Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
41st NOAA Annual Climate Diagnostics and Prediction Workshop
Orono, ME, 3-6 October 2016

Exploring the Impact of SST on the Extended Range NCEP Global Ensemble Forecast System

Christopher Melhauser[2], Wei Li[2], Yuejian Zhu[1], Xiaqiong Zhou[2], Malaquias Peña[2], and Dingchen Hou[1]

[1]*Environmental Modeling Center, NOAA/NWS/NCEP, College Park, Maryland*
[2]*IMSG at EMC, College Park, Maryland*

1. Introduction

The National Oceanic and Atmospheric Administration (NOAA) is accelerating its efforts to improve its numerical guidance and prediction capability for the extended range - the weeks 3 & 4 period that bridges the gap between weather and climate. Operational global numerical guidance for weeks 3 & 4 and monthly prediction are currently available from NOAA's National Center for Environmental Prediction (NCEP) Climate Forecasting System Version 2 (CFSv2) coupled (ocean, sea-ice, land, atmosphere) model (Saha *et al.*, 2006; Saha *et al.*, 2010; Saha *et al.*, 2014). Extending the NCEP Global Ensemble Forecasting System (GEFS) to cover the weeks 3 & 4 period provides additional benefits over the CFSv2 including a more frequent model upgrade cycle, higher model resolution, state-of-art flow-dependent initial perturbations from a hybrid data assimilation system, stochastic physics, and larger ensemble membership (80 perturbed members and 4 control runs for every 24-h period), all providing an improved sampling of forecast uncertainty.

In this study, an operational GEFS configuration is extended to 35 days and the forecast skill is evaluated. Various SST forcing experiments are performed to examine the impact of SST forcing on the extended-range forecast skill of global 2-m temperature, accumulated precipitation over the contiguous United States (CONUS), and Madden-Julian Oscillation (MJO; Madden and Julian, 1971) indices.

2. Methodology

The current operational configuration of GEFS uses the GFS Global Spectral Model v12.0.0 for integration four times per day (0000, 0600, 1200 and 1800 UTC) out to 16 days (Han and Pan 2011; Juang 2011, 2014). For days 0-8, GEFS has a spectral resolution of T_L574 (approximately 34 km) with 64 vertical levels and the horizontal resolution is reduced to T_L384 (approximately 52 km) for days 8-16. The horizontal resolution is further reduced to T_L254 (approximately 78 km) for days 16-35 for the extended GEFS runs in this study. The 20-member ensemble initial condition perturbations are selected from the operational hybrid NCEP Global Data Assimilation System (GDAS) 80-member Ensemble Kalman Filter (EnKF; Whitaker *et al.*, 2008; Wang *et al.* 2013; Kleist and Ide 2015) prior.

The SST configurations for this study consist of the operational GEFS 90 day e-folding of the observed RTG SST anomaly relaxed to climatology (CTL), an optimal Atmospheric Model Intercomparison Project (AMIP; Gates *et al.* 1999) configuration using the observed RTG SST analysis updated every 24-h during model integration (RTG), a 2-tier approach using the CFSv2 predicted SST updated every 24-h during model integration (CFS), and a 2-tier approach using biased corrected CFSv2 predicted SST updated every 24-h during model integration (CFS_BC). Detailed formulations for CTL and CFS_BC can be found in Appendix A.

All experiments span the fall and winter of 2013-14 and are initialized every 24 h starting 1 Sep 2013 and ending 28 Feb 2014. Over the experiment period, the MJO was weak or non-existent (Climate Prediction Center; http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/whindex.shtml) and ENSO neutral conditions persisted (Earth System Research Laboratory; http://www.esrl.noaa.gov/psd/enso/mei).

Correspondence to: Christopher Melhauser, IMSG at Environmental Modeling Center, NOAA/NWS/NCEP, 5830 University Research Court, College Park, MD; E-mail: christopher.melhauser@noaa.gov.

The forecast skill for 2-m temperature and accumulated precipitation are evaluated using a tercile (below-normal, normal, or above-normal with random chance being ⅓) probabilistic Heidke Skill Score (HKSS; *e.g.*, Wilks, 2011). The 2-m temperature is verified for land only against the 00 UTC GDAS analysis and the accumulated precipitation is verified for land only against the 00 UTC NCEP Climatologically Calibrated Precipitation Analysis (CCPA; Hou *et al.*, 2014) averaged or accumulated over days 8-14 (week 2) and days 15-28 (weeks 3 & 4).

The MJO is evaluated using the traditional real-time multivariate MJO (*i.e.* RMM) index (WH index; Wheeler and Hendon 2004, Gottschalck *et al.* 2010). The MJO forecast skill is defined as the bivariate anomaly correlation between the analysis and forecast RMM1 and RMM2. The long term climatology is calculated from the NCEP/NCAR Reanalysis 1 (http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html) and NCAR Interpolated Outgoing Longwave Radiation (http://www. esrl.noaa.gov/psd/data/gridded/data.interp_OLR.html) for the period of 1981-2010. The long term mean and average of the previous 120 days are removed from the forecast to eliminate long-term trends and seasonal variability.

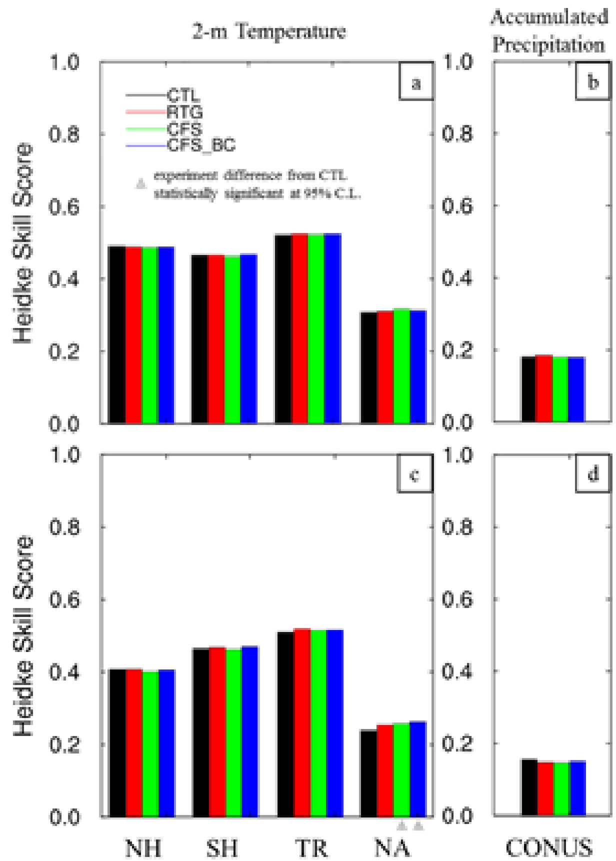

Fig. 1 Regional Heidke Skill Score for CTL (black), RTG (red), CFS (green), and CFS_BC (blue) calculated for week 2 (top row) and week 3 & 4 (bottom row) for 2-m temperature (a, c) and accumulated precipitation (b, d) averaged over the 6-month experiment period. The gray triangle indicates the difference of an experiment from CTL is statistically significant at the 95% confidence level.

3. Results and discussion

3.1 2-m temperature forecast skill

Over the fall and winter of 2013-14, the global land only 2-m temperature HKSS is regionally and lead time dependent. The tropics (TR) have the highest HKSS for both week 2 (Fig. 1a) and weeks 3 & 4 (Fig. 1b) with N. America (NA) having the lowest. Comparing between week 2 and weeks 3 & 4, the HKSS remains similar for the tropics and Southern Hemisphere (SH) with the Northern Hemisphere (NH) and NA dropping ~0.1-0.2. Within each region, the forecast skill for the SST forcing experiments is generally statistically indifferent from CTL for both week 2 and weeks 3 & 4. Both CFS and CFS_BC show a statistically significant improvement during weeks 3 & 4 over NA. It is interesting that RTG does not have a significant improvement over any land region compared to the other experiments, given this experiment is being forced the SST analysis. This suggests that there may be deficiencies in the forecast model which are limiting the spread of information from the ocean boundary to atmospheric land areas and the climatology of the SST analysis is most likely different than that of the model.

3.2 Accumulated precipitation forecast skill - CONUS

The CONUS accumulated precipitation HKSS shows no statistically significant difference between CTL and RTG, CFS, or CFS_BC for week 1 (not shown), week 2 (Fig. 1b), or weeks 3 & 4 (Fig. 1d). The magnitude of the HKSS falls off drastically after week 1 - approx. 0.55 at lead day 1 and 0.25 by lead day 7 (die off curves not shown). The aggregate accumulated week 2 HKSS is slightly higher than week 3 & 4, but overall, the results suggest minimal skill with the current model configurations, regardless of SST forcing.

3.3. MJO forecast skill

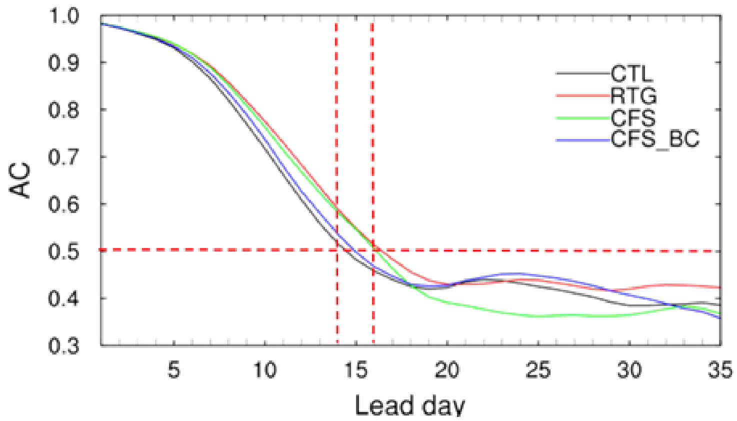

Fig. 2 MJO forecast skill (*i.e.* bivariate correlation between ensemble mean forecast and analysis data) as a function of lead time for the period of September 1, 2013 - February 28, 2014.

The forecast skill of MJO is a key metric when evaluating the capability of operational models for subseasonal forecasts (Kim *et al.* 2014; Shelly *et al.* 2014; Ling *et al.* 2014; Xiang *et al.* 2015). The MJO forecast skill in the operational version of GEFS is ~14.6 days (defined as the lead time when the bivariate anomaly correlation coefficient drops to 0.5) during the experimental period (Fig. 2). After week 2, MJO forecast skill quickly drops. Changing the prescribed SST to be closer to observations (RTG), the MJO forecast skill was improved up to ~2 days. For the weeks 3 & 4 range, the most skillful SST forcing is RTG with the CFS_BC being the most skillful scheme that could be practically used in operations. This implies that the MJO prediction skill is related to the accuracy of the representation of the SST. Therefore, without changing the model, it is found that improving the SST may potentially lead to an increase of the MJO skill.

4. Summary

The NCEP GEFS is being extended from 16 to 35 days to cover the subseasonal forecast period. The impact of SST forcing on the extended range land only global 2-m temperature, CONUS accumulated precipitation, and MJO indices forecast skill were examined using various SST forcing configurations.

Extending the GEFS showed forecast skill over weeks 3 & 4 for temperature, but minimal to no skill for accumulated precipitation. Forcing the GEFS with an optimal SST configuration showed minimal to no improvement in land only 2-m temperature and accumulated precipitation. Minimal improvements using more realistic SST over the current operational SST configuration were found when validating over the Tropics, Northern Hemisphere, Southern Hemisphere, and North America. The bias corrected CFS_BC SST performed the best over NA with statistically significant improvements for 2-m temperatures. The minimal differences in skill between SST forcing experiments suggests that systematic model errors dominate at the extended period with model boundary condition forcing having a secondary impact. The MJO skill in operational GEFS is 14.6 days. Using more realistic SST (RTG, CFS, and CFS_BC), MJO skill increase by 10%.

Observations indicate that the fall and winter of 2013-14 has a generally weak MJO. Future work will focus on a two-year span that covers a stronger MJO period spanning 1 May 2014 to 31 May 2016 providing insight into the predictability from strong MJO and its relationship with 2-m temperature and CONUS accumulated precipitation from global teleconnections.

This summary is a subset from a study in preparation for publication (Zhu *et al.* 2017).

APPENDIX

SST Forcing Calculations

Operational GEFS SST Forcing (CTL)

The GEFS v11 operational SST forcing uses a 90-day e-folding of the RTG analysis at initialization, relaxed to climatology, calculated as

$$\mathrm{SST}_f^t = \left[\mathrm{SST}_a^{t_0} - \mathrm{SST}_c^{t_0}\right] e^{-(t-t_0)/90} + \mathrm{SST}_c^t$$

where *f* is the forecast, *a* the analysis, *c* is climatology, *t* is forecast lead time, and t_0 is the initial time.

Bias Corrected CFSv2 Predicted SST Forcing (CFS_BC)

The CFS_BC SST forcing is a hybrid of a persisted RTG anomaly at short lead times and bias corrected CFSv2 predicted SST at longer lead times. The CFSv2 predicted SST is bias corrected using both the CFSR climatology and CFSv2 model climatology. The persisted RTG anomaly is linearly combined with the bias corrected CFSv2 predicted SST over the 35 d period, calculated as

$$\mathrm{SST}_f^t = (1-w)\left[\mathrm{SST}_a^{t_0} - \mathrm{SST}_{cfsrc}^{t_0} + \mathrm{SST}_{cfsrc}^t\right] + w\left[\mathrm{SST}_{cfs}^t - \left(\mathrm{SST}_{cfs\_c}^t - \mathrm{SST}_{cfsrc}^t\right)\right]$$

where *f* is the forecast, a the analysis, *cfsrc* is the CFSR reanalysis climatology, *cfs* is the CFSv2 model forecast, *cfs_c* is the CFSv2 model climatology, *t* is forecast lead time, t_0 is the initial time, and *w* is defined as

$$w = (t - t_0)/35$$

References

Gates, W. and coauthors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **80**, 29–55, doi: 10.1175/1520-0477(1999)080<0029:AOOTRO> 2.0.CO;2.

Gottschalck, J., and Coauthors, 2010: A framework for assessing operational Madden–Julian oscillation forecasts: A CLIVAR MJO working group project. *Bull. Amer. Meteor. Soc.*, **91**, 1247–1258, doi:10.1175/2010BAMS2816.1.

Han, J. and H. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. *Wea. Forecasting*, **26**, 520–533, doi: 10.1175/WAF-D-10-05038.1.

Hou, D., M. charles, Y. Luo, Z. Toth, Y. Zhu, R. Krzysztofowicz, Y. Lin, P. Xie, D. Seo, M. Pena, and B. Cui, 2014: Climatology-calibrated precipitation analysis at fine scales: Statistical adjustment of stage IV toward CPC gauge-based analysis. *J. Hydrometeor.*, **15**, 2542–2557, doi: 10.1175/JHM-D-11-0140.1.

Juang, H., 2011: A multiconserving discretization with enthalpy as a thermodynamic prognostic variable in generalized hybrid vertical coordinates for the NCEP Global Forecast System. *Mon. Wea. Rev.*, **139**, 1583–1607, doi: 10.1175/2010MWR3295.1.

Juang, H., 2014: A discretization of deep-atmospheric nonhydrostatic dynamics on generalized hybrid vertical coordinates for NCEP Global Spectral Model. NCEP Office Note 477, doi:10.7289/V5G73BM6.

Kim, H.-M., P. J. Webster, V. E. Toma, and D. Kim, 2014: Predictability and prediction skill of the MJO in two operational forecasting systems. *J. Climate*, **27**, 5364–5378, doi:10.1175/JCLI-D-13-00480.1.

Kleist, D. and K. Ide, 2015: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. *Mon. Wea. Rev.*, **143**, 452–470, doi: 10.1175/MWR-D-13-00350.1.

Ling, J., P. Bauer, P. Bechtold, A. Beljaars, R. Forbes, F. Vitart, M. Ulate, and C. Zhang, 2014: Global versus local MJO forecast skill of the ECMWF model during DYNAMO. *Mon. Wea. Rev.*, **142**, 2228–2247, doi:10.1175/MWR-D-13-00292.1.

Madden, R. and P. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708, doi: 10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

Saha, S. and co-authors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517, doi: 10.1175/JCLI3812.1.

Saha, S. and co-authors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi: 10.1175/2010BAMS3001.1.

Saha, S. and co-authors, 2014: The NCEP Climate Forecast System Version 2. *J. Climate*, **27**, 2185–2208, doi: 10.1175/JCLI-D-12-00823.1.

Shelly, A., P. Xavier, D. Copsey, T. Johns, J. M. Rodr´ıguez, S. Milton, and N. Klingaman, 2014: Coupled versus uncoupled hindcast simulations of the Madden-Julian Oscillation in the Year of Tropical Convection. *Geophys. Res. Lett.*, **41**, 5670–5677, doi:10.1002/2013GL059062.

Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, doi: 10.1175/MWR-D-12-00141.1.

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, doi: 10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Whitaker, J., T. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482, doi: 10.1175/2007MWR2018.1.

Wilks, D.S., 2011: Statistical methods in the atmospheric sciences. *Academic press*. Vol. 100, ISBN: 978-0-12-385022-5.

Xiang, B., M. Zhao, X. Jiang, S.-J. Lin, T. Li, X. Fu, and G. Vecchi, 2015: The 3–4-week MJO prediction skill in a GFDL coupled model. *J. Climate*, **28**, 5351–5364, doi:10.1175/JCLI-D-15-0102.1.

Zhu, Y., X. Zhou, M. Pena, W. Li, C. Melhauser, and D. Hou, 2017: Design and evaluation of sea surface temperature forcing on week 3&4 forecast skill using the NCEP Global Ensemble Forecasting System. *Wea. Forecasting*, in preparation.