A Comparison of Skill between Two Versions of NCEP Climate Forecast System (CFS) and CPC's Operational Short-Lead Seasonal Outlooks

Peitao Peng[1], Anthony Barston[2] and Arun Kumar[1]

[1]CPC/NWS/NOAA

[2]IRI, Columbia University

Based on

1. P. Peng, A. Arun, M. Halpert and A. Barnston: An Analysis of CPC's Operational 0.5-month Lead Seasonal Outlooks, *Wea. Forecasting*, 2012.
2. P. Peng, A. Barnston, A. Kumar: A Comparison of Skill between Two Versions of the NCEP Climate Forecast System (CFS) and CPC's Operational Short-Lead Seasonal Outlooks:. Submitted to Wea. Forecasting

Purposes of the Work

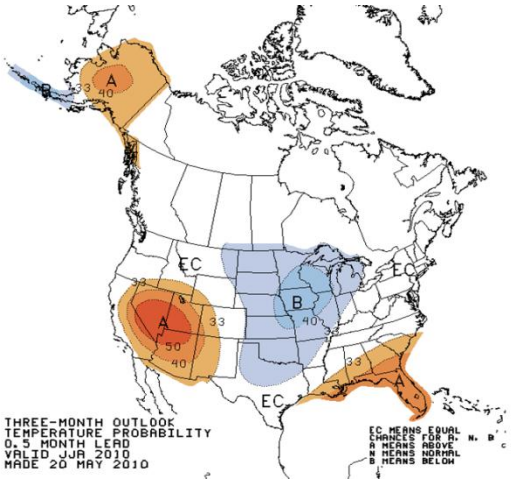
1. Document improvements in CFS skill for model developers and forecast producers;
2. Inform user community of CFSv2 skill to guide decision making processes.

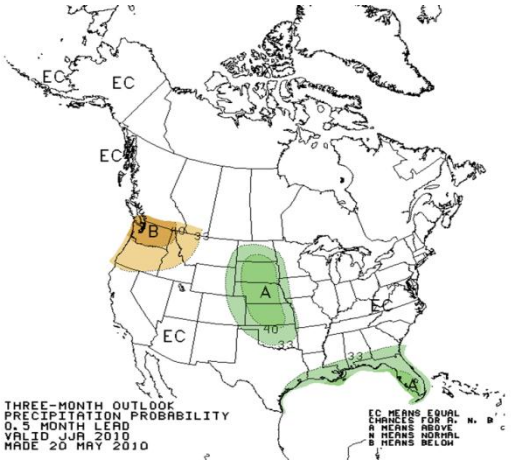
Outline

1. CPC's operational seasonal outlooks and CFSv1&v2 dynamical forecasts (hindcasts);
2. Data and methods of skill assessment;
3. Results of skill assessment and comparison;
4. Summary

CPC Operational Seasonal Outlooks for Temperature and Precipitation

- Rely on a combination of empirical and dynamical prediction tools, such as CCA, ECCA, OCN, EOCN, regression tool, partly coupled model (1995-2004), CFSv1(2004-2011), CFSv2(from 2011);
- Have been routinely made since December of 1994;
- Issued for 13 running seasons and released in the middle of each calendar month
- Probabilistic format: shifts in the probabilities away from their climatological values (of 1/3) for three equiprobable tercile-based categories (below, near and above normal)





CFS version-1 and version-2

- Fully coupled dynamical model (Saha and coauthors, 2006);
- Some basic specifications for CFSv1 and v2:

	CFSv1	CFSv2
Horizontal Resolution	T62 (~2°)	T126 (~1°)
Vertical Resolution	64 levels	64 levels
Ensemble Size	15	20
Initial Conditions for 0.5 month Outlook (example given is for DJF seasonal mean forecast, made in November)	Five initial conditions from each near the 1st and 11th of November, and the 21st of October	Four initial conditions from each of the 17th, 12th, 7th, 2nd of November, and the 27th of October
Climatological Base Period	1982-2004	1982-2004
Maximum Forecast Lead Time	9 months	9 months
Source of Initial Condition Data	NCEP/DOE Reanalysis	Climate Forecast System Reanalysis (CFSR)
Carbon Dioxide Concentration Setting	Fixed at 1988 level	Evolving

Forecast and verification Data: Temp and Prec

- 0.5-month lead seasonal forecasts for CONUS in the common period of 1995-2009.
- For CFS forecasts, the tercile boundaries are determined by ranking all ensemble members over the climatological base period (1982-2004); for CPC's they are determined based on the analyses over the WMO climatological periods (30-year long, updated every 10-year period);
- Verification data are from CPC analysis;
- Data resolution: 2x2 grid;

Verification methods

Categorical measures:

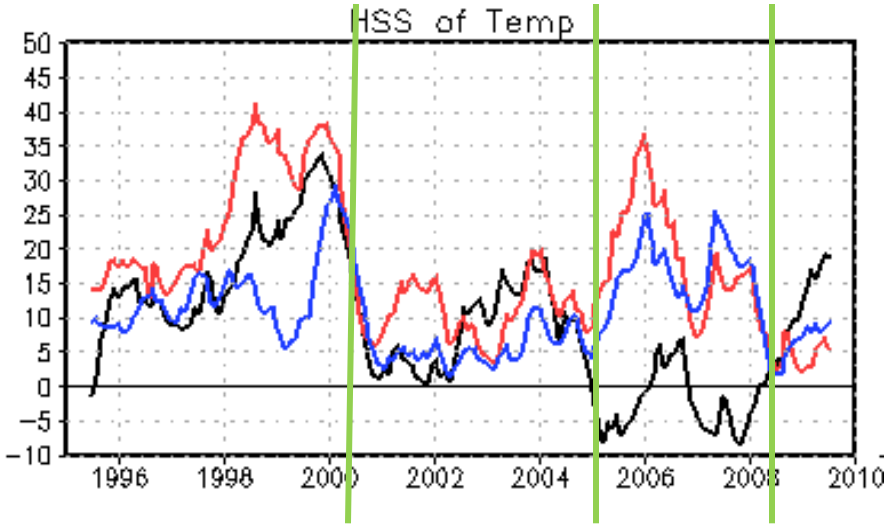
Heidke Skill Score (HSS):

HSS=(h-e)*100/(t-e), (range: -50 -> 100) where e=t/3

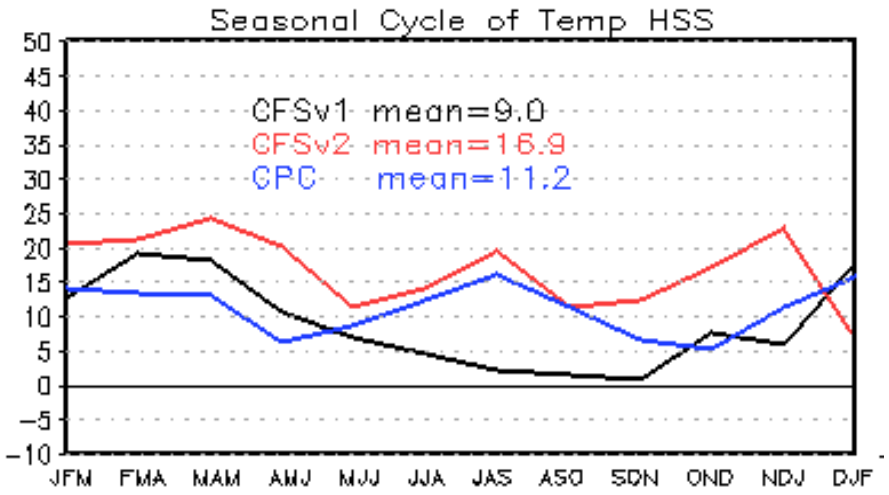
Probabilistic measures:

1) Ranked probability Skill Score (RPSS), a measure of the squared distance between the forecast and the observed cumulative probabilities (Kumar et al. 2001), ranges from minus infinity to 1;
2) Relative Operating Characteristics (ROC) (Mason 1982);
3) Reliability Diagrams, for comparing the forecast probabilities against their corresponding frequencies of observed occurrence, and examining the frequency distribution of the issued probabilities (sharpness).

Temporal Variation and Seasonal Cycle of Spatially Averaged HSS for Temp Forecast

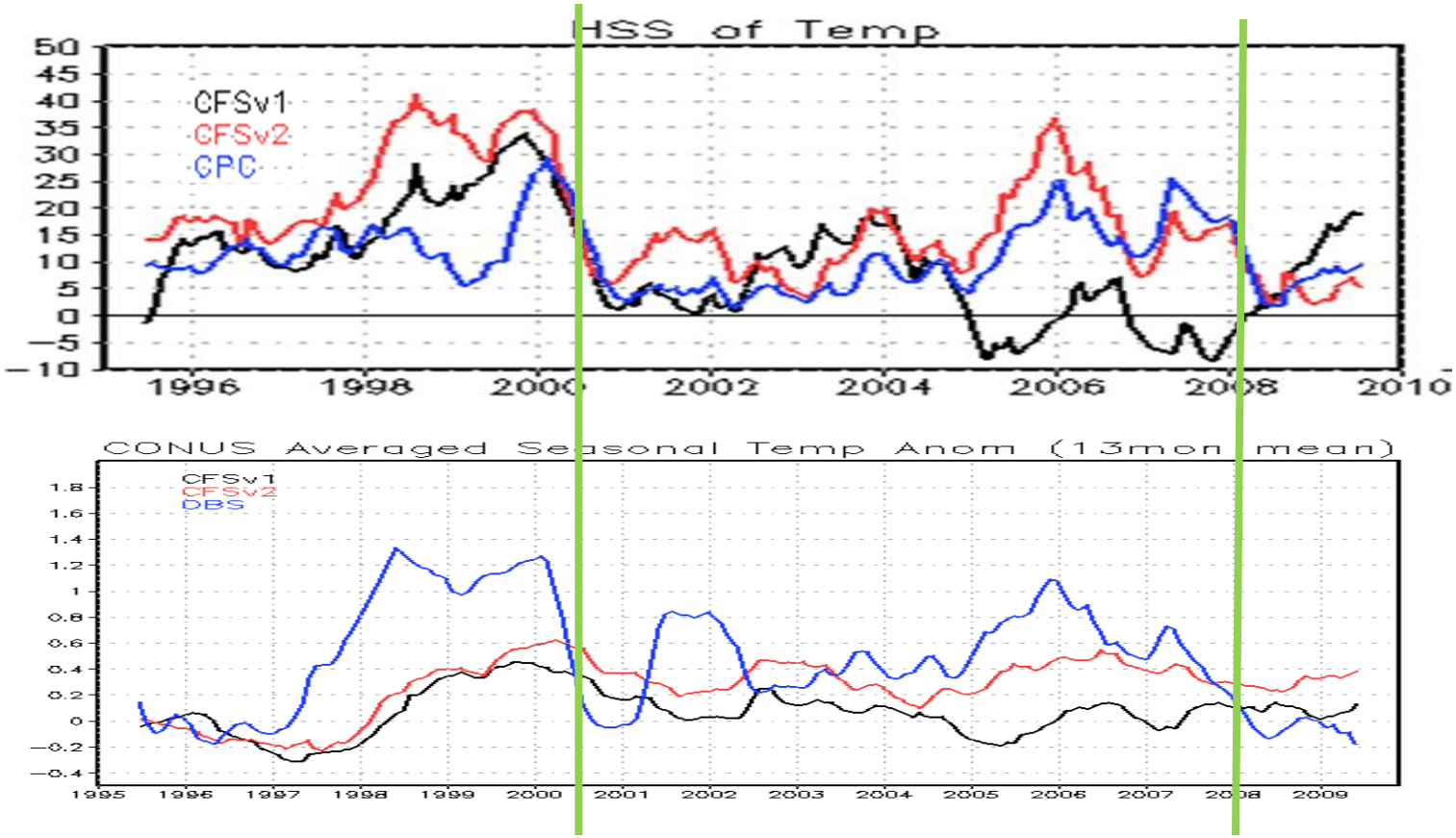


1. CFSv2 outperformed CFSv1 and CPC throughout most of the period;
2. CFSv1 significantly underperformed CFSv2 and CPC for 2005-2008, but turned the best afterwards;
3. CFSv1 and v2 appears well correlated during 1995-2000, but this correlation decreases thereafter;
4. Prior to 2007, CFSv2 is rarely outperformed by CPC.



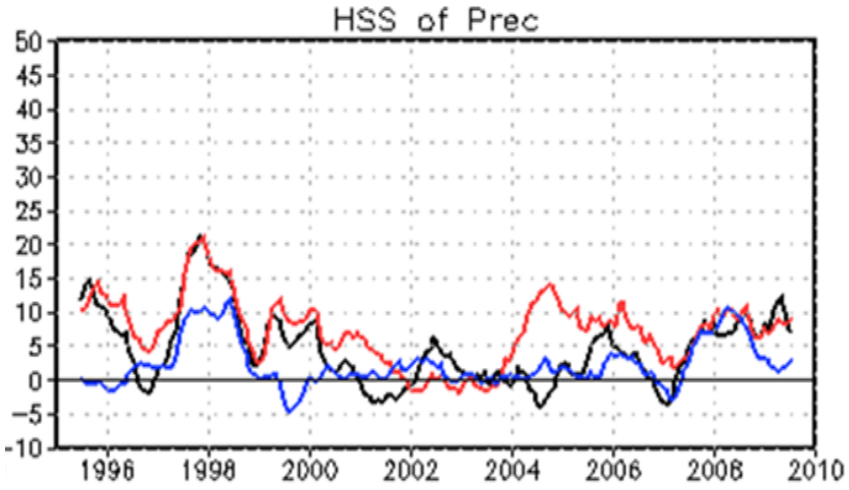
1. CFSv2 and CPC have peaks in northern winter and summer and minima in late spring and late fall;
2. CFSv2 has lowest skill in DJF;
3. CFSv1 is lack of a summer skill peak, probably due to a problem in initializing soil moisture (Wang et al 2009)

Interpretation of the change in the skill coherence between CFSv1 and CFSv2

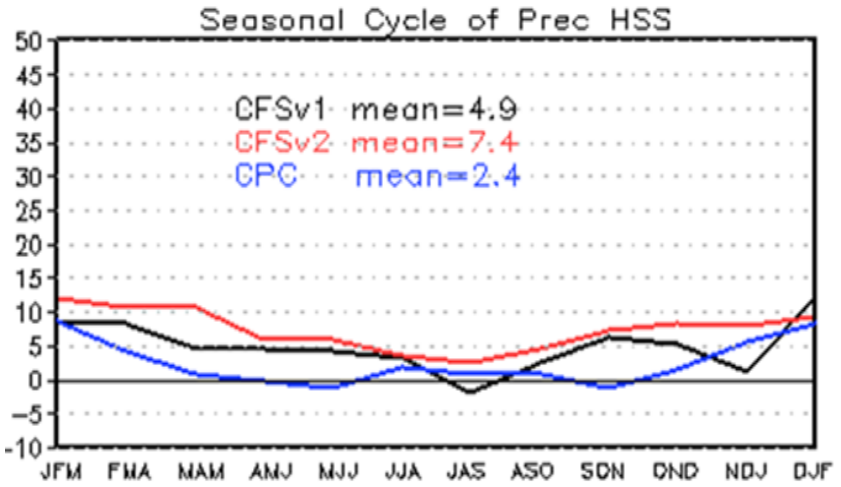


1. During 1995-2000, both CFS versions forecasted warm anomalies;
2. After 2000, the CFSv2 forecasted temp begins exceeding that of CFSv1, following observed warmer before 2008;
3. After 2008, observed temp turned to lower, makes the skill of CFSv1 higher;
4. The warmer temp in CFSv2 may be due in part to the specification of evolving CO_2 in CFSv2.

Temporal Variation and Seasonal Cycle of Spatially Averaged HSS for Prec Forecast

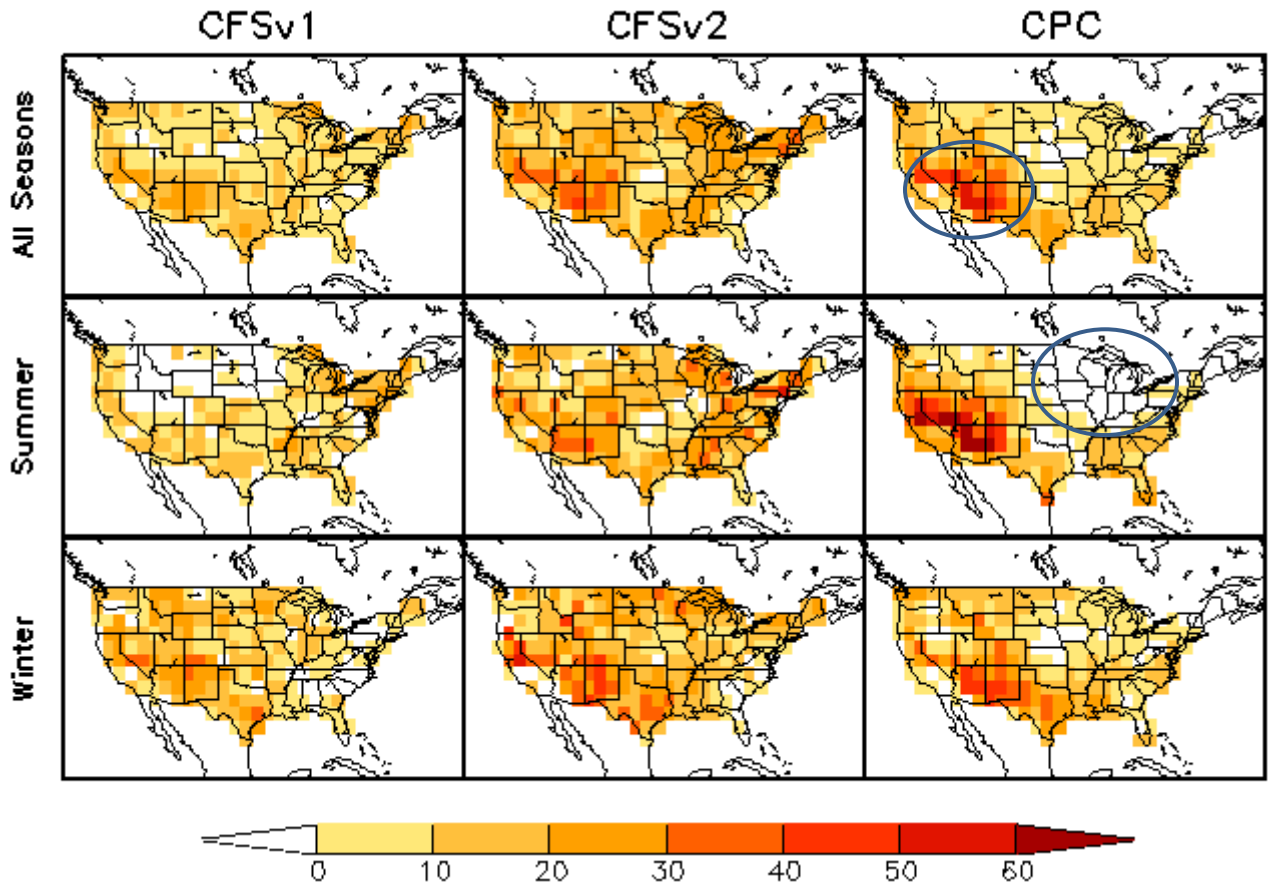


1. Prec skill is considerably lower than temp skill;
2. The skill is higher for CFSv2 than CFSv1;
3. The skill of two models are strongly correlated before 2000;
4. CPC underperforms both CFS versions



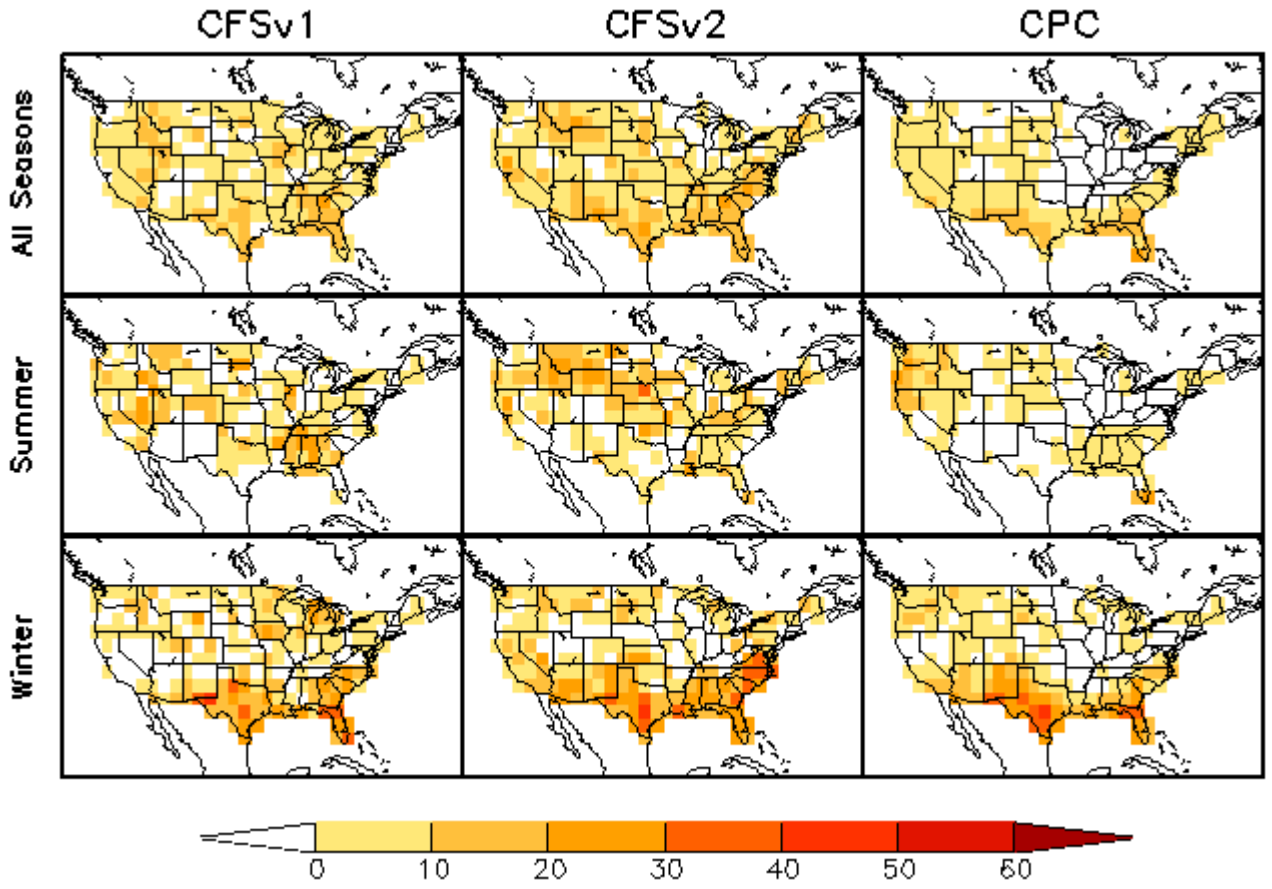
1. Essentially unimodal, lack of summer peak, due to robust ENSO tele-connections in the US being mainly limited to winter.

Geographical Distribution of Temporally Averaged HSS for Temp



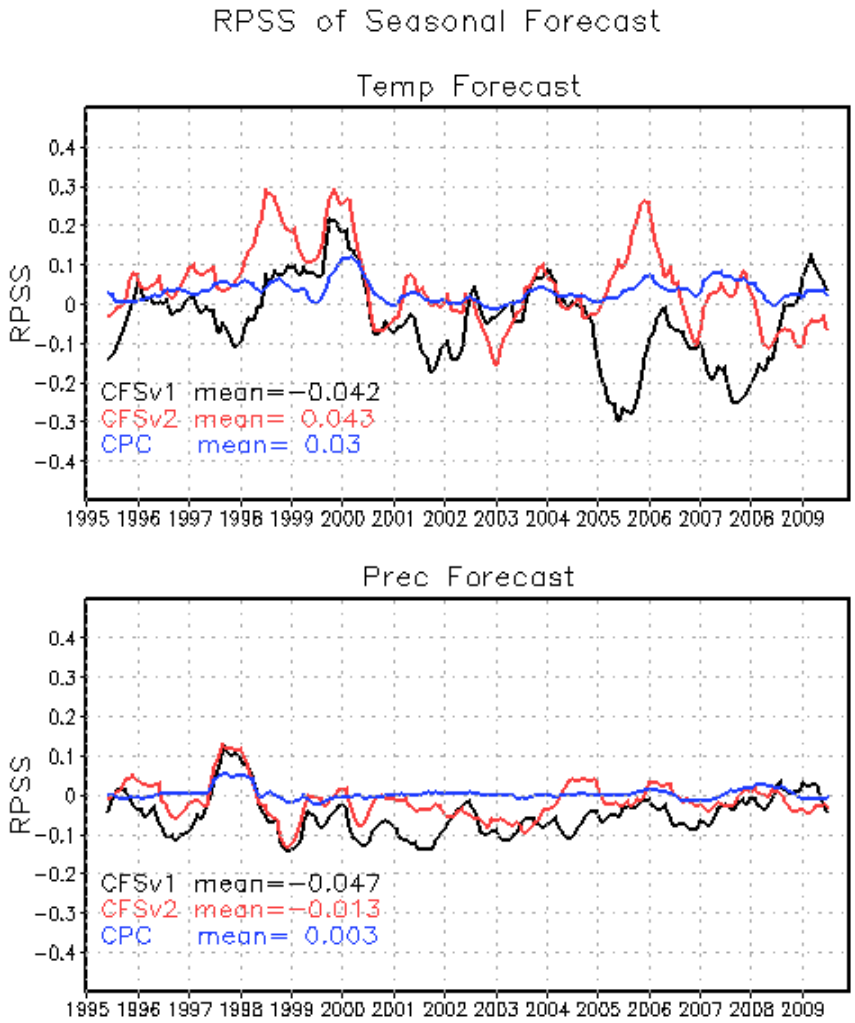
1. CFSv2 outperforms CFSv1 over much of the US;
2. A skill maximum in southwest region, where strong warming trend has been observed;
3. CPC's high skill over the southwest region relies on OCN tool;
4. CFSv2 forecasts the warming trend better than CFSv1, due in part to evolving CO_2 in it;
5. Both CFS versions performed better than CPC in far northeastern US.

Geographical Distribution of Temporally Averaged HSS for Prec



1. Both CFS versions have quite uniform skill distribution, CPC has skill holes in Rockies , eastern Great Pains and the Midwest, causing overall skill lower than CFS;
2. Skill in summer is generally low, there is no much geographical preference for CFS, but CPC's is noticeably lower in southwest and northeast;
3. In winter southern tier has higher skill due to ENSO;

Temporal Variation of Spatially Averaged RPSS



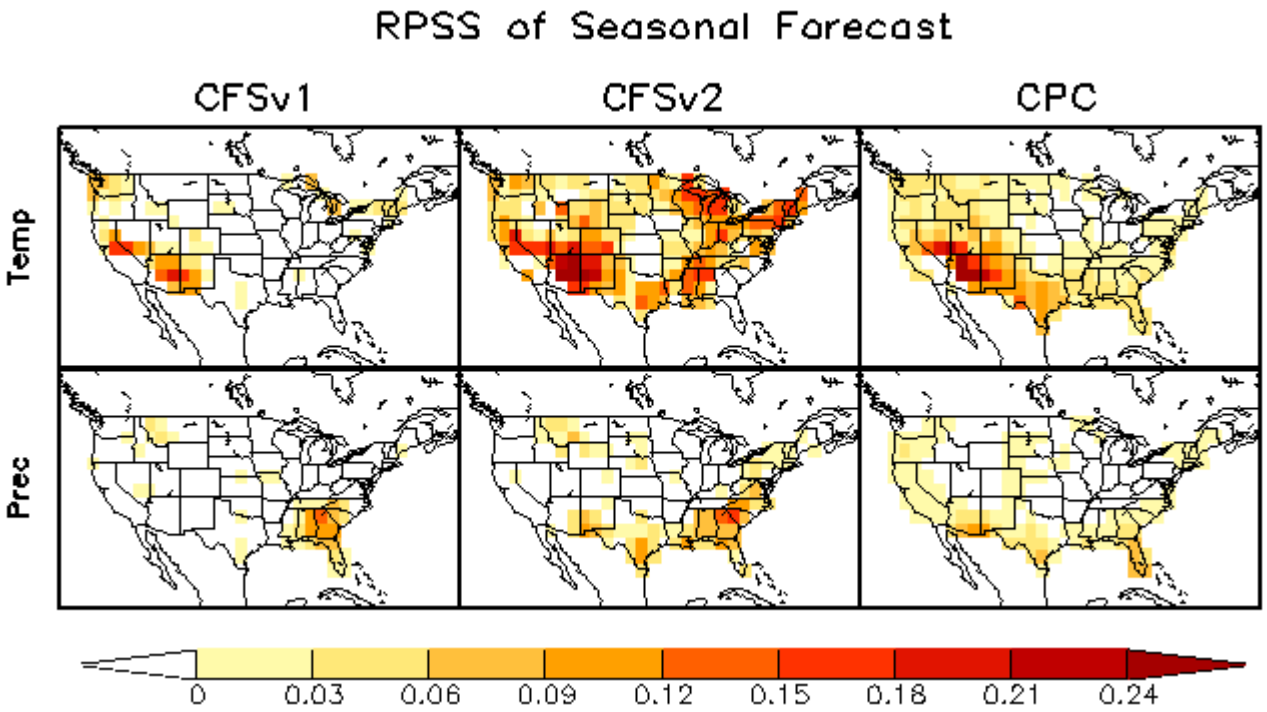
For temp skill:

1. CFSv2 is higher than CFSv1;
2. CFSv2 has similar average RPSS to CPC, but its temporal variation is much bigger.

For prec skill:

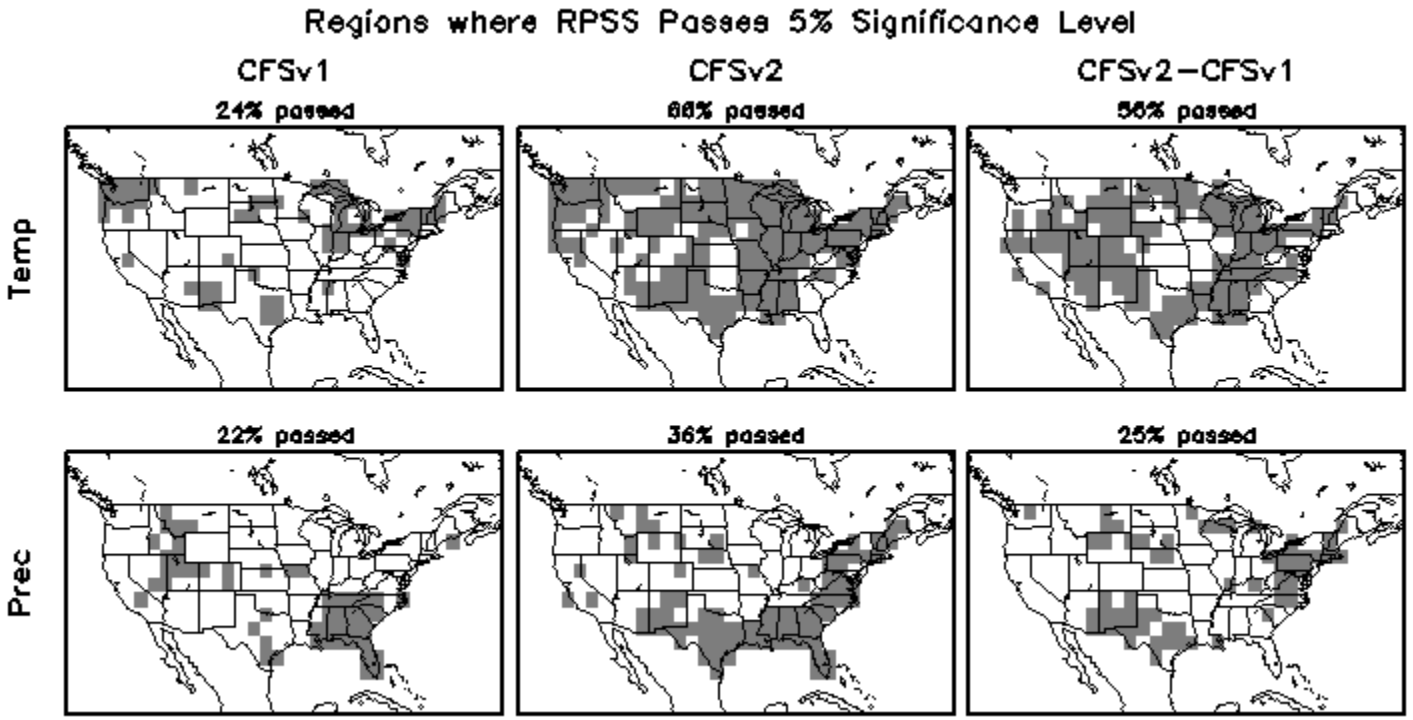
1. CFSv2 is still higher than CFSv1, but lower than CPC;
2. Overconfidence in CFS may reduce RPSS.

Geographical Distribution of Temporally Averaged RPSS



1. Patterns are roughly similar to those for HSS, but regions of relatively low positive HSS tend to have negative RPSS;
2. CFSv2 is obviously better than CFSv1.

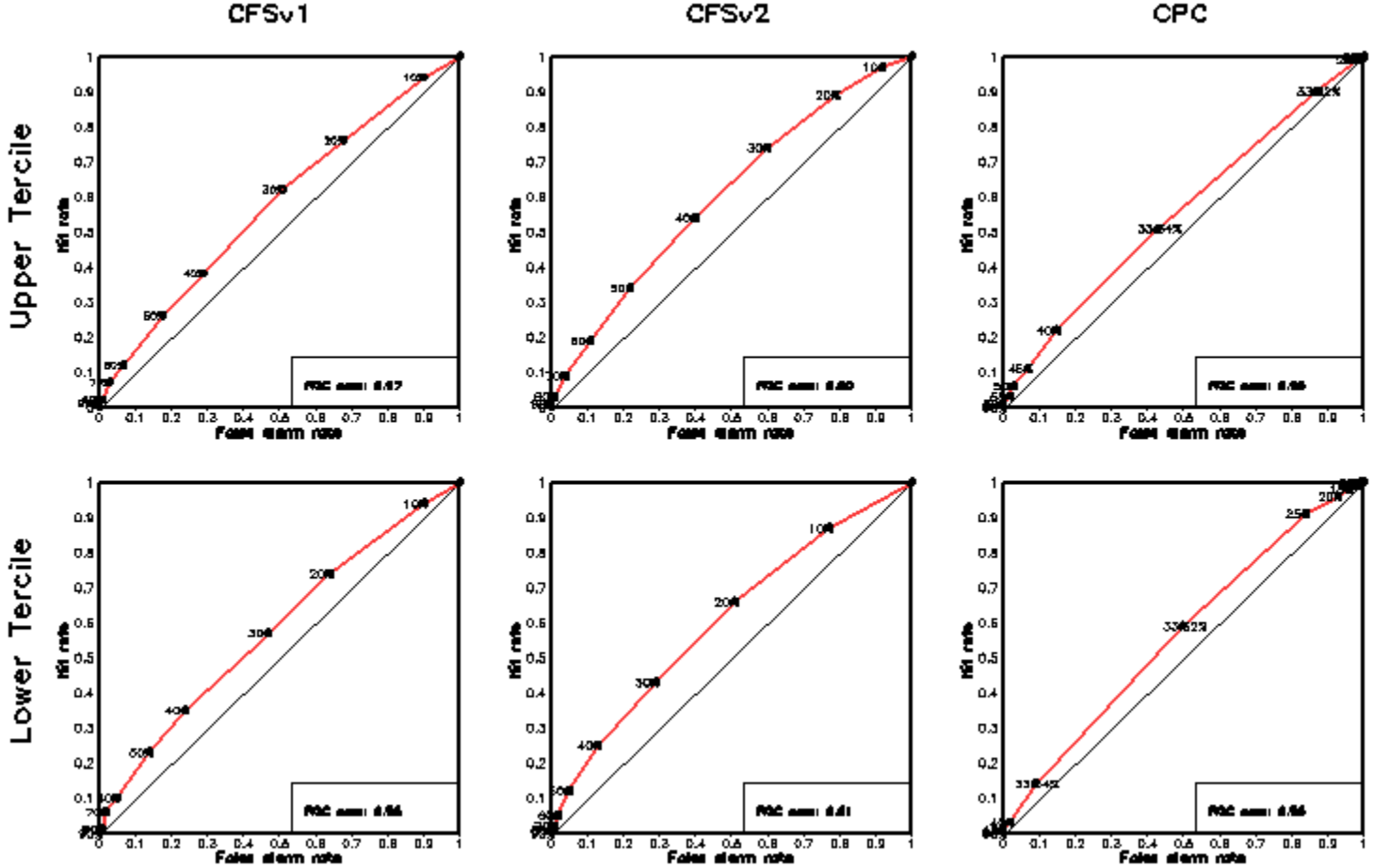
Significant Test of RPSS with Monte Carlo Approach



Field significance Test (with the average RPSS over the US as the test statistic):

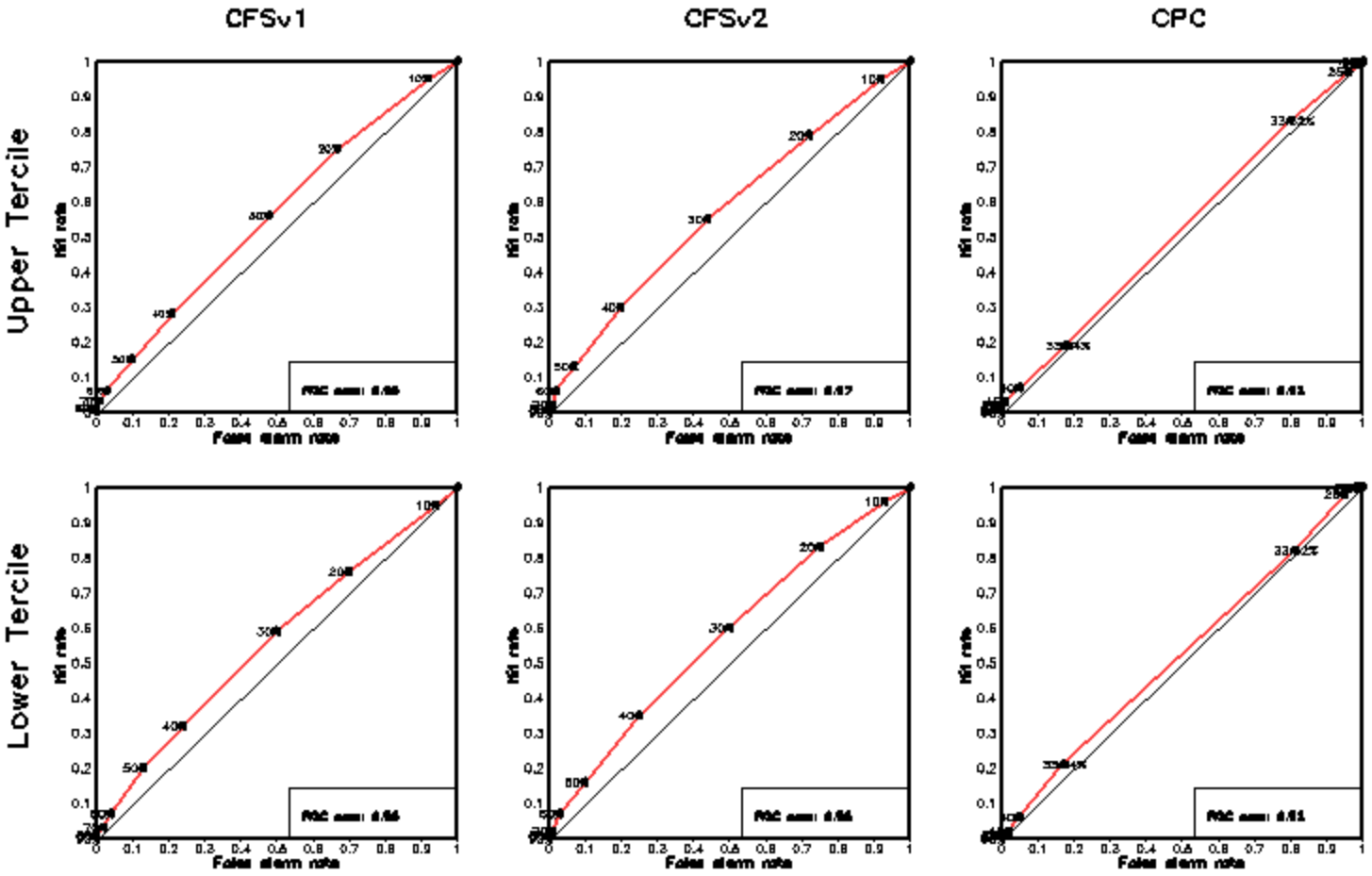
1. For prec, both CFS versions are at < 0.0001
2. For temp, CFSv1 < 0.0002; CFSv1 < 0.0001;
3. For RPSS difference between CFSv1 and v2, < 0.0001 for both temp and prec.

ROC Score of Temp Forecast

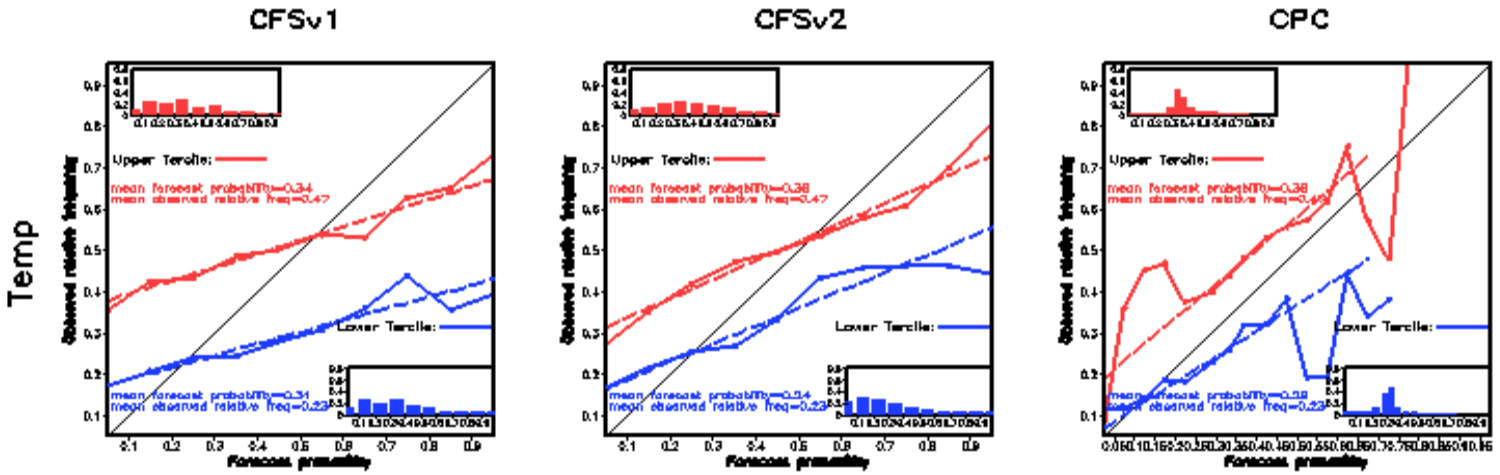


1. ROC area under the ROC curve shows the cumulative hit rate against false-alarm rate. ROC area=0.5 means no skill;
2. CFSv2 is slightly better than CFSv1;
3. CPC has the lowest skill because a sizable proportion of CPC forecasts are for equal chance (EC).

ROC Score of Prec Forecast

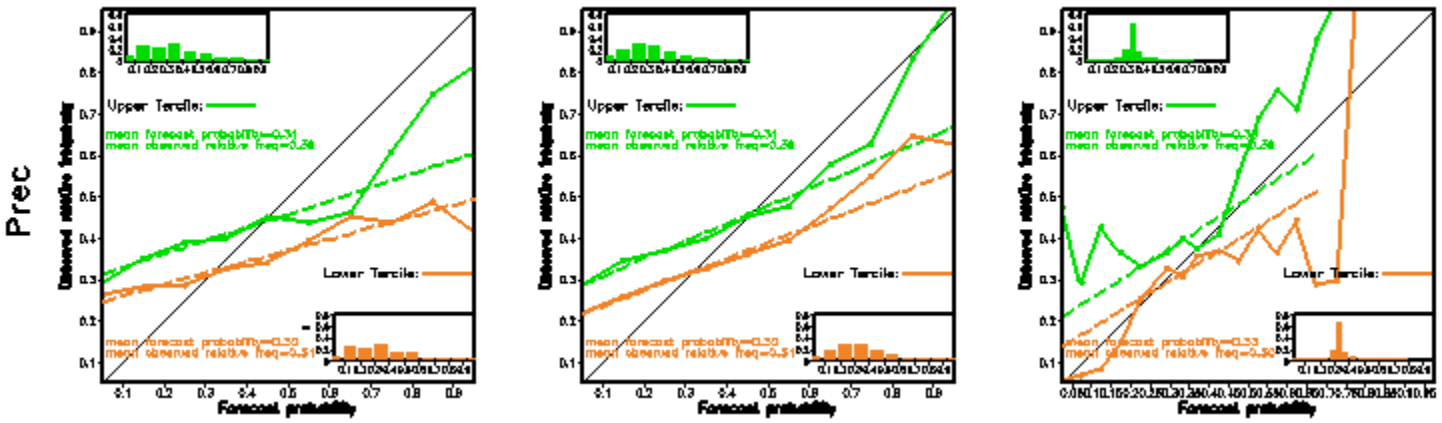


Reliability of Seasonal Temp Forecast



1. All three products show a bias of underpredicting above normal and overpredicting below normal;
2. CPC has the most appropriate level of probabilistic confidence with slope closer 1. Two CFS versions are overconfident (slope < 1);
3. Two CFS versions are much sharper than CPC, but not justified because of their overconfidence.

Reliability of Seasonal Prec Forecast



1. **Still have a bias of underpredicting above normal and overpredicting below normal, but has less extent than temp forecasts;**
2. **Still overconfidence in CFS versions, less in CPC;**
3. **Still greater sharpness in CFS versions than CPC.**

Summary/conclusions

- The predictive skill of CFSv2 clearly improved over that of CSFv1 by most verification metrics, over most US locations and over most seasons.
- The improvement is attributable to its improved physics and better data assimilation/analysis methods.
- Another reason for the improvement was the inclusion of time evolving greenhouse gas concentration in CFSv2.
- The average performance of the CPC's seasonal outlooks has been somewhat better than that of CFSv1 alone, but in many cases slightly lower than that of CFSv2