Seasonal Predictability of Weather Type Frequencies over the Contiguous US

Jun 13, 2023
CESM Earth System Prediction Working Group Meeting
Center Green, NCAR

Erin Towler (C3WE/MMM/NCAR)

With contributions from:

James Done, Ming Ge, Eric Gilleland, Danielle Touma, Andreas Prein

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

MESOSCALE & MICROSCALE METEOROLOGY LABORATORY

Part of an ongoing NSF project: **COEXIST** (Connected Extremes In Space and Time)

COEXIST group:

James Done (PI), Erin Towler, Ming Ge, Daniel Swain, Erin Evans, Manuela Brunner, Jennie Bukowski, Danielle Touma, Mari Tye

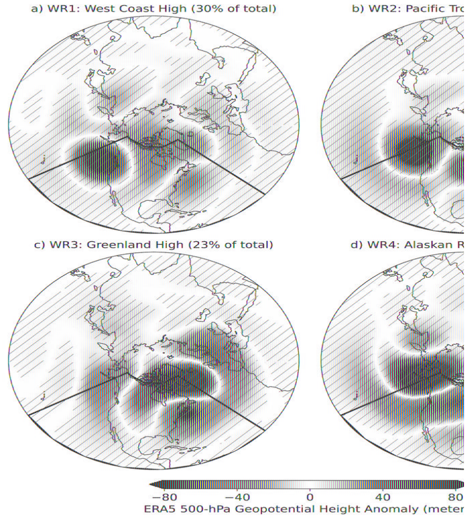
COEXIST Goals:

- Identify connected extremes
- Understand the drivers of spatiotemporally connected extremes
- Quantify the impact of connected extremes
- Explore sub-seasonal to seasonal predictability of connected extremes

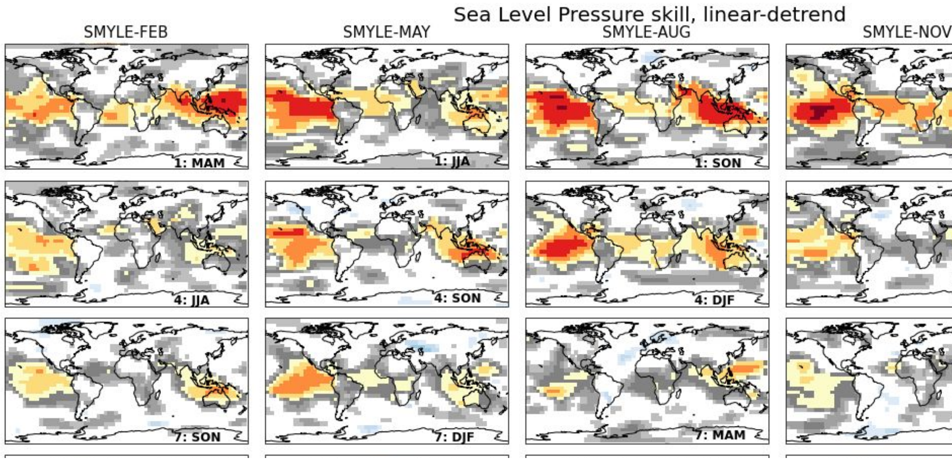
This talk

How well do **seasonal** forecast models predict large-scale weather patterns over the contiguous United States (CONUS)?

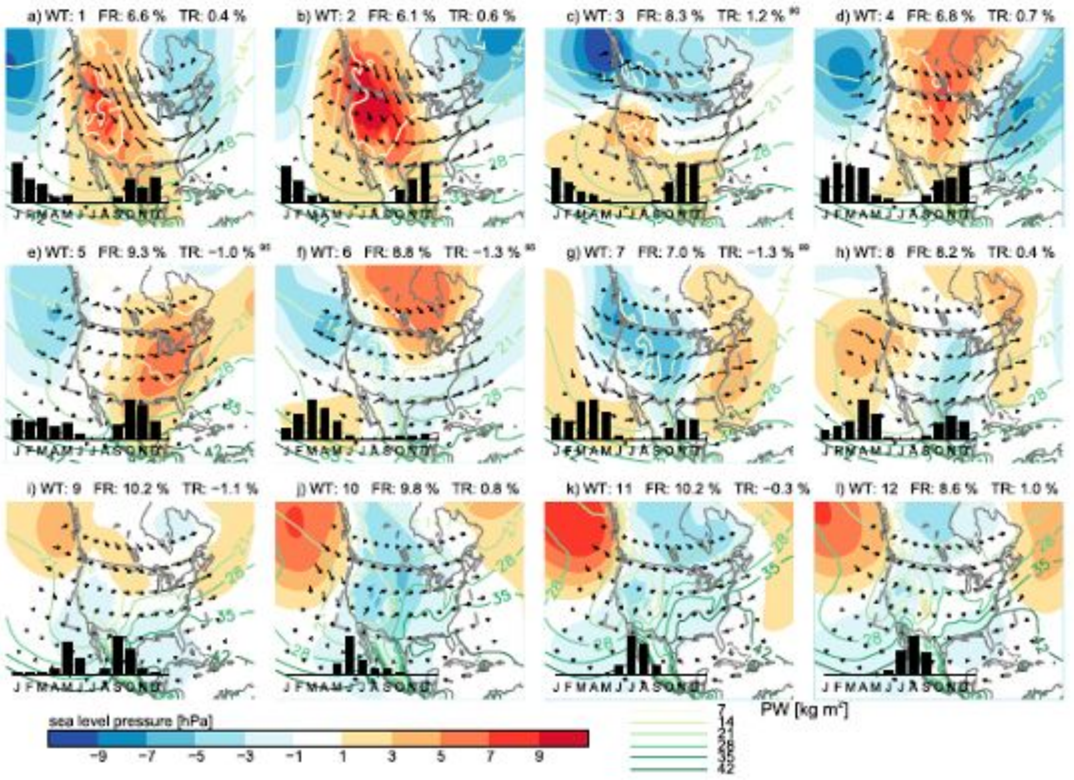
Molina et al (2023) examine ***subseasonal*** predictability of 4 Northern Hemisphere weather regimes using CESM2 weekly hindcasts (Richter, et al. 2022).



Yeager et al (2023) provide a broad overview of prediction skill for the CESM2 quarterly hindcasts at lead times ranging from ***1 month*** out to 2 years.



Recurring large-scale atmospheric patterns, or weather types (WTs) are identified using clustering*.

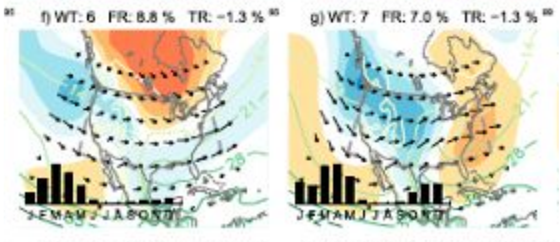


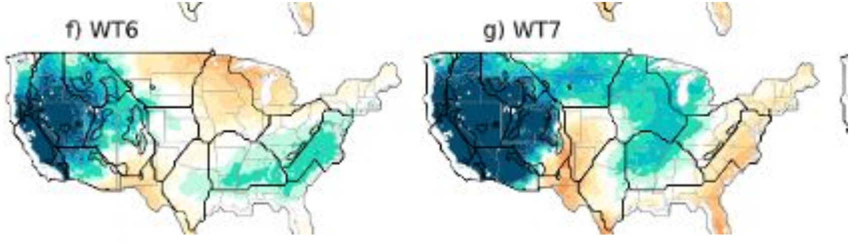
*Sea Level Pressure, Precipitable Water, and winds

Prein et al. 2016

Weather types (WTs) are associated with precipitation anomalies.

E.g., WT6 & WT7 associated with Western precipitation.





Prein et al. 2016

Step 1. Reproduce Historical CONUS Weather Types.

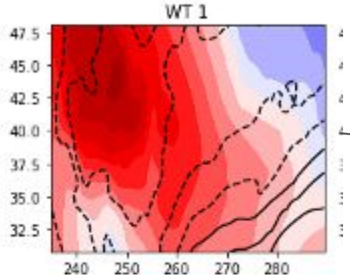
- ERA-Interim
  - also checked ERA5.

Generalized WT clustering algorithm* (thanks Ming!) to be reproducible.

*https://github.com/ming80302/WT

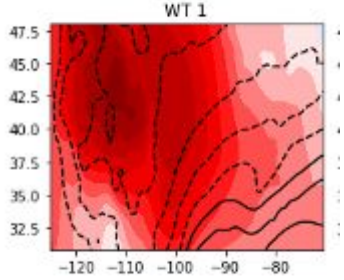
Andy's 2016

WT1 Centroid

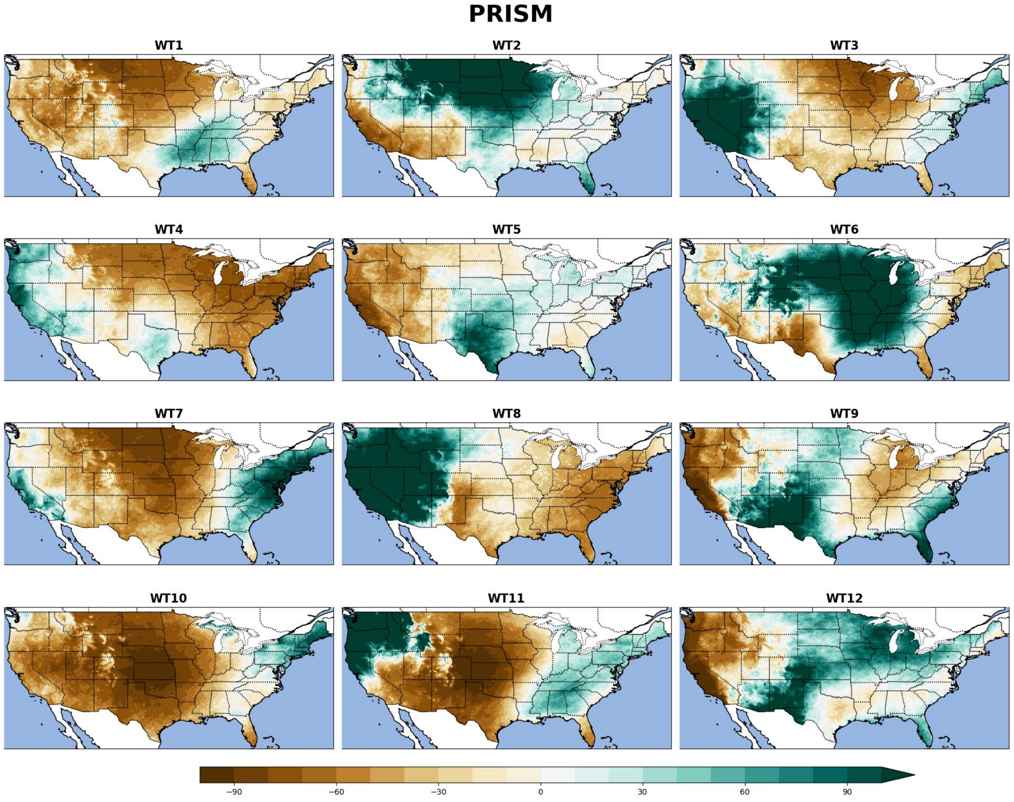


Ming's reproducible

WT1 Centroid



From reproduced CONUS WTs, associated PRISM precipitation anomalies show spatial coherence.



Courtesy Ming Ge.

Step 1. Reproduce Historical CONUS Weather Types.

Step 2. Apply Weather Type Clustering Algorithm to Seasonal Hindcasts.

- ECMWF: European Centre for Medium-Range Weather Forecasts (Integrative Forecasting System; IFS, Version 5 from the Copernicus Climate Change Service)
- SMYLE: Seasonal-to-Multiyear Large Ensemble using NCAR's CESM2 *(Yeager et al. 2022, GMD)*.

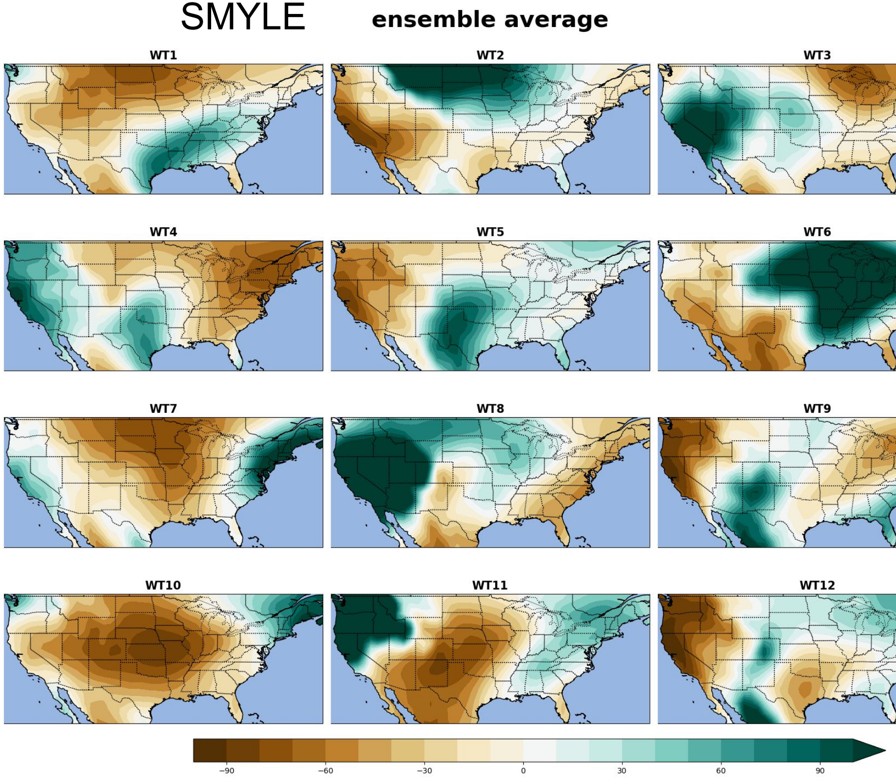
Analysis conducted on overlapping periods.

Model	Full Time Period	Overlapping period
ECMWF	1993-2022	1993-2019*
SMYLE	1970-2019	
ERA-Interim	1979-2021	

*n=104 seasons (26 years x 4 seasons).

Using the ensemble average, days are assigned to WT categories for SMYLE and ECMWF.

SMYLE



Courtesy Ming Ge.

Step 1. Reproduce Historical CONUS Weather Types.

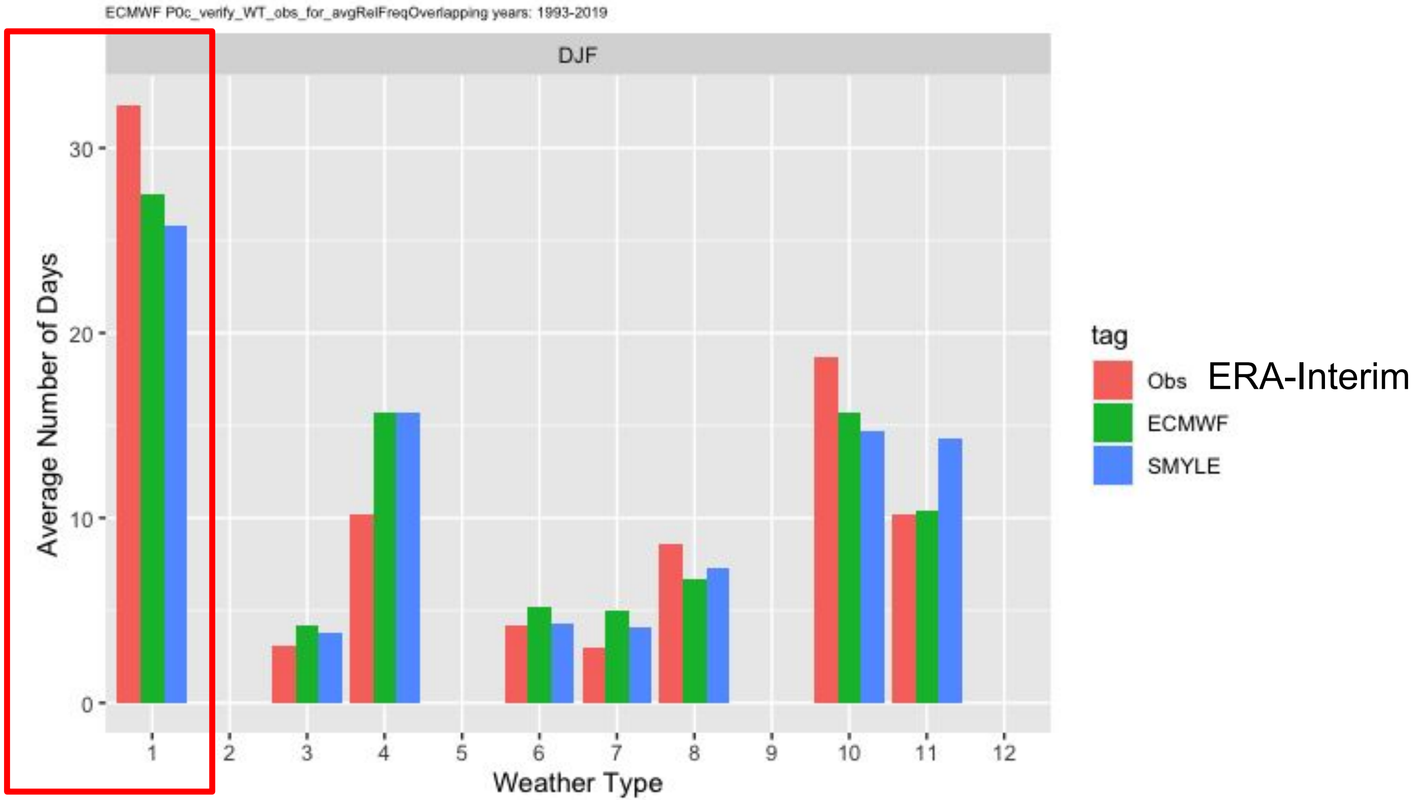
Step 2. Apply Weather Type Clustering Algorithm to Seasonal Hindcasts.

Step 3. Verify the hindcasts.

- WT frequencies
  - focus on “lead 1” forecast, (i.e., DJF forecast issued in November)

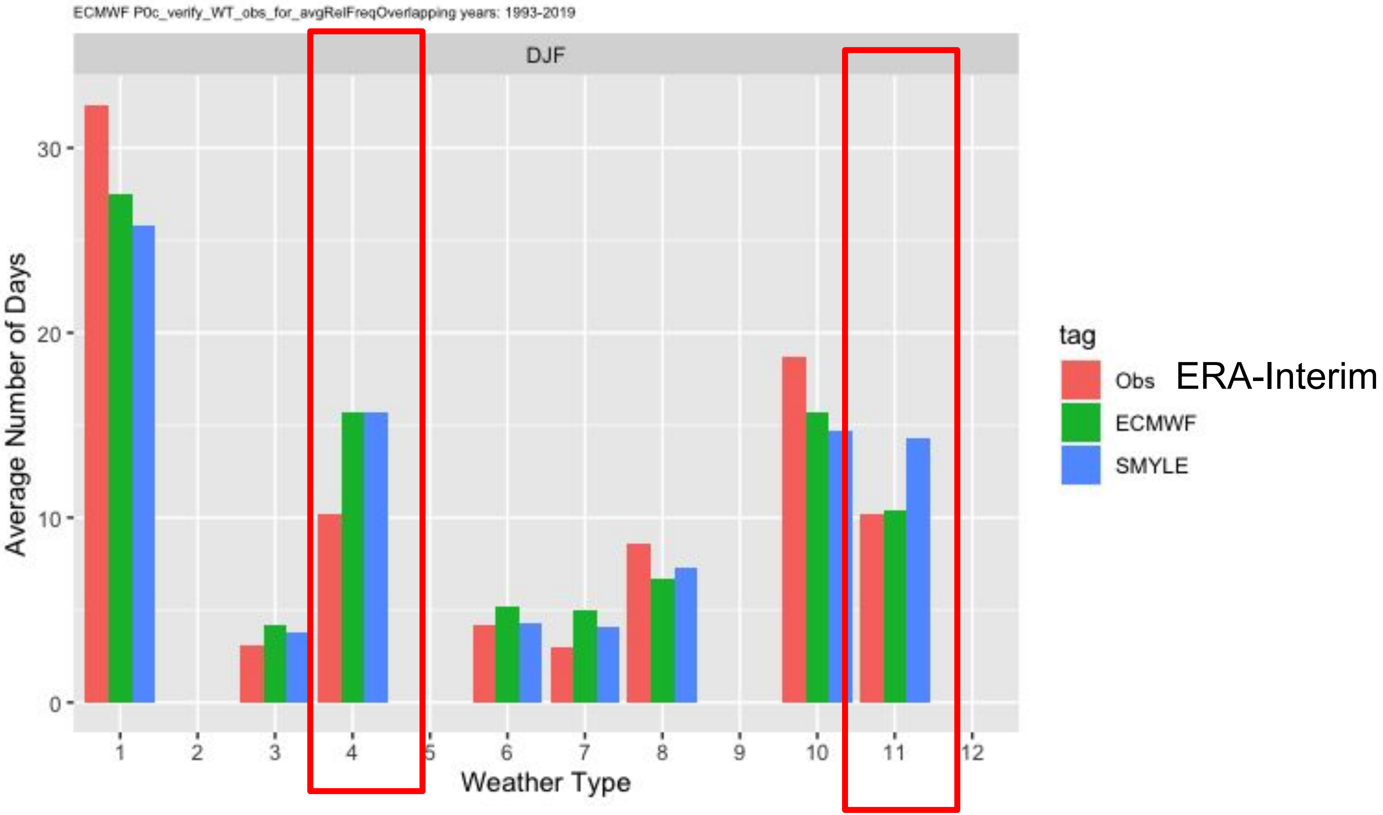
For DJF, both forecast models do a reasonable job at capturing the ***average*** number of days in each WT.

- E.g., WT1 is most common in ERA-Interim, ECMWF & SMYLE are low.

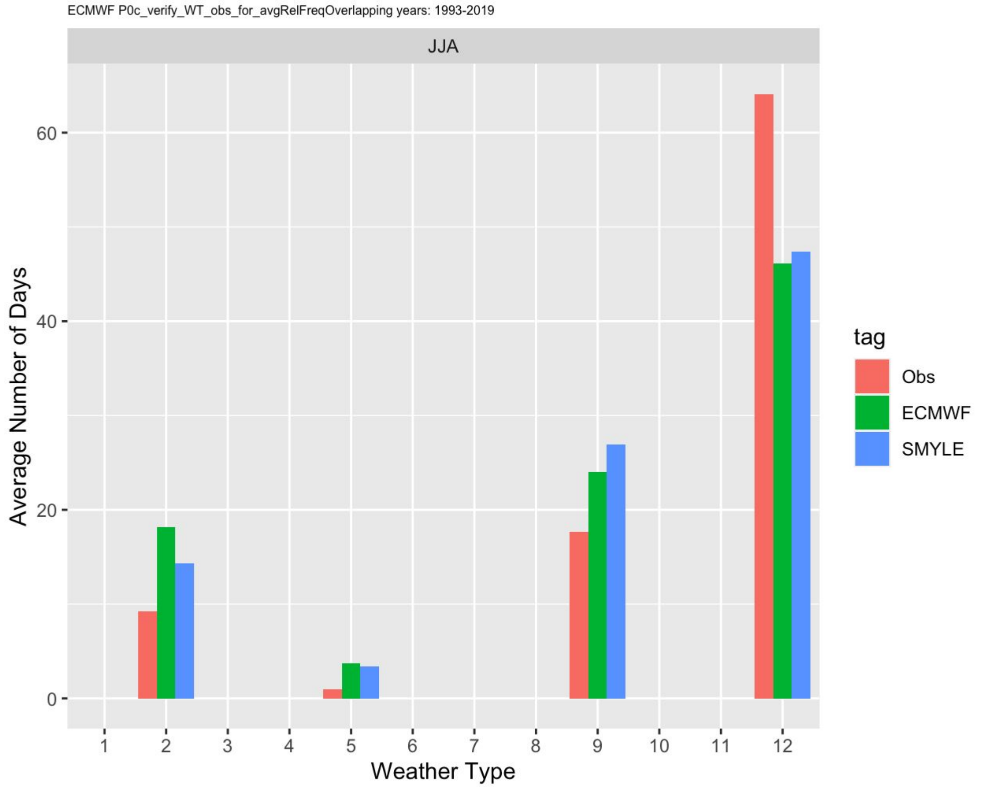


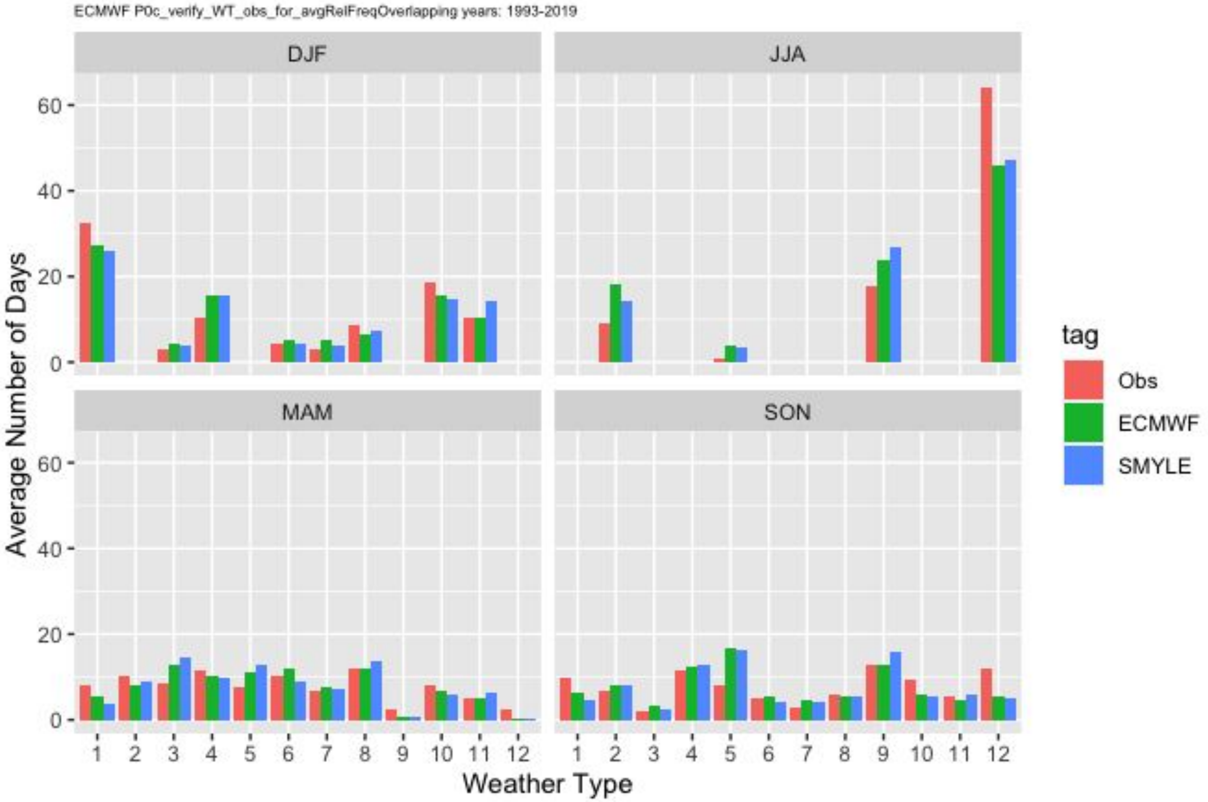
For DJF, both forecast models do a reasonable job at capturing the ***average*** number of days in each WT.

- E.g., WT1 is most common in ERA-Interim, ECMWF & SMYLE are low.
  - SMYLE & ECMWF are overestimating WT4 frequency

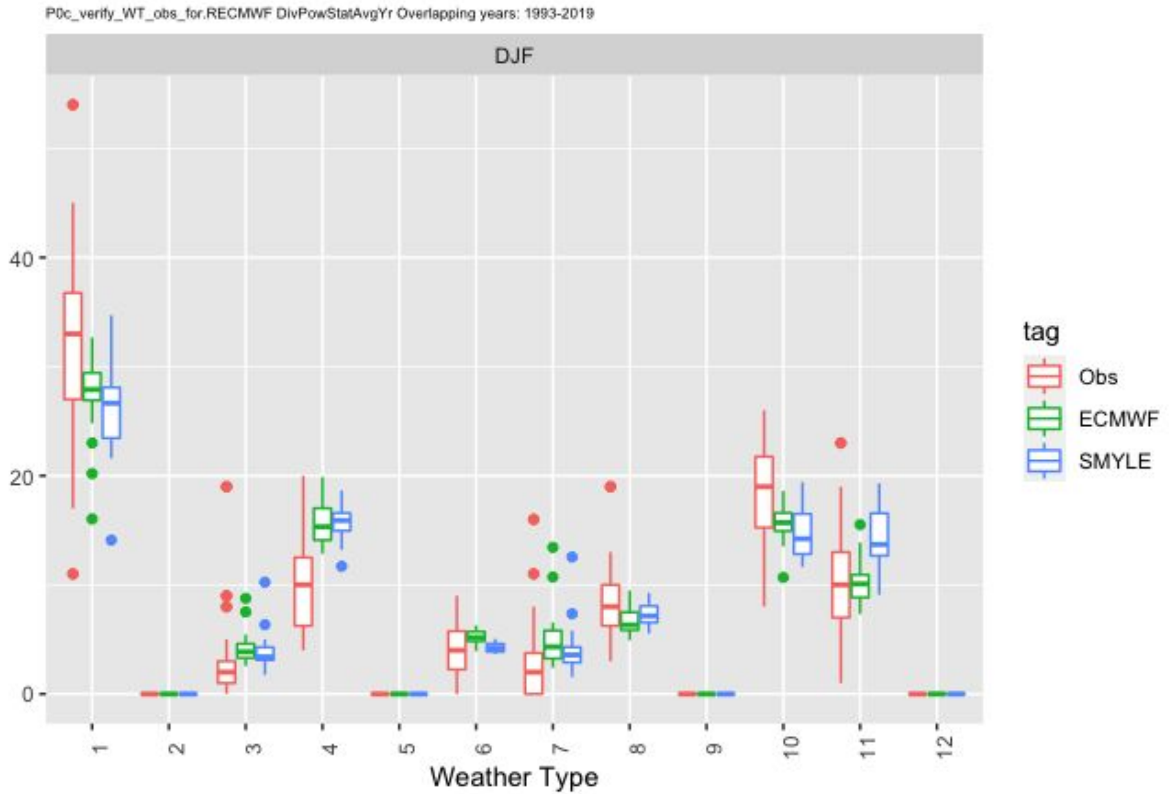


For the 4 WTs in JJA, neither forecast model performs well at capturing the ***average*** number of days in each WT.

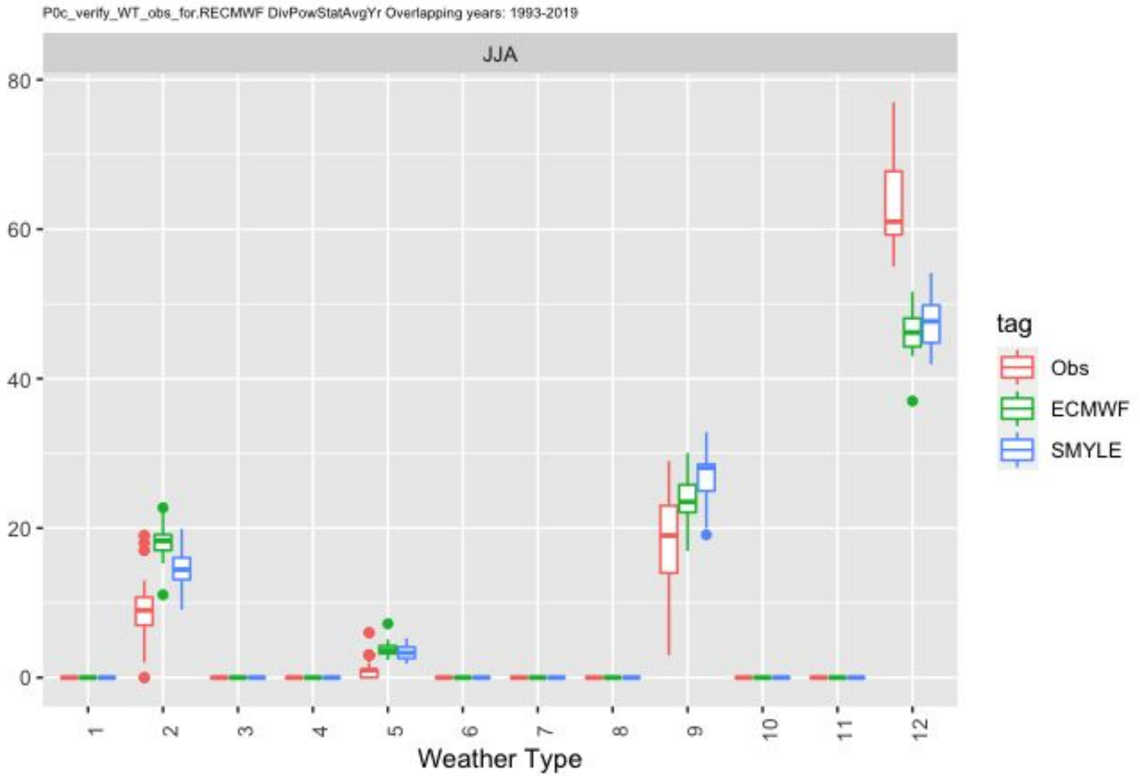



ECMWF P0c_verify_WT_obs_for_avgRelFreqOverlapping years: 1993-2019
DJF
JJA
MAM
SON
Average Number of Days
Weather Type
tag
Obs
ECMWF
SMYLE
0
20
40
60
1
2
3
4
5
6
7
8
9
10
11
12

For DJF, both forecast models tend to underestimate the variability of the WT frequencies.



For JJA, variability is being underestimated and forecast models are not performing well for the most frequent WT (WT12)



The power-divergence statistic can be used to compare observed and expected frequencies.

(Read and Cressie 1988; Gilleland et al. in review)

k=12, if we use all 12 WTs

\# observed days in WT *k*

$$\frac{2}{\lambda(\lambda+1)}\sum_{i=1}^{k} \text{observed}_i\left[\left(\frac{\text{observed}_i}{\text{expected}_i}\right)^{\lambda} - 1\right];$$

User selected parameter, lambda

\# forecasted days in WT *k*

- Best score is “0” (frequencies match perfectly)
- Bigger scores are worse (frequencies *diverge* more)
- Scores compared with test stat from Chi-squared distribution with degrees of freedom = k -1

```
>  sig_value = qchisq(.95, df=11)
> sig_value
[1] 19.7 # Needs to be less than 19.7
```

Looked at lambda = 1, 0, and ⅔, but only showing results from lambda =1.

The power-divergence statistic has two special cases where lambda = 1 and lambda -> 0.

(Read and Cressie 1988)

Two very important special cases of the power-divergence statistic are Pearson's X^2 statistic (put $\lambda = 1$)

$$\sum_{i=1}^{k} \frac{(\text{observed}_i - \text{expected}_i)^2}{\text{expected}_i}; \qquad (1.2)$$

and the loglikelihood ratio statistic G^2 (the limit as $\lambda \to 0$)

$$2 \sum_{i=1}^{k} \text{observed}_i \log\left[\frac{\text{observed}_i}{\text{expected}_i}\right]. \qquad (1.3)$$

The power-divergence statistic using lambda = 2/3 has some nice properties.

(Read and Cressie 1988)

The power-divergence statistic (1.1) provides an important link between the well-known statistics (1.2) and (1.3). This link provides a mechanism to derive more general results about the behavior of these statistics in both large and small samples (Chapters 4–8). As a result of this analysis, the power-divergence statistic with $\lambda = 2/3$ emerges; i.e.,

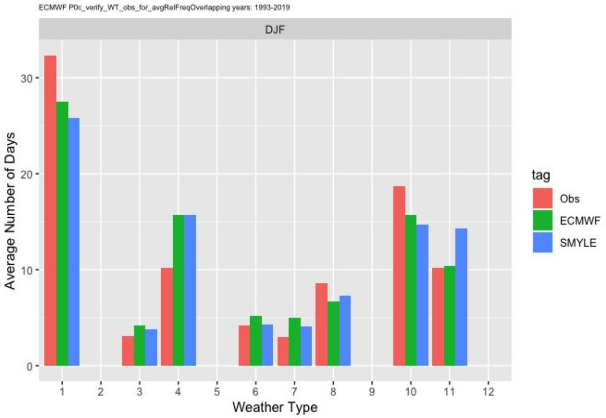
$$\frac{9}{5}\sum_{i=1}^{k} \text{observed}_i\left[\left(\frac{\text{observed}_i}{\text{expected}_i}\right)^{2/3} - 1\right].$$

This statistic lies "between" X^2 and G^2 in terms of the parameter λ, and has some excellent properties (Chapter 5). These properties are explained partially by examining the role of the power transformations in Sections 6.5 and 6.6.

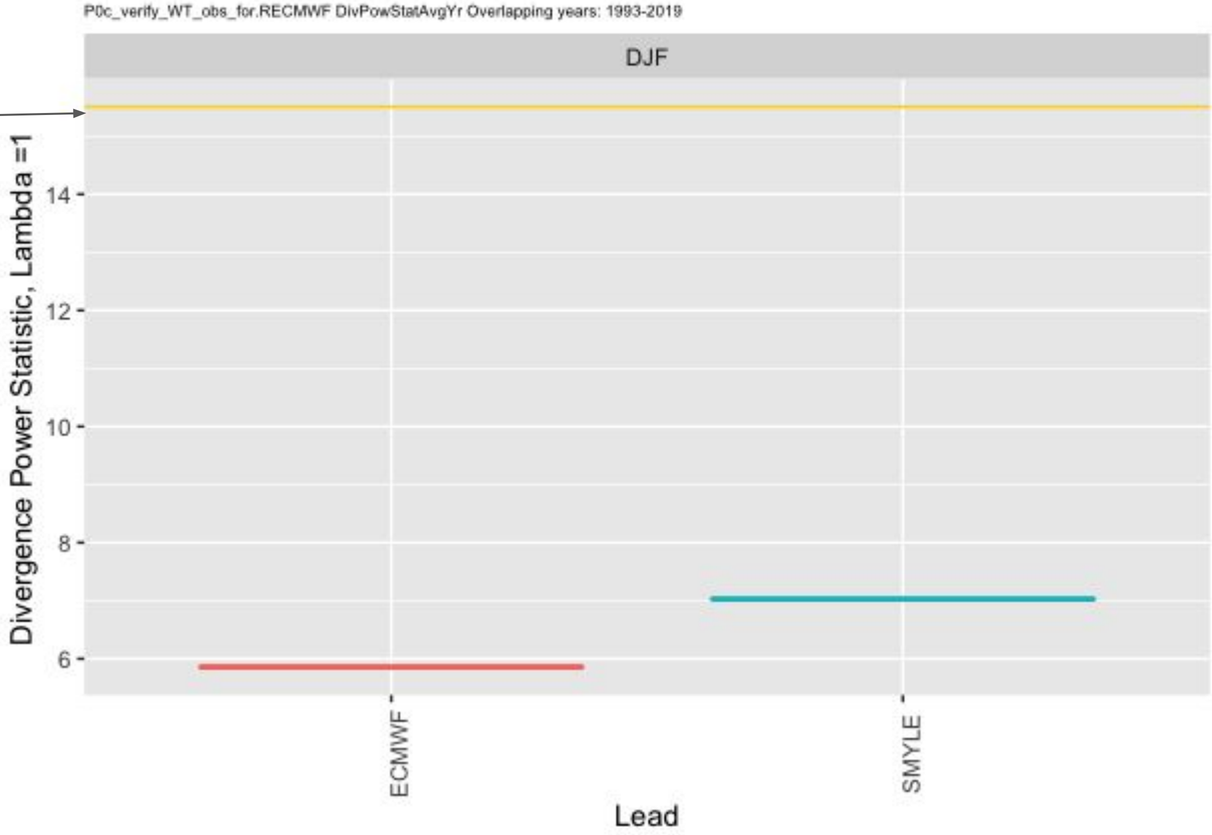
Looked at lambda = 1, 0, and ⅔, but only showing results from lambda =1.

For DJF, power divergence statistic indicates that frequency distributions of forecast models are not statistically different.

(i) Verify *average* number of days in each WT.

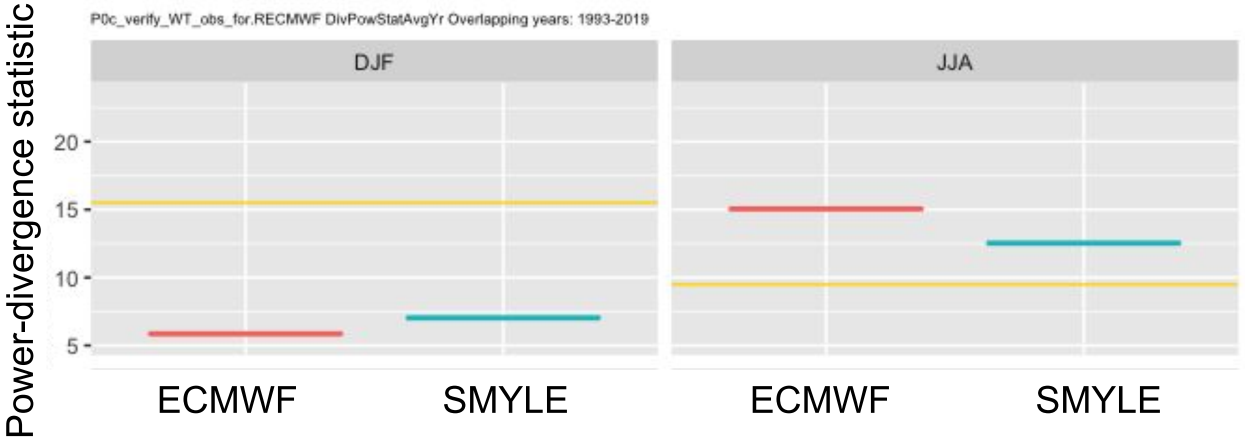


Test statistic



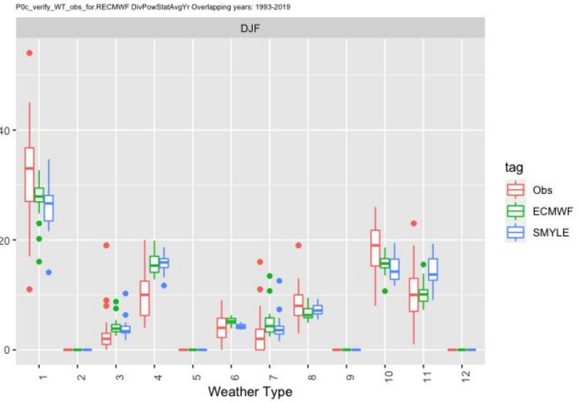
Seasonal differences are greater than forecast model differences.

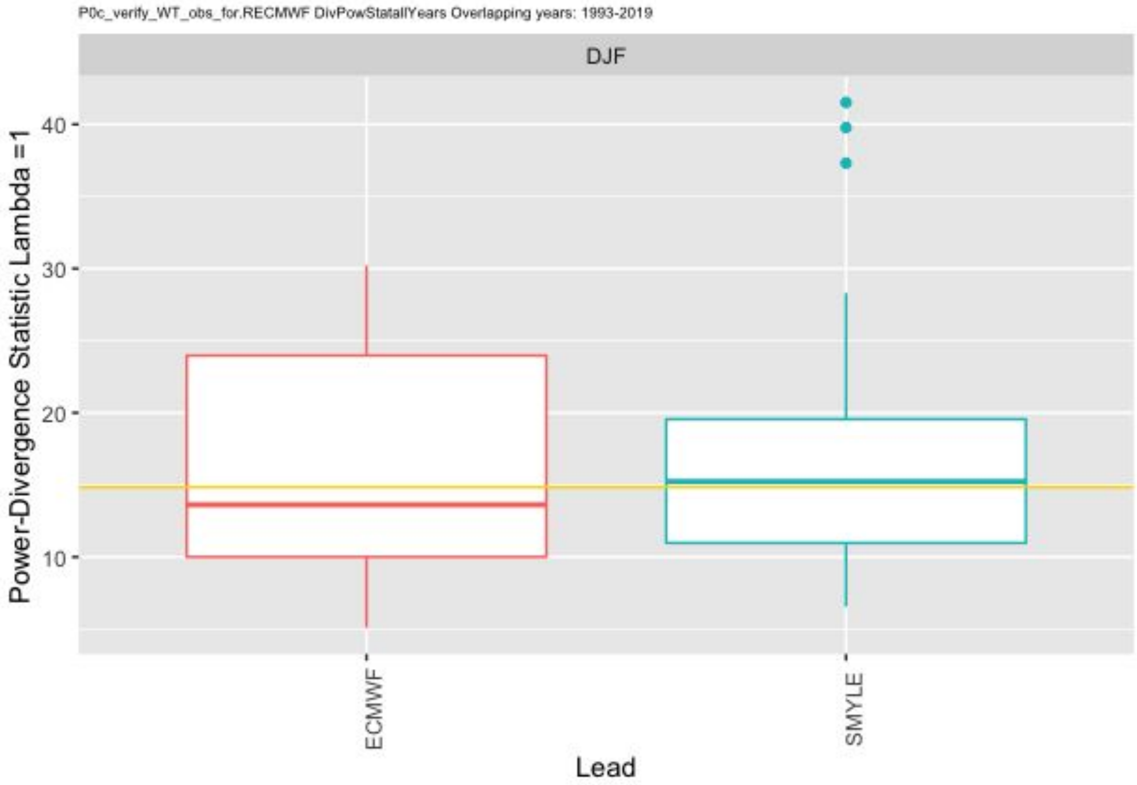
- DJF: both below ("good")
- JJA: both above ("bad")



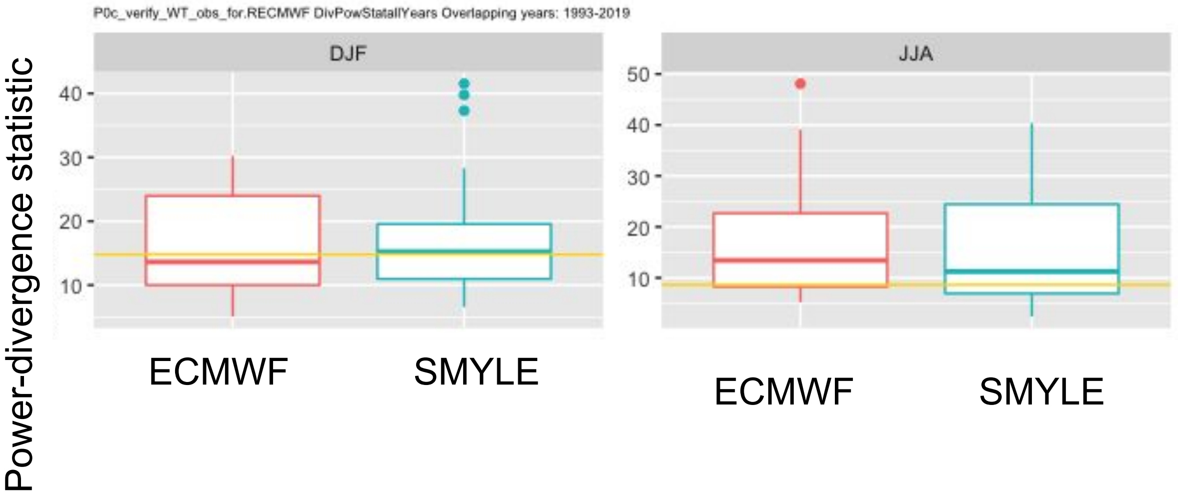
For DJF, about half the years have a distribution that is not statistically different than observed.

(ii) Verify number of days in each WT every year.





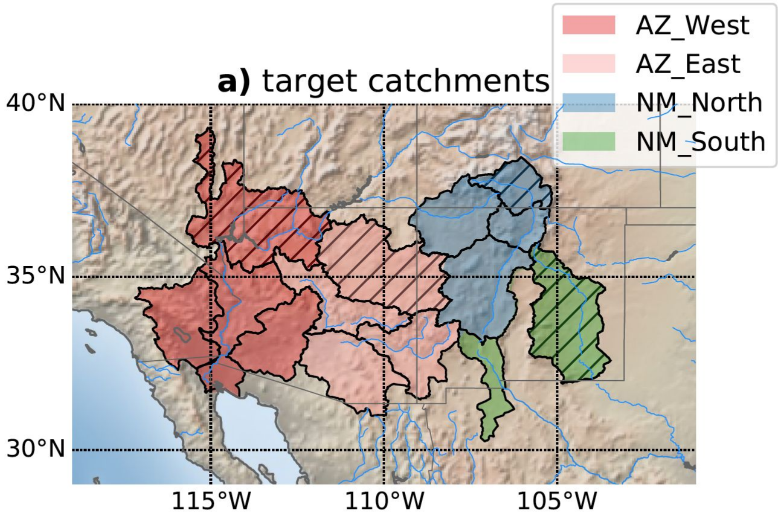
Year-to-year variability hard for the forecast models to capture; forecast models are similar.



Weather types were developed for all of CONUS and all days, could a more targeted approach improve predictability?

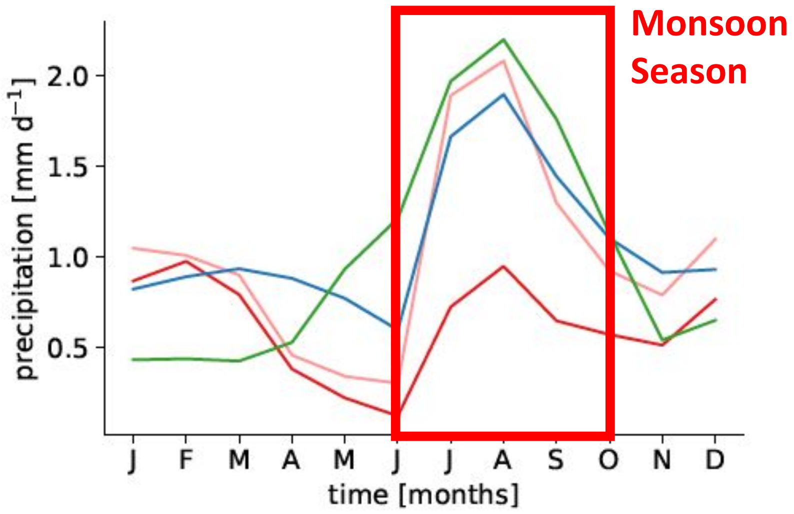
- Weather types are developed for Southwest domains over AZ and NM for monsoon season (June - Oct). Will this improve results?

4 regions identified based on Weather Type (WT) analysis for each HUC6.



Basins can be aggregated into regions that feature similar Weather Types. Those regions are

1. Arizona West (dark red)
2. Arizona East (light red)
3. New Mexico North (blue)
4. New Mexico South (green)

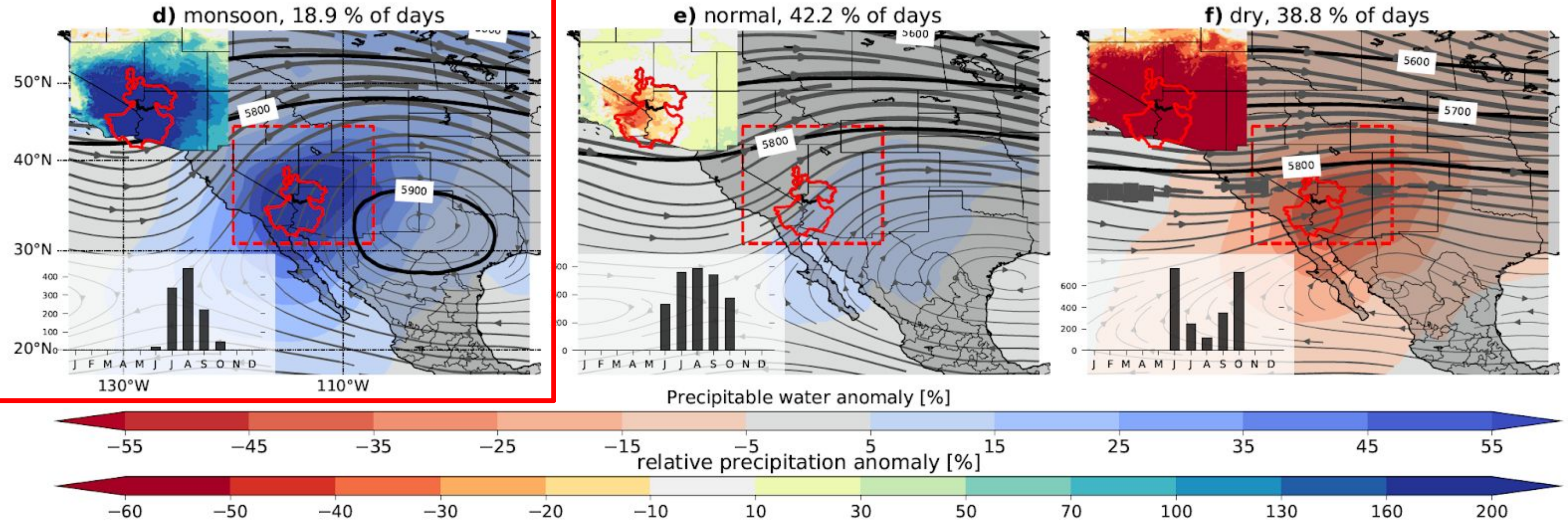


BUREAU OF RECLAMATION

Prein, Towler, et al. *GRL* (2022)

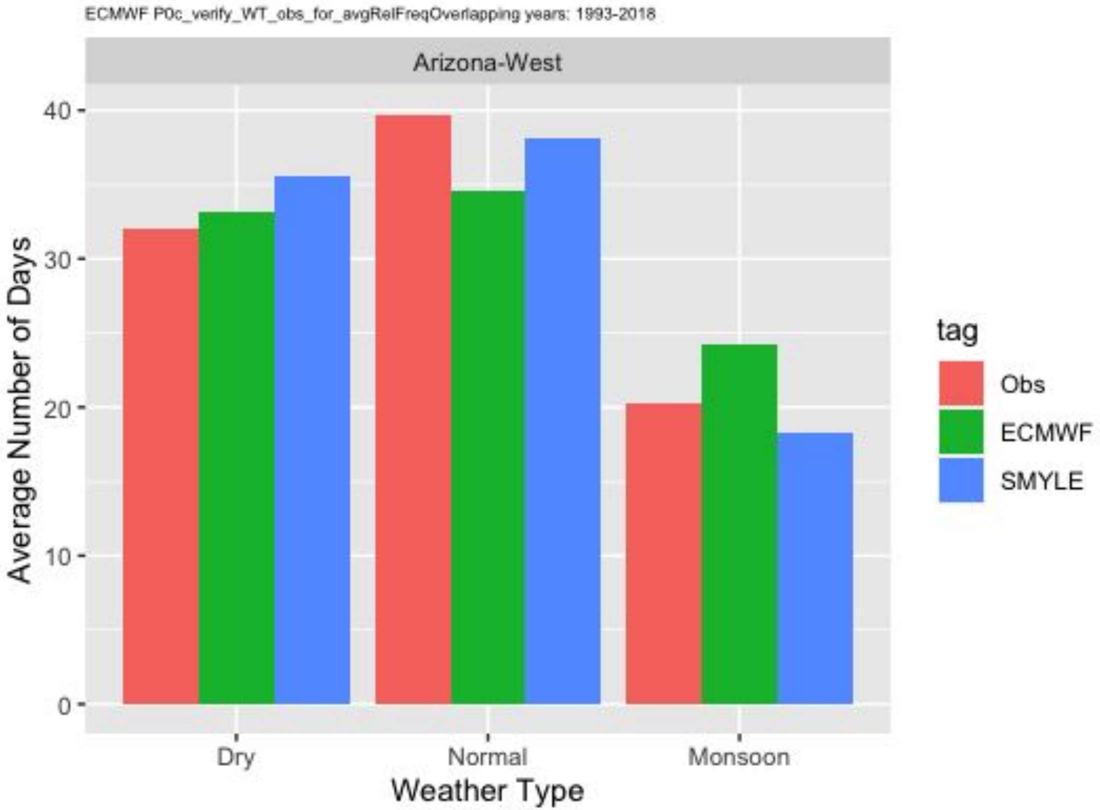
Each region has a monsoon, normal, and dry weather type.

Monsoonal Moisture Surge | **Normal Conditions** | **Dry Conditions**

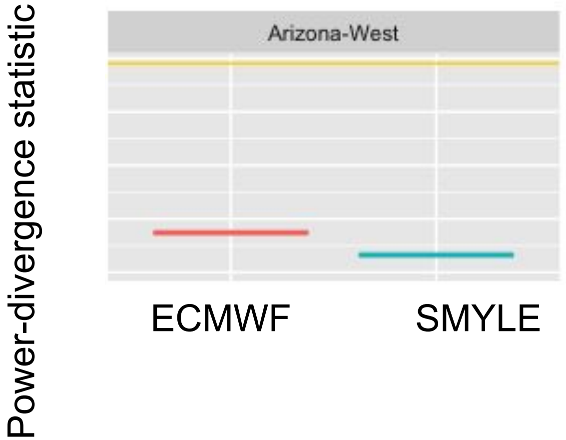


Average WT precipitable water anomalies (colored contour), 850 hPa wind speed (streamlines), and 500 hPa geopotential height (contours)

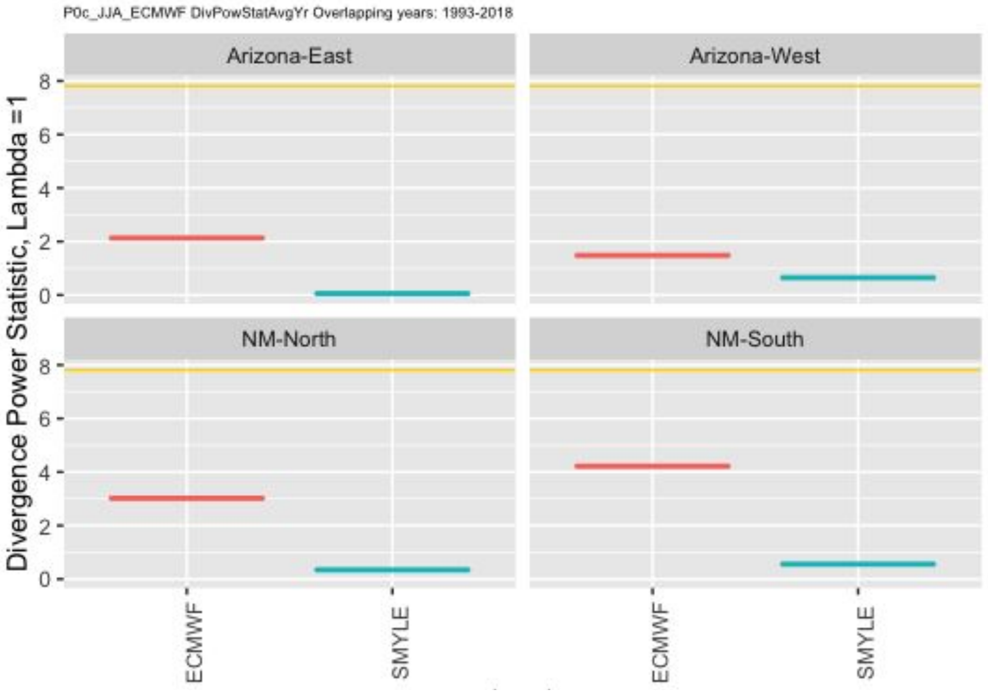
For the SW WTs, average number of days is well forecasted by both forecast models.



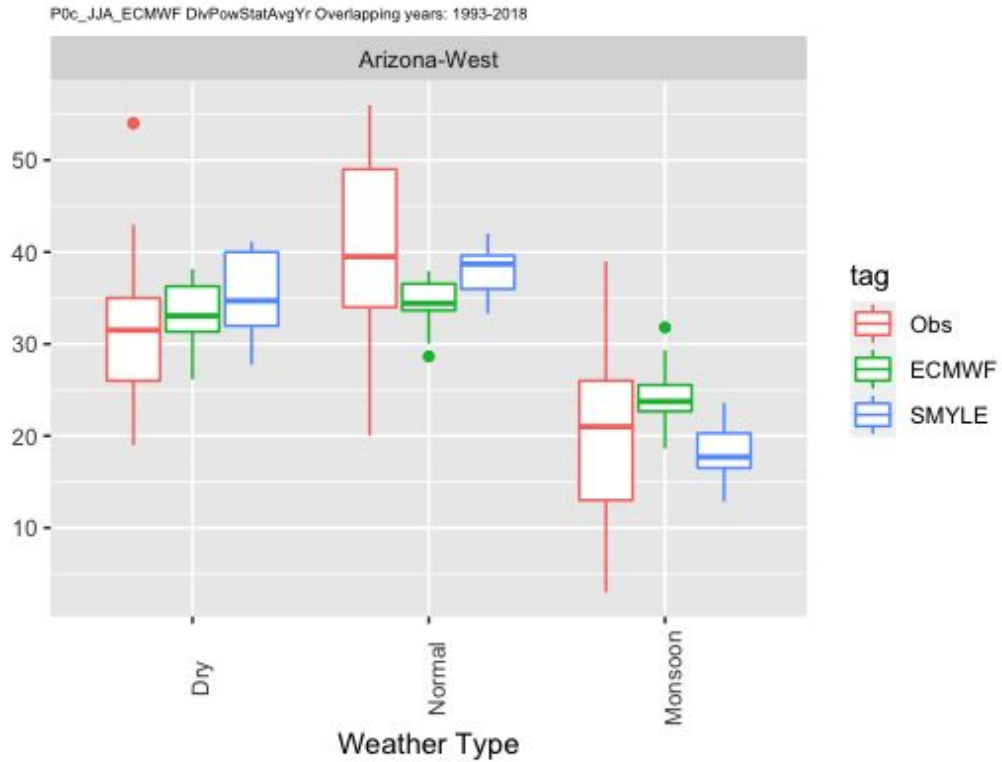
For Arizona-West annual average, power-divergence statistic is “good” (below threshold) for both models.



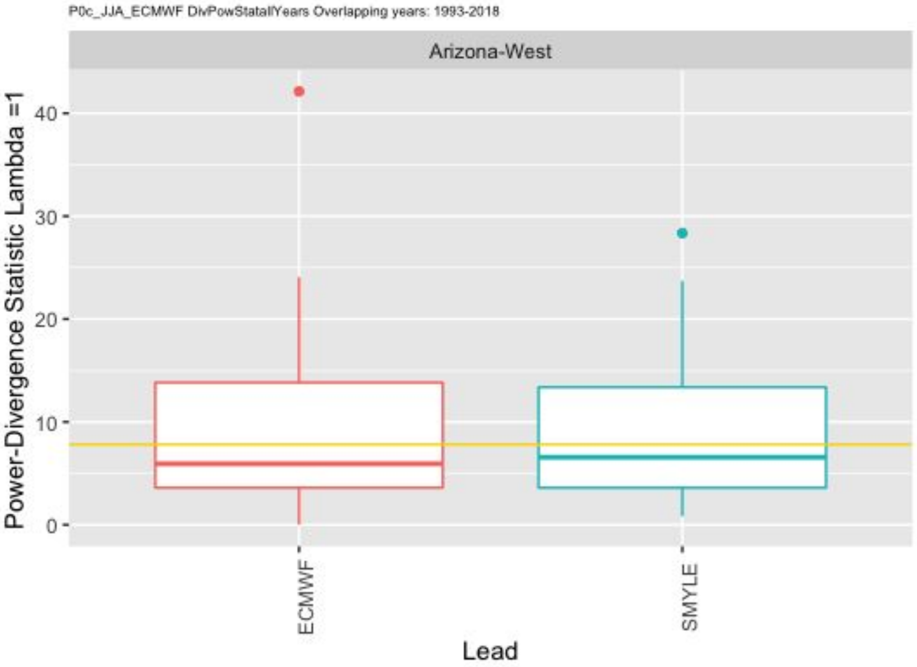
For annual average, power-divergence statistic is "good" (below threshold) for all regions. SMYLE slightly better than ECMWF.



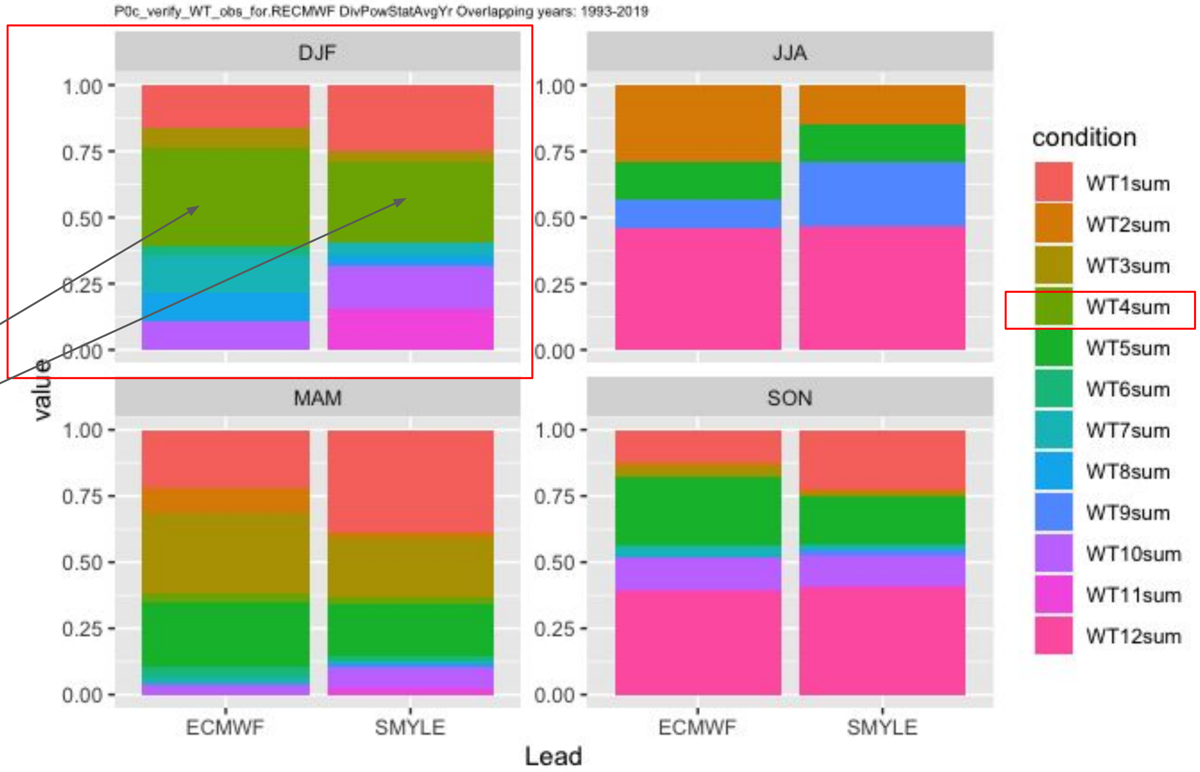
For AZ-West: forecast models capture climatology well but underestimates variability.



For Arizona-West: Power-divergence statistic is under the threshold a bit more than 50% of time.



For DJF, the largest fraction of the power-divergence statistic comes from WT4 for both ECMWF and SMYLE

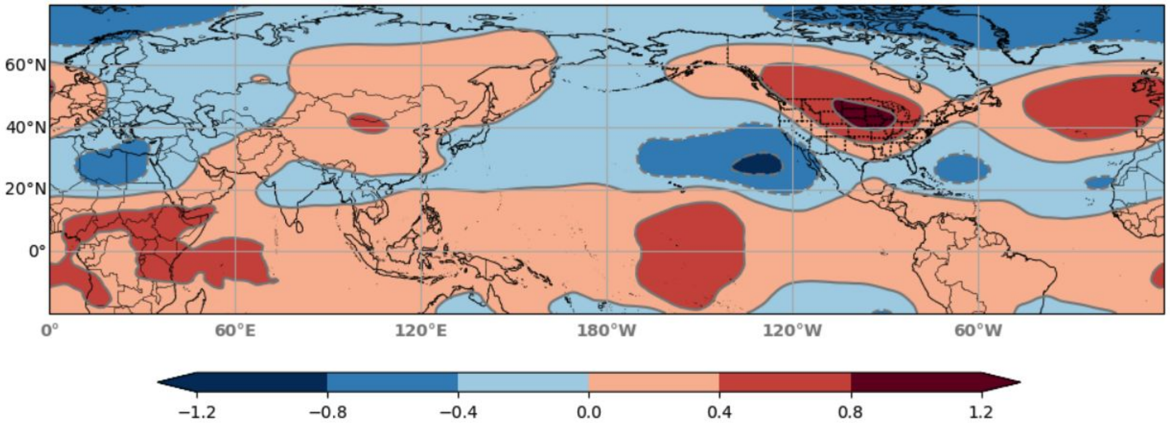


Which WTs are we missing the “worst”? Why?

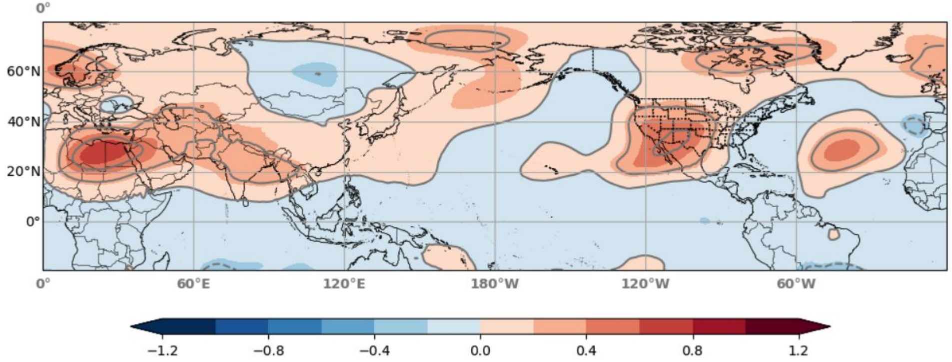
Wet California Winter WT4 pattern connects to the tropical Pacific

Z500 anomaly of years with most WT4 occurrences / stddev all years

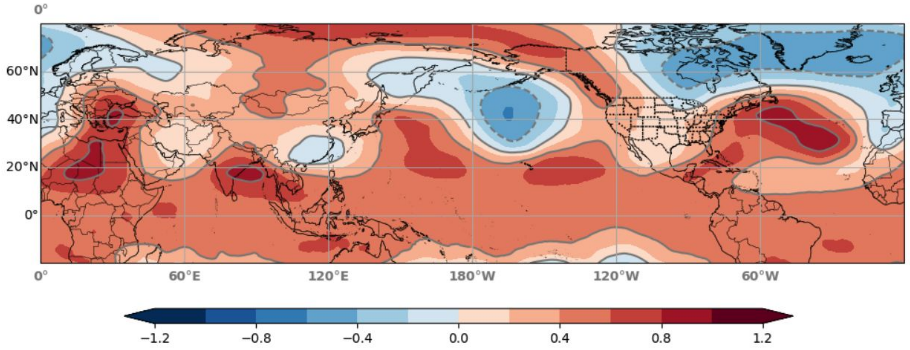
ERA-I



SMYLE 1 member



ECMWF 1 member



Recent stakeholder workshops* on co-producing useable S2S forecasts can help guide next steps.

*NSF-funded PRES^{2}iP (OU and NCAR; PI Lazrus)

VanBuskirk et al. 2021, *BAMS*

VanBuskirk et al. 2023, *BAMS*

Conclusions

- Seasonal forecast products can reproduce dominant weather patterns on average, but struggle with year-to-year variability
  - SMYLE is on par with operational ECMWF products.
  - CONUS-wide WTs work for DJF, but more targeted approach needed for JJA.
- Low variability could be due to taking ensemble average (and/or forecast products tending to be underdispersed).

Next Steps

- Identify alternate ways to verify for useability
- Process exploration to determine if forecasts are right for the right reasons (Look at wave patterns, teleconnections, etc).
- What types (i.e., wet/dry) of years are we getting correct?

Thank you!

towler@ucar.edu