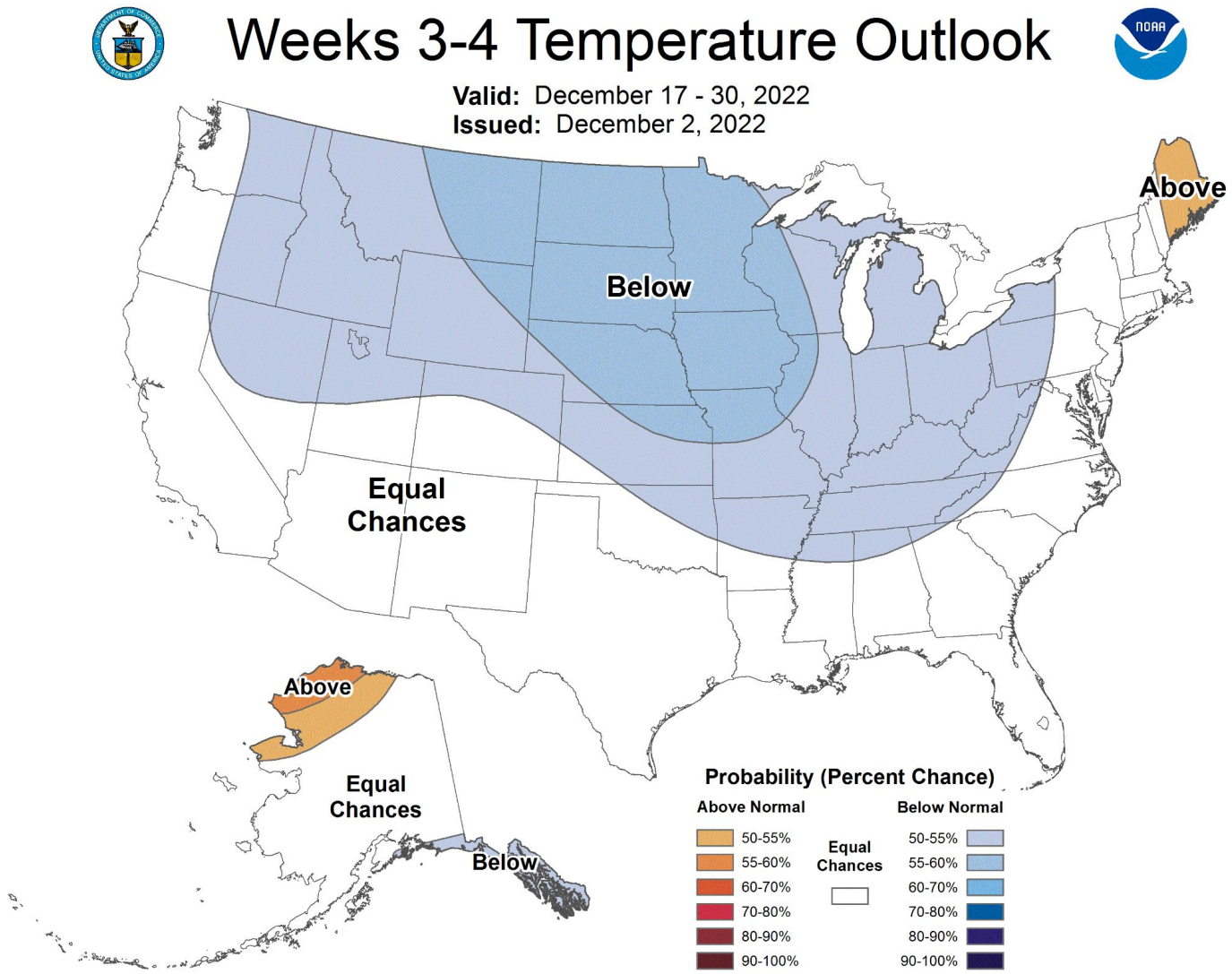
Weeks 3-4 Multi-Model Ensemble Subsampling: A Real-time Verification

Cory Baggett[1]

[1]Climate Prediction Center / NCEP / NWS / NOAA

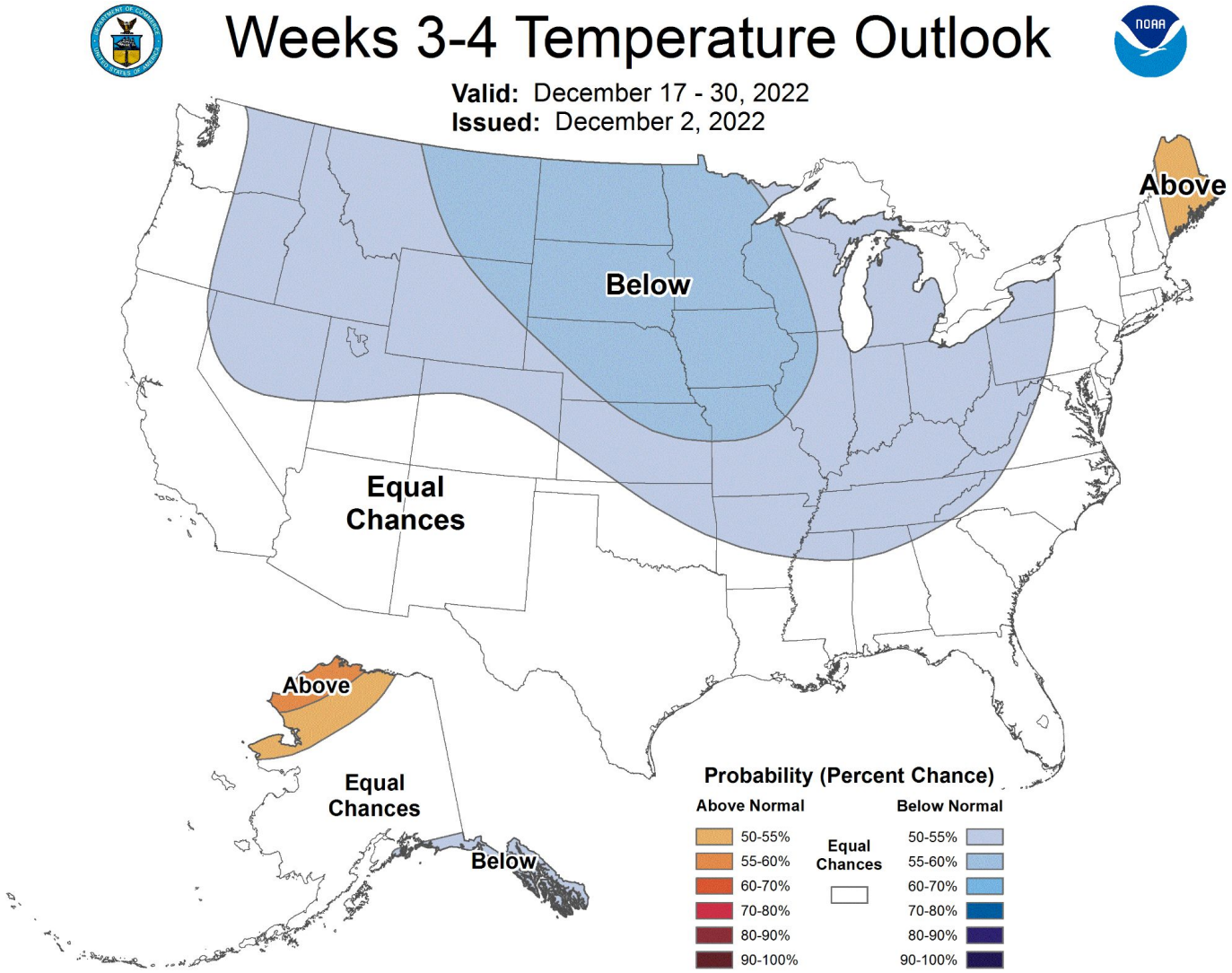
9th NOAA Ensemble Users Workshop

August 23, 2023



Motivation

- The Weeks 3-4 Temperature and Precipitation Outlooks are issued by the Climate Prediction Center each Friday.
- The outlooks provide stakeholders with probabilistic forecasts of the likelihood of above or below normal temperature and precipitation during the 2-week period beginning at lead Day 15.
- A large variety of dynamical and statistical model guidance are available to the forecaster to help produce the outlook.



Motivation

Dynamical model guidance comes primarily from 5 extended-range ensemble prediction systems provided by:

ECMWF	51 members	European Centre for Medium-range Weather Forecasting
GEFSv12	31 members	NCEP Global Ensemble Forecast System Version 12
CFSv2	32 members	NCEP Climate Forecast System Version 2
ECCC	21 members	Environment and Climate Change Canada
JMA	50 members	Japan Meteorological Agency

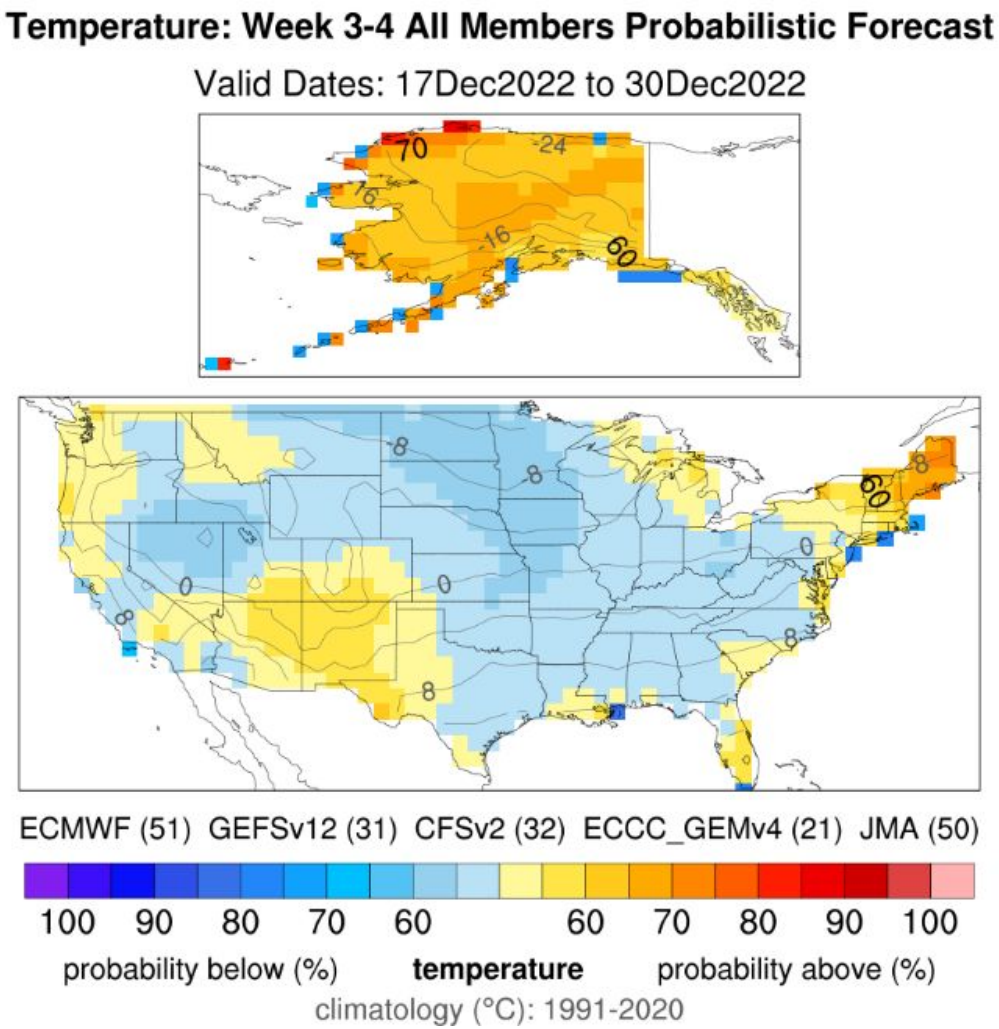
A total of 185 ensemble members in this multi-model ensemble suite.

Motivation

Question: Is there a subsample of these 185 ensemble members that can be chosen objectively in real-time that provides a more skillful forecast than the all member multi-model ensemble suite?

- It's possible such a subsample exists in a multi-model ensemble suite, where not all models are produced using the same protocols, physics, and/or perturbation schemes.
- Work for this project was funded as an OSTI NWS Fiscal Year 2020 Milestone.
- A promising subsampling method was implemented as an experimental real-time tool at the beginning of Fiscal Year 2021.
- This presentation highlights results from the experimental real-time tool.
- Work is ongoing to move the tool into operations.

Example Case - All Member Forecast



Weeks 3-4 All-Member Temperature Probabilistic Forecast

Method: All (185) Members
Issued: December 2, 2023
Valid: December 17-30, 2023

Synopsis: Above normal temperatures favored in Alaska; below normal temperatures favored throughout much of CONUS

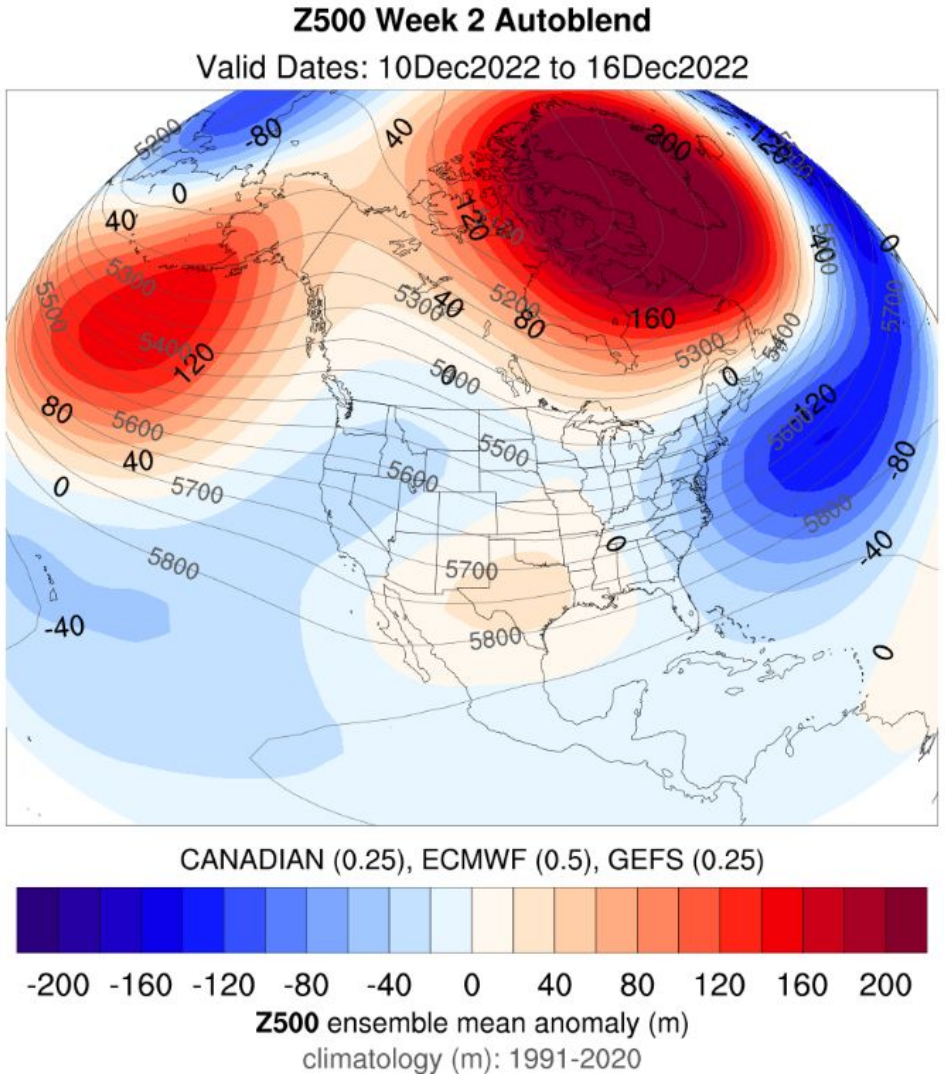
Example Case - Subsampling Method

Week 2 500-hPa Height Anomaly Autoblend

Method: Weighted-mean of the ensemble mean anomalies from the ECCC (25%), ECMWF (50%), and GEFS (25%)

Issued: December 2, 2023

Valid: December 10-16, 2023



Example Case - Subsampling Method

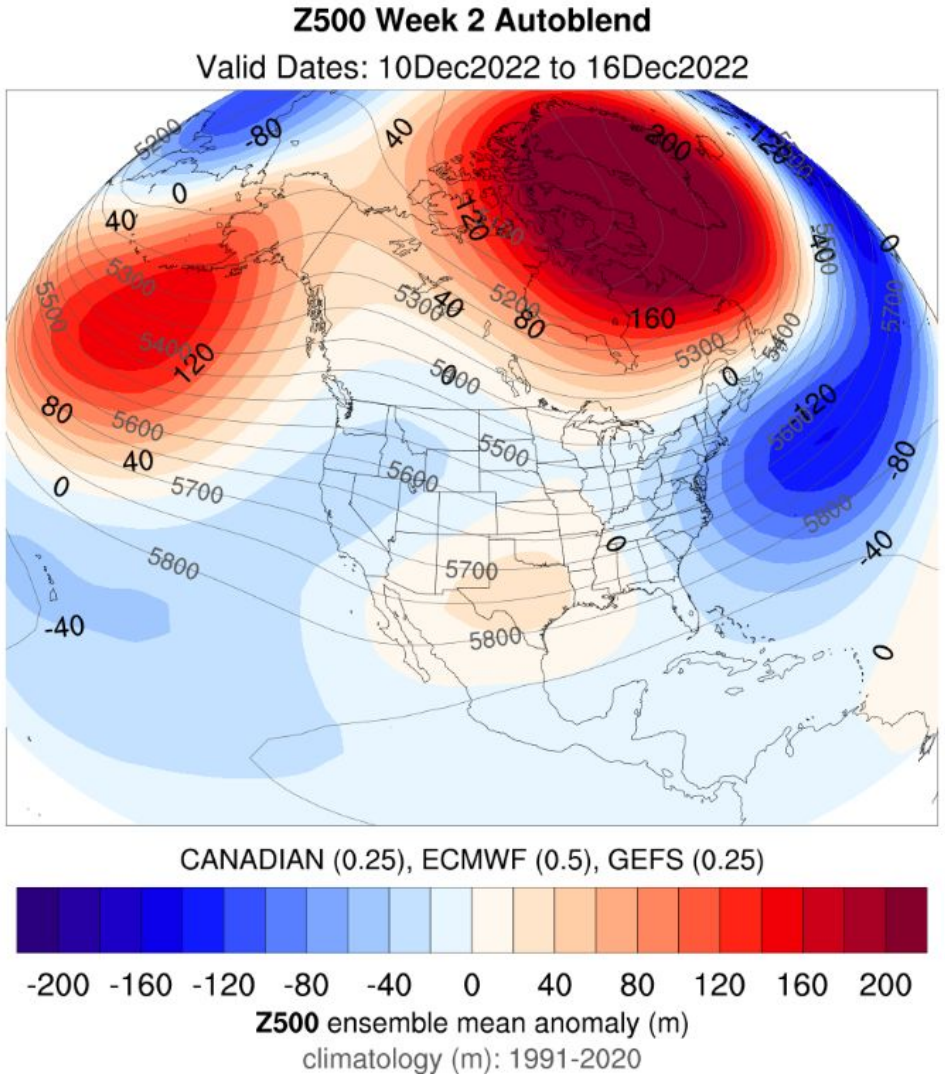
1. Calculate the pattern correlations between:

   the Week 2 500-hPa Height Anomaly Autoblend

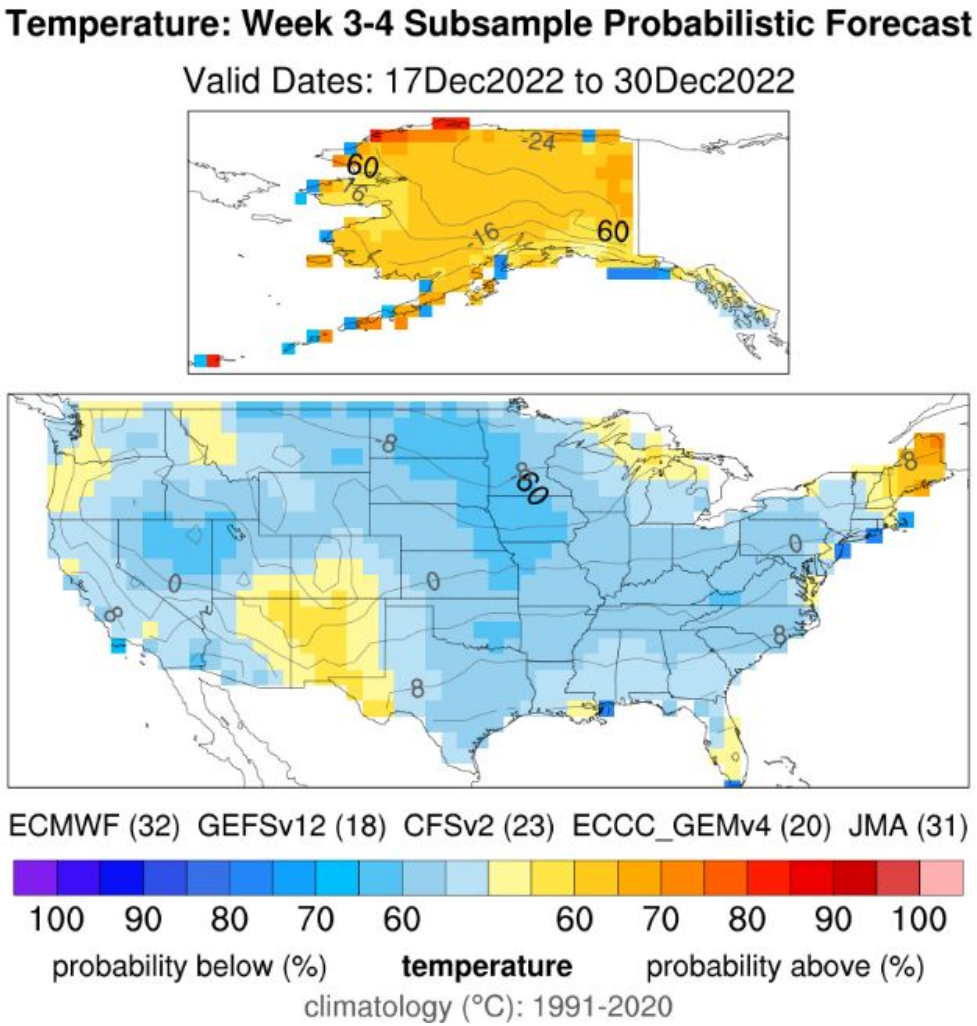
   and

   the Week 2 500-hPa Height Anomaly forecast from each individual member of the 185 member multi-model ensemble suite.

2. Rank these 185 pattern correlation values.
3. Subsample the members with the highest pattern correlations.
4. From this subsample, create :
   a. Weeks 3-4 500-hPa Heights Ensemble Mean Anomaly Forecast
   b. Weeks 3-4 Temperature Probabilistic Forecast
   c. Weeks 3-4 Precipitation Probabilistic Forecast



Example Case - 67% Subsample Forecast



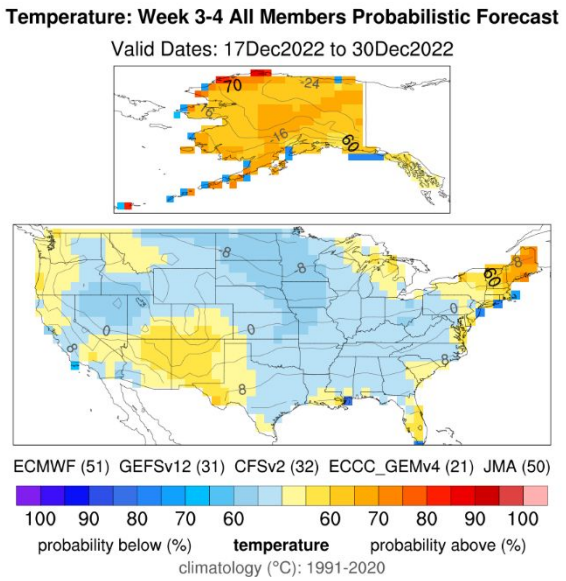
Weeks 3-4 67% Subsample Temperature Probabilistic Forecast

Method: 67% (124) Members
Issued: December 2, 2023
Valid: December 17-30, 2023

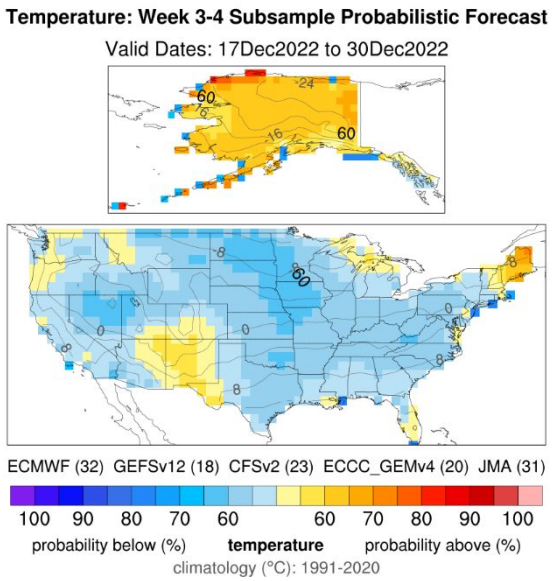
Synopsis: Above normal temperatures favored in Alaska; Below normal temperatures favored over much of CONUS.

Example Case - Difference between All Member and 67% Subsample Forecasts

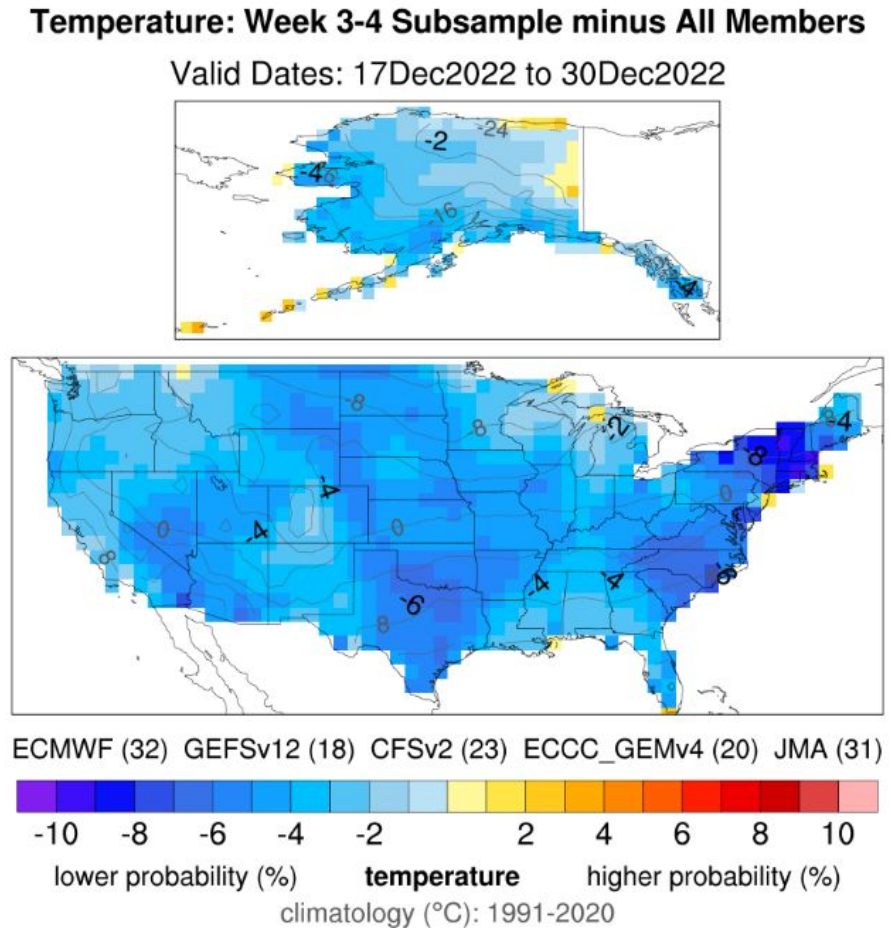
All Members



67% Subsample



Difference



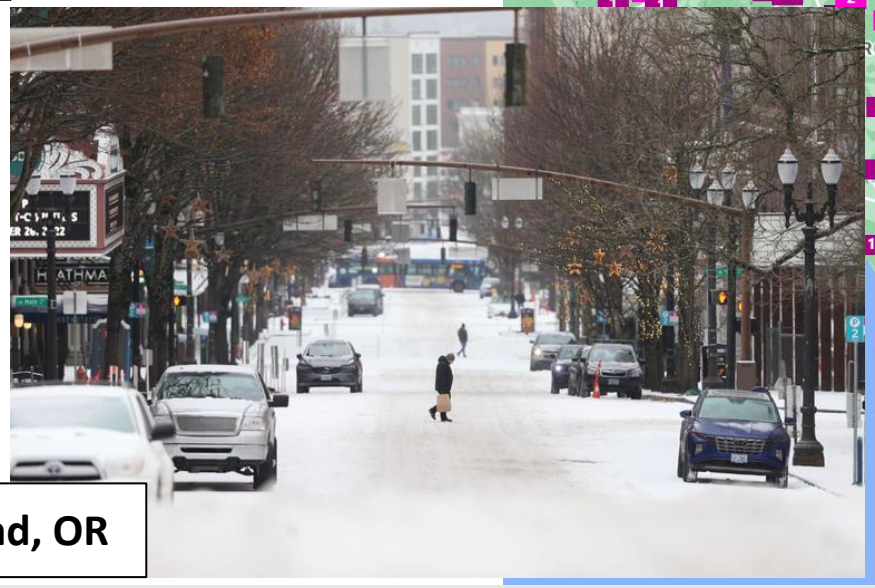
The subsample forecast increased the probabilities of below normal temperatures across nearly all of Alaska and CONUS.

Example Case - What verified?

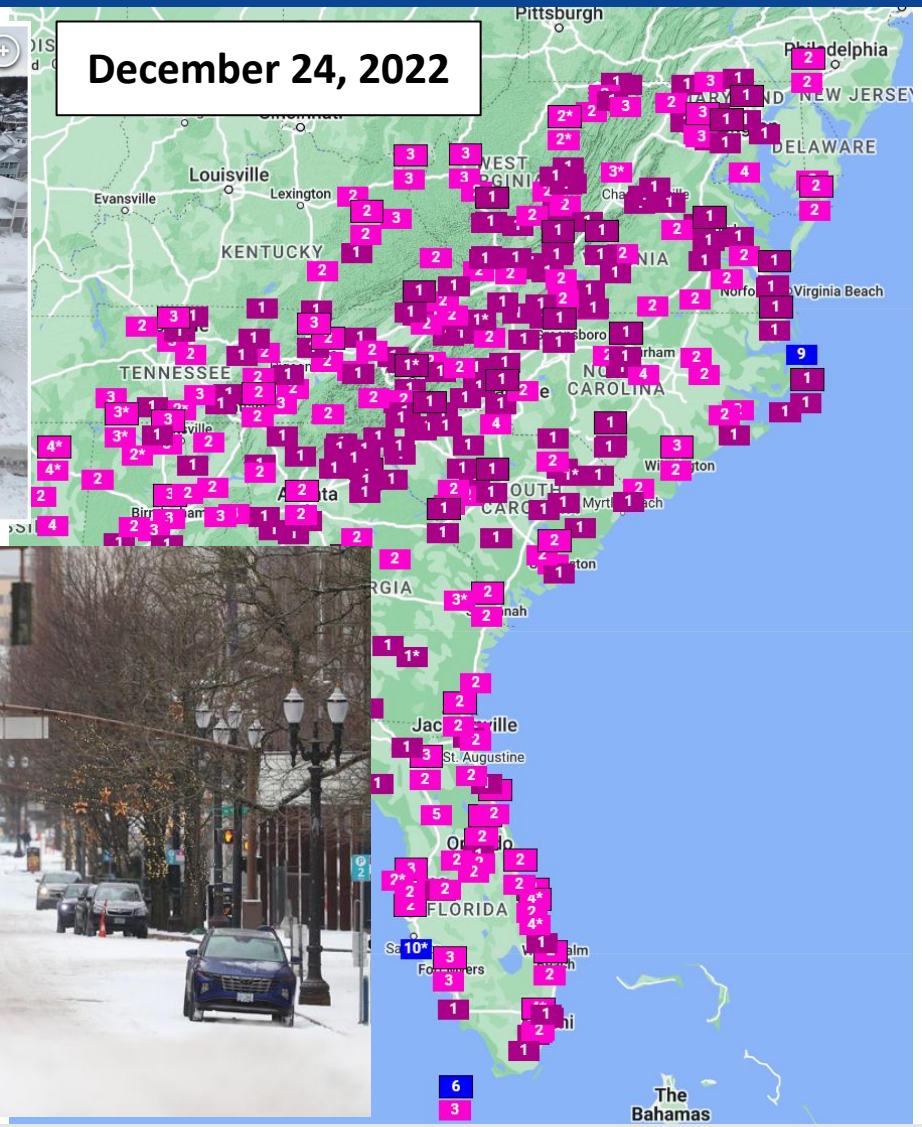
Ohio

Buffalo

Everywhere

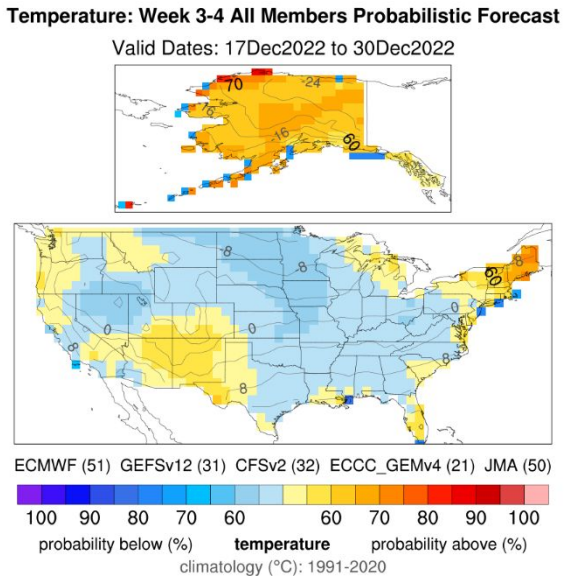


Portland, OR



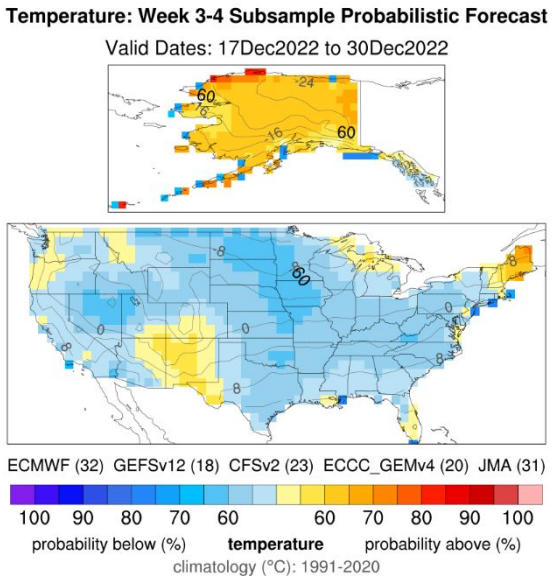
Example Case - Verification

All Members



HSS = 27.5

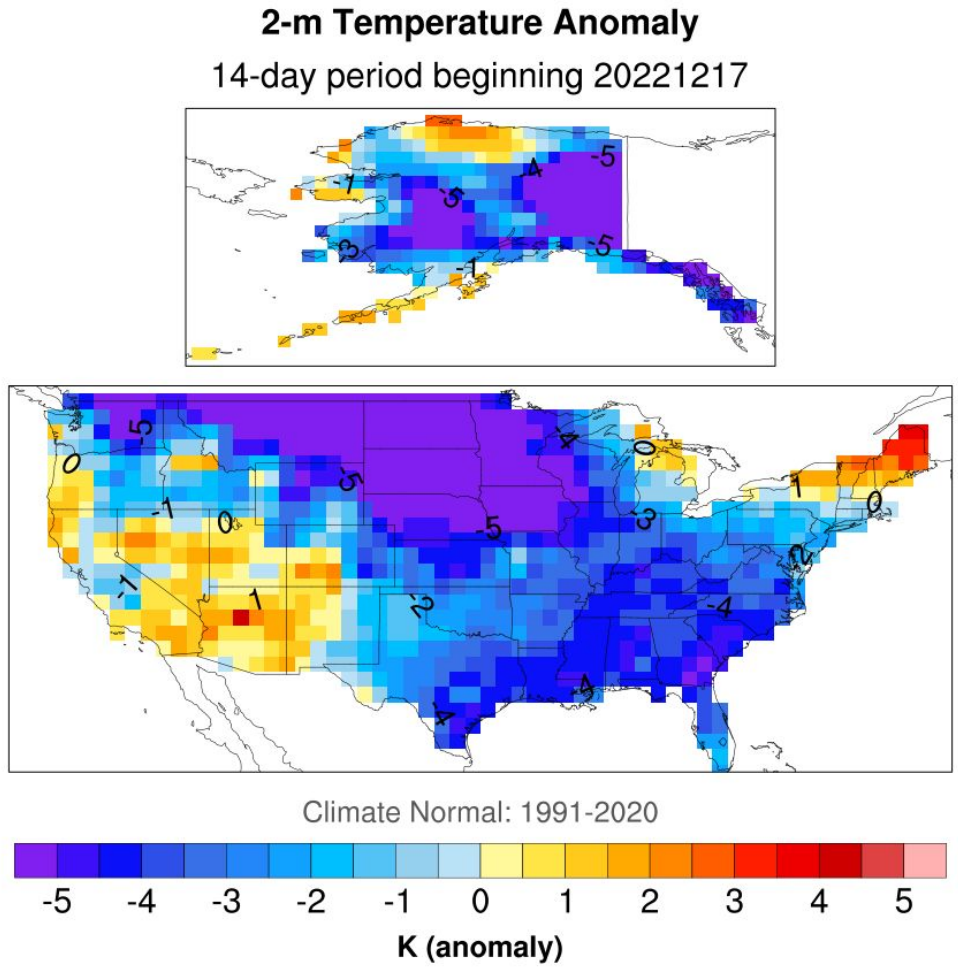
67% Subsample



HSS = 44.6

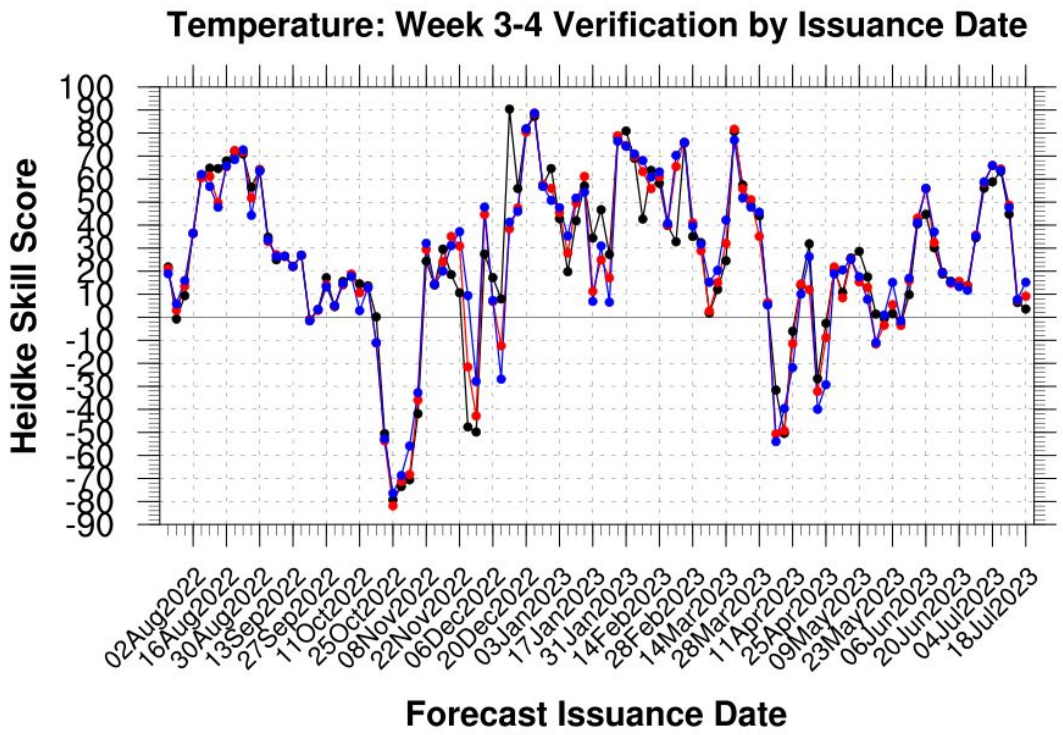
In this example case, the 67% subsample forecast has a higher skill score compared to the all member forecast.

Observed



Real-time Verifications

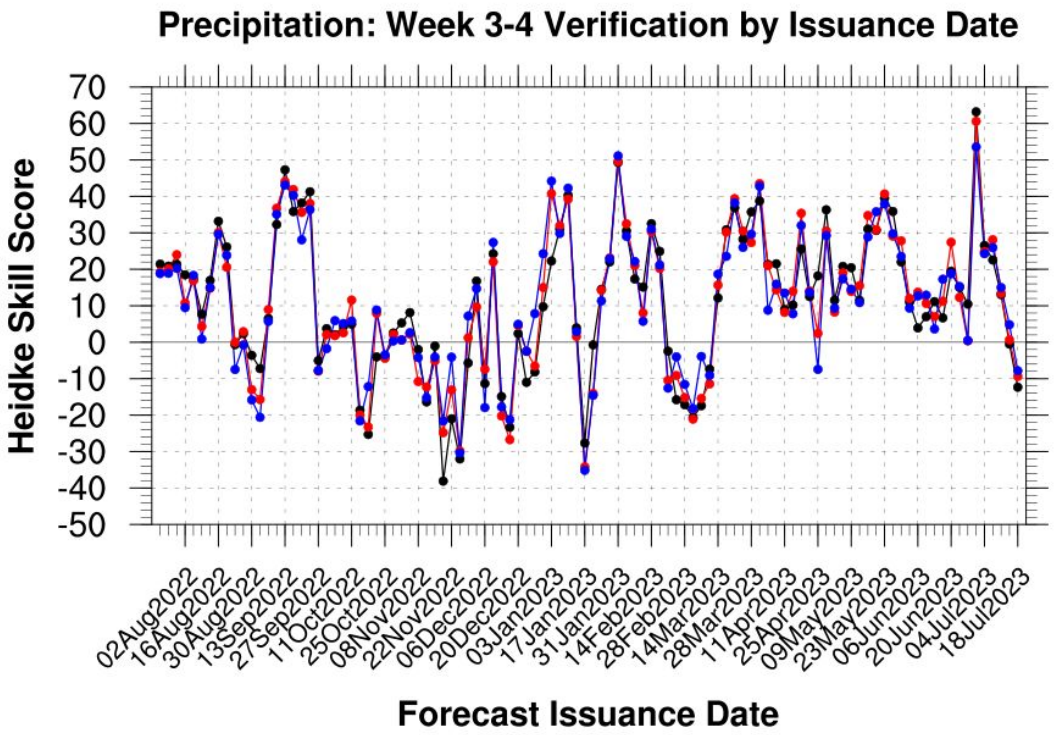
Temperature



All Members (avg: 25.2) | 67% of Members (avg: 24.6) | 50% of Members (avg: 25.3)

Domain: CONUS/AK; **# of Forecasts:** 104

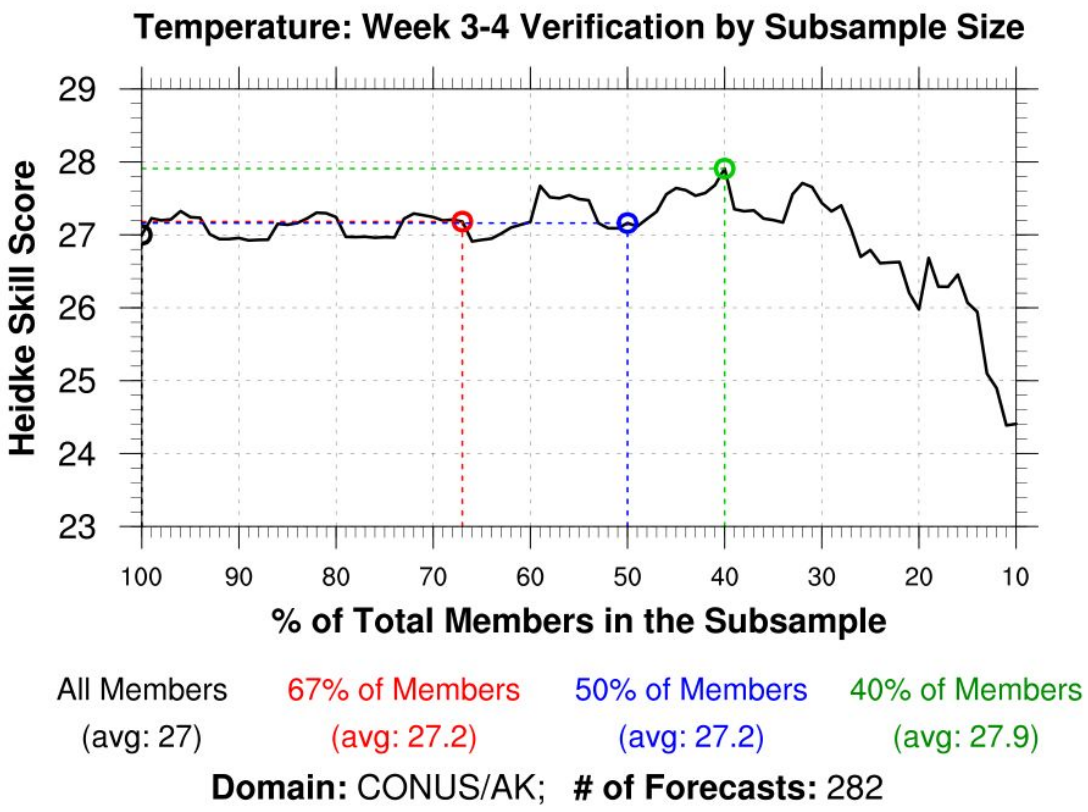
Precipitation

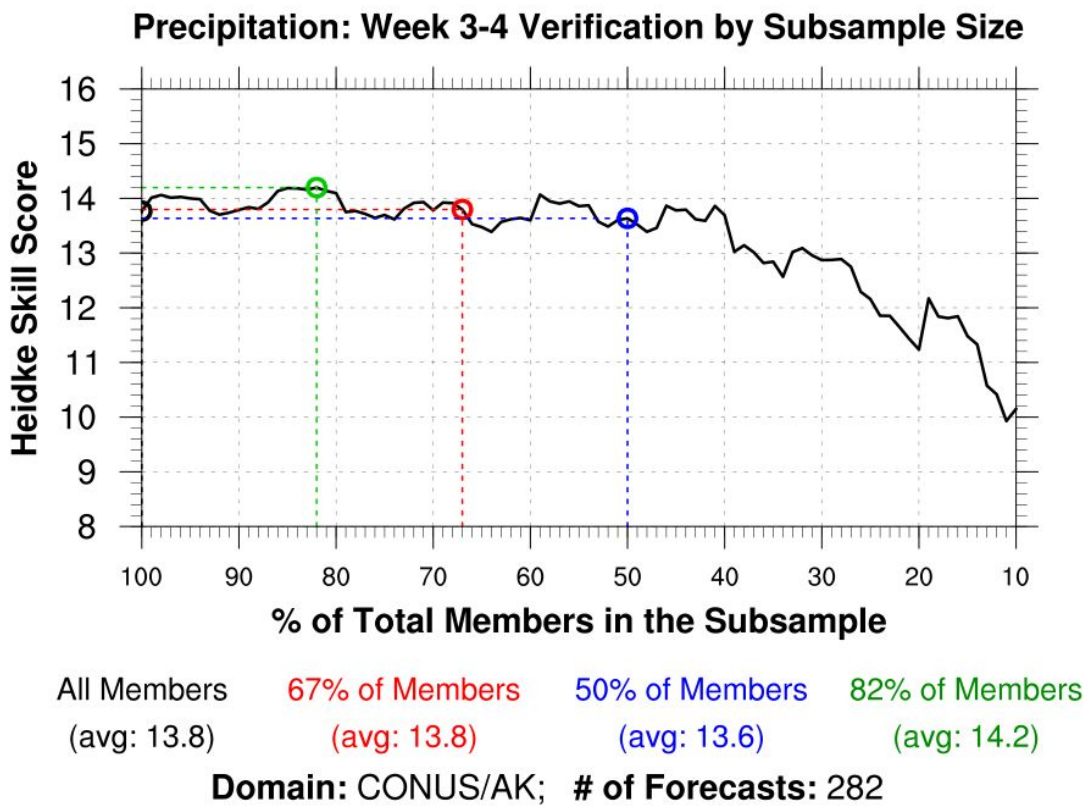


All Members (avg: 11.2) | 67% of Members (avg: 11) | 50% of Members (avg: 11)

Domain: CONUS/AK; **# of Forecasts:** 104

Real-time Verifications

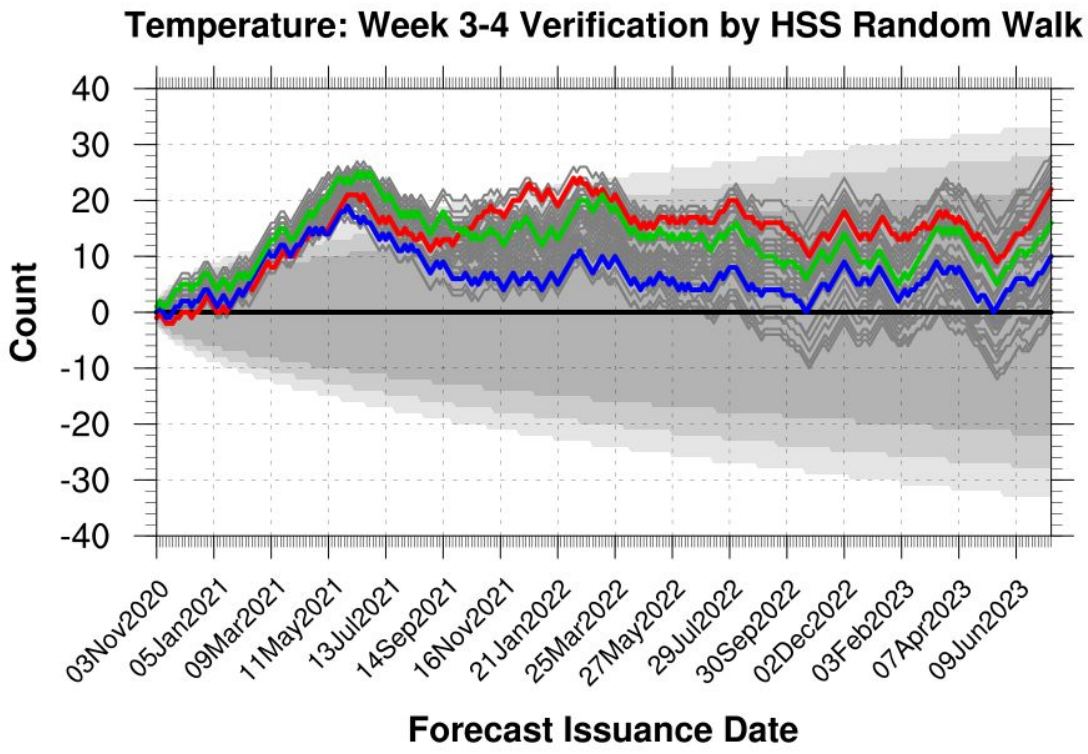




~3.3% improvement using 40% of members

~2.9% improvement using 82% of members

Real-time Verifications



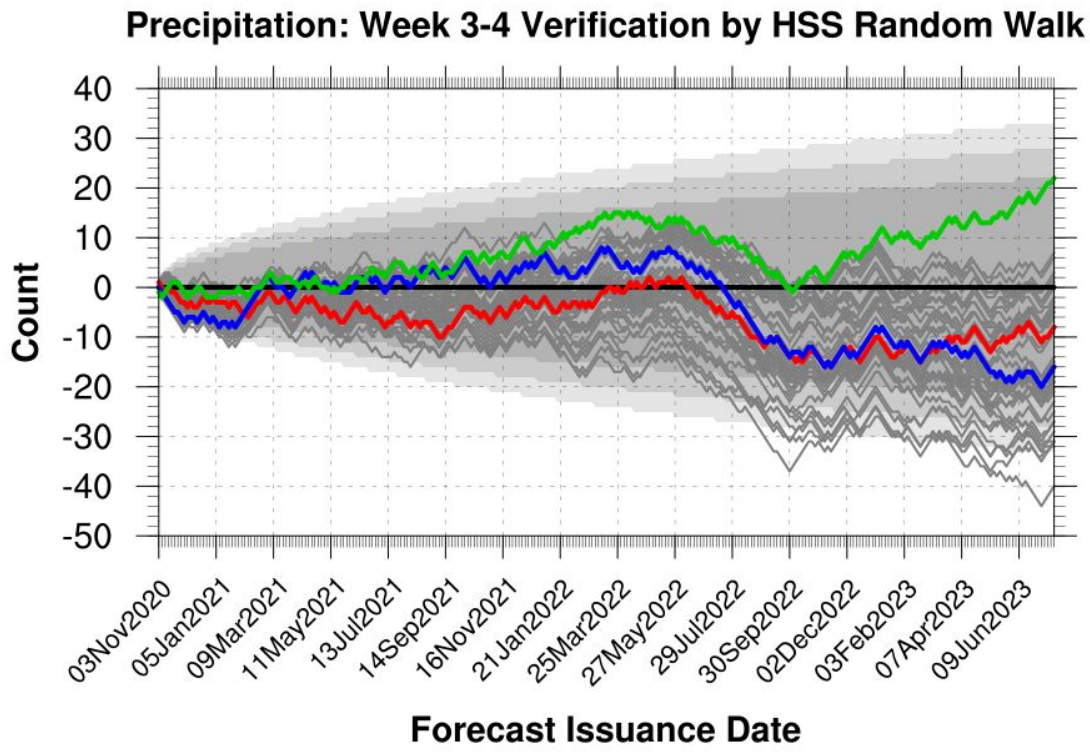
All Members	67% of Members	50% of Members	40% of Members
(Count: 0)	(Count: 22)	(Count: 10)	(Count: 16)

Gray Lines: 70% to 30% of Members

Gray Shading: α = 0.20, 0.10, 0.05 (dark to light)

Positive Count: Subsample Wins; **Negative Count:** All Members Wins

Domain: CONUS/AK; **# of Forecasts:** 282



All Members	67% of Members	50% of Members	82% of Members
(Count: 0)	(Count: -8)	(Count: -16)	(Count: 22)

Gray Lines: 70% to 30% of Members

Gray Shading: α = 0.20, 0.10, 0.05 (dark to light)

Positive Count: Subsample Wins; **Negative Count:** All Members Wins

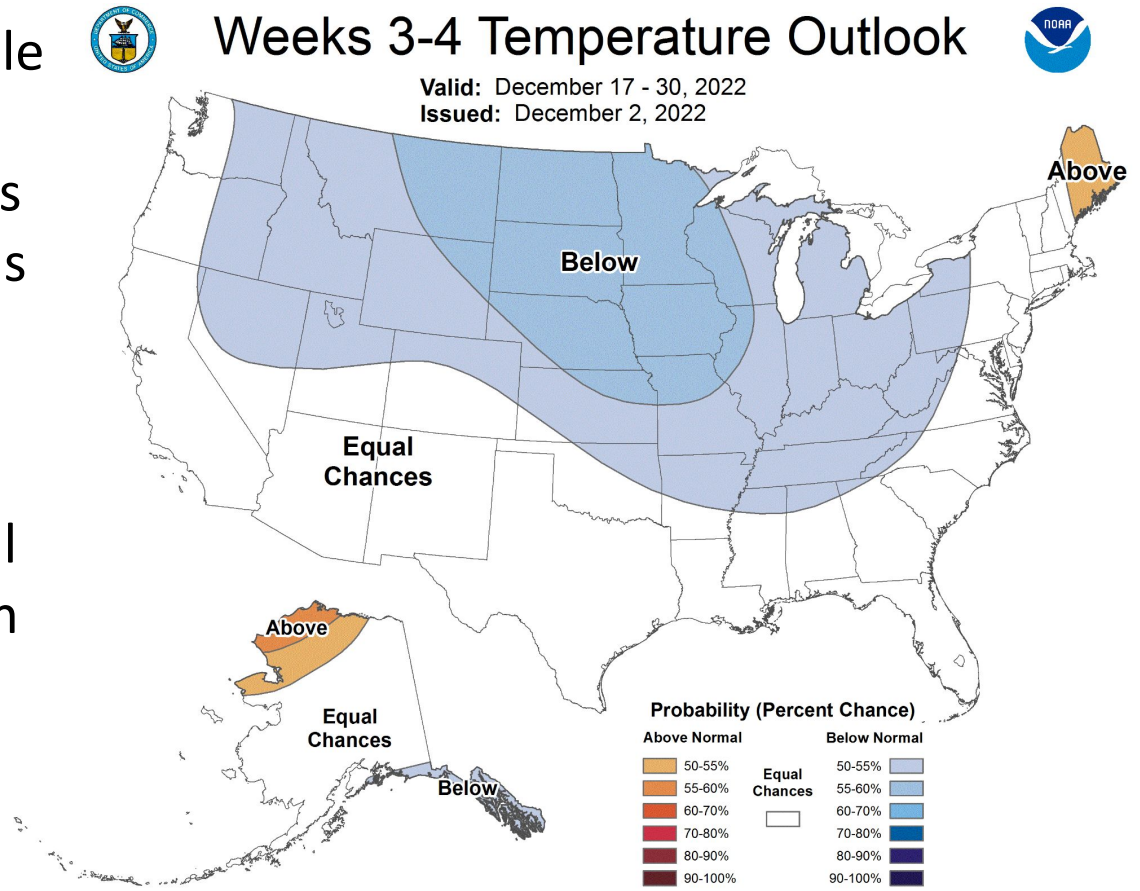
Domain: CONUS/AK; **# of Forecasts:** 282

Conclusions

- There is borderline statistically significant skill score improvement using subsampled multi-model ensemble suites.
- It is not entirely clear in real-time how many members should be subsampled, but the real-time verification is providing insight.
- The tool still provides value, as it 1) provides a link between CPC's Week 2 and Weeks 3-4 outlooks, 2) provides an additional interpretation of the dynamical models through its unique post-processing, and 3) can be used to test additional subsampling methods.

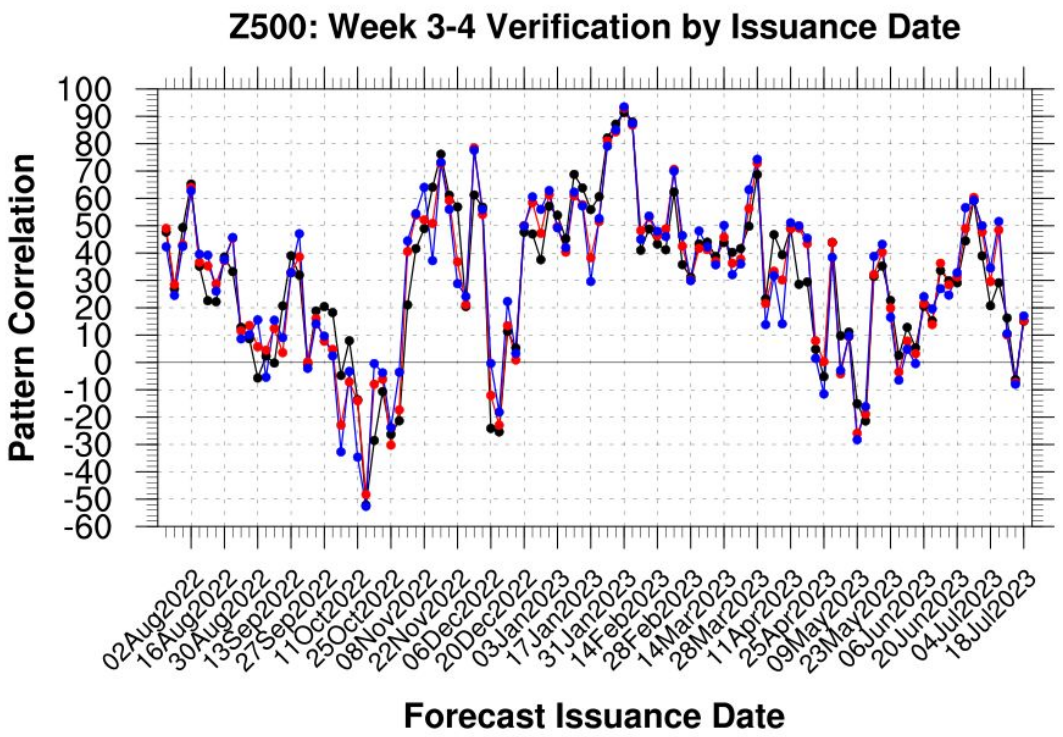
Thank you! Any Questions?

[Link to Climate Diagnostics and Prediction Workshop Extended Abstracts]



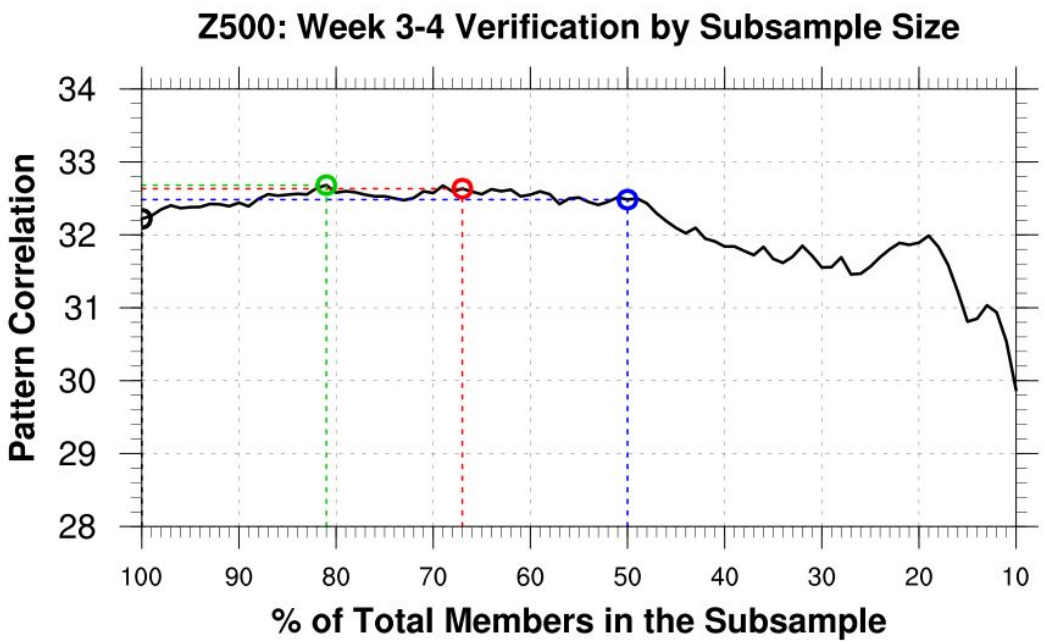
Additional Slides

Real-time Verifications



All Members (avg: 29.7) | 67% of Members (avg: 30.4) | 50% of Members (avg: 30.4)

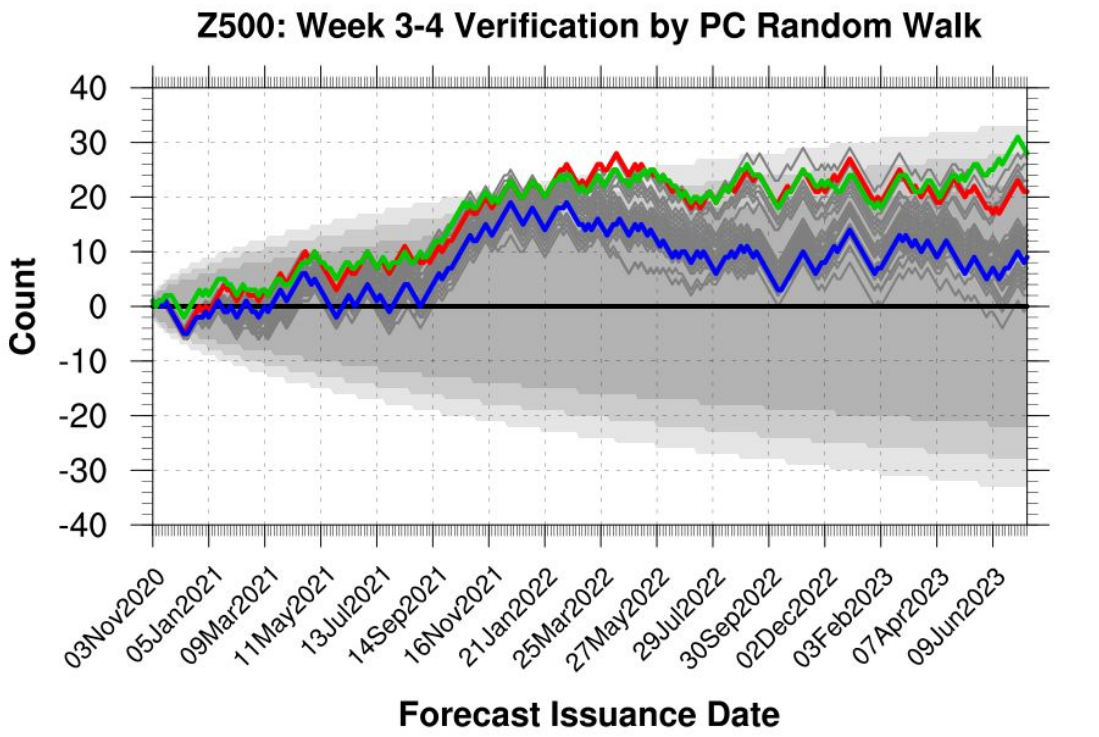
Domain: 20N to 87.5N and 180E to 330E; **# of Forecasts:** 104



All Members (avg: 32.2) | 67% of Members (avg: 32.6) | 50% of Members (avg: 32.5) | 81% of Members (avg: 32.7)

Domain: 20N to 87.5N and 180E to 330E; **# of Forecasts:** 282

Real-time Verifications



All Members (Count: 0) | 67% of Members (Count: 21) | 50% of Members (Count: 9) | 81% of Members (Count: 28)

Gray Lines: 70% to 30% of Members

Gray Shading: α = 0.20, 0.10, 0.05 (dark to light)

Positive Count: Subsample Wins; **Negative Count:** All Members Wins

Domain: 20N to 87.5N and 180E to 330E; **# of Forecasts:** 282

Subsampling Methods Tested

Name of Method	Brief Description of Method
Autoblend	Subsample members by removing individual members that least match the Week 2 autoblend forecast of Z500 derived from the ECMWF (50%), Canadian (25%), and GEFS (25%)
Ensemble Mean	Subsample members by removing individual members that least match the Week 3-4 ensemble mean forecast of Z500 derived from all members
Model	Subsample members by removing the entire suite from a given subseasonal model
ENSO-MJO Verifications	Subsample members by removing the entire suite from a given subseasonal model based on its historical verification given certain ENSO and MJO conditions at model initialization
Regime Transition Verifications	Subsample members by removing individual members that forecast regime transitions that have poor historical verifications
Regime Transition Frequencies	Subsample members by removing individual members that forecast regime transitions that have rarely occured in observations
MLR	Subsample members by removing individual members that least match the Week 3-4 forecast derived from the statistical multiple linear regression (MLR) tool
Trend Pattern	Subsample members by removing individual members that least match the expected pattern derived from long-term trends at Week 3-4
Week 1*	Subsample members based on their Week 1 verification
Week 2*	Subsample members based on their Week 2 verification
Tropical Precipitation*	Subsample members based on their Week 2 verification of tropical precipitation
Attribution*	Subsample members based on their Week 2 verification over regions attributed to skill at Week 3-4