Relative Merit of Model Improvement versus Availability of Retrospective Forecasts: The Case of Climate Forecast System MJO Prediction

QIN ZHANG AND HUUG VAN DEN DOOL

NOAA/NWS/NCEP/Climate Prediction Center, Camp Springs, Maryland

(Manuscript received 28 October 2011, in final form 22 February 2012)

ABSTRACT

Retrospective forecasts of the new NCEP Climate Forecast System (CFS) have been analyzed out to 45 days from 1999 to 2009 with four members (0000, 0600, 1200, and 1800 UTC) each day. The new version of CFS [CFS, version 2 (CFSv2)] shows significant improvement over the older CFS [CFS, version 1 (CFSv1)] in predicting the Madden–Julian oscillation (MJO), with skill reaching 2–3 weeks in comparison with the CFSv1's skill of nearly 1 week. Diagnostics of experiments related to the MJO forecast show that the systematic error correction, possible only because of the enormous hindcast dataset and the ensemble aspects of the prediction system (4 times a day), do contribute to improved forecasts. But the main reason is the improvement in the model and initial conditions between 1995 and 2010.

1. Introduction

It is commonplace to assume that changes made to numerical weather prediction (NWP) systems will lead to better predictions. After all, these changes are inspired by improved insights in physics, data assimilation, numerics, and/or by increases in computing power (i.e., higher spatial resolution). While this has to be generally true, the introduction of model changes into a real-time environment can be disappointing. The new model may not seem much better than the (now discontinued) old one, in part because interannual variability of predictability may obscure the improvement, if any. It is only when verification scores over many years are considered that the truly impressive general upward trend in forecast skill at all NWP centers is seen. The day 5 anomaly correlation for NH 500-mb height has steadily improved from the 0.6–0.7 range in 1985 to the 0.8–0.9 range in 2012. An up-to-date comparison of National Centers for Environmental Prediction (NCEP) and European Centre for Medium-Range Weather Forecasts (ECMWF) scores is available online (http://www.emc.ncep.noaa.gov/gmb/STATS/html/aczrnmn4.html). This great improvement amounts to a revolution accomplished inch by inch by dozens of changes in computing, models, and data assimilation over 25 yr.

From an improvement of 20 anomaly correlation points in 25 yr, and say one model upgrade per year, one may surmise that a new model is on average only order 1 point better than its immediate predecessor. That too makes it hard to see the improvement in real time with only one model running at any given time.

A different way of improving NWP forecasts is to generate a large set of retrospective forecasts (also called hindcasts or reforecasts) before implementing a new model. This is a considerable investment requiring initial states for past cases that suit a modern modeling system, and lots of computer time, and thus this method is not very popular, but it does allow practitioners to correct models in real time for known biases, whether it be in the mean, higher-order moments, or even spatial pattern offsets. Upfront one must distinguish two limiting situations: 1) the early leads with high forecast skill and a small bias and 2) the longer leads with low forecast skill and potentially very large biases. Obviously, when a model has only small biases (typically the early leads), the generation of hindcasts is not going to help much because there is little to correct. Equally obvious is that when a model has virtually no skill intrinsically (at the ultralong leads), the availability of hindcasts may help to correct embarrassing

Corresponding author address: Dr. Qin Zhang, NOAA/NWS/NCEP/Climate Prediction Center, 5200 Auth Rd., Camp Springs, MD 20746.
E-mail: qin.zhang@noaa.gov

DOI: 10.1175/WAF-D-11-00133.1

systematic errors but will still fail to deliver skillful forecasts. It is somewhere in between, at intermediate leads with nonzero intrinsic skill and significant systematic error, that hindcasts may be expected to help us the most. It is not always clear a priori where these ''intermediate'' leads really are. For precipitation it may be within 24 h, while for the Madden and Julian oscillation (MJO) it may be days 10–30. The MJO, a dominant 20–70-day phenomenon in the tropics, has been predicted by NWP with only modest skill out to about 10 days, and with limited progress in forecast skill over the years [see Gottschalck et al. (2010) and references therein). Intuitively, however, one assumes that the predictability horizon is related to the phenomenon's life time (Van den Dool and Saha 1990), which would be much more than 10 days for a 20–70-day oscillation.

We recently had a rare opportunity to compare two NCEP models of vintage ~1995 and ~2010, both of which were accompanied by an impressive set of hindcasts. We focus here exclusively on the challenging MJO prediction by studying the relative merits of model improvement (in this case between 1995 and 2010) and the availability of hindcasts for each. We will report here on rather large improvements of the NCEP Climate Forecast System, version 2 (CFSv2), versus version 1 (CFSv1), and good improvement for both systems thanks to the availability of hindcasts. Saying that CFSv2 is much better than CFSv1 does not mean the CFSv2 is the best model in the world; in fact, we will report levels of skill for CFSv2 that appear to be comparable to those reported by Rashid at al. (2010) for the Predictive Ocean–Atmosphere Model for Australia (POAMA).

In section 2 we describe the models and the data used, as well as the procedures used to focus on MJO, do systematic error correction, etc. In section 3 we present the results and we close with conclusions and discussion in section 4.

2. Data and analyses

a. Models

In the spring of 2010, NCEP finished a new global reanalysis, the Climate Forecast System Reanalysis (CFSR), for the period 1979–present (Saha et al. 2010). This new reanalysis, at T382L64 resolution in the atmosphere, was primarily motivated by the need to provide initial conditions for retrospective predictions that will go along with the implementation of the new CFSv2, which became operational at NCEP in March 2011. CFSv2 will replace the older CFSv1, which has been operational since mid-2004. The new CFSv2 has as much consistency as possible between the historical initial conditions and the hindcast model as well as with the real-time coupled prediction model. The CFSv2 reforecasts are described by S. Saha et al. (2012), unpublished manuscript). The CFSv1, implemented in 2004 at an atmospheric T62L64 resolution, was built around such diverse pieces as the NCEP/Department of Energy Global Reanalysis 2 (R2) atmospheric initial conditions (Kanamitsu et al. 2002), a global atmospheric model (Global Forecast System, GFS) of vintage 2003, a near-global ocean model (version 3 of the Modular Ocean Model), near-global ocean initial conditions produced by the stand-alone Global Ocean Data Assimilation System (GODAS), plus hindcasts from 1981 to 2004 thereof (Saha et al. 2006). Because of R2, the oldest and probably the weakest link, the vintage of CFSv1 is described here as mid-1990s. Both CFSv1 and CFSv2 are coupled to the ocean.

b. Data of model forecasts

For CFSv1 (CFSv2), we study 30- (45-) day forecasts during 1999–2009. CFSv2 begins forecasts every 6 h, at 0000, 0600, 1200, and 1800 UTC, 365 days a year. CFSv1 has starts at 0000 UTC in three groups of 5 days centered on the 1st, the 11th, and the 21st of the month, for a total of 15 days per month. The CFSv1 forecasts are treated and verified as 15 individual runs, not as a lagged ensemble (as we do for the seasonal CFSv1 forecasts). The CFSv2 forecasts are also treated and verified as individual runs, not as a lagged ensemble (as we do for the seasonal CFSv2 forecasts), except that in most plots we have averaged the 0000, 0600, 1200, and 1800 UTC forecast from the same initial date into an ensemble mean. Later (Fig. 3), we will compare the ''0000 UTC only'' predictions of both systems.

For MJO applications purposes, we retain only three fields: the zonal wind (u) components at 850 and 200 hPa, and the outgoing longwave radiation (OLR), all on a 2.5×2.5 grid. The meridional average between 15°S and 15°N is taken, yielding a one-dimensional array depending on longitude only, denoted as $X(t, \tau, \lambda)$ for any of the three variables, where λ is longitude, τ is lead time, and t is the starting day [t runs through time, both days (1–365) and years (1–11)]. The notation $X(t, \tau, \lambda)$ covers both initial states ($\tau = 0$) and forecasts ($\tau > 0$), although we have to understand that there is no model-generated or model-analyzed OLR at the initial time. Below, we will refer to $X_{v1}(t, \tau, \lambda)$ and $X_{v2}(t, \tau, \lambda)$ as the data from CFS versions 1 and 2, respectively.

We need to define a general purpose ''observed'' dataset $X(t, \tau = 0, \lambda)$ where the observations,

consistent with Climate Prediction Center (CPC) operational practice, are taken from a satellite dataset of OLR, and u200 and u850 are taken from R2. This observed dataset, to be considered the daily mean data, covers 1979–2010. CFSv1 and CFSv2 forecasts are verified against the same observed dataset to avoid ambiguities.

c. Climatologies

We need climatologies for $X(t, \tau, \lambda)$, to form observational and forecast anomalies, and to define the systematic error correction (SEC) as practiced here. Smooth climatologies of X are formed by fitting the annual mean and four harmonics to a time series of raw climatologies of any of the three variables at any given longitude for forecast lead τ, $\tau = 0...30$ (or 45), following the method of Schemm et al. (1997), described in great detail as it was applied to CFSv1 by Johansson et al. (2011). ["Raw" climatology is defined as the multiyear average for each day of the year; this forms a very noisy time series of length 365 days even when based on 24 (CFSv1) or 11 (CFSv2) yr] For the observations, $\tau = 0$, the period 1981–2004 is used to determine a smooth climatology of X, denoted as $C(t, 0, \lambda, 1981–2004)$ where the 4th argument reminds us of the range of years in the climatology, a very important detail. The variable t in C runs through the days of the years, but repeats itself each year.

The harmonic fit does not only smooth noisy climatologies, but also interpolates them to the dates for which no historical hindcasts are available; this comment applies particularly to CFSv1 because it has three groups of 5 days (15 per month) with forecasts and holes in between.

Observed anomalies are then given by $X(t, 0, \lambda) - C(t, 0, \lambda, 1981–2004)$. Verification will be done in terms of anomalies relative to 1981–2004. CFSv1 has its forecast climatology based on 1981–2004. CFSv2 has its forecast climatology based on 1999–2009 only. This discrepancy is addressed in section 2d. The forecast climatologies are denoted by $C_{v1}(t, \tau, \lambda, 1981–2004)$ and $C_{v2}(t, \tau, \lambda, 1999–2009)$.

d. Systematic error correction

The generic definition of the systematic error (SE) is the difference in the expected value of the forecasts and verifying observations—and this is estimated as the mean over as many cases (years) as possible. For instance, for CFSv1 we have $\mathrm{SE} = C_{v1}(t, \tau, \lambda, 1981–2004) - C(t, 0, \lambda, 1981–2004)$. In that case, subtracting from a forecast the lead dependent forecast climatology is enough to obtain forecast anomalies corrected for the systematic error in the mean; that is,

$$X_{v1}(t, \tau, \lambda) + \varepsilon[-C_{v1}(t, \tau, \lambda, 1981–2004) + C(t, 0, \lambda, 1981–2004)] - (1 - \delta)C(t, 0, \lambda, 1981–2004), \quad (1)$$

where we obtain full fields for $\delta = 1$ and anomalies for $\delta = 0$. For $\varepsilon = 1$ (or $\varepsilon = 0$), we perform (or do not perform) an SEC. For CFSv2, the situation is more complicated because the matching years are only 1999–2009, while anomalies relative to 1981–2004 are required for verification in the same way as CFSv1. This can be achieved by

$$X_{v2}(t, \tau, \lambda) + \varepsilon[-C_{v2}(t, \tau, \lambda, 1999–2009) + C(t, 0, \lambda, 1999–2009)] - (1 - \delta)C(t, 0, \lambda, 1981–2004), \quad (2)$$

where, again, we obtain full fields for $\delta = 1$ and anomalies for $\delta = 0$. Notice that we have an intermediate observed climatology denoted as $C(t, 0, \lambda, 1999–2009)$, which is based on a satellite dataset of OLR, and u200 and u850 are taken from CFSR.

Below in section 3 we compare scores for systematic error corrected forecasts ($\varepsilon = 1$) to raw uncorrected forecasts ($\varepsilon = 0$).

e. MJO and verification

Following Wheeler and Hendon (2004), a joint EOF analysis is performed on the observed OLR, u200, and u850 fields over 1979–2004, and two PCs of combined EOFs (real-time multivariate MJO series 1 and 2, RMM1 and RMM2) describing the propagating MJO are kept. They explain about 25% of the variance in filtered data and are used, unchanged, in all seasons. The EOF analysis has been done after the observed data are filtered as follows: (a) the climatology is removed, (b) ENSO variability is removed (by regression on Niño-3.4), and (c) the latest 120-day mean is removed. The forecasts, in anomaly form, are projected onto these EOF1 and EOF2 patterns for verification purposes. The verification is done by the bivariate anomaly correlation (BAC) given in Lin et al. [(2008), their Eq. (1)]:

$$\mathrm{COR}(\tau) = \frac{\sum_{i=1}^{N} [a_{1i}(t)b_{1i}(t) + a_{2i}(t)b_{2i}(t)]}{\sqrt{\sum_{i=1}^{N} [a_{1i}^2(t) + a_{2i}^2(t)]}\sqrt{\sum_{i=1}^{N} [b_{1i}^2(t) + b_{2i}^2(t)]}}, \quad (3)$$

where $a_{1i}(t)$ and $a_{2i}(t)$ are the observed RMM1 and RMM2 amplitudes at day t, and $b_{1i}(t)$ and $b_{2i}(t)$ are their

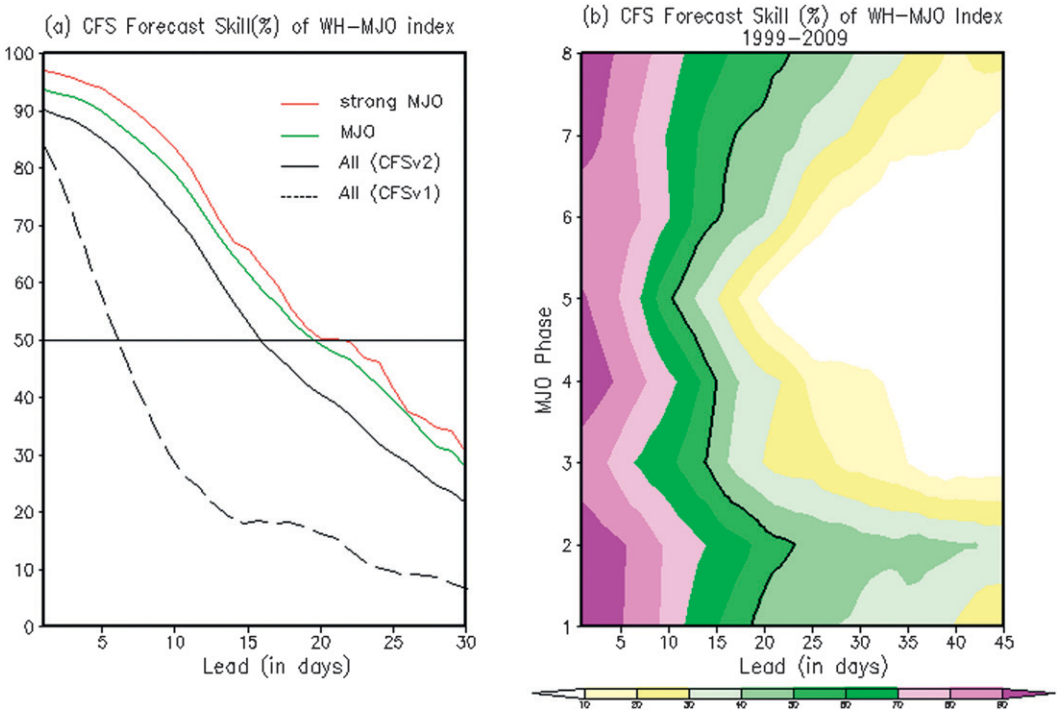

FIG. 1. (a) The BAC × 100 of bias-corrected CFSv2 (solid black line) and CFSv1 (dashed black line) in predicting the MJO during the period 1999–2009, as expressed by the WH index (two EOFs of combined zonal wind and OLR). The BACs of CFSv2 for selected MJO cases and strong MJO cases, respectively, are plotted in the green and red lines. (b) The BAC of the eight MJO initial phases of CFSv2 for all cases. The black lines [horizontal in (a), vertical in (b)] designates the 0.5 level for the BAC.

respective forecasts, for the ith forecast with a τ-day lead. Here, N is the number of forecasts. COR(τ) is equivalent to a spatial pattern correlation between the observations and the forecasts when they are reconstructed from the two leading combined EOFs.

The verification of the forecasts made from initial conditions during 1999–2009 and the forecast lead is measured in units of days; for instance, today's 0000 UTC forecast verifying tomorrow is counted as a 1-day lead. The CFSv2 forecasts are, in general, a lagged-average forecast, based on 0000, 0600, 1200, and 1800 UTC, and the starting point in that case is 1800 UTC. The difference in lead between CFSv1 (which has only a 0000 UTC forecast) and CFSv2 (which has four starts a day) is taken into account in the comparisons that follow. Note that we include (Fig. 3) a comparison between CFSv2 and CFSv1 using only the 0000 UTC data; that is, we forego in Fig.3 the advantage CFSv2 may have over CFSv1 from the ensemble of four starts a day.

3. Results

Although the new CFSv2 has been developed ostensibly as a seasonal prediction system, there will be viable applications to much shorter forecasts leads as well because the forecasts will be run in real time without delay. Below, we highlight the skill in MJO hindcasts in the range of 1–45 days and compare CFSv2 to CFSv1, ignoring the fact that CFSv1 was run with a delay.

Figure 1a shows the annually aggregated correlation of CFSv2 (black solid), and CFSv1 (dashed), for all cases. The forecast lead is horizontal, and the correlation × 100 is applied along the vertical axis. We calculate correlations as a function of lead for each initial day, that is, for given τ over 11 × 365 cases, one case for each day (labeled "All" in Fig. 1a), MJO cases (2590 cases) when the MJO amplitude [square root of ($RMM1^2$ + $RMM2^2$)] is greater than the MJO standard deviation, and strong MJO cases (663 cases) when the MJO amplitude (square root of ($RMM1^2$ + $RMM2^2$) is greater than 2 times the standard deviation.

Figure 1a shows the skill [as per anomaly correlation (AC)] of CFSv2 in predicting the MJO, as expressed by the Wheeler and Hendon (WH) index (two EOFs of combined zonal wind and OLR), which is subjected to SEC. It is quite clear that CFSv2 has much higher skill than CFSv1 for the "All cases" and the forecast skill for the MJO and strong MJO cases is out to 2 and 3 weeks,

respectively. An increase in the anomaly correlation (a signal-to-noise measure), when the (initial) signal is stronger (from all via MJO to strong MJO), is to be expected (see Van den Dool and Toth 1991). Figure 1a shows the improvement over half a generation (~15 yr of work by many people), taking into account that many physical processes have been improved and that CFSv1 has rather old R2 atmospheric initial conditions as its weakest component (S. Saha et al. 2012, unpublished manuscript). One rarely sees such a demonstration of improvement. This is because atmospheric NWP models are normally abandoned when a new model comes in. But in the application to seasonal climate forecasting, systems tend to have a longer lifetime. This gave us a rare opportunity to compare two models that are about 15 yr apart in vintage. The causes for the enormous improvement seen in Fig. 1 are probably very many and result from the model physics improvements, and from the improved initial states in the tropical atmosphere and the consistency of the initial state and the model used to make the forecasts play a role (S. Saha et al. 2012, unpublished manuscript). Further research should bring out the importance of coupling to the ocean and its quantitative contribution to skill.

We believe that the improved atmospheric initial conditions are the main reason for the large improvement seen in CFSv2 relative to CFSv1. Such improvements were anticipated in studies by Vintzileos and Pan (2008) and Fu et al(2011), who made a limited set of CFSv1 integrations starting from operational GDAS initial states and found a major improvement in skill relative to CFSv1 initialized by R2.

The right panel in Fig. 1 is the categorized correlation for the eight phases of MJO (Wheeler and Hendon 2004) against the forecast lead day for the all cases of CFSv2. The higher skill levels around phases 2 and 8 indicate the longer time predictability of MJO when it forms and redevelops over Africa and the western Indian Ocean with eastward propagating into the Western Hemisphere, respectively. The skill levels in phases 3 and 5 are lower, corresponding to the difficulties of predicting MJO over the Indian Ocean and the Maritime Continent, where convection involves complex air–sea interaction and has a strong seasonally varying component of northward propagation (Fu et al. 2003).

Figure 2 is plotted with the day of the year in the vertical (months are labeled for reference) and forecast lead horizontally, and the correlation $\times$ 100 is contoured. It shows the results with and without the benefit of SEC for CFSv2. In Fig. 2 we show the MJO scores for CFSv2 for raw model predictions (left), after systematic error correction (middle), and the gain (or difference) on the right. We see that SEC results in improvements for CFSv2 far more often than not, and overall the improvement is between 5 and 10 points, which could be the equivalent of several new model implementations. As expected the improvement depends on both lead and season. There is very clear improvement for leads of 10–30 days for starts in August and September, which means the model (CFSv2) has a large correctable bias in these months. For some long leads, the scores actually deteriorate when applying SEC. This happens when the error in the estimate of the SE is large in a relative sense, that is, relative to the true systematic error (which is unknown), or when skill cannot be salvaged (long leads).

Figure 2 also shows a seasonal cycle in forecast skill with maxima in May–June and November–December, respectively, and minima in between, something we also find for CFSv1 (not shown). This semiannual timing of higher skill may be related to those two times of the year when the eastward propagation of the dominant wavenumber 1 in the anomaly velocity potential along the equator was found to be climatologically at a maximum (~10–11 m s^{-1}) in May and November (Van den Dool and Saha 2002), with much smaller phase speeds (~6 m s^{-1}) in between in February and August–September. Additional reasons for (semi-) annual behavior may be related to the more northward propagation of the intraseasonal oscillation (Fu et al. 2003) in certain seasons and the inability of the season-invariant EOFs to handle this aspect properly.

As is the case with CFSv2, CFSv1 did benefit noticeably from the availability of its hindcasts (not shown). While the distribution of the improvement with lead and season is different for CFSv1, the overall improvement is rather comparable.

We summarize in Fig. 3 the comparison of both systems by taking an annual mean of the BAC for leads 1–45 days for CFSv2 and out to 30 days for CFSv1. Both with and without SEC, CFSv2 is some 11–12 days ahead of CFSv1, when measured at the 0.5 correlation level. Both CFSv1 and CFSv2 appear to gain about 2–3 days by applying an SEC. Obviously, the model and data assimilation improvements between 1995 and 2010 count for much more than the availability of the hindcasts, but the latter still corresponds to a few years of model improvement. There is third, smaller factor to consider: a small gain (at most a day) achieved by CFSv2 hindcasts through a lagged-average forecasting approach over forecasts from 0000, 0600, 1200, and 1800 UTC; see the dashed green line in Fig. 3, which indicates when only raw 0000 UTC data are used. Since CFSv1 has only 0000 UTC forecasts (the black and blue full lines), it is appropriate to compare its results to the CFSv2 0000 UTC scores (red and green dashed lines). The added advantage of CFSv2 having an ensemble is

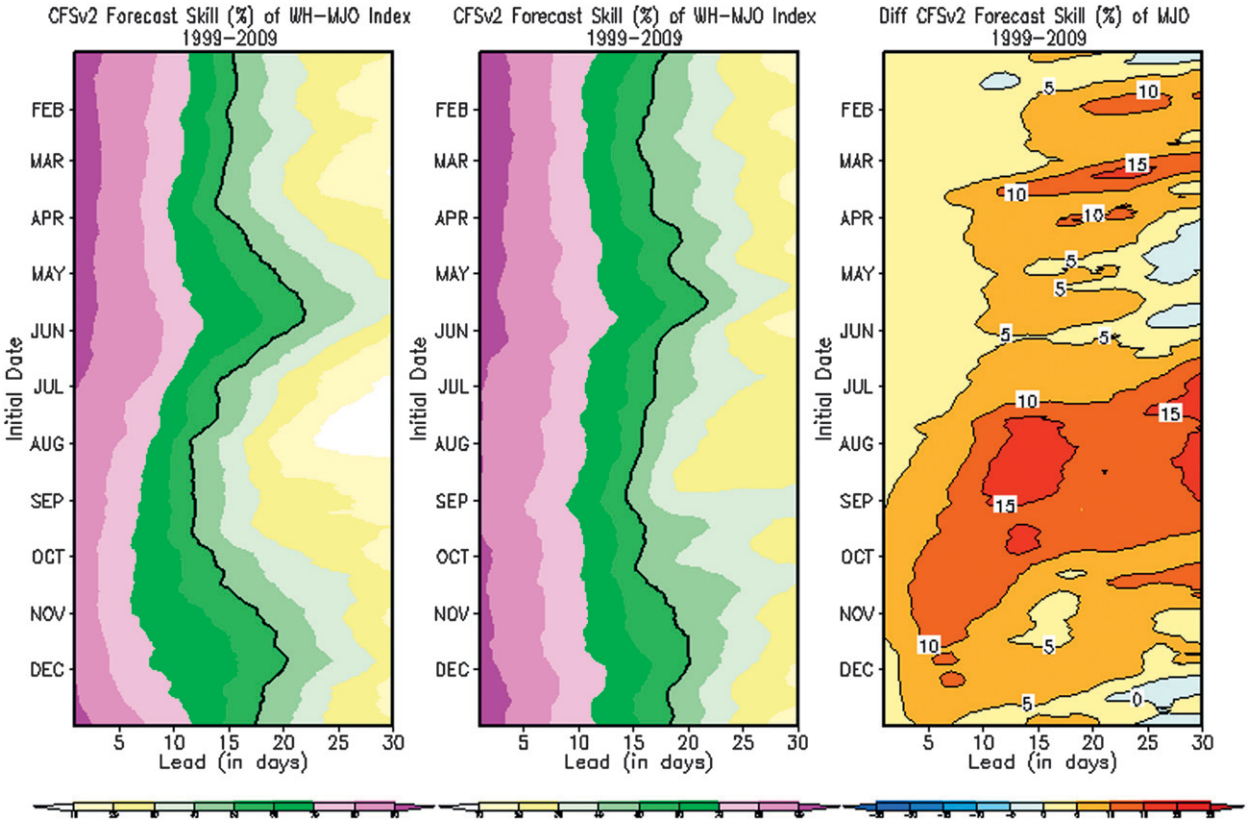

FIG. 2. The BAC calculated from CFSv2 forecasts (left) before and (middle) after systematic error correction. (right) The difference in the contouring of 0.05, with red (blue) areas of increased (decreased) skill. A slight smoothing is applied in the vertical (across adjacent start days).

at most a day, and this gain is so small partly because the ensemble is ''lagged'' in nature (causing a setback of 18 h). Note that in real time CFSv2 will have four members from 0000 UTC (as well as four from the previous 1800, 1200, and 0600 UTC), so the skill of CFSv2 in real time should be higher than what is shown here for the hindcasts.

4. Conclusions and discussion

Two conclusions suggest themselves. The first is that CFSv2 is markedly better than CFSv1 in predicting MJO. The credit goes to all the hard work done at NWP centers between 1995 and 2010. The second is that the availability of hindcasts really helps with MJO forecasting and, quantitatively, appears to be worth as much as several new model implementations. This has rarely been demonstrated, although it is often assumed to be true, especially for seasonal prediction where the SE can be embarrassingly large. We hasten to point out that we have only addressed the correction of the mean error. Other benefits that can be harvested from having gone through the trouble of making hindcasts may lie in correcting 1) higher-order moments and 2) offsets in pattern. This, however, takes a significant amount of data (large sample) and these issues have come up more often, naturally, within the context of having hindcasts in an

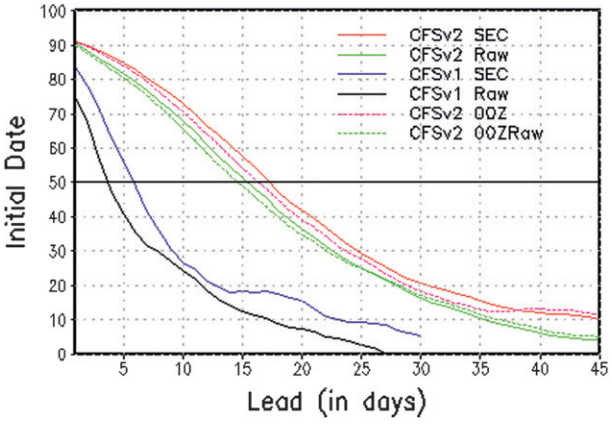

FIG. 3. Annual mean score (BAC) as a function of lead for CFSv2 (curves on the right) and CFSv1 (curves on the left). The dashed lines are for CFSv2 using only the 0000 UTC forecasts.

ensemble prediction setting (Hamill et al. 2006), which is yet another track for improving forecasts.

Acknowledgments. The authors thank Suranjana Saha, Patrick Tripp, and others at NCEP's Environmental Modeling Center for making these forecasts and providing the data. The authors gratefully acknowledge the support by a NASA-MAP proposal entitled Pathways to Predictability on Sub-Seasonal Time Scales: Assessing the Role of Tropical Forcing and Land Surface Conditions.

REFERENCES

Fu, X., B. Wang, T. Li, and J. P. McCreary, 2003: Coupling between northward-propagating, intraseasonal oscillations and sea surface temperature in the Indian Ocean. *J. Atmos. Sci.,* **60,** 1733–1753.

——, ——, J.-Y. Lee, W. Wang, and L. Gao, 2011: Sensitivity of dynamical intraseasonal prediction skills to different initial conditions. *Mon. Wea. Rev.,* **139,** 2572–2592.

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.

Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.,* **87,** 33–46.

Johansson, Å., C. Thiaw, and S. Saha, cited 2011: CFS retrospective forecast daily climatology. [Available online at http://cfs.ncep.noaa.gov/cfs.daily.climatology.doc.]

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83,** 1631–1643.

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.

Rashid, H. A., H. H. Hendon, M. C. Wheeler, and O. Alves, 2010: Prediction of the Madden–Julian oscillation with the POAMA dynamical prediction system. *Climate Dyn.,* **36,** 649–661, doi:10.1007/s00382-010-0754-x.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate,* **19,** 3483–3517.

——, and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.,* **91,** 1015–1057.

Schemm, J.-K. E., H. M. van den Dool, J. Huang, and S. Saha, 1997: Construction of daily climatology based on the 17-year NCEP–NCAR reanalysis. Proceedings of the First WCRP International Conference on Reanalyses, WMO/TD-876, WCRP-104, 290–293.

Van den Dool, H. M., and S. Saha, 1990: Frequency dependence in forecast skill. *Mon. Wea. Rev.,* **118,** 128–137.

——, and Z. Toth, 1991: Why do forecasts for ''near normal'' often fail? *Wea. Forecasting,* **6,** 76–85.

——, and S. Saha, 2002: Analysis of propagating modes in the tropics in short AMIP runs. *WCRP/WGNE AMIPII Workshop,* Toulouse, France, WMO, 87–90. [Available online at http://www.cpc.ncep.noaa.gov/products/people/wd51hd/mjo/AMIPIIdoolsaha.pdf.]

Vintzileos, A., and H.-L. Pan, 2008: On the importance of horizontal resolution and initial conditions to forecasting tropical intraseasonal oscillations: The maritime continent prediction barrier. *Proc. CTB–COLA Joint Seminar Series,* Camp Springs, MD, NOAA/NWS/Office of Science and Technology. [Available online at http://www.nws.noaa.gov/ost/climate/STIP/CTB-COLA/ctb-cola_seminar_summaries.pdf.]

Wheeler, M., 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.