Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
43rd NOAA Annual Climate Diagnostics and Prediction Workshop
Santa Barbara, CA, 23-25 October 2018

CPC's New Consolidated Hybrid Statistical/Dynamical Model for Seasonal Prediction of Temperature and Precipitation

Daniel Barandiaran
Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD
and
Innovim, LLC, Greenbelt, MD

1. Introduction

It has been known for some time that CPC could benefit from a new seasonal forecast consolidation that will serve as a 'first guess' for the forecaster, with the aim of improving forecast reliability and month-to-month consistency across forecast and forecasters. The prior consolidation, implemented in 2006, had the positive benefit of leading to increased forecast coverage and improved 'all forecasts' skill scores (Baxter 2016). There are, however, some limitations of this consolidation that reduce its usefulness to CPC forecasters. First, the consolidation uses climate division data (CD-102) with coverage of the continental United States only. Second, the consolidation makes use of decades-old statistical tools and only one dynamical model input, namely the Climate Forecast System (CFS). Finally, the consolidation process itself is something of a black box, for instance giving little information to the forecaster regarding the contribution of various components of the consolidation to overall forecast skill.

Since the implantation of the operational consolidation in 2006, there have been advances in model post-processing and calibration (*e.g.*, Unger *et al.* 2009; Ou *et al.* 2016; van den Dool *et al.* 2017) that have been implemented across many of CPC's operational forecast products and tools. An effort was therefore initiated to take advantage of such methodologies as well as to make use of newer statistical tools and a larger pool of dynamical models, therefore creating a robust forecast tool that can more easily be utilized by the forecaster.

2. Methods and data

The primary goal of this project was to apply a probability anomaly correlation (PAC) calibration to a new suite of empirical forecast tools, and consolidate those tools with the constituent models of the National Multi-model Ensemble (NMME) system, which have been PAC-calibrated in real-time since 2016. The PAC

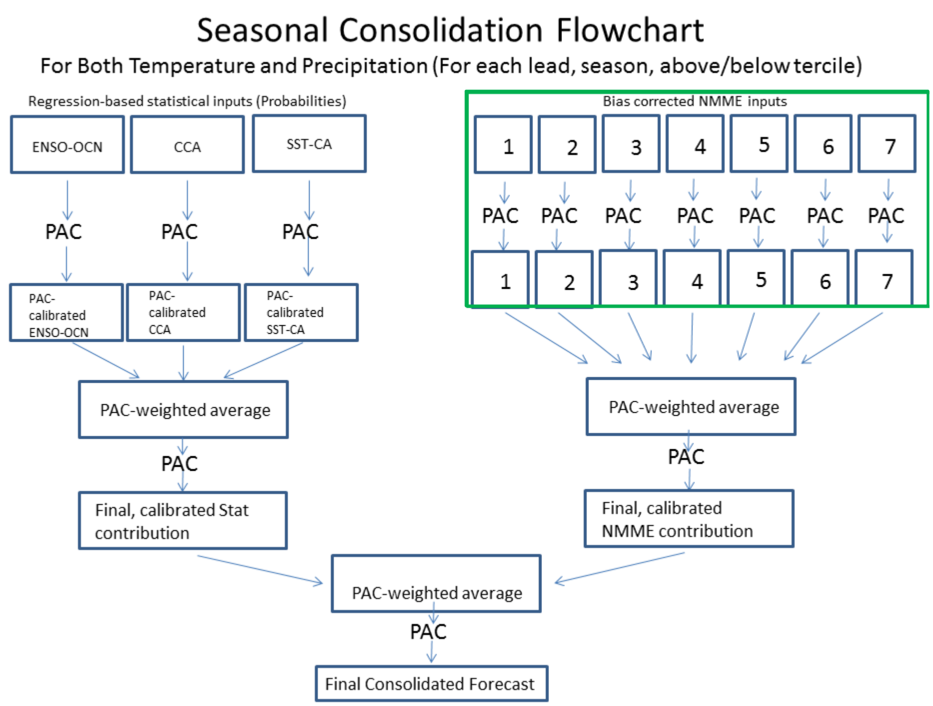

Fig. 1 Seasonal consolidation flowchart. The green box indicates process that currently executes operationally upstream of the consolidation. All other processes are included as part of this experimental consolidation process.

Correspondence to: Daniel Barandiaran, Climate Prediction Center, NOAA/NWS/NCEP, 5830 University Research Court, College Park, MD 20740, and Innovim, LLC, Greenbelt, MD; E-mail: daniel.barandiaran@noaa.gov

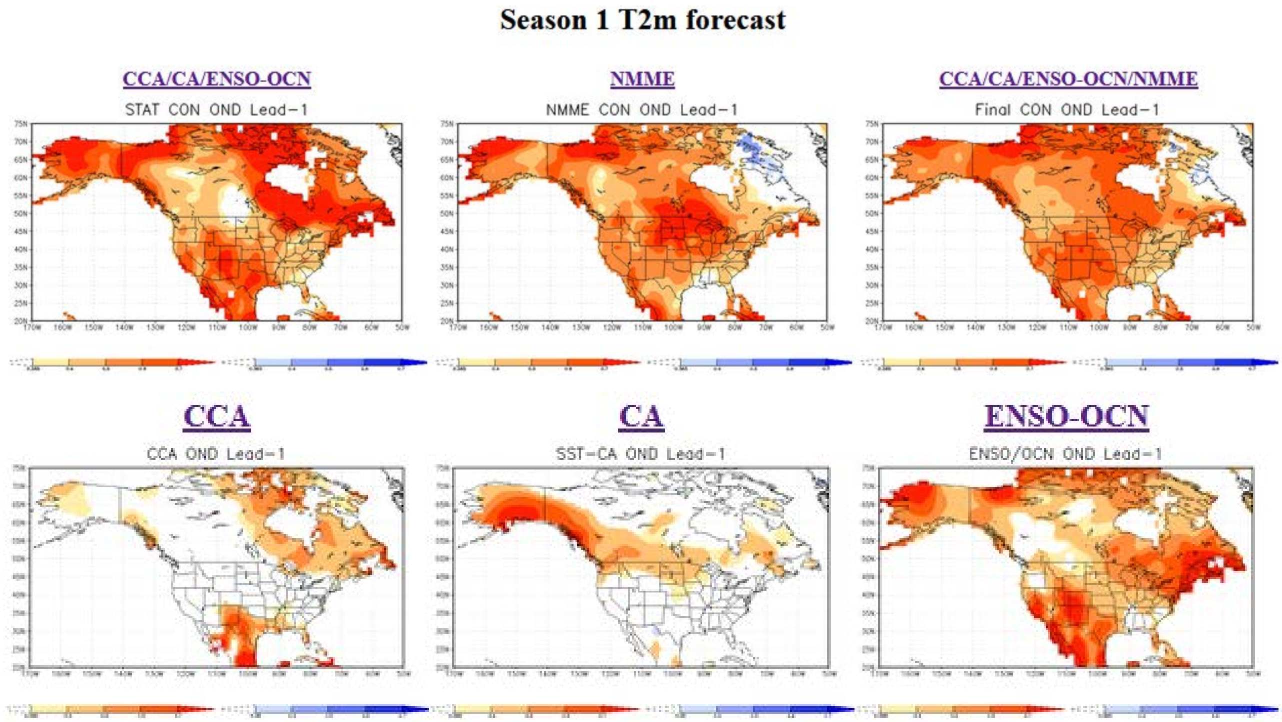

Fig. 2 Sample output graphics available to forecasters for the Lead-1 temperature forecast.

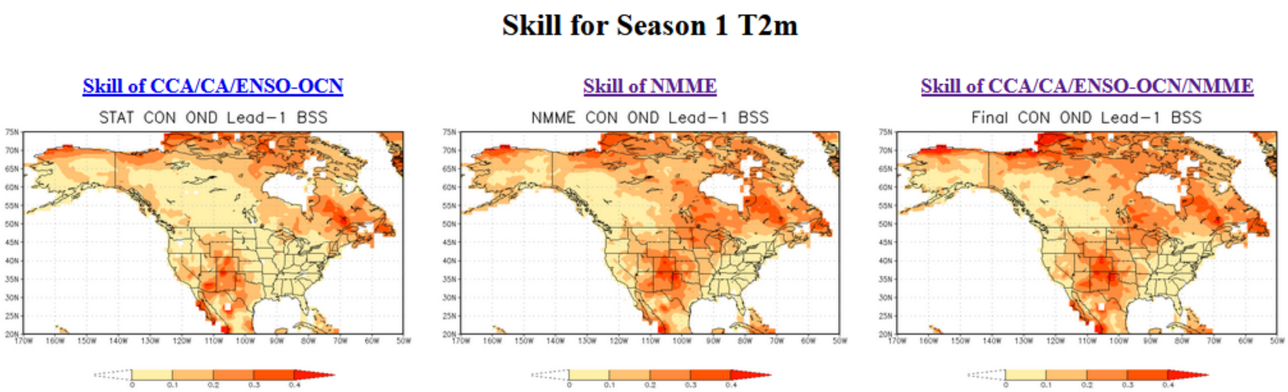

Fig. 3 Sample historical Brier skill score (BSS) graphics output alongside the forecast graphics.

methodology, acting on probability anomalies, is analogous to traditional linear regression acting on temperature and precipitation anomalies themselves; the former minimizes the Brier score, while the latter minimizes the mean squared error (van den Dool *et al*. 2017).

The suite of empirical forecast tools being used in the consolidation include a canonical correlation analysis-based model (CCA, Barnston and He 1996), constructed analog based on sea surface temperatures (SST-CA, van den Dool *et al.* 2003), and a hybrid El Niño-Southern Oscillation/long term trends forecast tool (ENSO-OCN, Barandiaran and Baxter 2017). These statistical models are all calibrated using GHCN (Global Historical Climatology Network)+CAMS (Climate Anomaly Monitoring System) for temperature (Fan and van den Dool 2008) and CPC's gridded precipitation reconstruction (Chen *et al.* 2002).

The new consolidation flow chart is shown in Fig. 1. The premise is to apply PAC calibration to each of the constituent models for both the statistical and dynamical model inputs, and then each stream, statistical (left) and dynamical (right), is combined by weighting based on the PAC coefficient (ranging from 0 to 1, with negative values set to 0). Because the combination of models is often more skillful than each model separately, the results at this point are expected to be underconfident. Therefore, a second pass PAC calibration will minimize the Brier score of the *combination* of forecast tools. This process is repeated to consolidate the statistical and dynamical forecast streams.

3. Results

Forecast probabilities and skill metrics are output and archived in real-time in both NetCDF and binary data formats, and forecast graphics are output and archived. A web interface was created where the forecaster can access the consolidation forecasts from both the NMME and statistical tools, and their final consolidation. An example of the graphics forecasters had access to for the September seasonal forecast cycle is shown in Fig. 2. Importantly, forecasters can see whether contributions to the forecast are coming from statistical models or the NMME. The statistical model stream is further broken down into its three constituent models. Associated skill maps are displayed as well, where the average of the hindcast Brier skill score for above- and below-normal temperature probabilities is plotted for that lead and target season (Fig. 3).

Evaluation of the consolidation was conducted by calculating the BSS for each lead and season as well as associated reliability statistics. The statistical and dynamical model components are compared to understand where the statistical guidance adds value to the

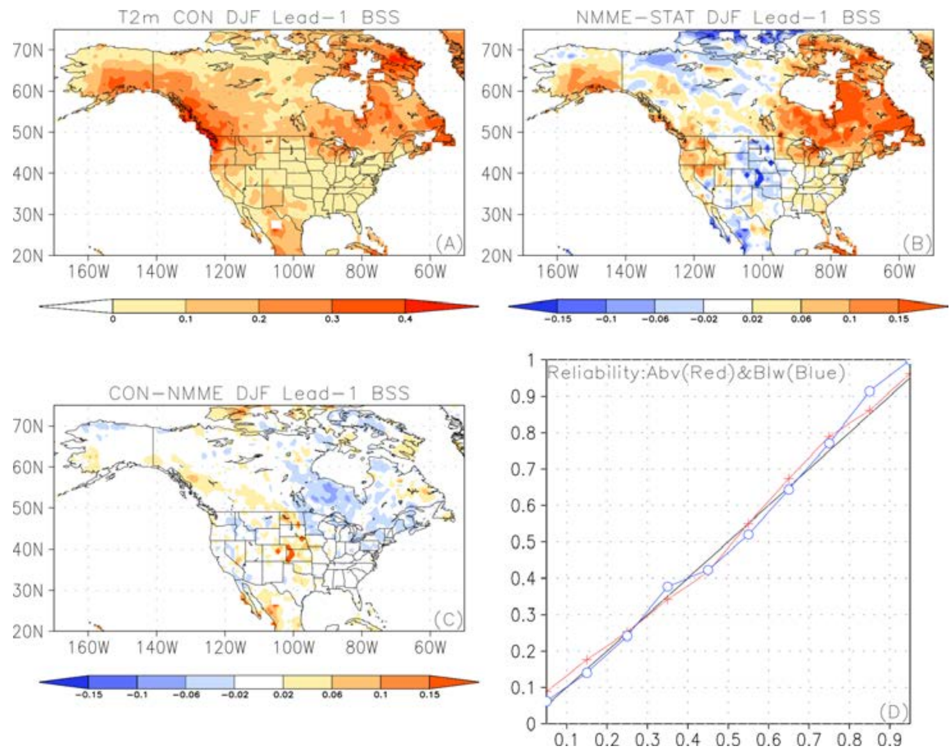

Fig. 4 Panel (a) shows the average Brier skill score (BSS) for Lead-1 above- and below-normal temperature forecasts for December-February (DJF). Panel (b) shows the BSS difference between the NMME stream and the statistical stream. Panel (c) shows the BSS difference.

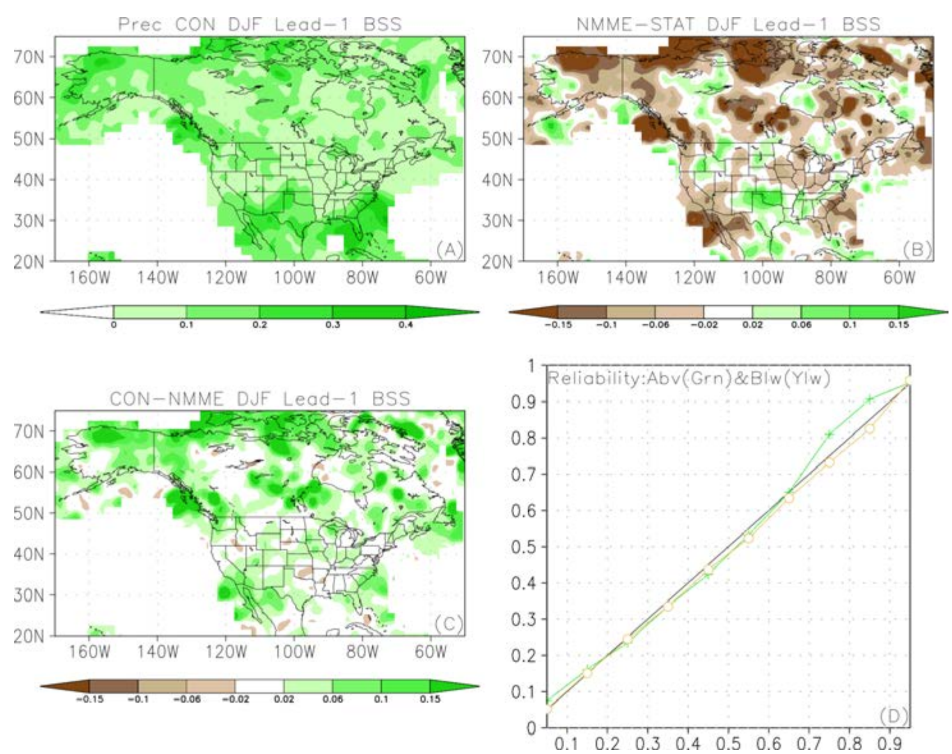

Fig. 5 Same as Fig. 4 except for Lead-1 DJF precipitation forecasts. In this case, the statistical forecast tools generally enhance the skill of the forecast (Panel C).

state-of-the-art dynamical model guidance. Finally, some comparison is made to the current NMME guidance utilized by forecasters.

The average BSS for Lead-1 temperature forecasts of December-February (DJF) is shown in Fig. 4a. As expected, skill is modest across much of the CONUS, except where ENSO and long-term trends are most important. The difference between the average BSS of NMME model consolidation and the statistical model consolidation is shown in Fig. 4b; the statistical models outperform the NMME only in low-skill areas over the central CONUS. Figure 4c shows the difference in BSS between the final consolidation and the NMME constituent; this can be thought of as the value added by the inclusion of the statistical guidance. There are areas where the statistical guidance clearly adds value, but it is mostly mixed. Figure 4d shows the reliability of above- and below-normal temperature forecasts from the final consolidation, respectively. As expected given this established methodology, the final consolidation is reliable across forecast probabilities. Figure 5 shows the same except for DJF Lead-1 precipitation forecasts. In this case, an obvious ENSO skill signature is seen, with the highest forecast skill over regions where seasonal precipitation is known to be more correlated to ENSO.

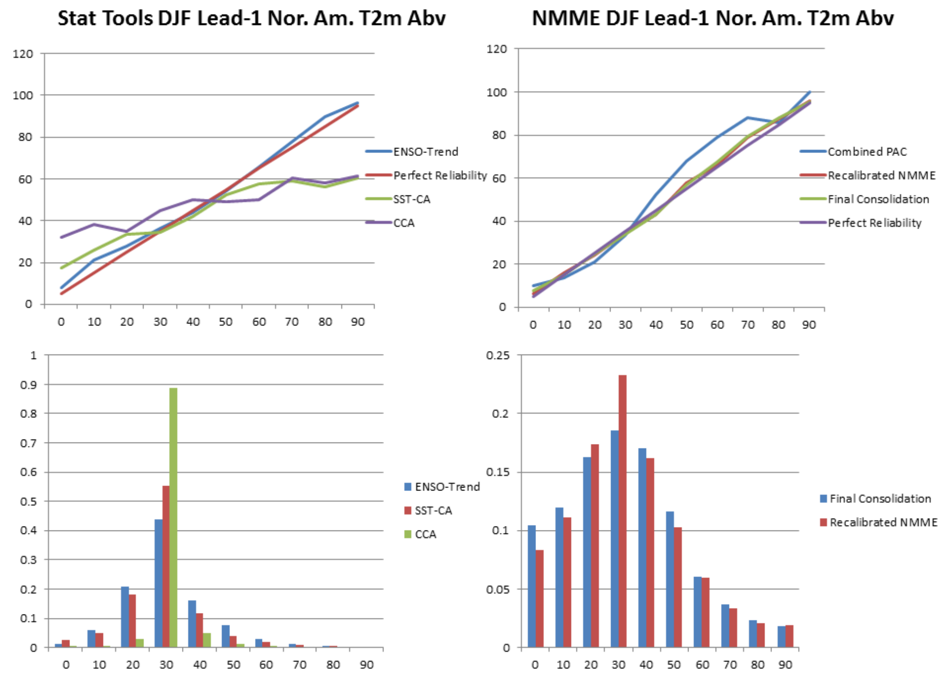

Fig. 6 This figure highlights the reliability and frequency of Lead-1 forecasts of above-normal temperatures for December-February from the statistical models (left) and the NMME and final consolidation (right). Importantly, the NMME as used by seasonal forecasters (blue line, upper right) is notably underconfident.

Finally, Fig. 6 shows a more in-depth breakdown of tools for the Lead-1 DJF temperature forecast. This reveals that the addition of the statistical models maintains reliability while adding resolution (increasing the frequency with which larger probabilities are forecast). Additionally, it shows that the NMME as currently used by CPC forecasters is quite under confident. The second pass PAC calibration in this case increases the probabilities to match forecast skill.

4. Summary

- The latest seasonal forecast tools, including constituent models from the NMME and newly derived empirical models, are consolidated and recalibrated using the probability anomaly correlation (PAC) methodology.
- The forecast consolidation occurs in two phases: the first in which statistical and dynamical tools are consolidated separately, and the second in which these two streams are consolidated (Fig. 1).
- Real-time forecast graphics are available to forecasters, along with associated skill metrics (Figs. 2 and 3).
- The inclusion of statistical tools improves the forecast skill for precipitation in all seasons. For temperature the impact is less notable, though there is some evidence that forecast resolution improves (Figs. 4, 5, and 6).

References

Barandiaran, D., and S. Baxter, 2017: Benchmark statistical model for seasonal prediction of temperature and precipitation. *42nd NOAA Annu. Clim. Diagn. Predict. Workshop*, Norman, OK, 23-26 October 2017, https://www.nws.noaa.gov/ost/climate/STIP/42CDPW/posters/IMG_4825.JPG.

Barnston, A. G., and Y. He, 1996: Skill of canonical correlation analysis forecasts of 3-month mean surface climate in Hawaii and Alaska. *J. Climate*, **9**, 2579-2605, doi:10.1175/1520-0442(1996)009<2579:SOCCAF>2.0.CO;2

Baxter, S., 2016: Evaluating CPC's operational seasonal temperature forecasts: Why aren't we beating a categorically warm forecast? *NWS Sci. Technol. Infusion Clim. Bull., 41st Annu. Clim. Diagn. Predict. Workshop*, Orono, ME, October 3-6 2016, 61-63, doi:10.7289/V5JS9NH0.

Chen, M., P. Xie, J. E. Janowia, and P. A. Arkin, 2002: Global land precipitation: A 50-yr monthly analysis based on gauge observations. *J. Hydrometeor.*, **3**, 249-266, doi:10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2

Fan, Y., and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948-present. *J. Geophys. Res*., **133**, D01103, doi:10.1029/2007JD008470.

Ou, M. H., M. Charles, and D.C. Collins, 2016: Sensitivity of calibrated week-2 probabilistic forecast skill to reforecast sampling of the NCEP Global Ensemble Forecast system. *Wea. Forecasting*, **31**, 1093-1107, doi:10.1175/WAF-D-15-0166.1.

Unger, D. A., H. van den Dool, E. O'Lenic and D. Collins, 2009: Ensemble regression. *Mon. Weather Rev.*, **137**, 2365-2379, doi:10.1175/2008MWR2605.1.

van den Dool, H., E. Becker, L. Chen, and Q. Zhang, 2017: The probability anomaly correlation and calibration of probabilistic forecasts. *Wea. Forecasting*, **32**, 199-206, doi:10.1175/WAF-D-16-0115.1.

—, J. Huang, and Y. Fan, 2003: Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981–2001. *J. Geophys. Res.*, **108**(D16), 8617, doi:10.1029/2002JD003114.