Development and testing of a multi-model ensemble prediction system for sub-monthly forecasts

PI: Andrew W. Robertson (IRI, Columbia)
co-PI: Michael K. Tippett (APAM, Columbia)
co-PI: Arun Kumar (CPC)

co-Investigators: Nicolas Vigaud, Lei Wang (IRI)

International Research Institute
for Climate and Society
EARTH INSTITUTE | COLUMBIA UNIVERSITY

NGGPS FFO PI Meeting
August 2-3, 2016 NCWCP, College Park, MD

NGGPS Columbia University Summary

- **Major Accomplishment in FY16:**
  - Archived subset of S2S forecast/re-forecast database in IRI Data Library (ECMWF, NCEP, CMA model subsets).
  - Evaluated ECMWF and NCEP week 1–4 re-forecast performance in weekly-averaged precip, 500hPa geopotential and 2m temperature (anomaly correlation, RME error).
  - Developed extended logistic regression calibration of sub-seasonal precip forecasts for ECMWF, NCEP and CMA models, together with a multi-model combination.

- **Priority Focus for FY17**
  - Improving extended logistic regression model for week 3-4 precipitation forecasts.
  - Diagnostics of ECMWF vs NCEP performance differences over the U.S.
  - Tailoring post-processing codes for NCEP usage.
- **Key Issue**
  - Developing well-calibrated sub-seasonal (week 3-4) probabilistic forecasts over the U.S.

Objectives

- **Quantify the sub-monthly hindcast skill of the CFSv2 and selected other individual models over the U.S.** in terms of: gridded fields of precipitation and temperature, as well as atmospheric indices such as the NAO and PNA; lead time and averaging range, including weekly averages in weeks 2–4; deterministic and probabilistic forecast skill metrics; and diagnostics of predictability.

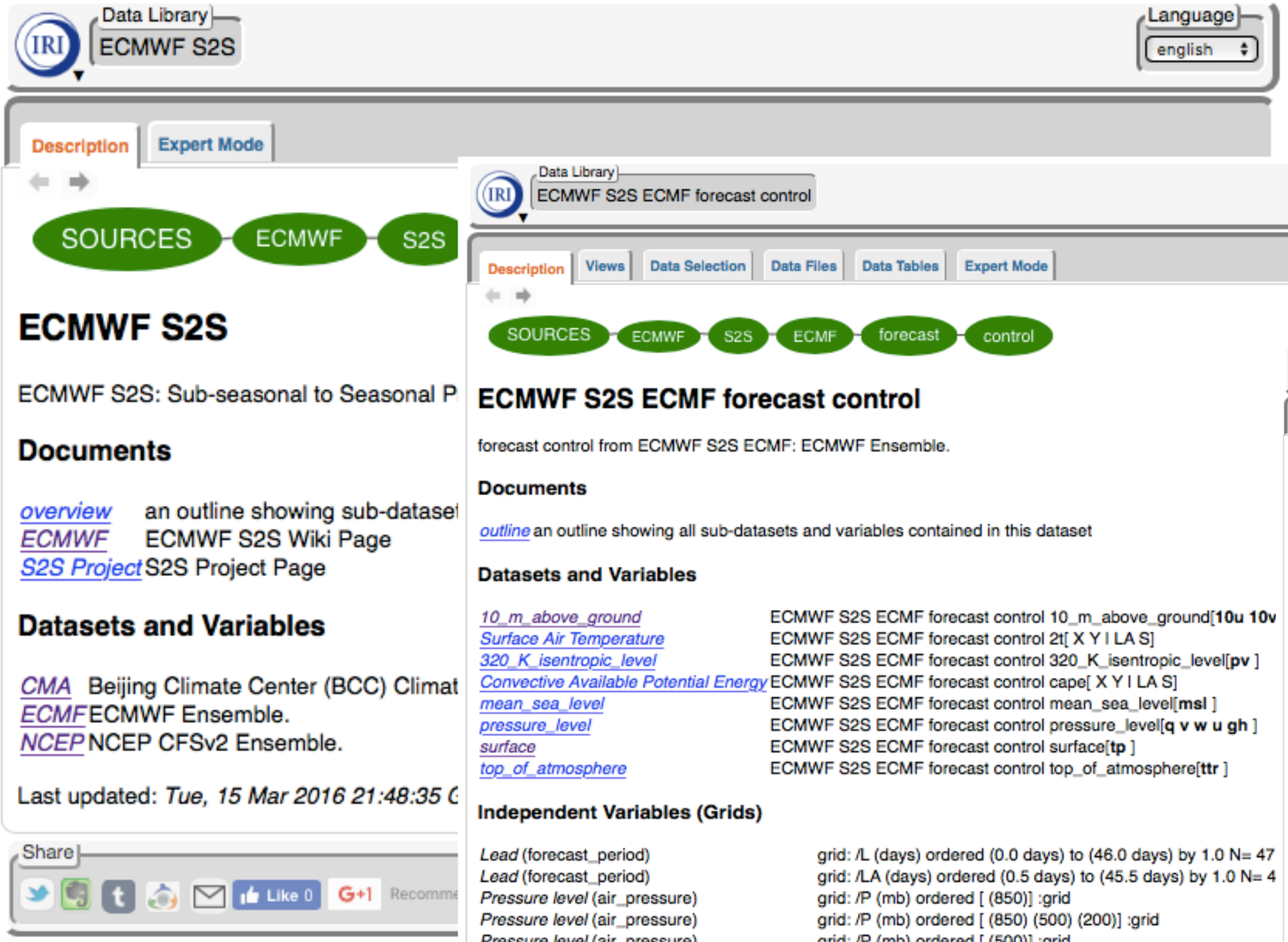
- **Develop MME methodology and evaluate the benefit of including an additional 1–3 models in a multi-model ensemble**, with focus over the U.S.

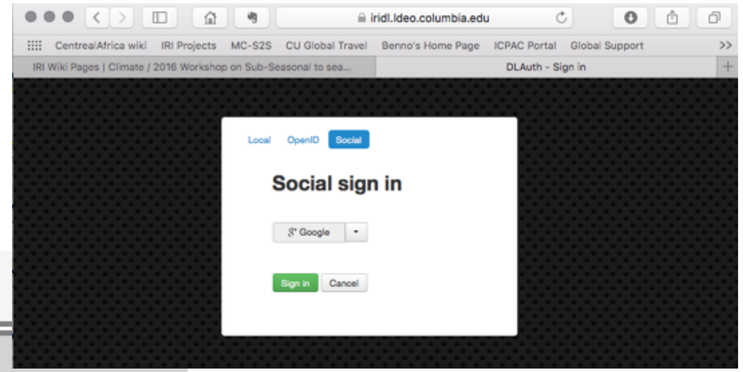
- **Improve physical understanding** of sub-monthly predictability over the U.S.

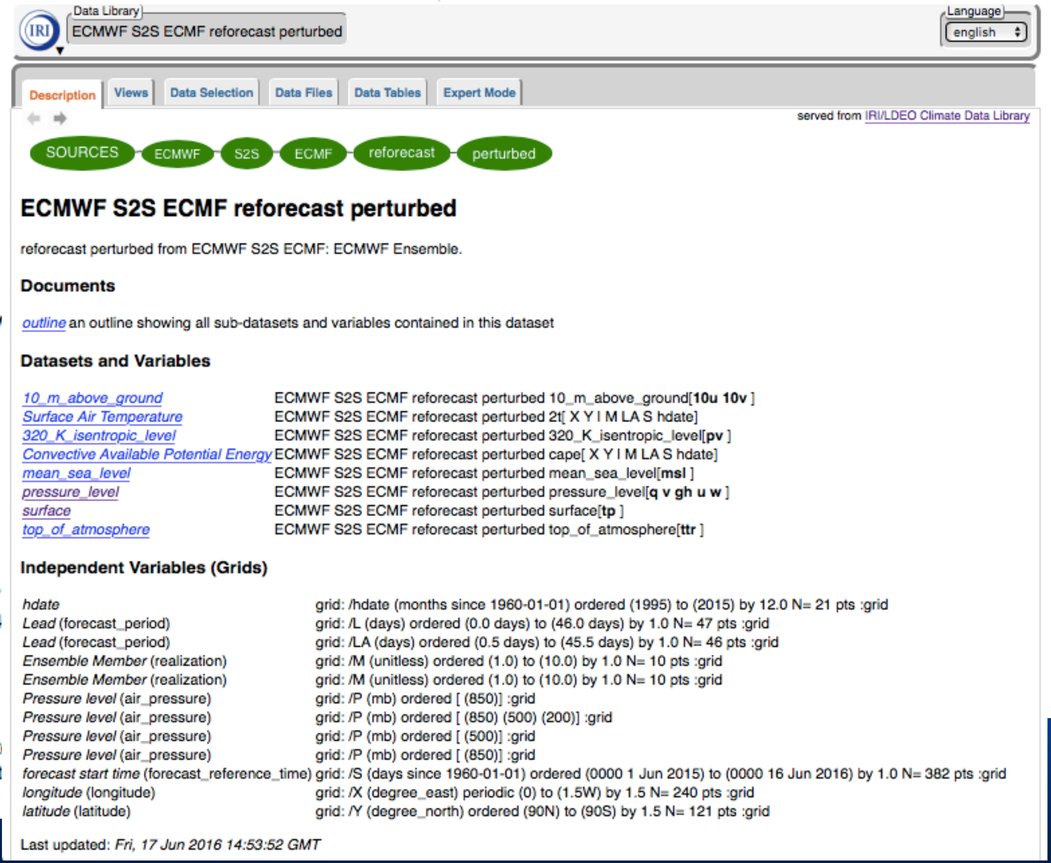
- **Establish the applicability of MME methods developed for weather/seasonal forecasts to the sub-monthly scale.**

- (Implement a real-time S2S MME at CPC, built using the most skillful and models that are available to CPC in real time.)

S2S Data @ IRIDL

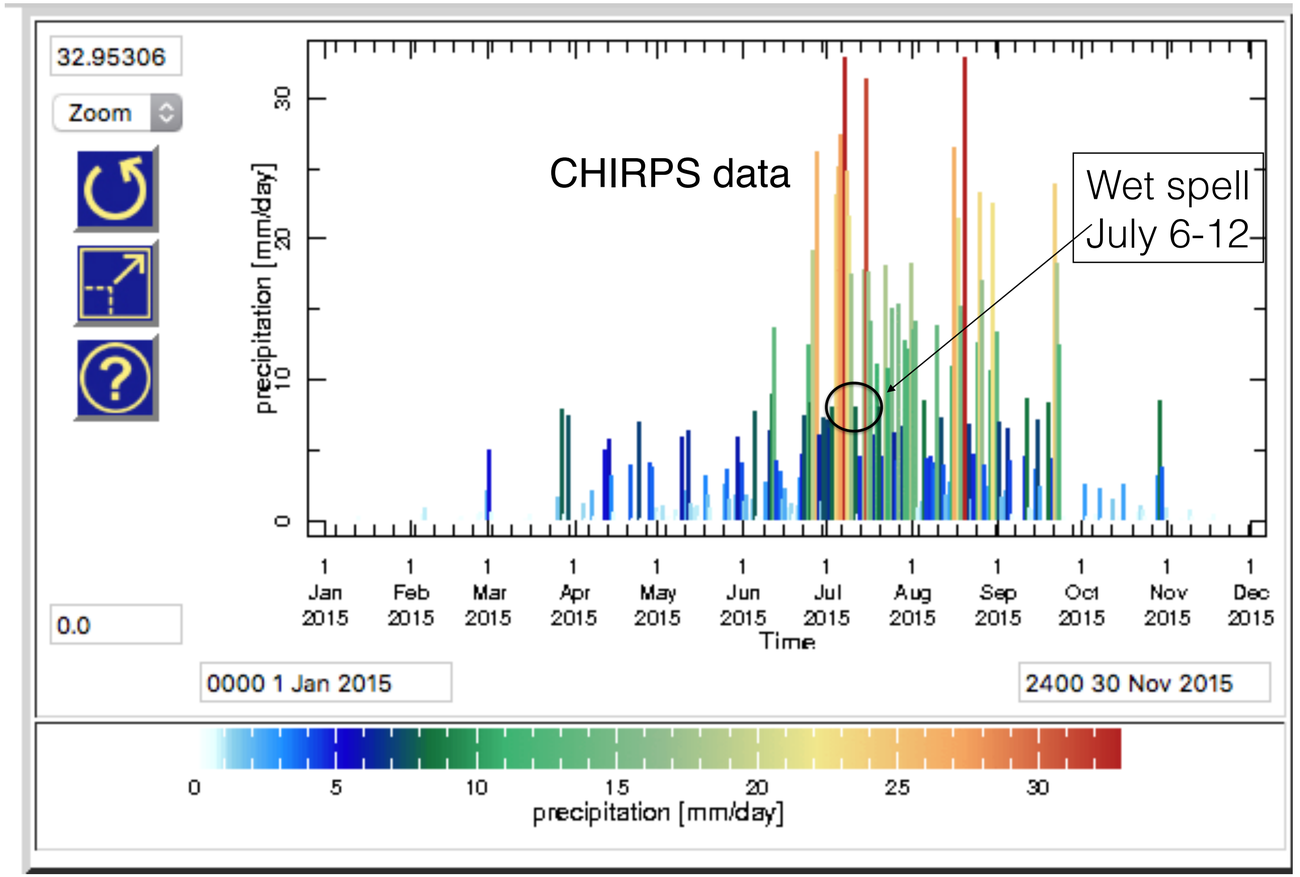


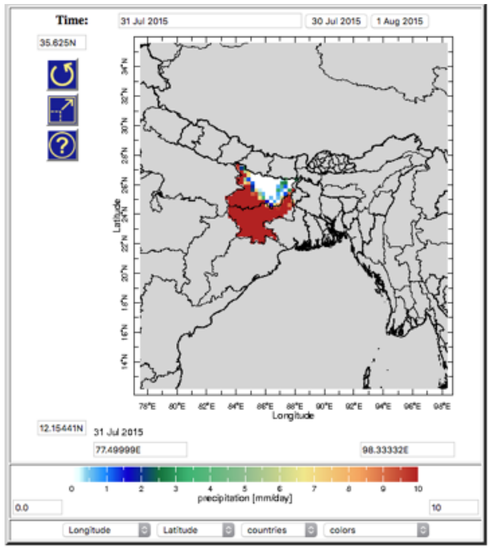




Example of S2S Data Base Use: Heavy rainfall over Bihar in 2015

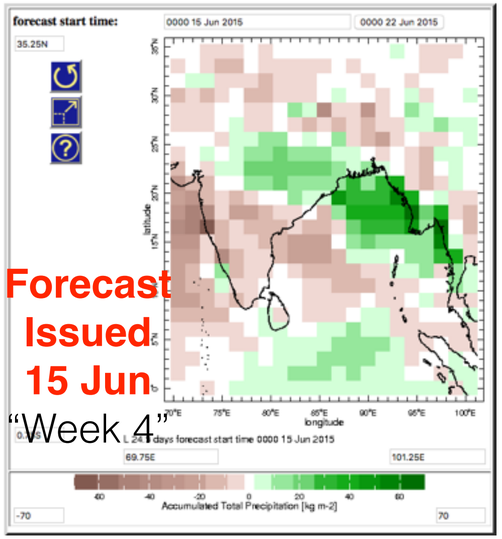
Can S2S Forecasts capture it?





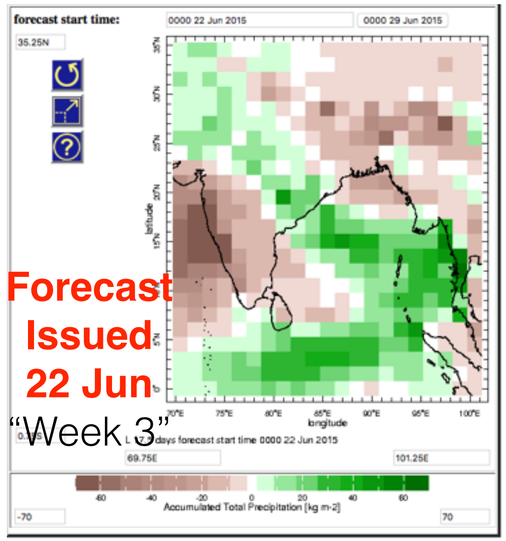
IRI Data Library

Diagnostics with S2S Database



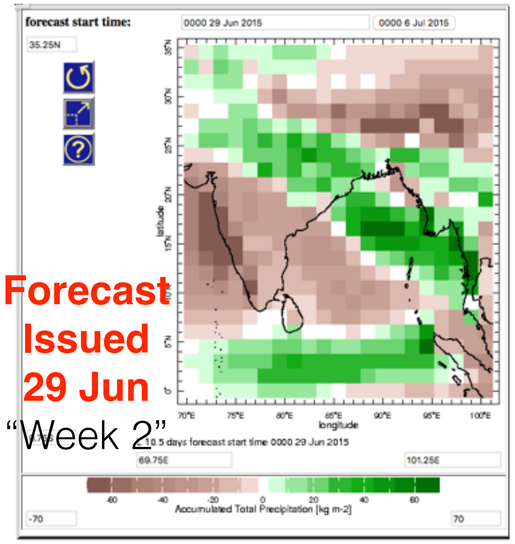
Forecast Issued 15 Jun

"Week 4"



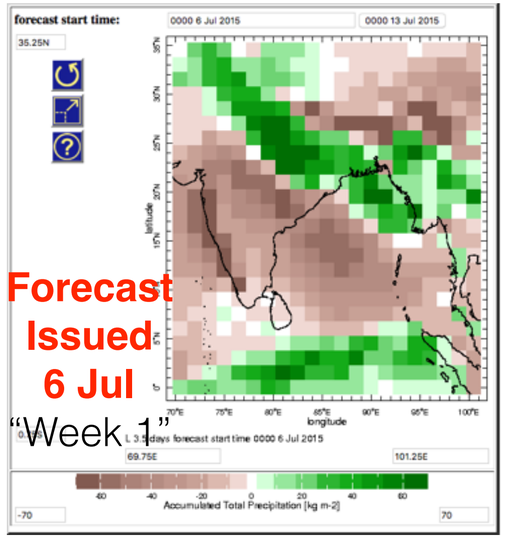
Forecast Issued 22 Jun

"Week 3"



Forecast Issued 29 Jun

"Week 2"



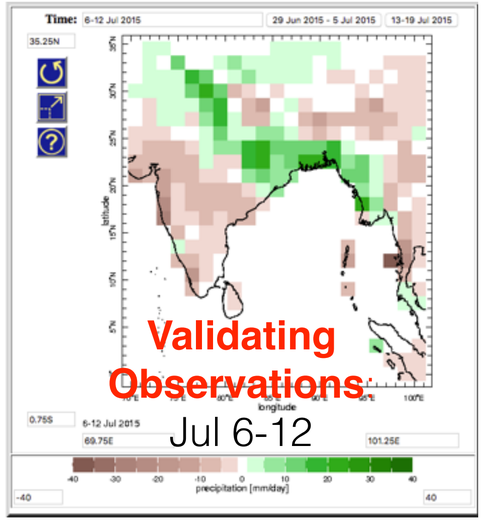
Forecast Issued 6 Jul

"Week 1"

ECMWF Forecasts

valid for Jul 6-12, 2015

Weekly average precip anomalies



Validating Observations:

Jul 6-12

IRI Data Library

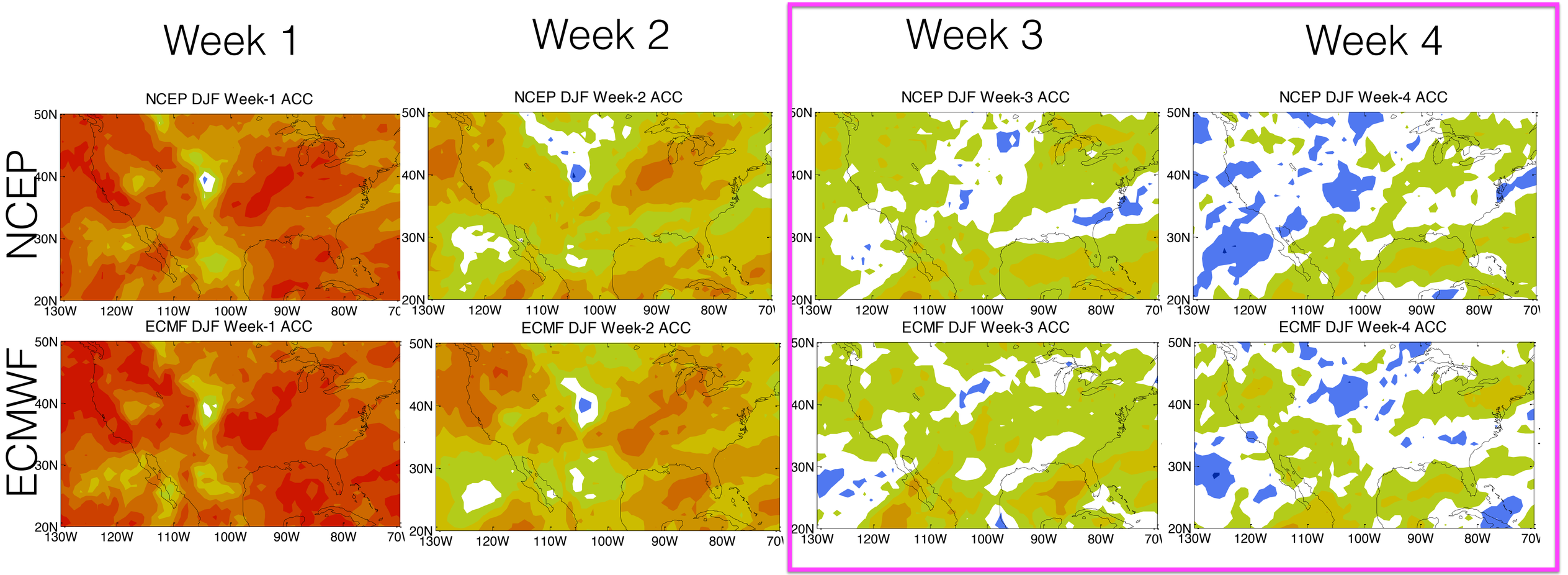
Model Configurations

Status on 10th March 2016	Time range	Resolution	Ens. Size	Frequency	Re-forecasts	Rfc length	Rfc frequency	Rfc size
BoM (ammc)	d 0-62	T47L17	33	2/week	fix	1981-2013	6/month	33
CMA (babj)	d 0-60	T106L40	4	daily	fix	1994-2014	daily	4
CNR-ISAC (isac)	d 0-31	0.75x0.56 L54	41	weekly	fix	1981-2010	every 5d	1
CNRM (lfpw)	d 0-32	T255L91	51	weekly	fix	1993-2014	2/monthly	15
ECCC (cwao)	d 0-32	0.45x0.45 L40	21	weekly	on the fly	1995-2014	weekly	4
ECMWF (ecmf)	d 0-46	Tco639/319 L91	51	2/week	on the fly	past 20 years	2/week	11
HMCR (rums)	d 0-61	1.1x1.4 L28	20	weekly	on the fly	1985-2010	weekly	10
JMA (rjtd)	d 0-33	T319L60	25	2/week	fix	1981-2010	3/month	5
KMA (rksl)	d 0-60	N216L85	4	daily	on the fly	1996-2009	4/month	3
NCEP (kwbc)	d 0-44	T126L64	16	daily	fix	1999-2010	day	4
UKMO (egrr)	d 0-60	N216L85	4	daily	on the fly	1996-2009	4/month	3

Evaluation of sub-monthly hindcast skill of ECMWF & CFSv2

- Anomaly correlation and RMS error for 1999–2010 common hindcast period
- By trimester of forecast **starts**: JJA, SON, DJF, MAM
- ECMWF Mon & Thurs starts
- NCEP 3-day lagged ensembles (Sat-Mon, Tue-Thu)

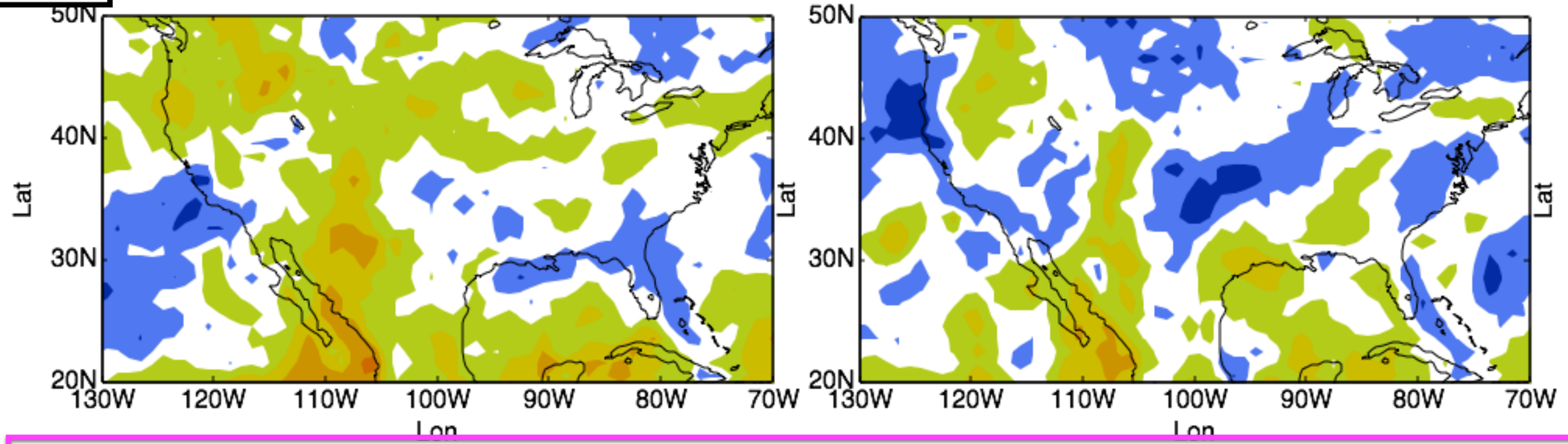
Weekly Precip Skill vs Lead time – DJF



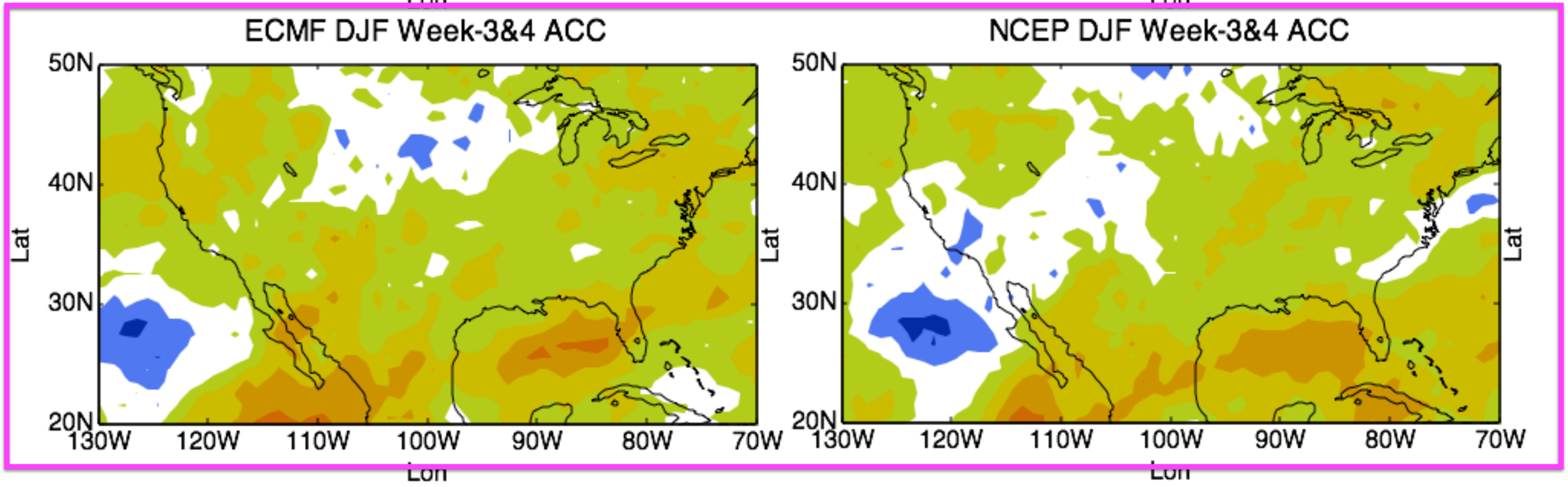
Weeks 3+4

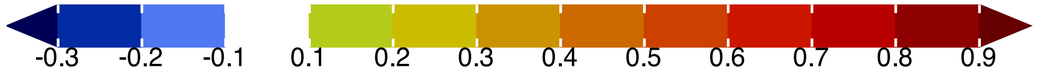
ECMWF vs **NCEP**

JJA



DJF

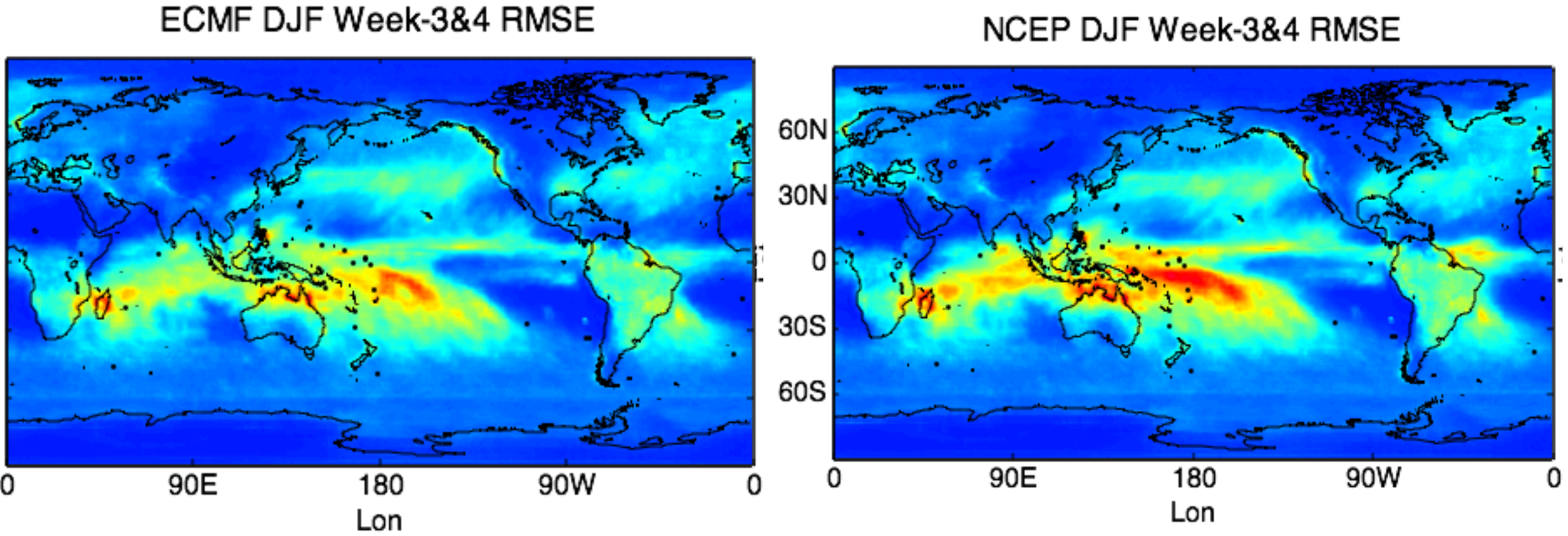




Prcp Refcst Anomaly Correlation 1999-2010

Gain in skill over Wks 3 & 4 individually

RMS Precip Error - DJF



Trop precip errors are larger in NCEP

Seasonality of W3+4 skill

Precip, 500hPa Geopotential, 2m Temperature

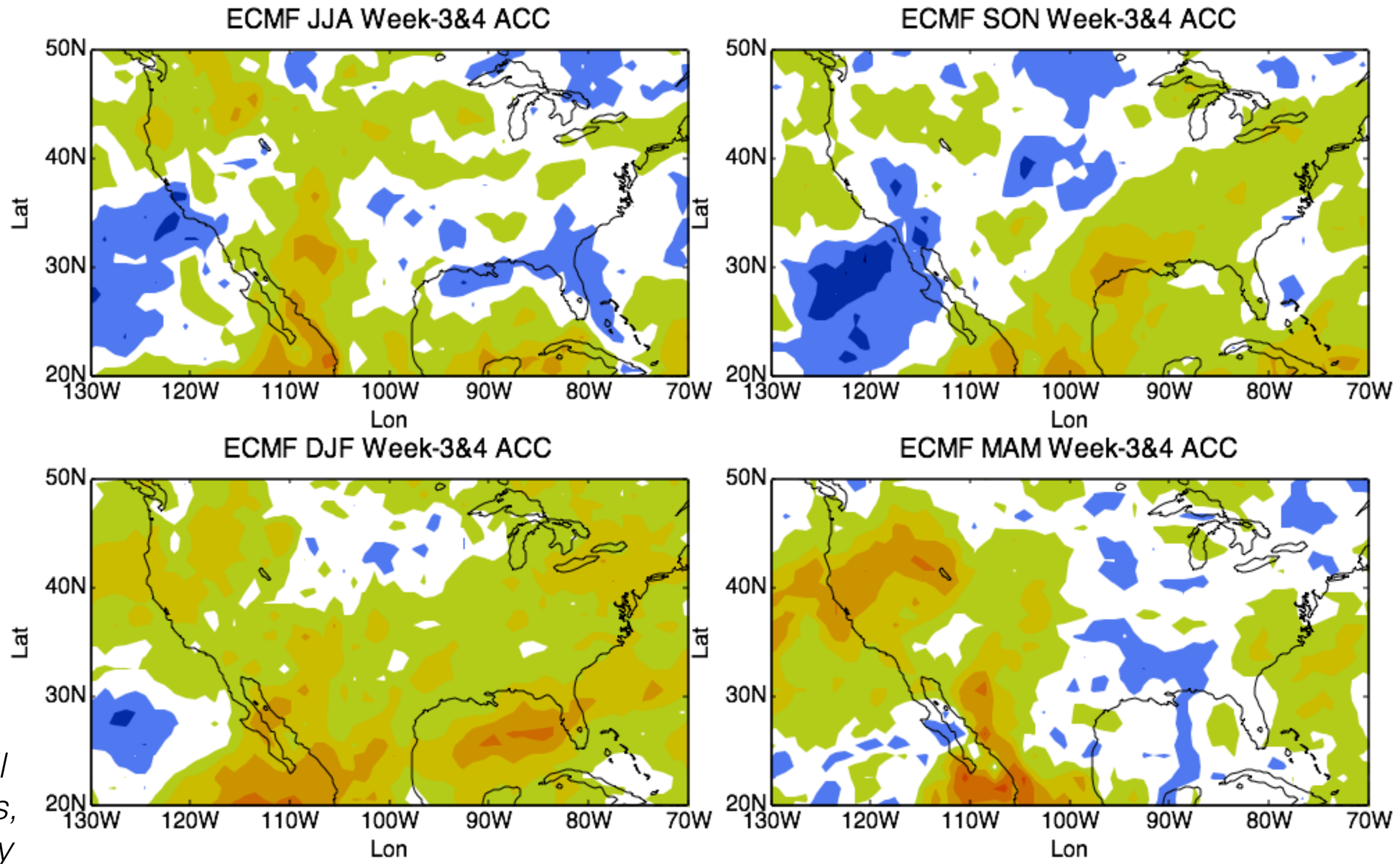
ECMWF Precip ACC Seasonality

JJA

SON

DJF

MAM



Areas of skill in all seasons, although only reach ~0.4.

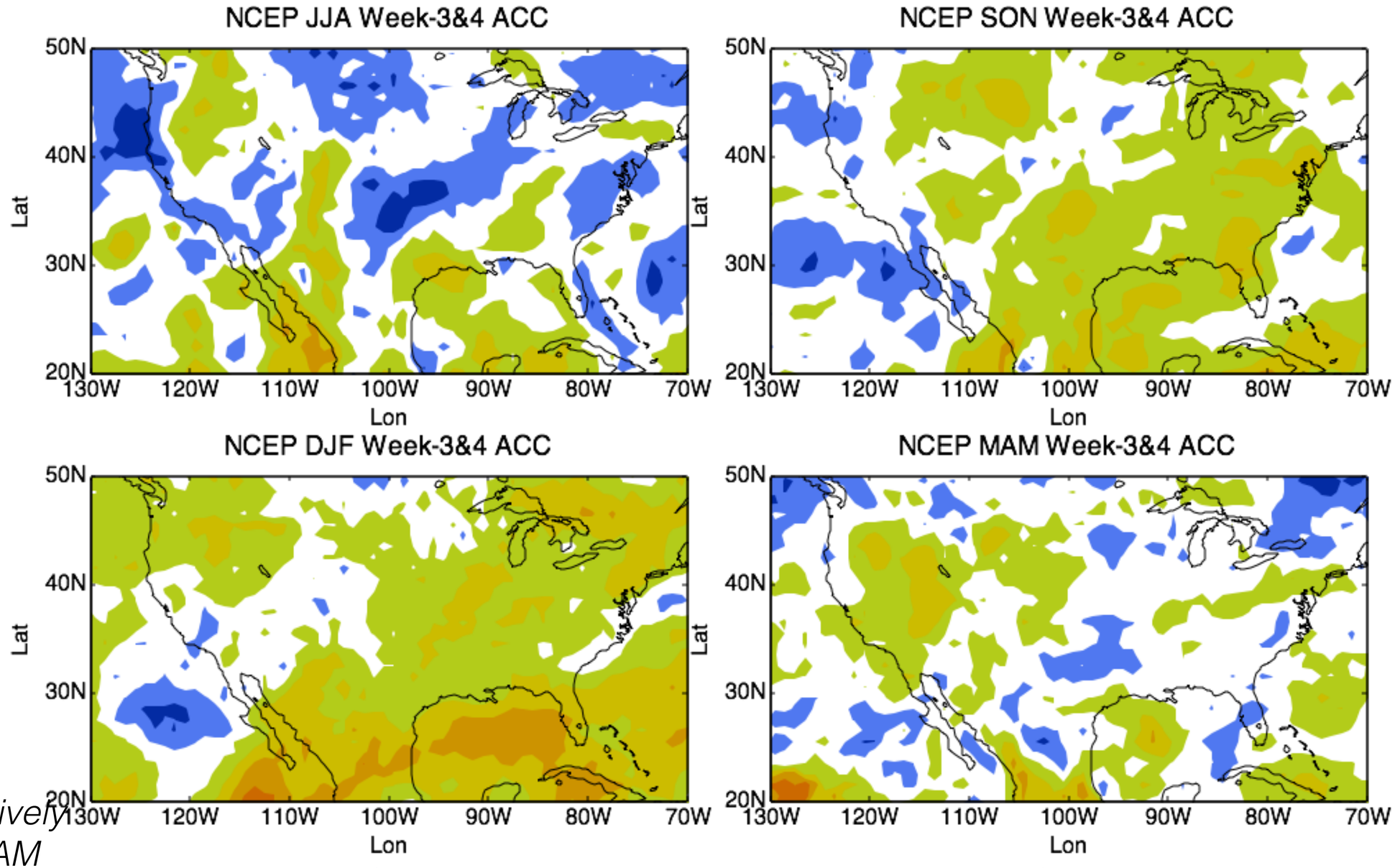
NCEP Precip ACC Seasonality

JJA

SON

DJF

MAM



NCEP is relatively worse in MAM

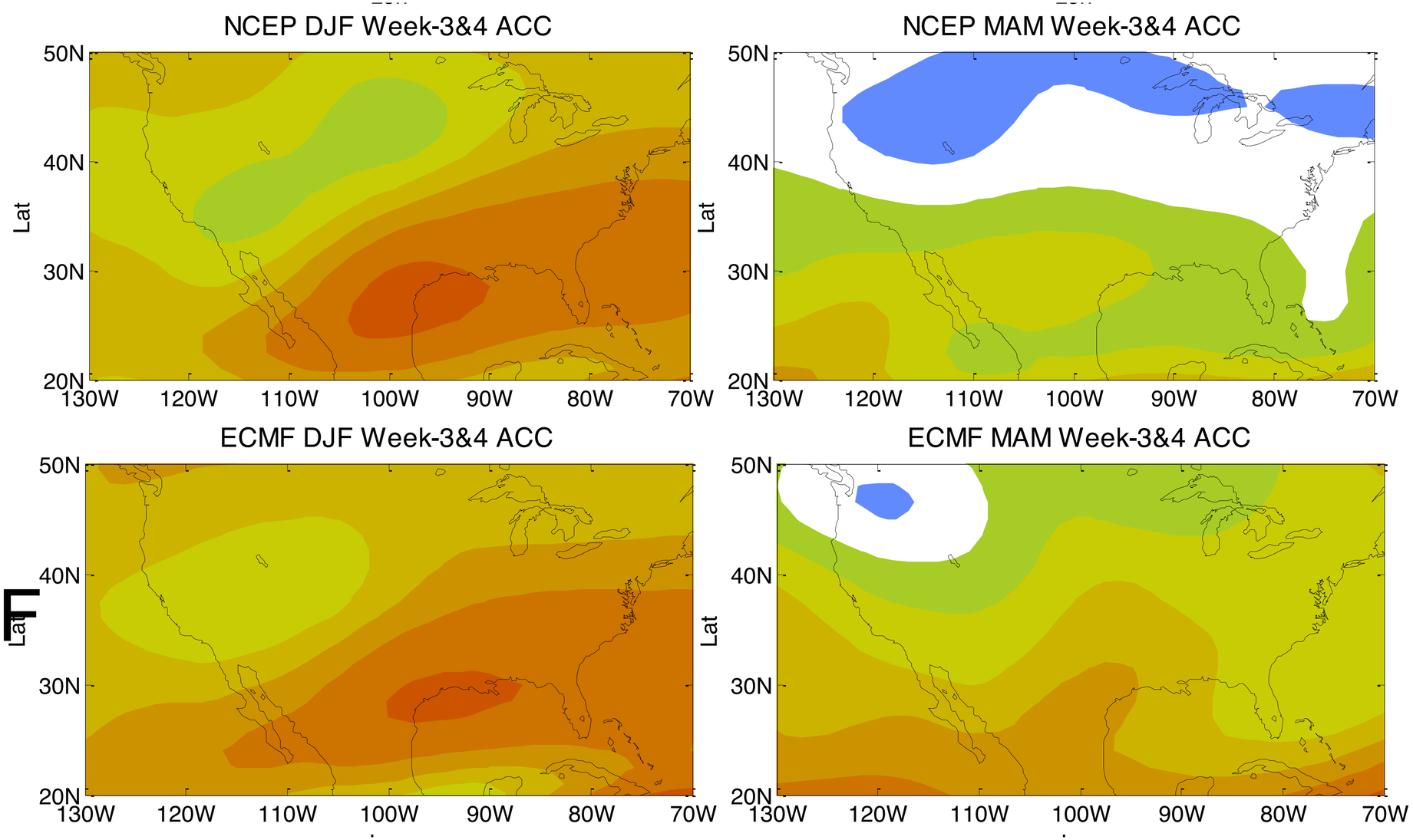
NCEP vs ECMWF 500 hPa Geopotential

DJF

MAM

NCEP

ECMWF



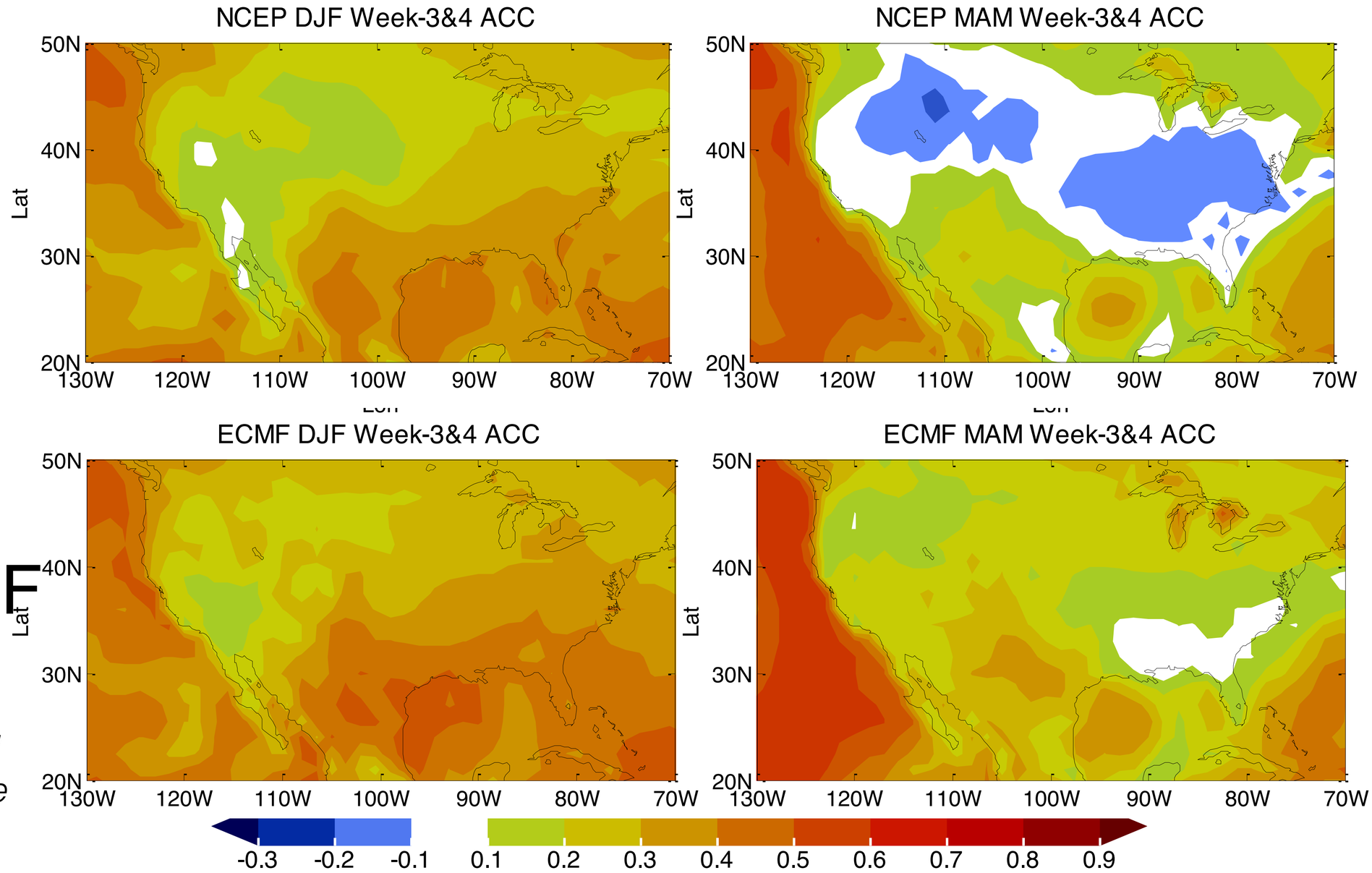
NCEP vs ECMWF 2m Temperature

DJF

MAM

NCEP

ECMWF



NCEP MAM performance flagged

Probabilistic Forecast Calibration using Quantile Regression

- **Distributional Regression** aims to provide the **conditional distribution** of the response variable given a set of explanatory variables

- **Quantile regression** is a reduced form in which the predictand is a quantile of the forecast PDF. **Logistic regression** is well suited to predicting a probability rather than a measurable physical quantity

$$\ln\left[\frac{p}{1-p}\right] = f(\mathbf{x}) \qquad p = \Pr\{V \leq q\}$$

p is the probability of not exceeding quantile q
This equation is linear on the logistic, or log-odds scale

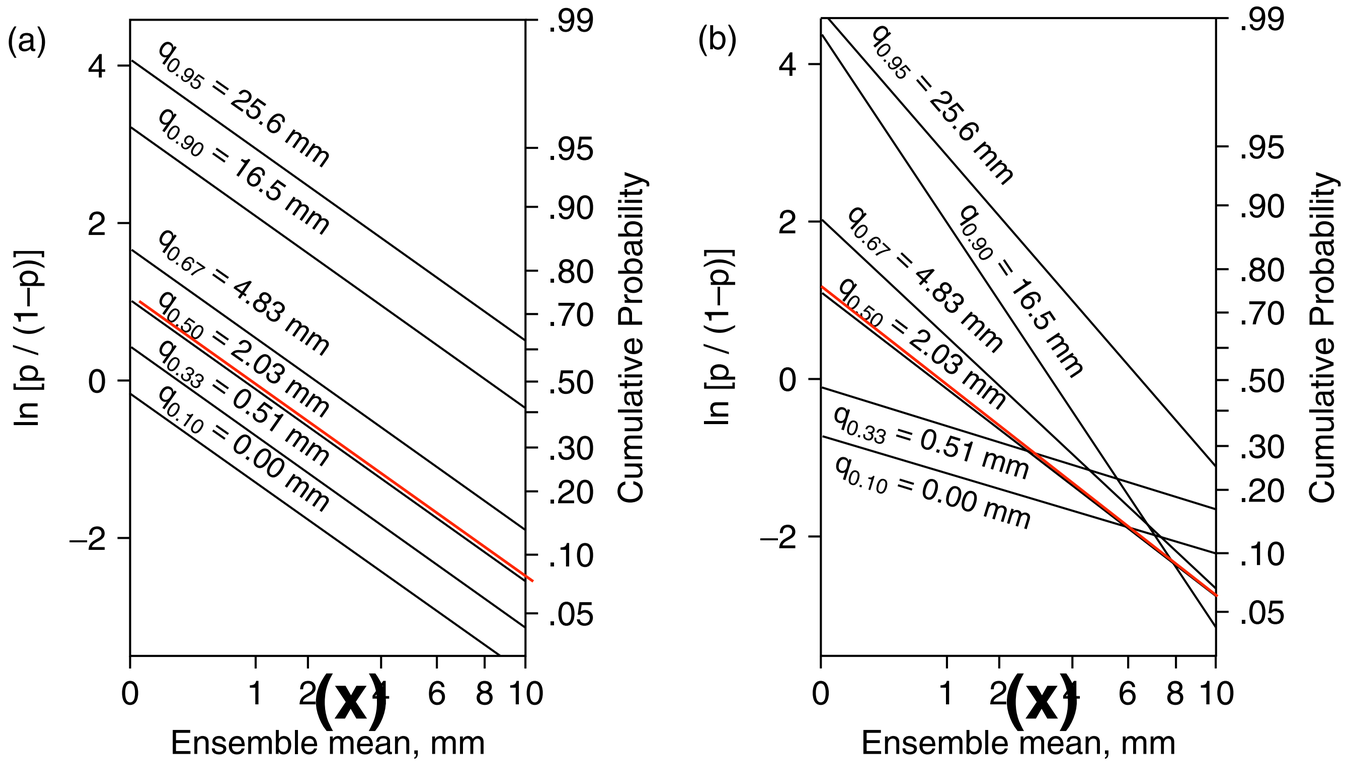
Extended Logistic Regression (ELR)

GFS Day 6–10 Precip
Forecast for Minneapolis
28 Nov – 2 Dec 2001

Wilks (2009)

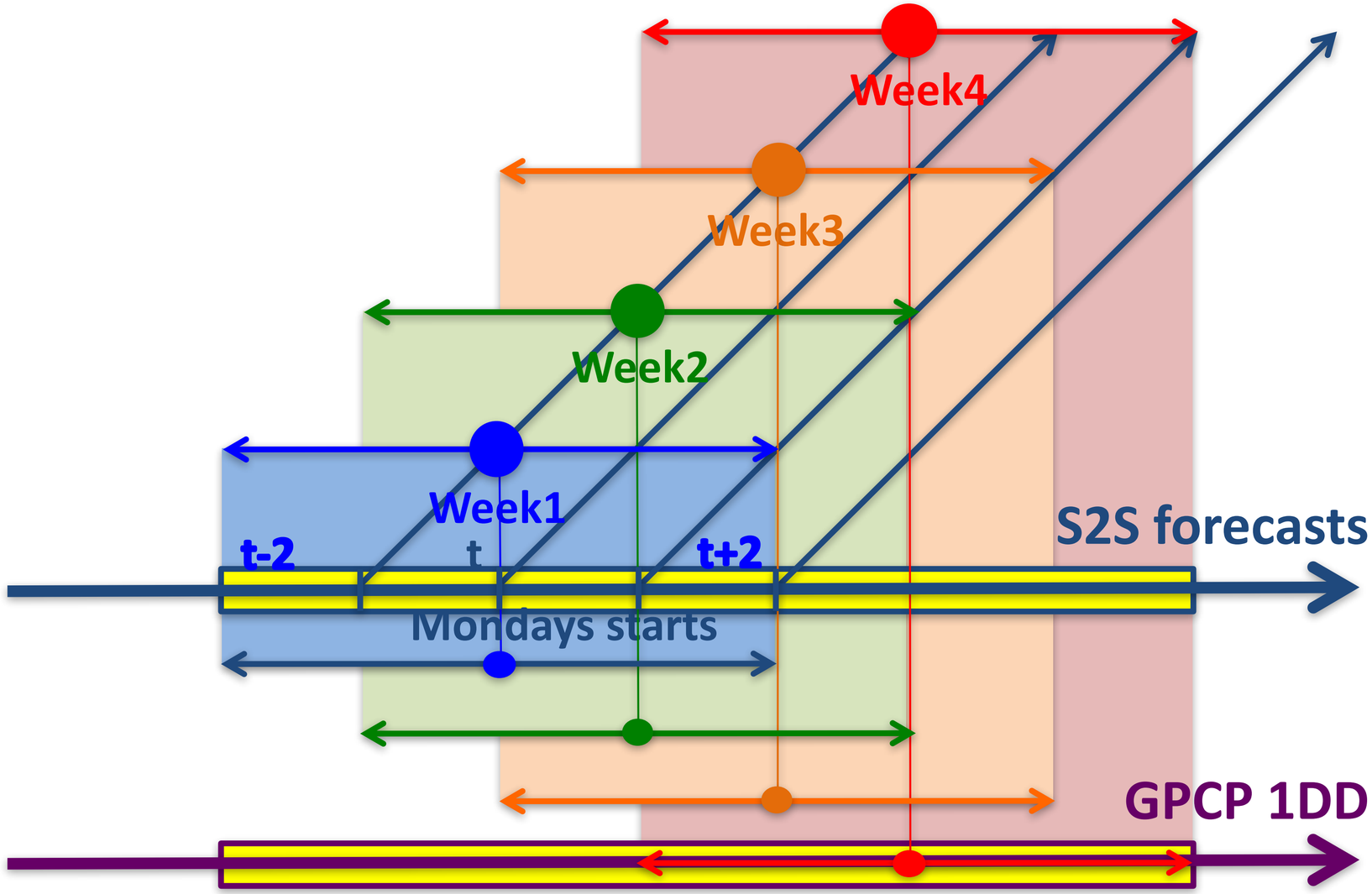
$$\ln\left[\frac{p(q)}{1-p(q)}\right] = f(\mathbf{x}) + g(q)$$

$$\ln\left[\frac{p}{1-p}\right] = f(\mathbf{x})$$



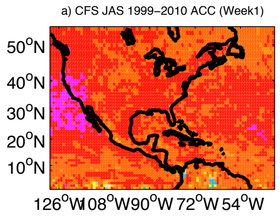
Extended Logistic Regression Set-up

1. Separate ELR for each grid point of the 10x10 GPCP 1° grid over N Am. (y)
2. Ensemble mean forecasts as x
3. Trained on weekly terciles computed from 5-weeks multi-year averages
4. Leave-one-year-out cross-validation, 1999-2010
5. Done for 1x weekly starts and from each GCM separately (11 ECMF, 4 NCEP, 4 CMA members)

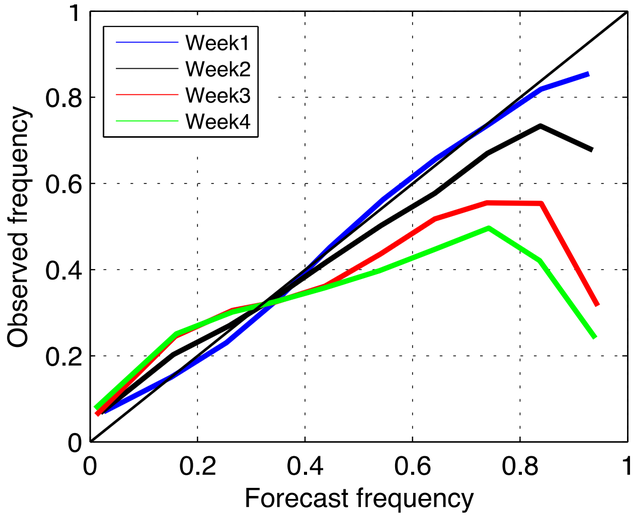


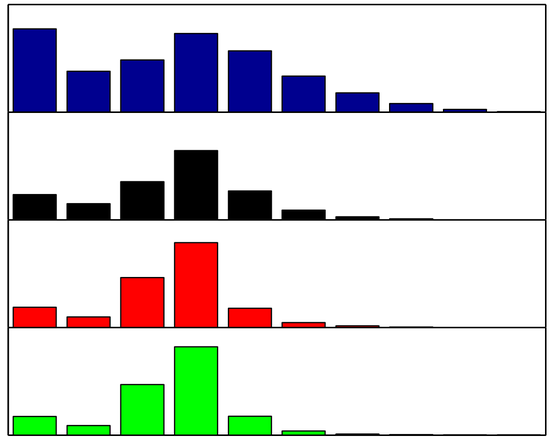
ECMF Reliability test: JFM 1999-2010

Diagrams below have been computed for all weekly starts of the JFM seasons from the 1999-2010 period using all gridpoints within N American sector (ocean+land points)

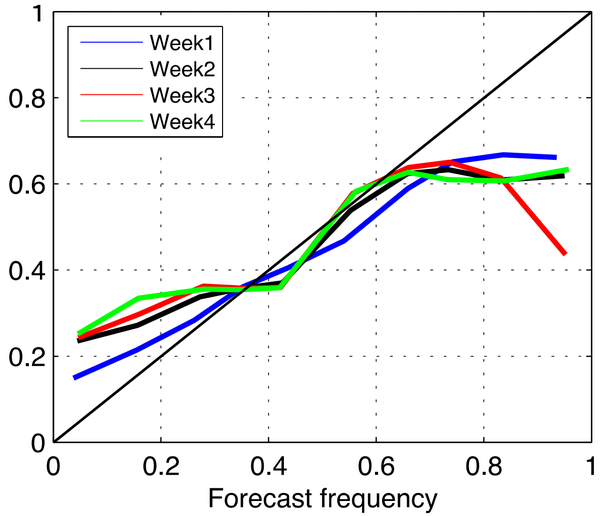


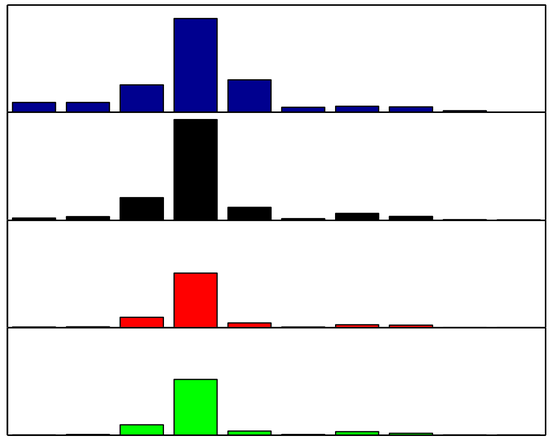
Below Normal



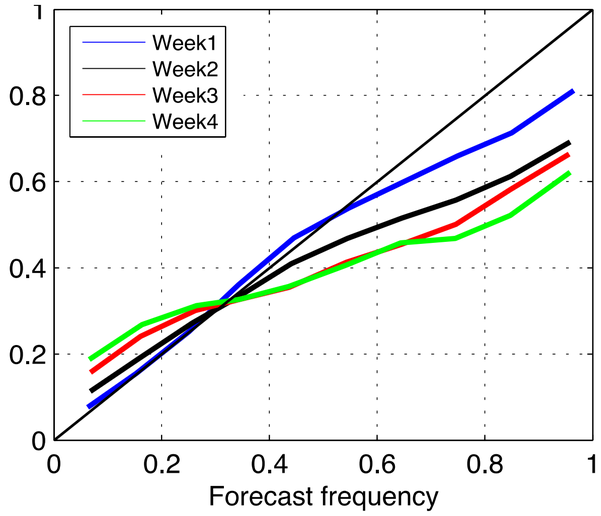


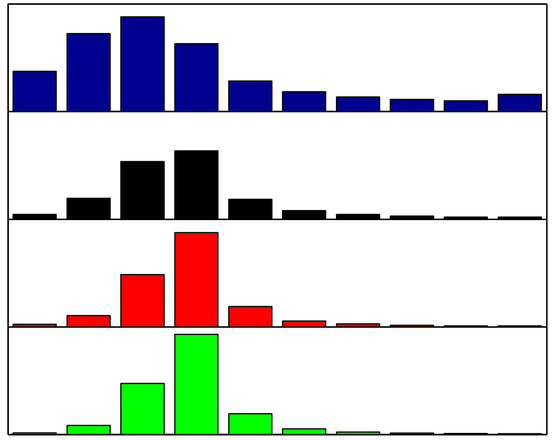
Normal





Above Normal

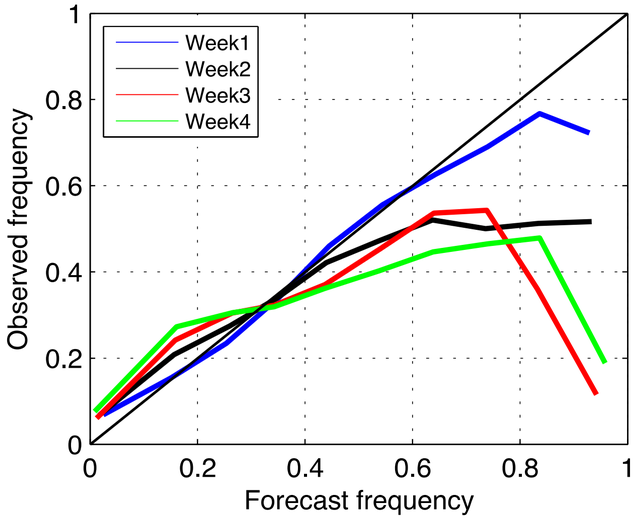


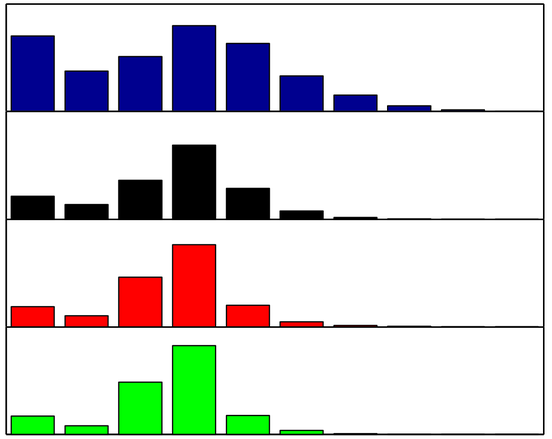


CFS Reliability test: JFM 1999-2010

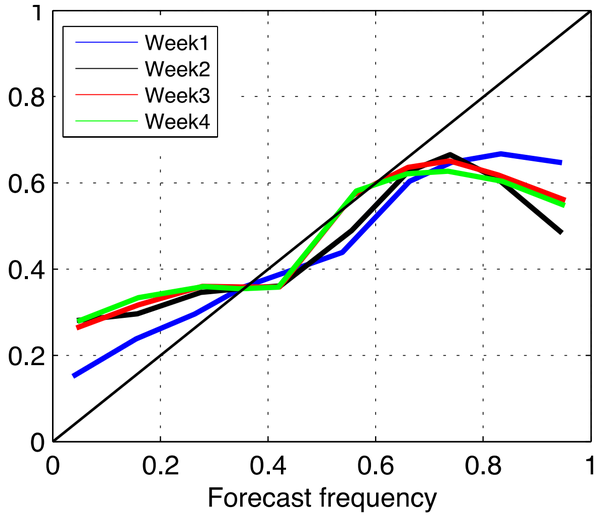
Diagrams below have been computed for all weekly starts of the JFM seasons from the 1999-2010 period using all gridpoints within N American sector (ocean+land points)

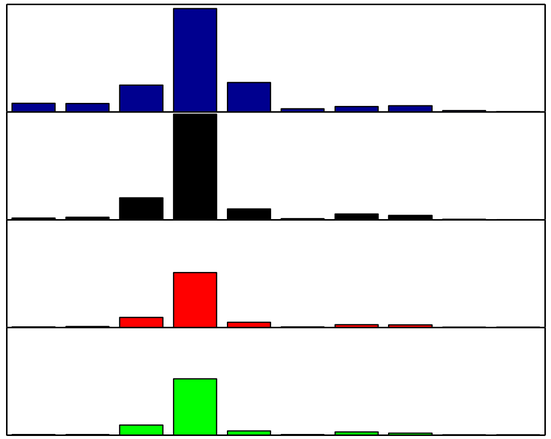
Below Normal



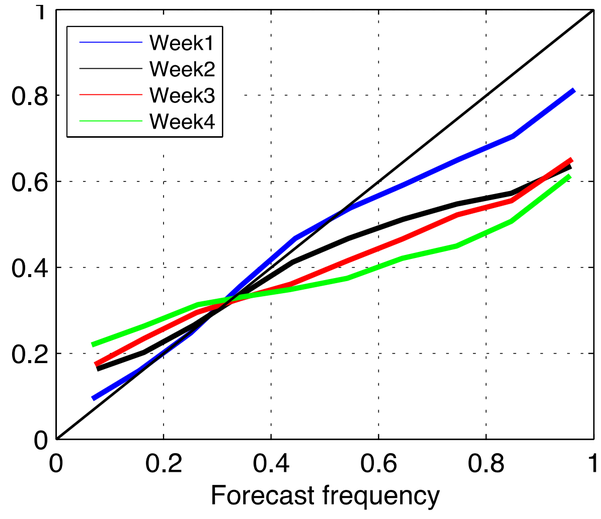


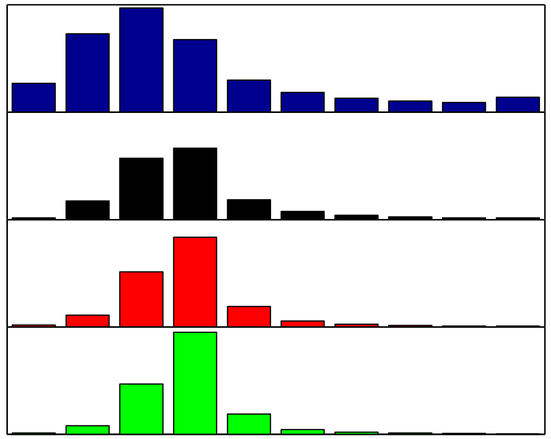
Normal





Above Normal

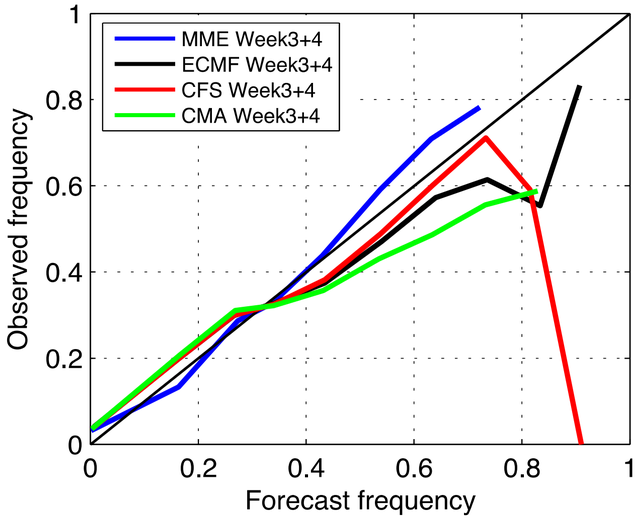


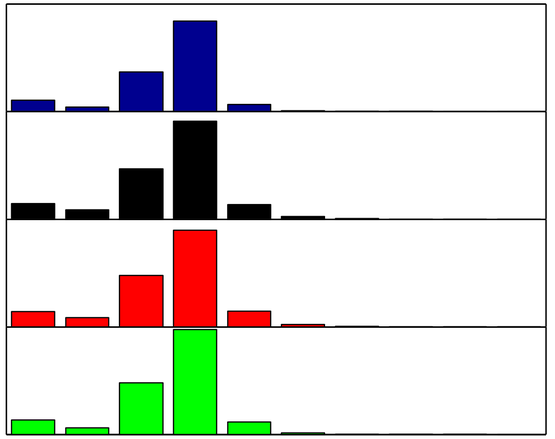


MME Reliability test Week3+4: JFM 1999-2010

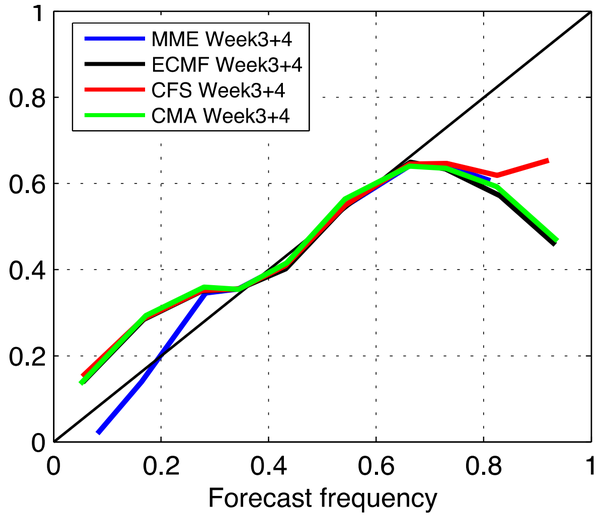
Diagrams below have been computed for all weekly starts of the JFM seasons from the 1999-2010 period using all gridpoints within N American sector (ocean+land points)

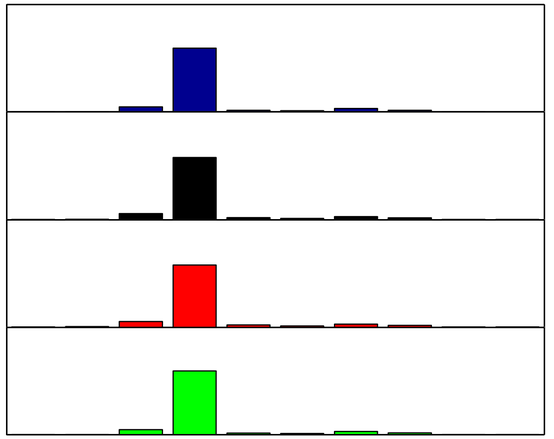
Below Normal



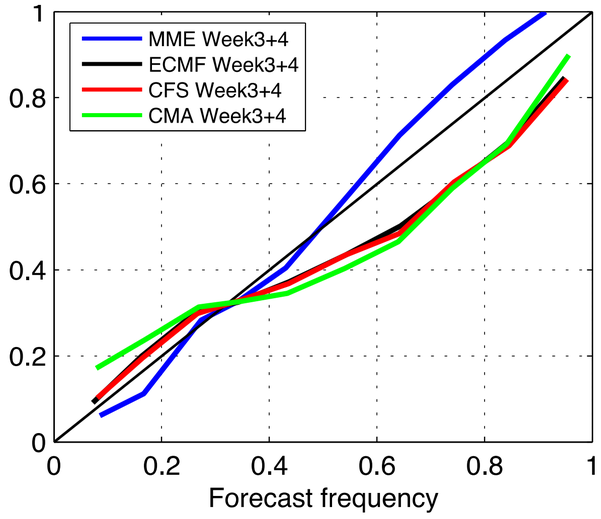


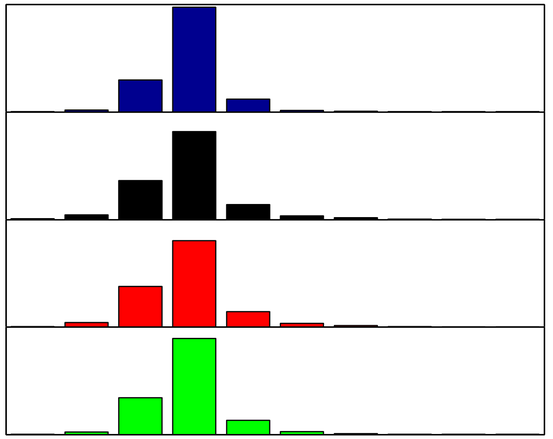
Normal





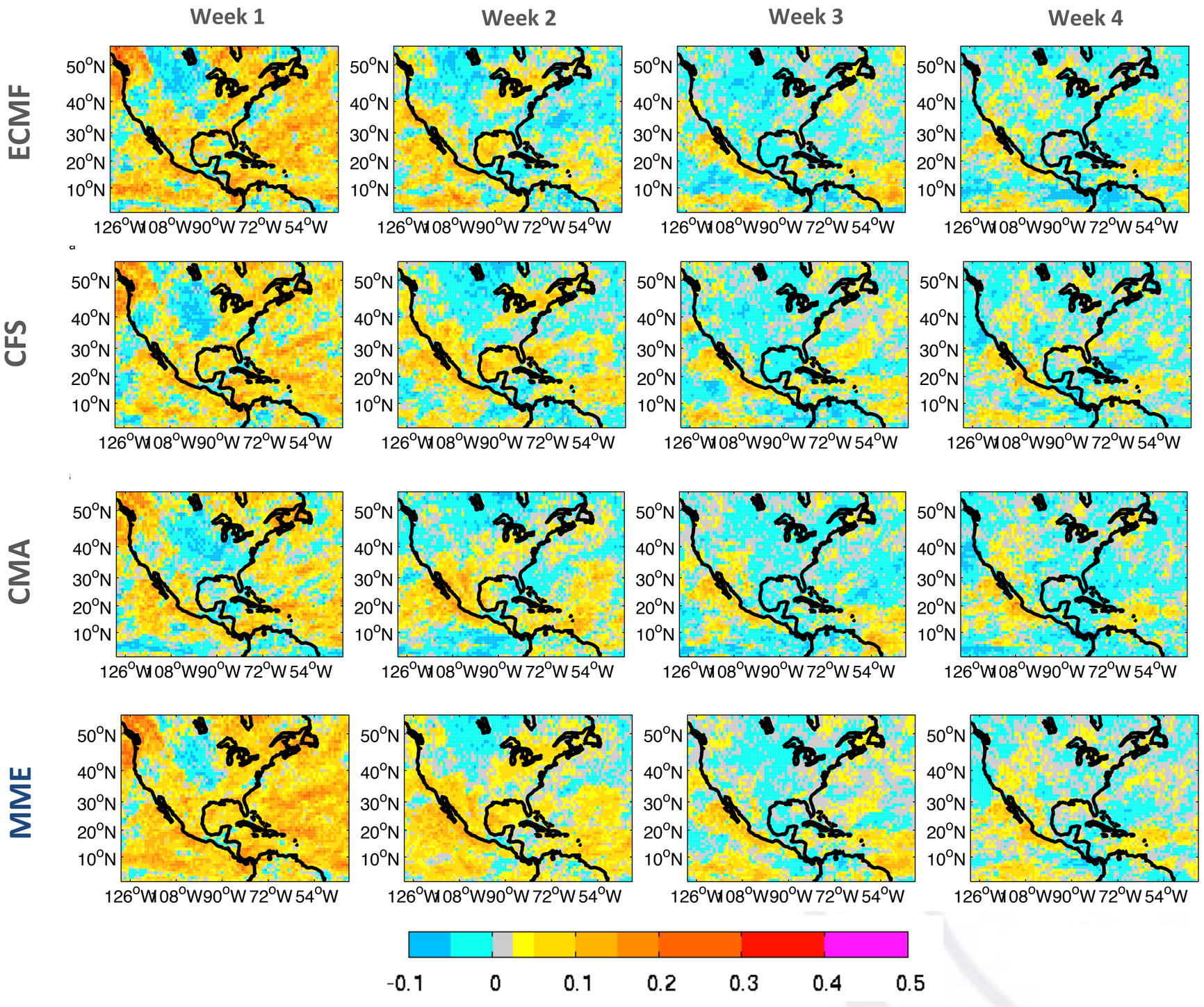
Above Normal





RPSS

Precip, JFM
1999–2010



Summary of Main Results

- ECMWF and NCEP models have fairly comparable precip correlation skill, strongest in winter
- NCEP performance is comparatively poor in spring
- Extended Logistic Regression produces good probabilistic reliability in precipitation but sharpness and RPSS remain low, especially over land
- Multi-model combination of ECMWF+NCEP+CMA appears to improve reliability of tercile-category forecasts

NGGPS Columbia University Summary

- **Major Accomplishment in FY16:**
  - Archived subset of S2S forecast/re-forecast database in IRI Data Library (ECMWF, NCEP, CMA model subsets).
  - Evaluated ECMWF and NCEP week 1–4 re-forecast performance in weekly-averaged precip, 500hPa geopotential and 2m temperature (anomaly correlation, RME error).
  - Developed extended logistic regression calibration of sub-seasonal precip forecasts for ECMWF, NCEP and CMA models, together with a multi-model combination.

- **Priority Focus for FY17**
  - Improving extended logistic regression model for week 3-4 precipitation forecasts.
  - Diagnostics of ECMWF vs NCEP performance differences over the U.S.
  - Tailoring post-processing codes for NCEP usage.
- **Key Issue**
  - Developing well-calibrated sub-seasonal (week 3-4) probabilistic forecasts over the U.S.

Relevance to the priorities of R2O Initiative

- Area 2b. Service impacts: Weeks 3–4 forecast development
- Area 5. Advances in ensemble development.
- Area 6. Advances in post-processing

Planned Deliverables

- A peer-reviewed manuscript documenting individual S2S model performance (includingCFSv2) over the U.S., with emphasis on weekly averages in weeks 2–4, gridded fields of precipitation and temperature, atmospheric indices such as the NAO and PNA, and diagnostics of predictability.
- A peer-reviewed manuscript describing the development and testing of the new MME methodology for sub-monthly forecasts.
- Implementation of a real-time S2S MME at CPC, built using the most skillful and models that are available to CPC in real time.
- Maintenance and updating of the North American Multi-Model Ensemble (NMME) archive of seasonal forecast model data in the IRI Data Library.