Exploring Spatial Distributions of Systematic Errors in the NCEP's Global Ensemble Precipitation Forecast Products

Yan Luo[1,2] and Jason Levit[1]

[1]NOAA/NCEP/Environmental Modeling Center

[2]I. M. Systems Group, Inc.

November 18, 2020

Outline

- Motivation for this work
- Introduction to NCEP GEFS ensemble precipitation products
- Spatial evaluation methods
- Precipitation product evaluation
- Summary and future plan

Motivation for this Work

- Despite recent progress in numerical weather prediction, the ensemble precipitation forecasts are still prone to systematic biases, remaining a challenge for NWP model guidance products

- Understanding such a persistent problem, how much spatial variations of systematic errors exist in global ensemble precipitation forecast products has been an ongoing and interesting research topic

- Assessing such performance of precipitation forecast is important for future research-to-operations activities and for forecasters to better understand NWP output

- Moreover, bias correction to precipitation forecast is hopefully a necessary post-processing step in the operational global ensemble forecasting

Objective

Provide a spatial view of the precipitation forecast performance from the operational NCEP ‘s global ensemble forecast system (GEFS)

- Explore useful information to identify model limitations and weaknesses

- Explore diagnostic metrics for improving model and ensemble forecast performance

- Investigate the usefulness and effectiveness of bias-correction approach currently applied to the ensemble forecast products

NCEP/GEFS Operational Precipitation Products

Direct model outputs

GEFSv11

- Raw forecasts for GFS, GEFS control forecasts, and 20 perturbed members
- 6hr accumulated, every 6hrs, out to 384hrs (16days)
- All four cycles: 00Z, 06Z, 12Z and 18Z
- 1 degree global QPF
- 3 hourly out to 8 days then 6 hourly 0.5 degree global QPF
- 0.5 degree 24 hour global PQPF with 13 thresholds

GEFSv12 (recently implemented on Sep. 23, 2020)

- 30 perturbed members
- 3 hourly out to 10 days at 0.25 degree
- 6 hourly beyond 10 days at 0.5 degree (out to 35 days)
- All Low-Res GEFS products are discontinued

Post-processed products

- Implemented in July 2018
- CCPA precipitation analysis used as proxy truth for bias correction
- Bias-correction Method: frequency match and decaying average (Zhu and Luo, 2015)
- Decaying weight W=1/50 ~ 2%
- 12 RFC CDFs, 9 thresholds (0.2,1,2,3.2,5,7,10,15,25mm/6hrs)
- Bias correction for GFS, GEFS/Control, and 20 perturbed members
- All four cycles: 00Z, 06Z, 12Z and 18Z
- Bias corrected products: 6hr &24 hr 0.5 degree global QPF, PQPF
- Downscaled products: 24 hr 2.5m NDGD CONUS QPF, PQPF

Spatial Evaluation Methods

- 24-hour precipitation forecasts in NCEP's GEFS operational raw and bias-corrected products are evaluated against CCPA at 0.5 degree grid
- Ensemble control and ensemble mean forecast are both evaluated for this study
- Evaluated for selected metrics and selected periods and conditioned at different lead-times and thresholds
  - Metrics:

    Mean Value , Mean Error (Bias), Frequency, Frequency Bias (Bias Score)
  - Lead times:

    Day 1 (12-36h), Day 2 (36-60h), … out to Day 10 (228-252h)
  - Thresholds:

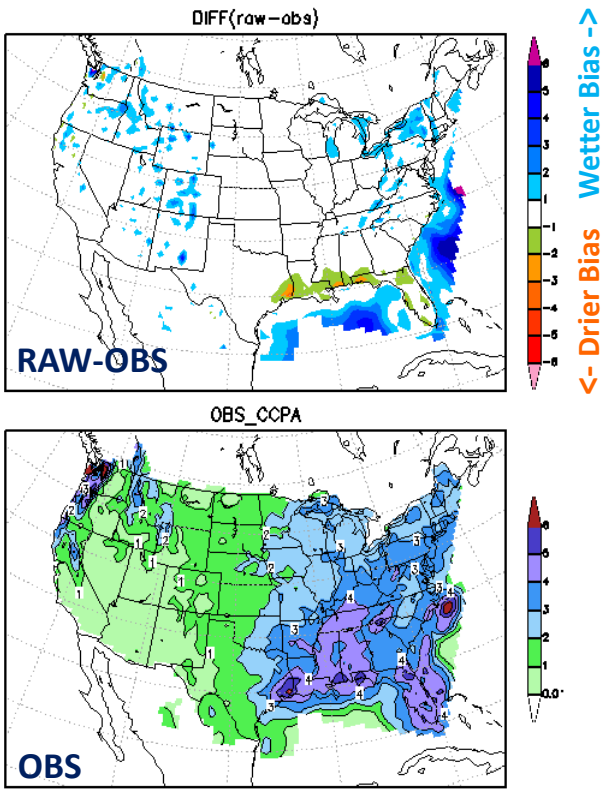
    1 mm, 6.35 mm (0.25 inch), 12.7 mm (0.5 inch), 25.4 mm (1 inch), 50.8 mm (2 inches)
  - Periods: Select one year (1 June 2017 – 31 May 2018) and another one year (1 Dec 2018 -30 Nov 2019) for examples
  - Domain: Continental US (CONUS)
- Side by side map comparison for a given period

Validation Dataset

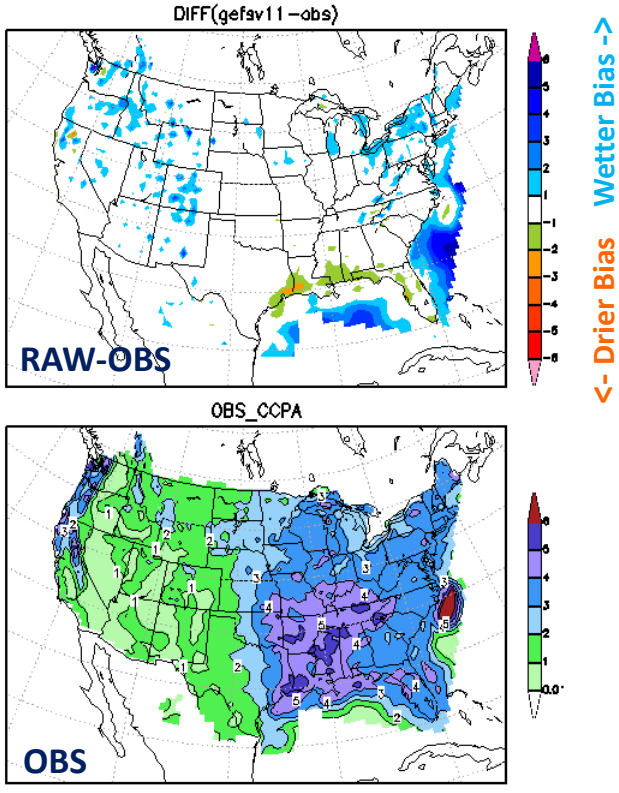
- Climatology-Calibrated Precipitation Analysis (CCPA)
  - A dataset of precipitation analysis, over CONUS at 6h, ~4km resolution
  - Statistical adjustment of Stage IV data toward CPC analysis
  - Simple linear regression at 0.125 degree and 24h accumulation
  - Keep the fine scale structures of Stage IV
  - Closer to CPC Unified Precipitation Analysis, in the sense of climatology
- Application: Provide a proxy of truth for precipitation forecast calibration and downscaling
- Developed and distributed by NCEP/EMC for operation
- First operational implementation on July 13, 2010
- Product period: 2002 - present
- Product grids:
  - HRAP (primary)
  - 2.5km & 5km NDGD, 0.125, 0.5 and 1 degree resolutions (byproducts)
  - 1 hour, 3 hour and 6 hour accumulations
- CCPA websites:
  - Introduction: http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-11-0140.1
  - Image: http://www.emc.ncep.noaa.gov/gmb/yluo/ccpa/ccpa.php

Spatial Distributions of Mean Error (Bias) in Different Years

One Year Period
(20170601 - 20180531)



Another One Year Period
(20181201 - 20191130)

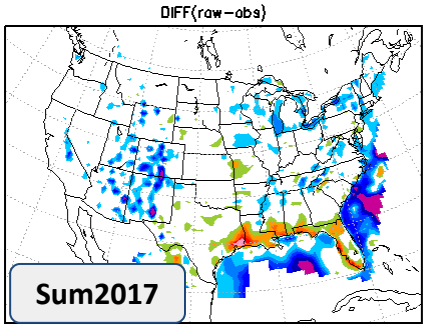


- **Very consistent mean errors repeated in different years**
- **Such systematic bias can be calibrated using bias correction methods**

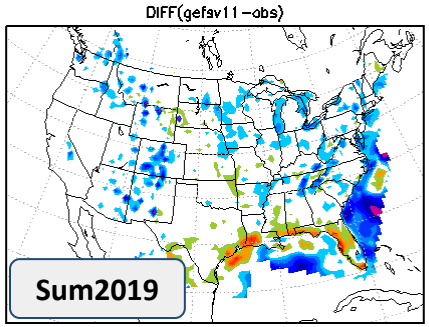
raw = gefsv11
Unit: mm/24hr

Seasonal Variation of Mean Errors in different years (fhr=12-36)

Summer Bias



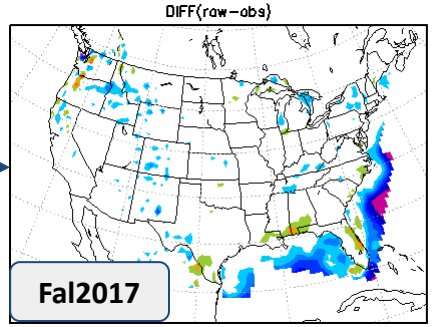
Sum2017



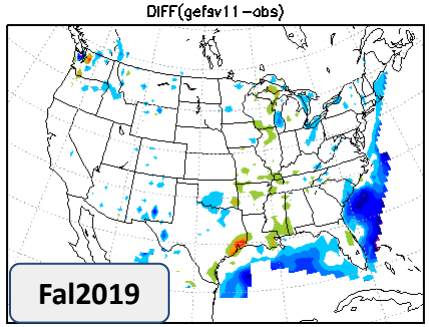
Sum2019

<- Drier Bias Wetter Bias ->

Fall Bias

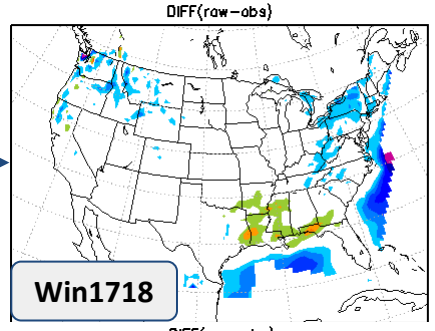


Fal2017

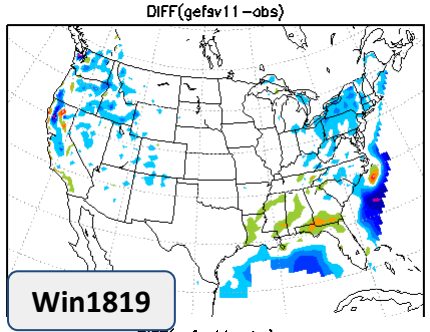


Fal2019

Winter Bias

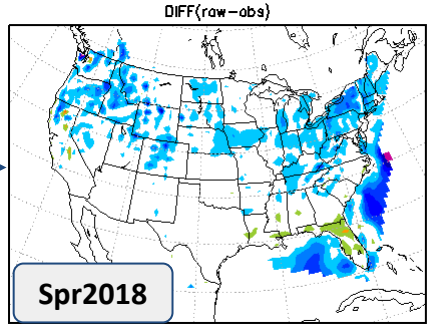


Win1718

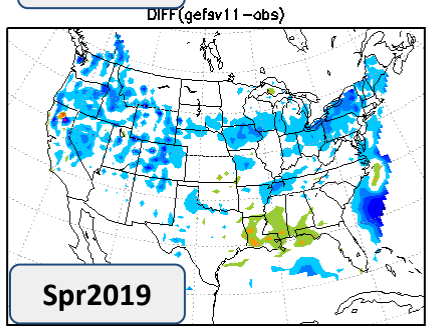


Win1819

Spring Bias



Spr2018



Spr2019

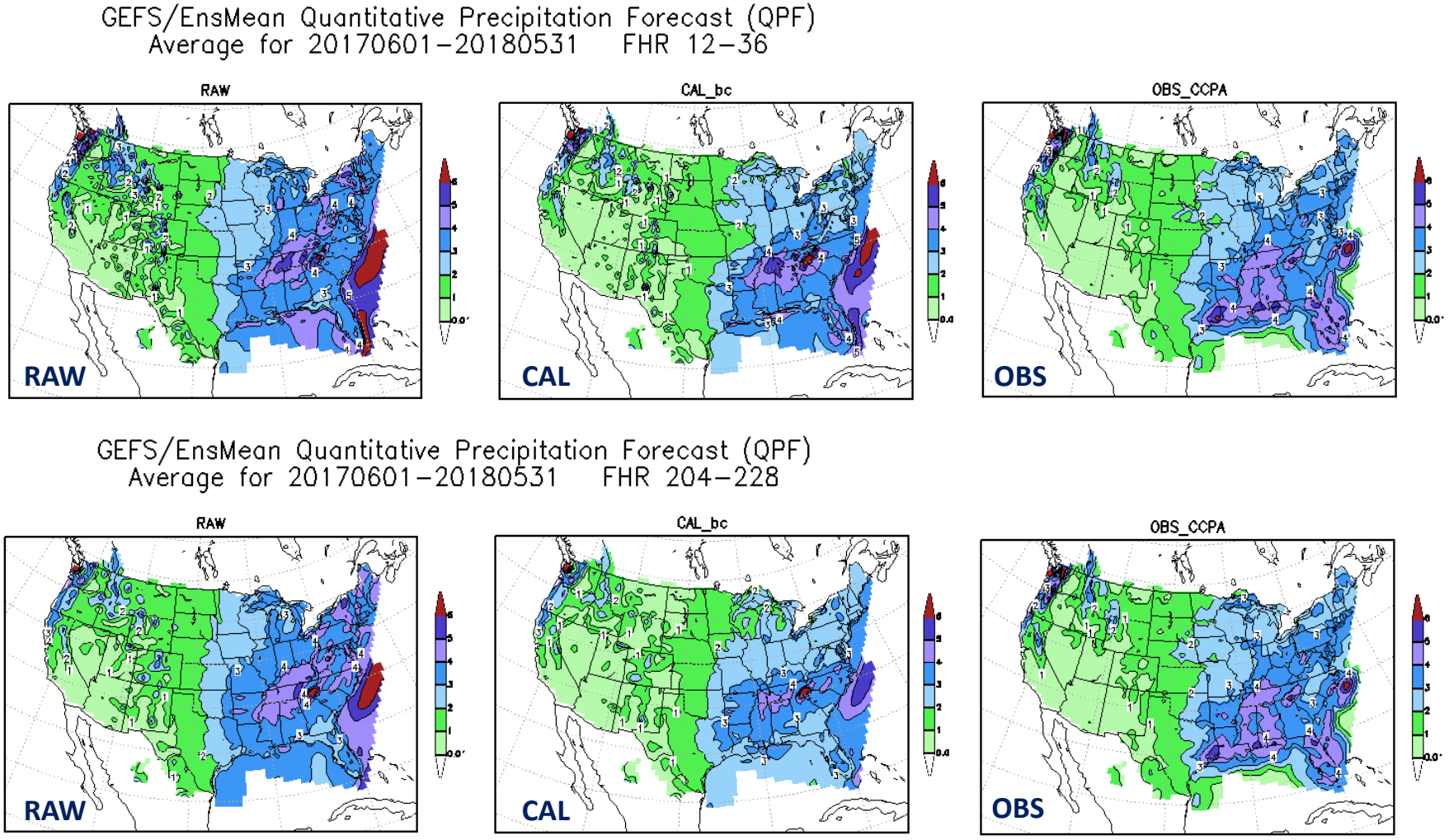
- Very clear seasonal variation of mean errors
- Even season has consistent mean errors persist in different years

raw = gefsv11
Unit: mm/24hr

Spatial Distributions of Mean Value: Raw Fcst vs. Cal Fcst

Shorter lead

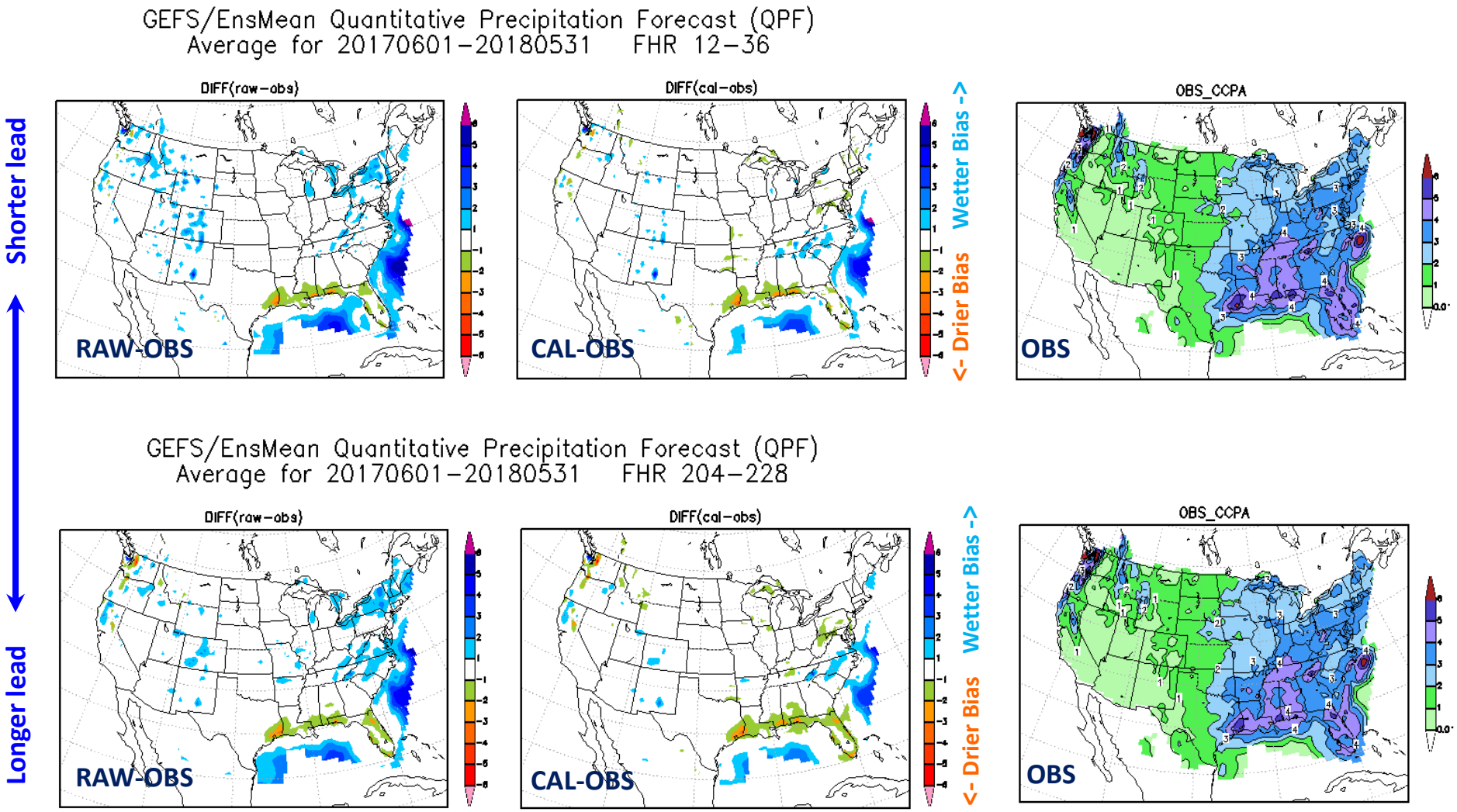
Longer lead



- **Roughly similar rain pattern**
- **Cal Fcst appears closer to OBS**

Unit: mm/24hr

Spatial Distributions of Mean Error (Bias): Raw vs. Cal



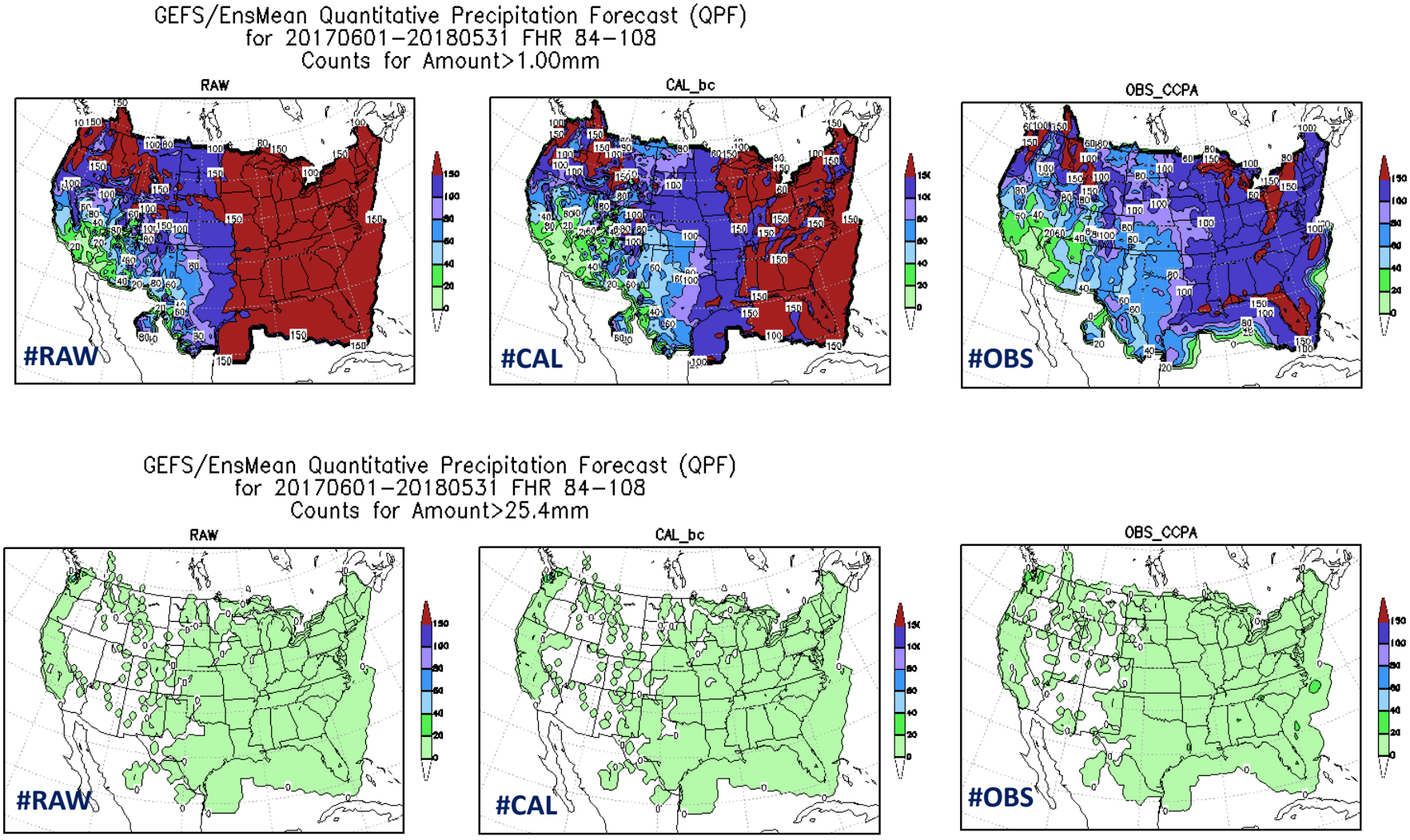
- Much reduced wet bias both for shorter and longer lead times
- Struggled with dry bias reduction

Unit: mm/24hr

Spatial Distributions of Frequency: Raw vs. Cal

Lower threshold

Higher threshold

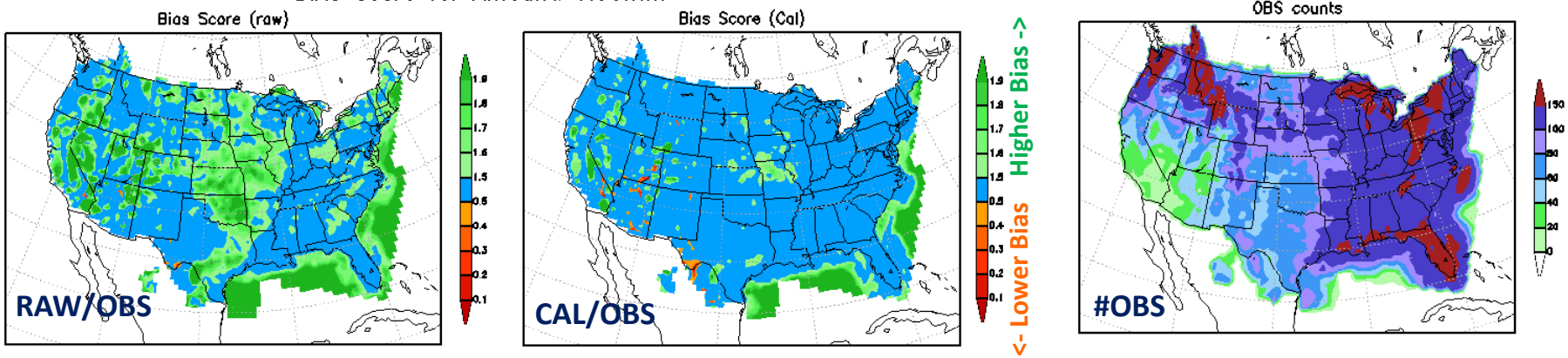


- Too much frequent light rain in Raw Fcst
- Less frequent light rain in Cal Fcst, closer to OBS
- No big difference in both Fcsts for heavy rain, appear less frequent than OBS

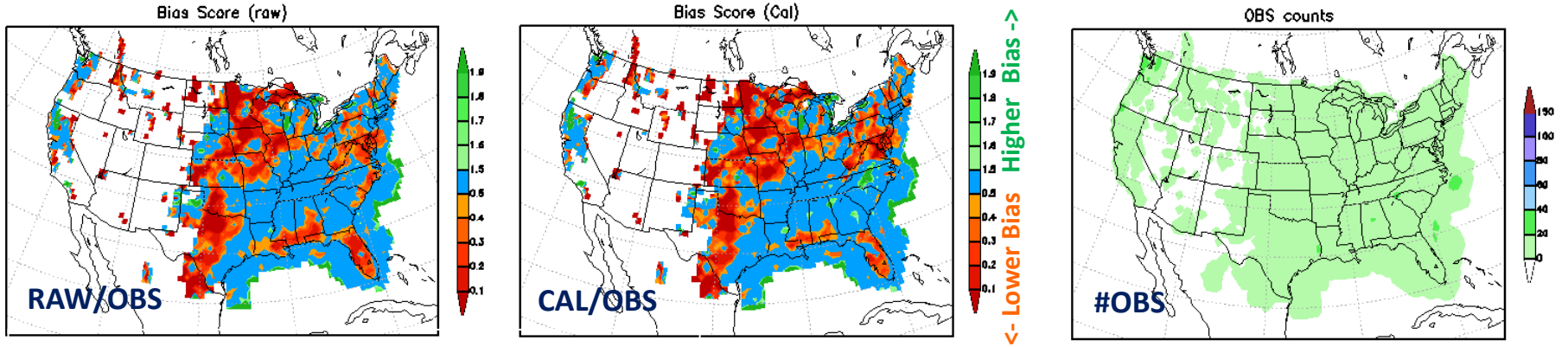
Spatial Distributions of Frequency Bias: Raw vs. Cal

Lower threshold

GEFS/EnsMean Quantitative Precipitation Forecast (QPF)
for 20170601–20180531 FHR 84–108
BIAS score for Amount>1.00mm



GEFS/EnsMean Quantitative Precipitation Forecast (QPF)
for 20170601–20180531 FHR 84–108
BIAS score for Amount>25.4mm



Higher threshold

FB(0.5, 1.5) ≈> 1 Perfect bias score

FB(>1.5) Higher bias for low amts

FB(<0.5) Lower bias for high amts

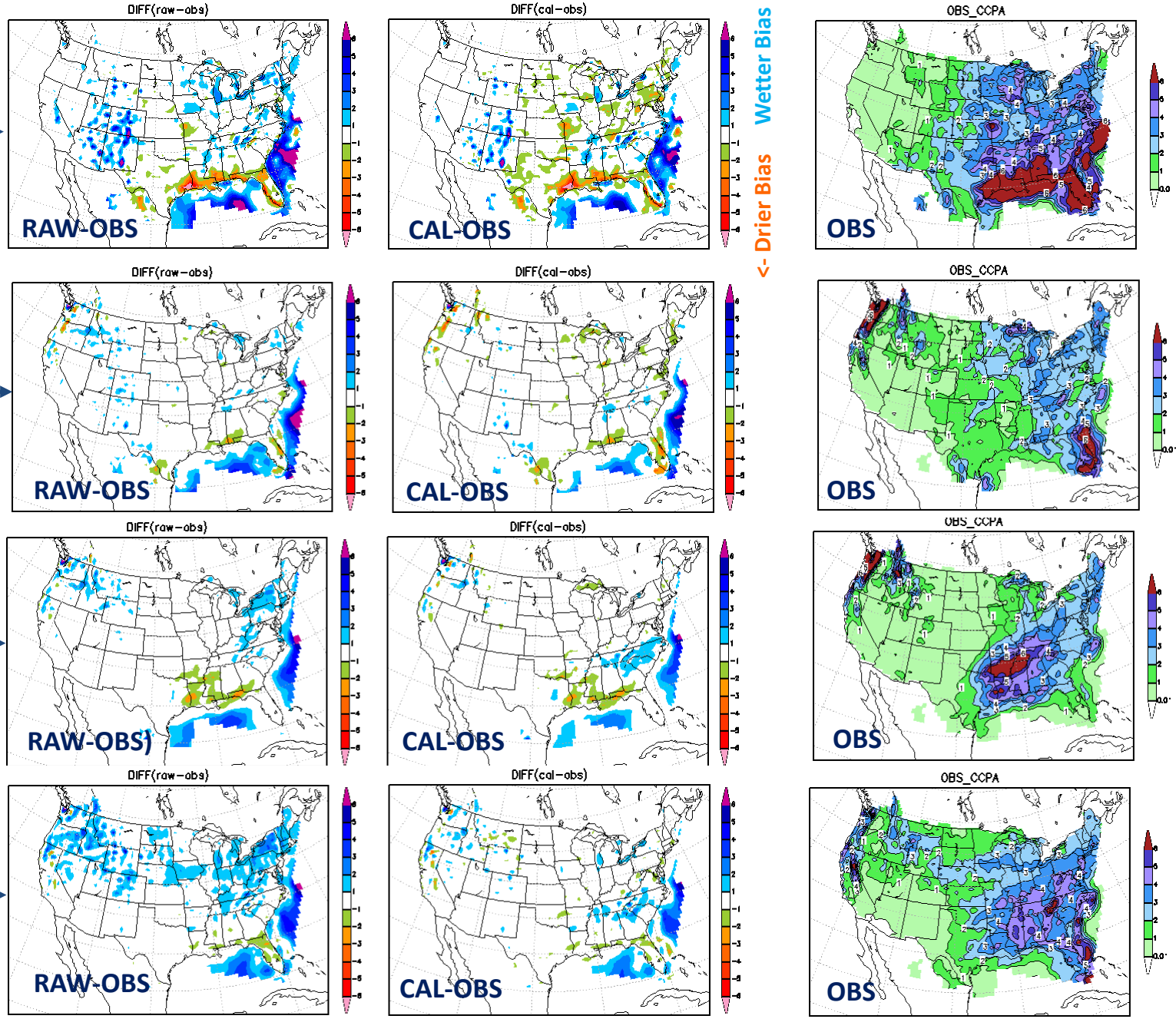
Seasonal Variation of Mean Errors: Raw vs. Cal (fhr=12-36)

Sum2017

Fal2017

Win1718

Spr2018



Unit: mm/24hr

Mean Error (Bias) for Day 1 (F12-36h)

BIAS	Raw Fcst	Cal Fcst
Sum2017	**0.090**	-0.364
Fall 2017	**0.087**	-0.272
Win1718	0.220	**-0.121**
Spr2018	0.564	**-0.036**
Full year	0.243	**-0.199**

Mean Error (Bias) for Day 5 (F108-132h)

BIAS	Raw Fcst	Cal Fcst
Sum2017	**-0.034**	-0.375
Fall 2017	**0.056**	-0.236
Win1718	0.308	**-0.04**
Spr2018	0.587	**0.070**
Full year	0.231	**-0.145**

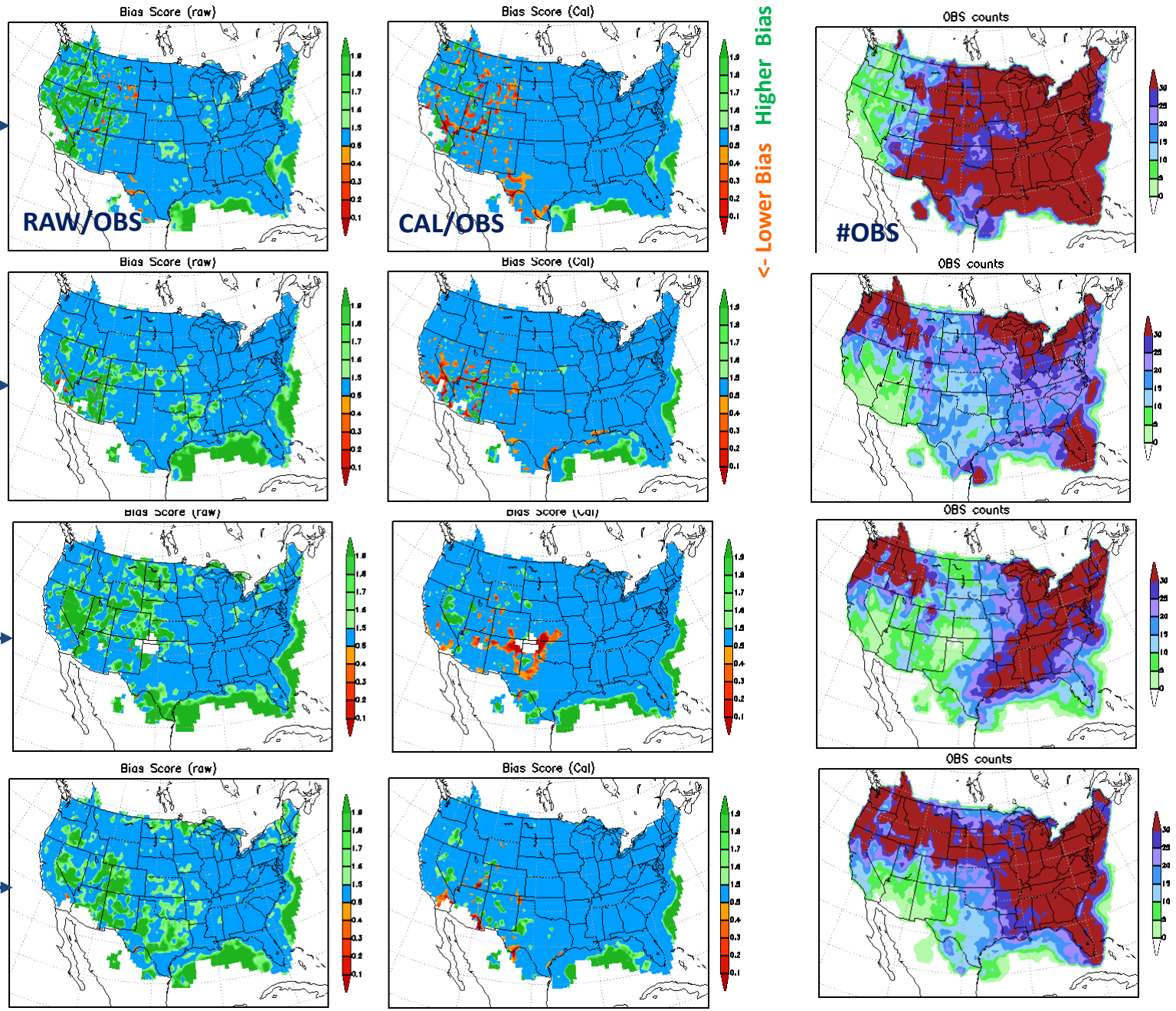
Seasonal Variation of Frequency Bias for Amount > 1mm, FHR 12-36

Sum2017

Fal2017

Win1718

Spr2018



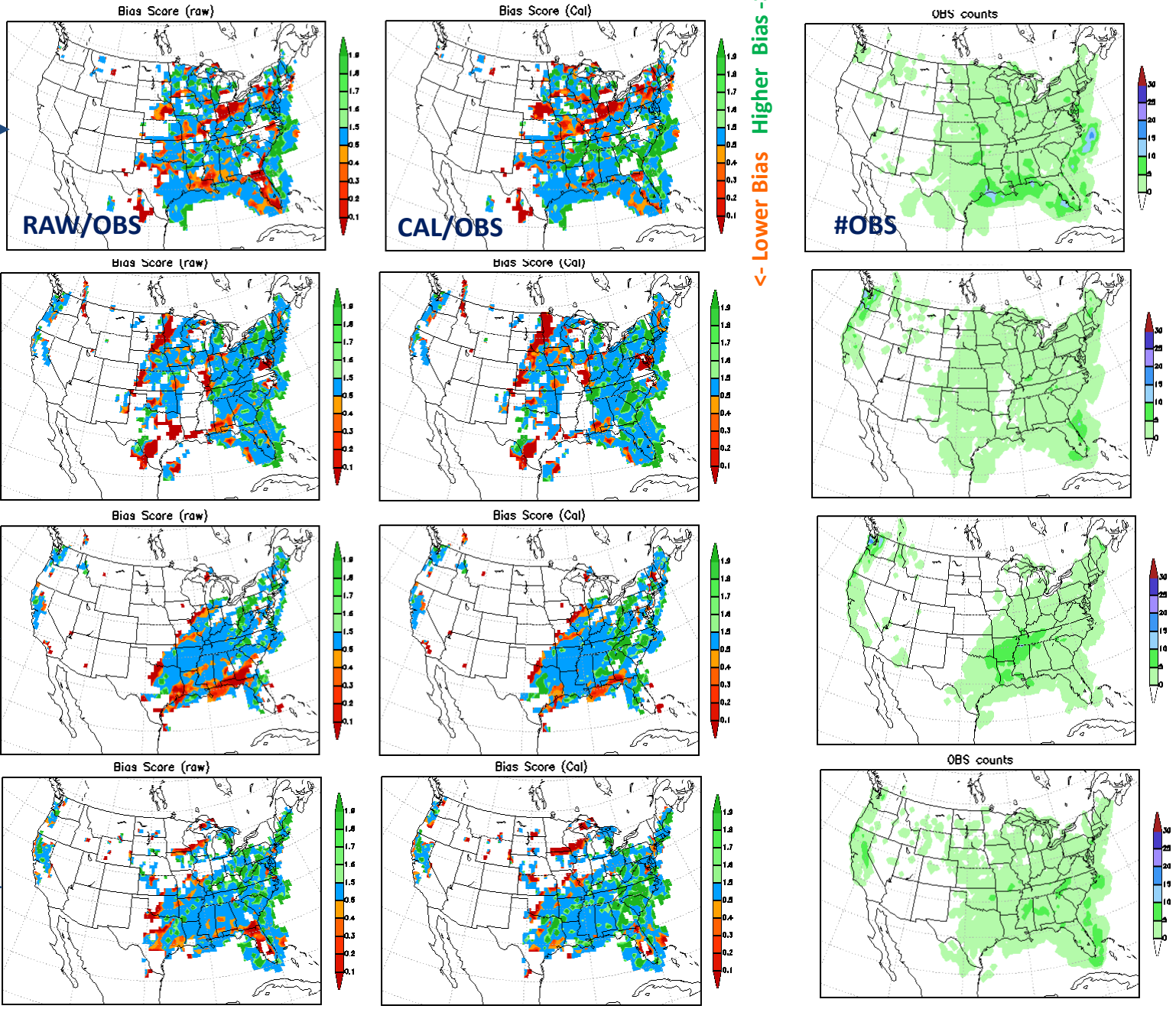
Seasonal Variation of Frequency Bias for Amount> 25.4mm, FHR 12-36

Sum2017

Fal2017

Win1718

Spr2018



Summary

- Spatial verification against CCPA for CONUS at 0.5 degree resolution for GEFS raw and calibrated forecasts
- Side by side map comparisons for selected metrics and selected full year periods and conditioned on different lead-times and thresholds
- Generally raw forecasts are dominated by wet biases with broad area coverages, mostly appear over western mountain terrains and the Northeast, while strong dry biases persist along Gulf Mexico area.
- Raw forecasts have fairly consistent bias for a year, even for same season at different years; the similar spatial pattern is repeated very well
- Bias-corrected forecasts show much effectively reduced wet bias, but struggle with correction of dry bias from non-precipitation cases and limited samples for high thresholds
- Bias correction works better in cold seasons than in warm seasons

Future Plan

- Ongoing efforts toward development using the MET/METplus tool
  - Transition the current framework to a new METplus version
  - Add new verification metrics
  - Enhance global ensemble verification capabilities

- Contribution to improving ensemble forecasts through verification using this METplus based framework
  - Investigate the impact of GEFS upgrade on the spatial distribution of the systematic errors in the precipitation products
  - Inter-compare with ensemble forecasts from other world centers

Extra Slides

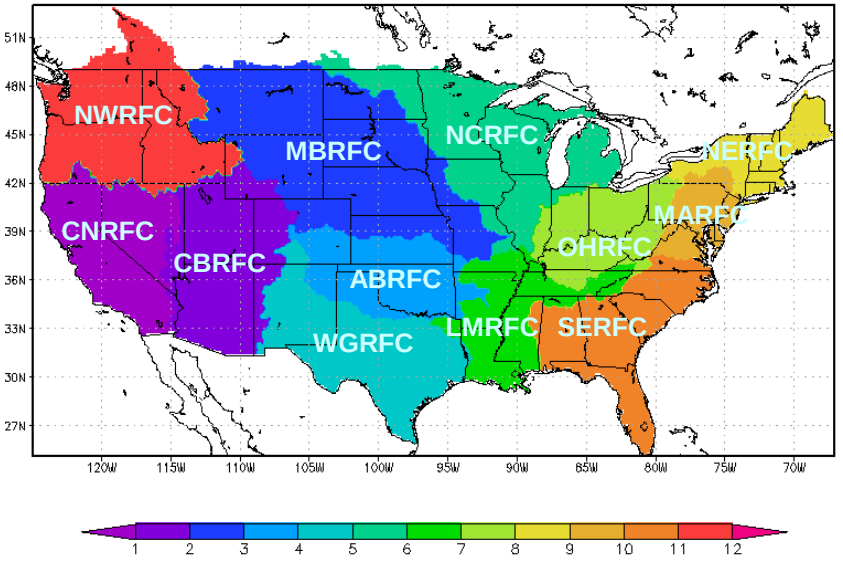
Precipitation Calibration Based on Frequency Matching Method (FMM)

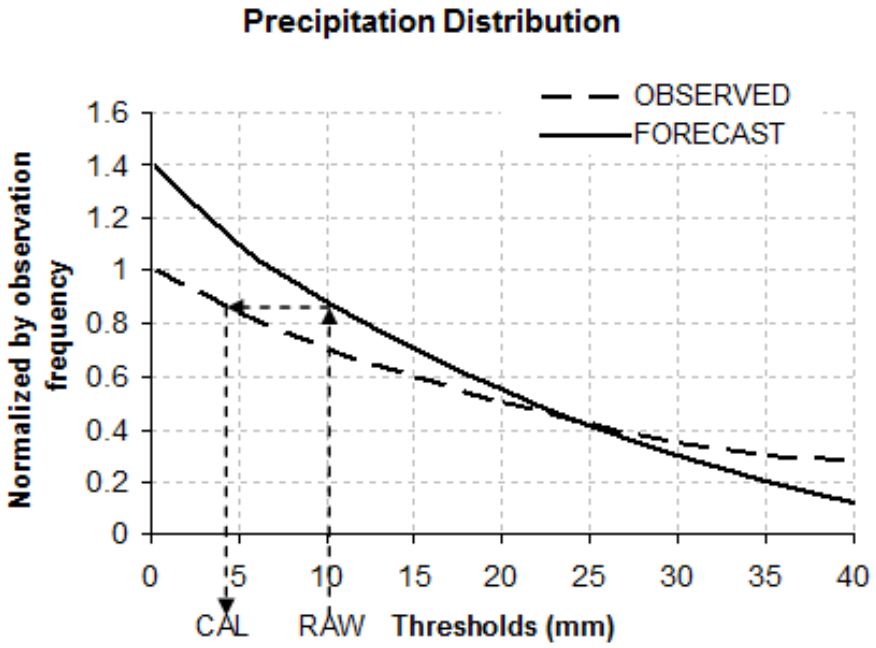
(Ref: Zhu and Luo, 2015: Weather and Forecasting)

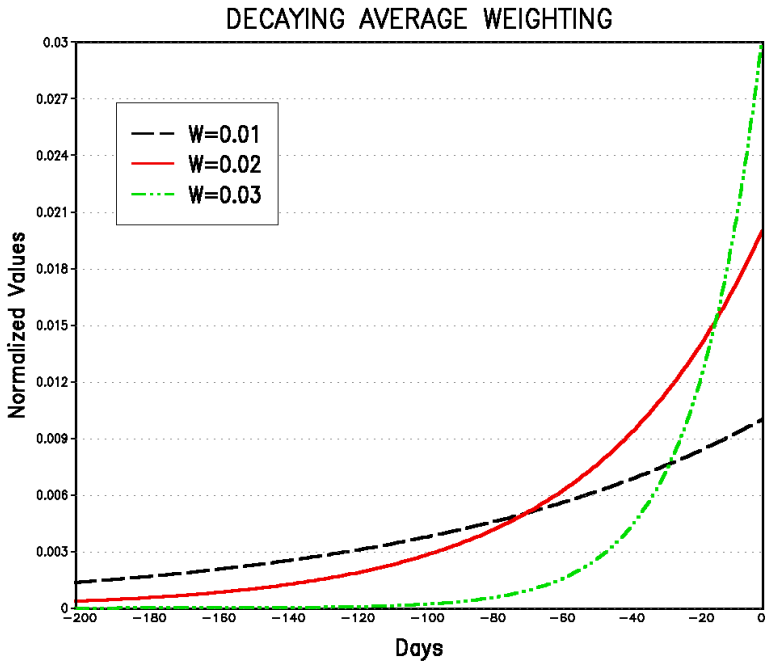
Calculate for Obs and Fcst respectively

$$\overline{CDF}_j = (1-W) * \overline{CDF}_{j-1} + W * CDF_j$$

W is weight to accumulate CDF



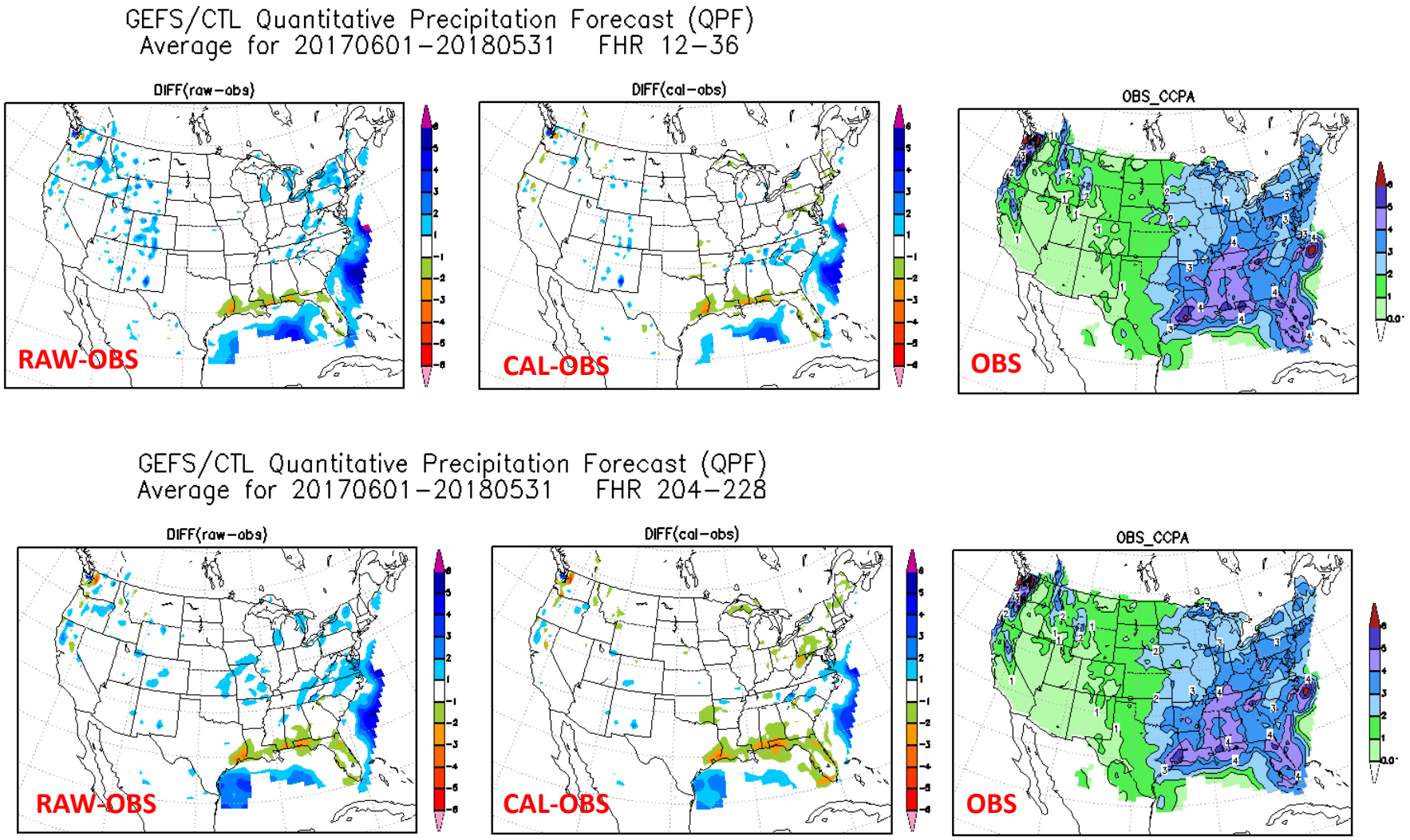




Spatial Distributions of Mean Error (Bias)

Shorter lead

Longer lead



Unit: mm/24hr