Verification of the NCEP/EMC Unified Forecast System for Subseasonal to Seasonal Timescales

Lydia Stefanova[1*], Jessica Meixner[2], Avichal Mehra[2], Partha Bhattacharjee[1], Robert Grumbine[2], Bin Li[1], Shrinivas Moorthi[2], Jiande Wang[1], Denise Worthen[1]
[1]IM Systems Group at NOAA/NWS/NCEP/EMC, [2]NOAA/NWS/NCEP/EMC
[*]email: Lydia.B.Stefanova@noaa.gov

1. The UFS system for seasonal to subseasonal prediction

The NCEP Environmental Modeling Center Unified Forecast System (UFS) is a community-based modeling system designed for weather and climate forecasting on global or regional scales. The configuration of UFS for seasonal to subseasonal timescales is currently under development; at present it consists of atmosphere, ocean, and sea ice component models, coupled through a NEMS mediator. The addition of a coupled global wave model (WAVEWATCH III) is planned in the near future. The atmospheric model in this system is composed of the Finite Volume Cubed Sphere (FV3) dynamical core with GFS physics and GFDL microphysics parameterization. The oceanic model is the Modular Ocean Model (MOM6), and the sea ice model is the Los Alamos Sea Ice Model (CICE5). Upgrades and bug fixes to the system components are constantly incorporated as model components are updated by community effort.

2. UFS prototypes and benchmark framework

As the system grows in maturity and complexity, a systematic monitoring of performance is needed to ensure its quality. As part of this monitoring, sequential system prototypes (identified by specific components, settings, and initial conditions) specified in the course of development are validated and verified within a fixed “benchmark” framework. This benchmark framework is designed to test system performance for each new prototype with a consistent structure and fixed metrics.

The consistent structure is provided by requiring each prototype to produce a set of 35-day coupled forecasts initialized on the first and fifteenth day of every month between April 2011 and March 2018 for a total of 168 forecasts. The length of this dataset is a balance between providing a sufficient length for statistical analysis and limiting the strain on computing resources. The chosen period for benchmark verification spans varying climate conditions, as it includes several El Niño and La Niña events, as well as recent years of both high and low Arctic ice extent.

Since the model components are constantly being upgraded, it is not feasible to conduct a benchmark evaluation after every change. Instead, benchmark testing is performed at specific milestones of system development. To date, four prototypes, primarily targeting the impact of changing the source of initial conditions, have been defined and fully evaluated. The first prototype, UFS_p1, consisted of model components as described above, with the component versions current as of Oct 2018, and CFSR initial conditions for atmosphere, ocean, and sea ice initialization. For Prototype 2 (UFS_p2), the model components were updated to their then-current Mar-2019 states, and the initial conditions for the ocean were replaced with the 3Dvar from the NCEP Climate Prediction Center (CPC) GODAS. For Prototype 3 (UFS_p3), the model components were updated to their Jun-2019 states, and additionally the sea ice initial conditions were replaced with an ice analysis developed by CPC. An additional intermediary Prototype 3.1 (UFS_p3.1) was created to assess the impact of the coding changes implemented between Jun 2019 and Jan 2020, with the same initial conditions as in UFS_p3. A separate prototype (UFS_p3.2) is in the process of being run for tests to document the impact of atmospheric initial conditions only, while holding the code base the same as in UFS_p3.1.

3. Metrics

The main benchmark verification metrics consist of bias, RMS errors and anomaly correlations (AC) for a set of surface and upper air fields by lead week. Anomalies are calculated with respect to a smoothly interpolated climatology calculated by fitting the 7-year time series to a sine wave of period 365.25 days plus three harmonics. The smoothly interpolated climatology is calculated in the same way for both forecast and verification fields, separately for each grid point and lead time. Verification is performed against the CPC global 0.5-degree Unified Rain Gauge data (for precipitation over land), 6-hourly analysis guess 6-hr predictions from operational CFSv2 CDAS (for precipitation over ocean and upper air fields), CPC global 0.5 degree daily 2-meter temperatures, daily 0.25-degree OSTIA SST analysis, 500-hPa geopotential 6-hourly analyses from the operational CFSv2 CDAS. Model and verification data sets are interpolated to a common resolution prior to anomaly calculations. In addition, MJO index RMM1 & RMM2, and bivariate correlation skill are calculated following Wheeler and Hendon (2004) and Lin et al. (2008).

4. Results

For brevity, we focus here on week 3 and 4 AC scores, as this is the lead time for which subseasonal forecasts hold the most unrealized potential. Skill at shorter lead times is larger for all benchmark comparisons, but the conclusions regarding relative performances are similar. Beginning with the first prototype, the UFS system offers an improvement over the operational CFSv2 in terms of the week 3 and 4 AC scores for most fields (Fig. 1, left panel). The replacement of ocean initial conditions between UFS_p1 and UFS_p2 provided an additional skill improvement. The subsequent replacement of sea ice initial conditions between UFS_p2 and UFS_p3 did not have a beneficial impact for these scores for the fields shown here; it did however result in more accurate ice concentration threat scores (not shown). Little change was seen between UFS_p3 and UFS_p3.1. The lead time until the MJO bivariate correlation falls to 60% in the benchmark comparison went from 12 to 16.5 days between CFSv2 and UFS_p1, 19 days in UFS_p2, and 18 days in both UFS_p3 and UFS_p3.1 (Fig. 1, right panel, colored bars). These results are encouraging when compared to the individual models from the WWRP/WCRP sub-seasonal to seasonal prediction (S2S) project (Fig. 1, right panel, grey bars; Vitart 2017)

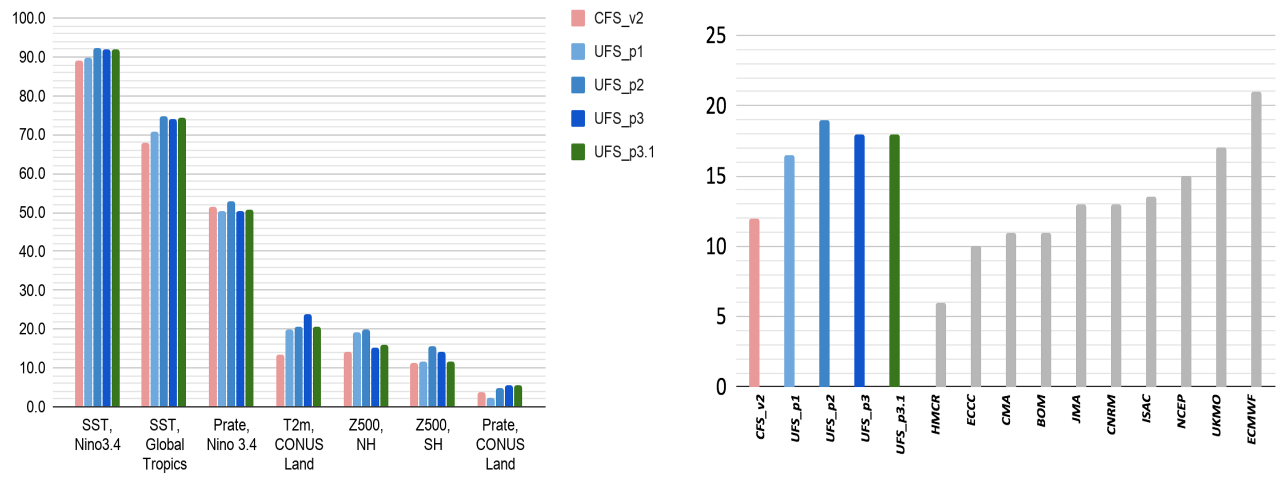

Figure 1. *Left*: Weeks 3 and 4 anomaly correlation (%) for select fields from benchmark runs (Apr 2011-Mar 2018) of CFSv2 and UFS prototypes 1 through 3.1. *Right*: Forecast lead time (days) at which the MJO bivariate correlation falls to 60%. Colored bars represent benchmark runs (Apr 2011-Mar 2018) of CFSv2 and UFS prototypes 1 through 3.1. Grey bars represent control runs (i.e., not ensembles) from various S2S models for 1999-2010 (based on Vitart, 2017).

Additional evaluations across prototypes demonstrate that ongoing developments have not altered the overall pattern of biases for most fields, and the prototypes are generally biased warm and wet. Across the lineup of prototypes, the largest boost in AC skill was associated with changing the ocean initial condition from CFSR to the 3Dvar CPC. AC scores for subsequent prototypes are comparable and remain an improvement over the operational CFSv2. This provides confidence in the system as components are refined; as system complexity increases, the reduction of biases and further skill improvement remain a target. It is likely that the greatest benefit for future performance is to be gained from planned component physics improvements and tuning, and advances in initializations, e.g., via land DA.

References

Lin, H., G. Brunet, and J. Derome, 2008: Forecast skill of the Madden–Julian Oscillation in two Canadian atmospheric models. *Mon. Wea. Rev.*, **136**, 4130–4149.

Saha, S., S. Moorthi, X. Wu, J. Wang, S. Nadiga, P. Tripp, D. Behringer, Y. Hou, H. Chuang, M. Iredell, M. Ek, J. Meng, R. Yang, M.P. Mendez, H. van den Dool, Q. Zhang, W. Wang, M. Chen, and E. Becker, 2014: The NCEP Climate Forecast System Version 2. *J. Climate*, **27**, 2185–2208.

Vitart, F., 2017: Madden–Julian Oscillation prediction and teleconnections in the S2S database. *Q.J.R. Meteorol. Soc*, **143**, 2210-2220.

Wheeler, M.C. and H.H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932.