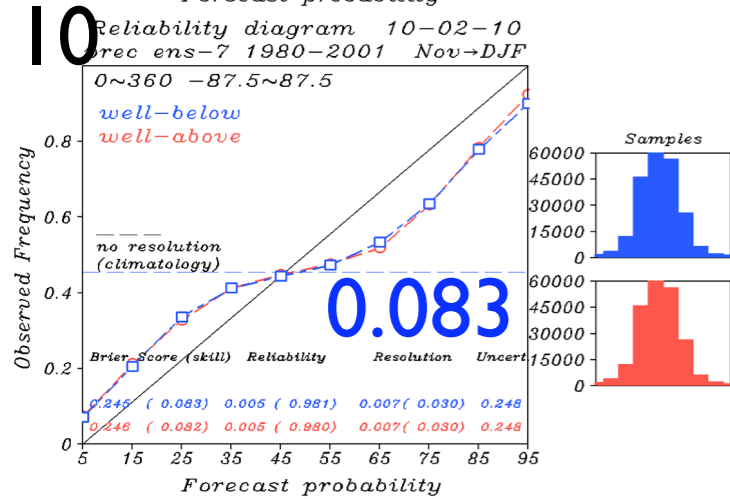
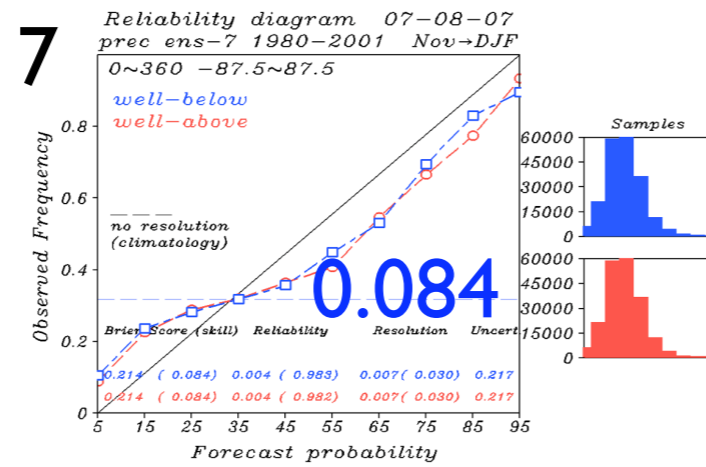
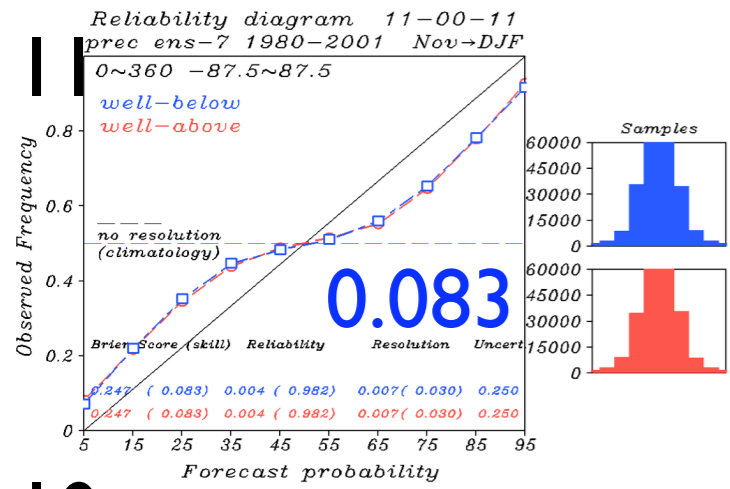
Brier skill score,
Nov → DJF, Tropics
Below Normal → Well Below Normal



Reliability diagram & freq. distribution

Brier skill score, Nov → DJF, Global

Below Normal → Well Below Normal



Concluding remarks (using DEMETER ensemble)

- For ETS (which category is more likely), forecast skill for relatively extreme seasonal condition in less than 20% of total distribution gradually decrease. But the threshold has seasonal and regional dependence.

Concluding remarks (using DEMETER ensemble)

- For ETS (which category is more likely), forecast skill for relatively extreme seasonal condition in less than 20% of total distribution gradually decrease. But the threshold has seasonal and regional dependence.
- For ROC score (resolution, whether forecast can discriminate events and non-events), forecast skill for relatively extreme seasonal condition gradually increase when the threshold is less than 15%. After that, skill start to drop down. Similar behaviors can be found in different regions and seasons.

Concluding remarks (using DEMETER ensemble)

- For ETS (which category is more likely), forecast skill for relatively extreme seasonal condition in less than 20% of total distribution gradually decrease. But the threshold has seasonal and regional dependence.
- For ROC score (resolution, whether forecast can discriminate events and non-events), forecast skill for relatively extreme seasonal condition gradually increase when the threshold is less than 15%. After that, skill start to drop down. Similar behaviors can be found in different regions and seasons.
- For Brier skill score (improvement of the probabilistic forecast relative to a reference forecast), forecast skill for relatively extreme seasonal condition remain similar up to 15-20% threshold, after that skill also start to diminish.