Subseasonal Forecasting with an Icosahedral, Vertically Quasi-Lagrangian Coupled Model. Part I: Model Overview and Evaluation of Systematic Errors

SHAN SUN

Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/OAR/ESRL/Global Systems Division, Boulder, Colorado

RAINER BLECK

Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/OAR/ESRL/ Global Systems Division, Boulder, Colorado, and NASA Goddard Institute for Space Studies, New York, New York

STANLEY G. BENJAMIN

NOAA/OAR/ESRL/Global Systems Division, Boulder, Colorado

BENJAMIN W. GREEN

Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/OAR/ESRL/Global Systems Division, Boulder, Colorado

GEORG A. GRELL

NOAA/OAR/ESRL/Global Systems Division, Boulder, Colorado

(Manuscript received 4 January 2018, in final form 2 April 2018)

ABSTRACT

The atmospheric hydrostatic Flow-Following Icosahedral Model (FIM), developed for medium-range weather prediction, provides a unique three-dimensional grid structure-a quasi-uniform icosahedral horizontal grid and an adaptive quasi-Lagrangian vertical coordinate. To extend the FIM framework to subseasonal time scales, an icosahedral-grid rendition of the Hybrid Coordinate Ocean Model (iHYCOM) was developed and coupled to FIM. By sharing a common horizontal mesh, air-sea fluxes between the two models are conserved locally and globally. Both models use similar adaptive hybrid vertical coordinates. Another unique aspect of the coupled model (referred to as FIM-iHYCOM) is the use of the Grell-Freitas scale-aware convective scheme in the atmosphere. A multiyear retrospective study is necessary to demonstrate the potential usefulness and allow for immediate bias correction of a subseasonal prediction model. In these two articles, results are shown based on a 16-yr period of hindcasts from FIM-iHYCOM, which has been providing real-time forecasts out to a lead time of 4 weeks for NOAA's Subseasonal Experiment (SubX) starting July 2017. Part I provides an overview of FIM-iHYCOM and compares its systematic errors at subseasonal time scales to those of NOAA's operational Climate Forecast System version 2 (CFSv2). Part II uses bias-corrected hindcasts to assess both deterministic and probabilistic subseasonal skill of FIM-iHYCOM. FIM-iHYCOM has smaller biases than CFSv2 for some fields (including precipitation) and comparable biases for other fields (including sea surface temperature). FIM-iHYCOM also has less drift in bias between weeks 1 and 4 than CFSv2. The unique grid structure and physics suite of FIM-iHYCOM is expected to add diversity to multimodel ensemble forecasts at subseasonal time scales in SubX.

1. Introduction

Over the past decade, there has been a significant effort (e.g., Brunet et al. 2010; Kirtman et al. 2014; Vitart

et al. 2017) to predict the state of the atmosphere at subseasonal time scales (3–4 weeks to 2 months). This time scale falls into a gap between numerical weather prediction (NWP) and climate prediction. NWP time scales fall within the theoretical deterministic predictability limit of \sim 2 weeks at midlatitudes (Lorenz

DOI: 10.1175/MWR-D-18-0006.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Corresponding author: Shan Sun, shan.sun@noaa.gov

1969), and large-scale external forcings to the atmosphere (such as from the ocean) generally evolve much more slowly than the rapid upscale error growth within the atmosphere. In contrast, the main concern with seasonal, interannual, and longer-term climate prediction of the atmosphere is on the evolution of the external forcings.

The challenge with forecasting at subseasonal time scales is stated well by Vitart et al. (2017, p. 164): "... the lead time is sufficiently long that much of the memory of the atmospheric initial conditions is lost and it is too short for the variability of the ocean to have a strong influence." Naturally, then, there have been two approaches toward subseasonal forecasting, both of which are found in the models contributing to the Subseasonalto-Seasonal Project (established by the World Meteorological Organization as a joint project between the World Weather Research Programme and World Climate Research Programme; Vitart et al. 2017). The first is "climate down," in which seasonal or longer-range climate models are simply run with higher-frequency (daily) output typically for shorter times. Examples include, but are not limited to, the National Oceanic and Atmospheric Administration (NOAA) operational Climate Forecast System version 2 [CFSv2; Saha et al. (2014, hereafter S14)]; the forecast-oriented low ocean resolution version of the Geophysical Fluid Dynamics Laboratory (GFDL) Coupled Model version 2.5 (FLOR; Vecchi et al. 2014); and the Met Office Global Seasonal forecast system version 5 (GloSea5; MacLachlan et al. 2015). The second approach, which for reasons discussed in the next paragraph has not been as common, is "weather up," in which forecasts from NWP models are extended beyond 2 weeks. Examples of "weather up" subseasonal forecasts from NWP models include the Canadian Global Environmental Prediction System (GEPS; Lin et al. 2016); the Global Ensemble Forecast System (GEFS; Zhu et al. 2017); and the GLOBO model run at the ISAC Institute of the Italian National Research Council (Mastrangelo et al. 2012).

The quasi-uniform horizontal-grid coupled model presented here has its origins in NWP and thus is a good example of the "weather up" approach. Typically, NWP models do not have a dynamical forward model to predict sea surface temperature (SST) because SST—which is a strong forcing to the atmospheric system—generally exhibits little to no change on NWP time scales. Instead, NWP models often start with observationally constrained SST fields. Over the course of the simulation, these models usually either (i) retain the initial SST *anomaly* on top of a time-dependent climatology (e.g., GEPS; Lin et al. 2016) or (ii) relax the initial SST toward climatology (e.g., GLOBO; Mastrangelo et al. 2012). Zhu et al. (2017) tested several approaches to provide SST information to the atmosphere-only GEFS and found the best performance, in terms of subseasonal forecast skill, from a "two-tiered" method in which biascorrected SST forecasts from CFSv2 were used.

In general, the main challenge for "weather up" subseasonal prediction is determining the level of complexity needed for prognostic SST. To more fully address Earth system processes likely important for subseasonal time scales, it was decided to couple the atmospheric model [the Flow-Following Icosahedral Model (FIM); Bleck et al. 2015] to a full-fledged three-dimensional ocean model [an icosahedral-grid rendition of the Hybrid Coordinate Ocean Model (iHYCOM); cf. Bleck 2002].

Two relevant aspects of the ocean model used here are that it has the same icosahedron-based quasiuniform horizontal grid as FIM, thereby allowing for perfect local conservation of air-sea fluxes and completely avoiding coastline discrepancies between atmosphere and ocean, and that it employs a similar kind of quasi-Lagrangian vertical coordinate as FIM.

The coupled FIM-iHYCOM system is currently providing real-time subseasonal forecasts out to 4 weeks to NOAA's Subseasonal Experiment (SubX; NOAA 2017); see section 2 below. There has already been some research into subseasonal prediction with this model. Green et al. (2017) found that a version of FIM-iHYCOM using an older variant of the Grell and Freitas (2014) convective scheme yielded similar predictive skill as CFSv2 in two bivariate indices representing the Madden-Julian oscillation (MJO; Madden and Julian 1971, 1972) and smaller errors than CFSv2 in zonal winds at 850 and 200 hPa in the tropics. The MJO is responsible for most of the intraseasonal (30-90 day) tropical variability (Zhang 2005). It impacts the entire Earth system (Zhang 2013) and thus is seen as a key to subseasonal predictive skill.

A comprehensive evaluation of the overall performance of FIM-iHYCOM is necessary to determine its long-term prospects as a competitive subseasonal prediction model. Such an evaluation is presented in two articles: this study (Part I) focuses on systematic errors of the model compared to CFSv2. Sun et al. (2018, hereafter Part II) evaluates forecast skill, from both FIM-iHYCOM and CFSv2, of various subseasonal phenomena and introduces the methodology for the associated necessary model bias correction. The remainder of Part I is organized as follows. Section 2 describes the key features of the FIM-iHYCOM version participating in SubX, along with data processing methods for this model, CFSv2, and reanalysis products. Section 3 shows model biases for selected fields (including SST, precipitation, and 2-m temperature) as a function of forecast lead time. A discussion and conclusions are provided in section 4.

2. Data and experimental methods

a. Model descriptions

1) FIM-IHYCOM

FIM (iHYCOM) employs a vertical coordinate that transitions from terrain following in the lower atmosphere (isobaric in the near-surface ocean) to one that adheres to surfaces of potential temperature (potential density). In the absence of diabatic forcing, isentropic (isopycnic) surfaces are approximately material, hence the term "flow following." The transition between the different coordinate types is gradual in space and time; in a nutshell, each coordinate layer perpetually attempts to follow a "target" isentrope (isopycnal), but in doing so is subject to minimum layer thickness constraints.

This article and Part II describe results from the FIMiHYCOM version used in NOAA's SubX (NOAA 2017); this version has been referred to as "FIMr1.1" to denote release version 1.1. Many of the features of FIMr1.1 are identical to those of the version detailed in Green et al. (2017). Specifically, the hindcast consists of a four-member time-lagged ensemble initialized at 1200 and 1800 UTC every Tuesday, plus 0000 and 0600 UTC every Wednesday for the period 1999–2014. The horizontal resolution is ~60 km for both FIM and iHYCOM. FIM has 64 vertical layers, whereas iHYCOM has 32.

All of the atmospheric model physics, except for convective parameterization, come from the 2015 Global Forecast System (GFS) physics package, as described in Bleck et al. (2015). There is no stochastic treatment of any aspect of model physics. Parameterization of shallow, midlevel, and deep convection is all done through a version of the Grell and Freitas (2014) scheme (hereafter "GF") from September 2016. This is a newer version than that used in Green et al. (2017), where an older version of GF deep convection and the simplified Arakawa-Schubert (SAS; Han and Pan 2011) scheme (for shallow convection) found in the 2015 GFS physics were used, and no midlevel convection was included. The only other noteworthy change to FIM-iHYCOM between Green et al. (2017) and FIMr1.1 is a reduction of the eddy viscosity coefficient, which had a slight but beneficial effect on the simulated frequency of midlatitude blocking (not shown).

All components of FIM-iHYCOM (atmosphere, land surface, ocean, and sea ice) are initialized from the CFS Reanalysis [CFSR; Saha et al. (2010, hereafter S10)] for January 1999 through March 2011 and from the CFSv2 operational analyses (S14) for April 2011 through December 2014.¹ The methods for horizontal and vertical interpolation to the FIM-iHYCOM grid are described in Bleck et al. (2010, 2015). Soil temperature and soil moisture are specified at the nearest grid point from the CFSR data. However, it is important to note that FIM-iHYCOM uses the Moderate Resolution Imaging Spectroradiometer (MODIS)-derived classification of land-surface properties, whereas CFSR (and CFSv2) use a different land-surface classification (S10, p. 1038; S14) based on the United States Geological Survey land use; see Smirnova et al. (2016) for a comparison of different land-surface classifications within short-term regional NWP modeling. More details of the iHYCOM ocean model, sea ice prediction, and FIM-iHYCOM's freshwater budget closure are provided in appendixes A, B, and C, respectively.

2) CFSv2

Simulations from CFSv2 were also examined to compare with FIM-iHYCOM. As detailed in S14, CFSv2 is a coupled system. The hydrostatic atmospheric model is run at ~100-km resolution with 64 hybrid vertical layers, while the ocean is simulated with version 4 of the Modular Ocean Model on a different horizontal grid finer than the atmospheric model (ocean resolution of 0.5° or higher; S10, 1031–1032). Recall that this contrasts with FIM-iHYCOM, which is laid out on matching atmospheric and oceanic horizontal grids. CFSv2 employs a cycling data assimilation scheme, namely, CFSR (S10). Because FIM-iHYCOM also uses these same initial conditions, subject to the unavoidable horizontal and vertical remapping, it is convenient to perform verification against CFSR-despite the fact that CFSv2, in contrast to FIM-iHYCOM, has (nearly) identical model physics to CFSR (cf. S14, p. 2187).

b. Verification and postprocessing

1) COMMON DATA PERIOD FOR BOTH MODELS

We use (re)forecasts from FIM–iHYCOM and CFSv2 performed over identical time periods, at the same initialization frequency, and with the same number of ensemble members in order to facilitate a controlled comparison. Since many of the CFSv2 hindcasts

¹For simplicity, both CFS Reanalysis and CFSv2 operational analyses will be referred to as CFSR henceforth.

with initial dates from 1 January to 31 March 2011 are partially or completely missing from public NOAA data sources, it was decided that for this article, two common forecast initialization periods be used: January 1999–December 2010 and April 2011–December 2014 (i.e., FIM–iHYCOM reforecasts from January through March 2011 were not used²). Therefore, there are 14 complete boreal winter [December–February (DJF)] "targeted" verifications from 1999/2000 to 2009/10 and 2011/12 to 2013/14 and 16 complete boreal summer [June–August (JJA)] verifications used in this article.

Following, in general, the SubX protocol, the reforecast schedule for FIM–iHYCOM (four times weekly) introduced additional limitations to a common data period with CFSv2 (scheduled at least four times daily). Thus, as in Green et al. (2017), only a subset of CFSv2 data is considered here: the 1200 and 1800 UTC Tuesday, plus 0000 and 0600 UTC Wednesday, control initializations.

2) VERIFICATION

As noted above, because both FIM-iHYCOM and CFSv2 use the same CFSR initial conditions, it is convenient to verify all reforecasts against CFSR. It could be argued that this provides an unfair advantage to CFSv2 because CFSR is derived from that model's cycled data assimilation, which uses the same model physics including for convection. There is some validity to this argumentespecially for very short forecast lead times (i.e., a few days)-because analyses of many fields, including precipitation and near-surface temperature, depend on parameterization schemes, especially over oceans and other data-sparse regions. Nevertheless, only performing verification against an independent reanalysis, such as ERA-Interim (Dee et al. 2011), is not without its own limitations. For example, different surface layer physics in ERA-Interim can yield 2-m temperature analyses that are noticeably different from CFSR in certain places such as Antarctica-to the point where both FIM-iHYCOM and CFSv2 forecasts consistently have much larger "errors" when verified against ERA-Interim rather than CFSR (not shown). We do not claim that CFSR provides analyses closer to the "truth" than ERA-Interim; rather, verification against CFSR gives a much better sense of how the climatologies of the models drift away from the climatology of the initial conditions. Thus, most of the results presented in this article and in Part II focus on verification against CFSR, although other analysis products are used on occasion (discussed as appropriate).

Following widespread practice for subseasonal forecasts (e.g., Vitart 2004; Zhu et al. 2014; Lin et al. 2016), verification in this article is based on weekly averages as functions of both forecast lead week and target season. Forecast lead week is straightforward, given that the common period for FIM–iHYCOM and CFSv2 described above has initializations on a weekly basis centered on midweek. Thus, lead day 1 is defined as the first Wednesday, and lead day 7 is the following Tuesday; lead week 1 is the average of lead days 1–7 (Wednesday– Tuesday). Lead weeks 2, 3, and 4 are also Wednesday– Tuesday averages. For verification, Wednesday–Tuesday averages of daily analysis fields are taken to get weekly analyses.

Compositing verification by target season is best illustrated by example. Consider the target season to be DJF. For the forecasts starting on Wednesday, 14 November 2012, the midweek dates of weeks 1, 2, 3, and 4 are, respectively, the Saturdays of 17 November, 24 November, 1 December, and 8 December. Because the Saturdays of weeks 3 and 4 fall into December, these forecasts are placed into the week 3 and 4 bins of DJF, whereas the week 1 and 2 forecasts are not binned into DJF. By performing this process over 14 DJF seasons, a fairly large set of forecasts with lead weeks 1-4 that verify in DJF can be obtained. The result is that the DJF seasonal climatology from analyses can be compared with the DJF seasonal climatologies for both FIM-iHYCOM and CFSv2 for forecast lead weeks 1 through 4. As expected (and shown later in section 3), differences between the DJF analysis climatology and the weekly DJF forecast climatologies grow as a function of forecast lead week.

3) MODEL POSTPROCESSING

As mentioned earlier, the common initialization times for FIM–iHYCOM and CFSv2 are 1200 and 1800 UTC every Tuesday, plus 0000 and 0600 UTC every Wednesday. The NOAA SubX protocol requires data to be provided as daily averages from 0000 to 2359 UTC (i.e., an effective 24-h period that does not use 0000 UTC from 2 consecutive days). FIM–iHYCOM was configured to output daily averages starting at every 0000 UTC regardless of the model initialization time;³ this is different from the hindcast dataset used by Green et al. (2017), in which output was saved every 6h. A

² The only exception is for the hindcast used to build the FIM–iHYCOM model climatology, in which the full 16-yr period was included (see Part II).

³ It was decided that the lead day 1 forecast from the 0000 UTC Wednesday initialization (m03 in Fig. 1) be copied to the 0600 UTC Wednesday initialization (m04 in Fig. 1) because m04 does not have any information for the first 6 h of day 1, and the resulting discrepancy between lead days 1 and 2 can be ignored in the context of subseasonal prediction.



FIG. 1. Schematic of the weekly four-member time-lagged ensemble for FIM-iHYCOM: "m01," "m02," "m03," and "m04" are initialized at 1200 UTC Tuesday, 1800 UTC Tuesday, 0000 UTC Wednesday, and 0600 UTC Wednesday, respectively. Note m01 and m02 discard the first 12 and 6 h, respectively, and begin archiving at 24-h intervals starting at 0000 UTC Wednesday. As described in footnote 3, the day 1 output from m03 is copied over to m04 (green arrows) because m04 is missing data for the first 6 h of the initial Wednesday.

schematic is shown in Fig. 1. While this archiving strategy may not be ideal for a time-lagged ensemble, the focus on weekly averages means that there will only be a minimal impact on the week 1 forecast, especially when the four-member time-lagged ensemble average is considered, as is done for the entirety of this article. Moreover, for consistency with FIM–iHYCOM, the same archiving procedure (Fig. 1) was performed for the CFSv2 dataset. Finally, following the SubX protocol (cf. Pegion 2017), all data sources were converted to a common $1^{\circ} \times 1^{\circ}$ grid on which all verification was performed.

3. Seasonal biases

Most of the results presented here are for different lead weeks with a target season of DJF; these results, especially when compared with JJA target season results, are often more meaningful than showing annual biases because seasonal biases can be of opposite sign and, thus, cancel out over the whole year. Here, the focus is on fields that are important for subseasonal forecasting, namely, SST, 2-m temperature (hereafter T2m), and precipitation. Moreover, at subseasonal time scales, there can be global influences on regional weather, so even though the focus of SubX is on North America, it is imperative to investigate model biases worldwide.

a. Boreal winter

1) LARGE-SCALE OVERVIEW

SST, whose daily average evolves slowly relative to most atmospheric fields, can be quite important for subseasonal prediction: for example, Green et al. (2017) showed that running FIM in atmosphere-only mode (i.e., uncoupled to iHYCOM, albeit with slowly evolving SST based on monthly observations) yielded significantly worse predictive skill of two MJO indices. A multitude of other studies (e.g., Woolnough et al. 2007; Pegion and Kirtman 2008; Kim et al. 2008, 2010; Fu et al. 2013; Seo et al. 2014; Klingaman and Woolnough 2014; DeMott et al. 2015) have investigated the role of air-sea coupling on the MJO and its prediction skill. Figure 2 shows SST biases (verified against CFSR) for the DJF target season at forecast lead weeks 1 and 4 for both FIM-iHYCOM and CFSv2. The Arctic Ocean, generally speaking, is completely frozen over with sea ice during this period. For areas with sea ice, SST is actually a surface skin temperature for CFSR and both models. Regardless, in sea ice areas, there are noticeable discrepancies in skin temperature between CFSR and both models, but particularly FIM-iHYCOM (not shown, although signs of this issue are evident in Fig. 2 just north of the Aleutian Islands, as well as in Hudson Bay). S14 describe in detail issues with sea ice, albeit with a focus on boreal summer, in both CFSv2 and its initial conditions. As discussed in appendix B, iHYCOM has a simplistic single-layer sea ice model in which skin temperature during the freezing season is diagnosed by assuming zero vertical heat flux divergence at the air-ice interface. The primary cause of the warm ice temperature bias in FIM-iHYCOM is the use of ice thermal conductivity throughout the frozen layer without any consideration of a low-conductivity snow layer on top of the ice; this was addressed after the release of FIMr1.1 (appendix B). But for FIMr1.1, which is the version of the model used here and in SubX, all that can be done is to note sea ice as a known issue.

In areas without sea ice and outside southern oceans adjacent to Antarctica, DJF SST biases are different in spatial distribution and sign, but similar in overall magnitude, for FIM-iHYCOM and CFSv2. Overall, the tropical and subtropical SST biases by week 4 are cold in FIM-iHYCOM and warm in CFSv2, especially in the marine stratus regions off the west coasts of South America and southern Africa. This warm SST bias in CFSv2 is also evident in the annual bias plot for 3-month lead times in S14 (bottom right of their Fig. 9). Both models have a warm SST bias around the Maritime Continent-more so with CFSv2. Because of the nonlinear interactions between air-sea fluxes and convection, errors in both the mean (and variability) of SST could have implications for relative MJO forecast skill (e.g., Green et al. 2017).

T2m biases for DJF are shown in Fig. 3; here, the bottom row (Figs. 3e,f) shows the difference between model weeks 4 and 1 to better illustrate model drift. Some of the largest biases in these figures are in areas with high and steep terrain (such as the Himalayas and Andes); these



FIG. 2. SST biases (K) for target season DJF for (a),(c) FIM–iHYCOM and (b),(d) CFSv2, both verified against CFS (re)analysis. (top) Forecast lead week 1 bias; (bottom) forecast lead week 4 bias.

locally large biases—even in week 1—are likely a consequence of CFSR having a higher native resolution [T382 (~38 km); S10] than both FIM–iHYCOM and CFSv2. Globally averaged temperature biases for all points, and for land points only, are shown in Table 1. The strong relationship between biases in SST and T2m over oceans (cf. Fig. 3 with Fig. 2) is not surprising. Over land, there is an early and persistent warm bias in CFSv2 over the Sahara and Arabian Deserts and much of Asia. In contrast, FIM shows little T2m bias in week 1, trending to a cold bias by week 4 (Table 1). This week 4 cold bias in FIM extends across the entire Northern Hemisphere landmass, except for Europe, where there is a warm bias. In general, however, the trend from week 1 to week 4 in both models is cooling over landmasses (Figs. 3e,f; Table 1).

Biases and drift in DJF precipitation are shown in Fig. 4. The right column of this figure shows results from a separate experiment: FIM/SAS, which is



FIG. 3. As in Fig. 2, but for T2m biases (K). Bottom panel shows differences between forecast lead weeks 4 and 1.

TABLE 1. Weekly averaged global T2m biases (verified against CFSR, in K) for FIM-iHYCOM and CFSv2 forecasts [combined land

and ocean outside parentheses; land-only points inside parentheses] verifying in the DJF target season [CFSR climatology: 286.02 K (276.68 K)] and the JJA target season [CFSR climatology: 289.51 K (287.71 K)].

 DJF
 JJA

	DJF		J	JA
	FIM	CFSv2	FIM	CFSv2
Lead week 1	-0.01(-0.18)	0.20 (0.43)	-0.15 (0.01)	0.12 (0.15)
Lead week 2	-0.13(-0.51)	0.05 (0.12)	-0.29(0.04)	0.02(-0.02)
Lead week 3	-0.20(-0.70)	-0.05(-0.12)	-0.34(0.05)	-0.01(-0.08)
Lead week 4	-0.23 (-0.81)	-0.10 (-0.27)	-0.37 (0.05)	-0.25 (-0.13)

identical to FIMr1.1 (the configuration for SubX and the primary focus of these articles), except that convection was parameterized using SAS from the 2015 GFS physics. It was decided that analyzed precipitation from the Global Precipitation Climatology Project (GPCP; Huffman et al. 2001) would be used for verification purposes rather than CFSR. This decision was made because CFSR is a cycling data assimilation system, and the resultant precipitation is strongly influenced by the imbalances introduced by assimilation. Issues associated with CFSR precipitation are also described in Zhang et al. (2012). Tests that used CFSR precipitation for verification (not shown) do not change our most general conclusions about the performance of FIM (both FIMr1.1 and FIM/SAS) and CFSv2 precipitation forecasts. Looking at Fig. 4, FIMr1.1 appears to be verifying closer to GPCP than CFSv2 does, and has less drift from week 1 to week 4. Both results are somewhat surprising because CFSv2 has cycling data assimilation (cf. Fig. 1 of S10), whereas FIMr1.1 is "cold started" (initial conditions taken directly from CFSR with no FIM-native data assimilation) and uses a different convection scheme (GF) than the model providing the initial conditions (SAS in CFS). While FIMr1.1 precipitation is far from perfect, it does not have nearly the same magnitude of biases across much of the Pacific Ocean or the substantial wet biases over the northern half of the Western Hemisphere that CFSv2 does. All three runs have a global area-averaged wet bias (Table 2) relative to GPCP: FIMr1.1 has the smallest biases, followed by FIM/SAS, with CFSv2 exhibiting the largest biases. From Fig. 4, the drift in FIMr1.1 precipitation (week 1 to 4) is smaller than that of CFSv2, which is consistent with the global area-weighted averages shown in Table 2: the DJF global precipitation increase in FIMr1.1 is ~ 0.08 mm day⁻¹, compared with an increase of ~ 0.15 mm day⁻¹ for CFSv2. The drift of FIM/SAS is even smaller ($\sim 0.03 \text{ mm day}^{-1}$).



FIG. 4. As in Fig. 3, but for precipitation verified against GPCP (mm day⁻¹). Note that the right column contains results from the FIM/SAS experiment described in the text.

TABLE 2. As in Table 1, but for global (land and ocean) precipitation biases (mm day⁻¹) verified against GPCP (DJF climatology: 2.65 mm day^{-1} ; JJA climatology: 2.69 mm day^{-1}). The FIM/SAS columns are based on a parallel experiment identical to FIMr1.1 [with Grell and Freitas (2014)-based convection as described in the text], but with convection parameterized instead by the simplified Arakawa–Schubert scheme.

	DJF			JJA		
	FIMr1.1	CFSv2	FIM/SAS	FIMr1.1	CFSv2	FIM/SAS
Lead week 1	0.29	0.40	0.41	0.36	0.52	0.50
Lead week 2	0.33	0.47	0.44	0.40	0.58	0.51
Lead week 3	0.36	0.53	0.45	0.42	0.62	0.51
Lead week 4	0.37	0.55	0.44	0.41	0.63	0.50

So far, no momentum or mass fields (i.e., circulationrelated fields) have been examined. Figure 5 shows the DJF biases in 500-hPa geopotential height (H500) poleward of 20°N for weeks 1 and 4. In week 1, CFSv2 has larger areas of bias. By week 4, both models have an overall negative height bias in the Northern Hemisphere, somewhat stronger in CFSv2. These may be related to the cold bias in these regions (cf. Fig. 3). FIM-iHYCOM has strong negative height biases in the general areas of the Aleutian and Icelandic lows; CFSv2 also has a fairly strong negative height bias between Alaska and Hawaii, albeit centered at a lower latitude. These biases have implications for both the predictive skill and relative frequency of tropospheric blocking events (D'Andrea et al. 1998). A more thorough investigation of blocking-which can impact weather on subseasonal time scales (e.g., Matsueda 2011)—is left for the Part II companion paper.

2) ERRORS OVER NORTH AMERICA

Given the focus of SubX on subseasonal prediction over North America, it is worthwhile to show in more detail the systematic errors in T2m and precipitation over this region (Figs. 6, 7, respectively). By week 4, both FIM-iHYCOM and CFSv2 show a cold bias for DJF. FIM-iHYCOM has a somewhat smaller drift from week 1 to 4, whereas CFSv2 shows a strong trend toward colder conditions. Because of a mistreatment of the land-surface type in the version of FIM-iHYCOM (FIMr1.1) used for SubX (real-time and retrospective) and for these articles, large lakes, including the Great Lakes, were erroneously classified as dry land. Consequently, the Great Lakes are much too cold in the winter and too warm in the summer. Although the inconsistency has since been removed, we did not deem this error to be substantial enough to warrant an immediate rerun of the entire 16-yr hindcast, especially before the start of SubX in July 2017. Moreover, systematic bias correction will remove most of the direct local impact of the erroneous classification, although it is likely that there are some secondary impacts that cannot be removed. For the time being, this will be documented as a known issue for SubX users.

For DJF precipitation biases over North America (Fig. 7), FIMr1.1 generally has smaller errors both in magnitude and spatial coverage by week 4 than CFSv2. In FIMr1.1, there is a dry bias in both weeks 1 and 4 in the central United States, a wet bias in Central America (Figs. 7a,c), and an increasingly wet bias over the southeastern United States and nearby bodies of water (Fig. 7e). CFSv2, by contrast, has a strong and growing wet bias over Mexico, extending into much of the United States and the western Caribbean. Again, in terms of precipitation, FIMr1.1 is closer to GPCP and exhibits less change as a function of lead time than does CFSv2.

b. Boreal summer

We next present biases during the boreal summer (JJA) but with less detail than for DJF. The boreal summer SST biases (Fig. 8) are slight only for week 1. By week 4, there are substantial warm biases in CFSv2 over the marine stratus regions west of the Americas, as well as in the North Pacific and Atlantic regions and near the Maritime Continent. FIM–iHYCOM shows some cold SST biases in week 4 equatorward of \sim 45° latitude in both hemispheres.

Figures 9 and 10 focus on T2m and precipitation biases over most of the globe (following Figs. 3, 4) for JJA. For week 1, CFSv2 again has a warm bias over northern Africa and the Arabian Peninsula; FIM has a warm bias in the Middle East and a cold bias in South America. The trend (week 1 to 4) in T2m bias (Figs. 9c,e) for JJA is opposite that of DJF in the Northern Hemisphere (Figs. 3c,e), with cooling over western Europe and warming over other landmasses. Over land, CFSv2 has a cooling trend in JJA (Table 1), which is most noticeable in the Southern Hemisphere and Africa. For both models, the SST biases shown in Fig. 8 begin to appear in T2m over the ocean by week 4. The inland lake problem with FIM–iHYCOM discussed earlier is also evident (best seen over the Caspian Sea in Figs. 9a,c) but does



FIG. 5. As in Fig. 2, but for H500 biases (m) poleward of 20°N latitude.

not seem to contribute to other errors in T2m or precipitation. Overall, the T2m biases in week 4 for both models are generally similar to those in week 1, but stronger in magnitude.

For JJA precipitation, FIM–iHYCOM (both FIMr1.1 and FIM/SAS) again appears to be less biased than CFSv2 (Fig. 10). CFSv2 shows a strong moist bias, compared to GPCP, with excessive precipitation in the intertropical convergence zone (week 1) and the equatorial Indian Ocean (week 4). Both FIM–iHYCOM runs have similar spatial distributions of wet/dry bias as CFSv2, but not nearly of the same magnitude (particularly for

FIMr1.1). This is reflected in the area-averaged precipitation shown in Table 2: while all three models have a wet bias compared to GPCP in JJA, that of CFSv2 is the largest and also increases much more between lead week 1 and lead week 4. Again, FIMr1.1 has a smaller wet bias than FIM/SAS.

The results in section 3 overall indicate that the biases in FIM-iHYCOM can be considered comparable to those of CFSv2: neither model has uniformly smaller biases for all fields examined here. FIM-iHYCOM—particularly FIMr1.1—shows an advantage over CFSv2 in precipitation bias (Figs. 4, 7, 10), T2m, DJF



FIG. 0. As in Fig. 5, but focused over North Americ

despite the fact that FIM-iHYCOM is "cold started" and FIMr1.1 uses a different convection scheme than the model providing initial conditions. Some of the improvement in precipitation biases comes from the use of the Grell and Freitas (2014) scale-aware convective parameterization, as evidenced in the FIM/SAS test (Figs. 4, 10; Table 2). FIM-iHYCOM does have larger SST biases than CFSv2 in the Southern Ocean near Antarctica (Figs. 2, 8), an issue that will be addressed in future studies investigating cloudiness and oceanic mixed layer parameterizations. Over the cold current areas where marine stratus dominates (i.e., west of Africa and the Americas), FIM-iHYCOM has improved SST and T2m results (i.e., smaller biases), compared to CFSv2.

4. Discussion and conclusions

This is the first of two articles on subseasonal prediction in the FIM-iHYCOM coupled modeling system. Part I (this article) introduces the coupled model and focuses on systematic biases in some key fields that are then compared with NOAA's operational CFSv2 model. The origins of this new coupled model are in the atmospheric model that was developed for mediumrange NWP, meaning that for subseasonal purposes, a "weather up" rather than a "climate down" approach was taken. This article also describes the icosahedral version of the HYCOM ocean model used in the coupled system. FIM-iHYCOM has strong potential to provide diversity to subseasonal forecasts, as it is





FIG. 7. As in Fig. 6, but for precipitation biases (mm day $^{-1}$).

endowed with attributes not found in other modeling systems used for subseasonal to seasonal forecasting: (i) a different parameterization for convection (following Grell and Freitas 2014) and a different land classification; (ii) finite-volume as opposed to finite-difference or spectral methods; (iii) quasi-uniform grids that, despite requiring indirect addressing, have tolerable computational overhead and can be efficiently parallelized on distributed-memory machines; (iv) perfectly matched atmosphere–ocean horizontal grids that allow for a consistent coastline and flawless local and global conservation of fluxes; and (v) a near-Lagrangian ("flow following") vertical coordinate in both atmosphere and ocean. Given the importance of convection, especially in the tropics, on the general circulation, using a variant of Grell and Freitas (2014) convection will be key to FIM-iHYCOM's role in adding diversity to subseasonal prediction efforts.

Despite these inherent differences, the current version of FIM-iHYCOM has systematic errors that are comparable to those of CFSv2. "Comparable" is used in the broadest sense: there are some fields—namely, precipitation—in which FIM-iHYCOM (FIMr1.1) exhibits larger areas of near-zero bias than CFSv2. One area in which FIM-iHYCOM has larger systematic errors than CFSv2 is in skin temperature over oceans both for the skin temperature of sea ice and for SST in open waters. The errors over sea ice are primarily



FIG. 8. As in Fig. 2, but for a target season of JJA.

related to how the single-layer sea ice model employed by iHYCOM (appendix B) diagnoses skin temperature, rather than errors in sea ice extent (not shown). The open-water SST biases also warrant more attention during ongoing and future model investigations.

Overall, systematic errors can be removed through model bias correction. The method we chose for calculating the model climatology needed for bias correction is described in Part II; it is used extensively there for evaluations of model skill, and for an assessment of model variability. Future work includes continued coupled model development with a focus on the ocean and sea ice components using the NOAA Next Generation Global Prediction System framework. In addition, advanced physics packages are currently being tested for future seamless weather-to-climate prediction. The model development team is also working with the broader research community to examine physical processes critical to subseasonal prediction, including, but not limited to, the MJO, blocking, and sudden stratospheric warming. Finally, by making FIMr1.1 hindcasts



FIG. 9. As in Fig. 3, but for a target season of JJA.

Precip, JJA



FIG. 10. As in Fig. 4, but for a target season of JJA.

available to the subseasonal research community through SubX, the benefits of including this model in a multimodel ensemble can be determined.

Acknowledgments. This project was supported by NOAA OAR funding for week 3-4 forecast improvement and the Earth System Prediction Capability program. Coauthors Sun, Bleck, and Green are also supported by funding from NOAA OAR Award NA17OAR4320101. The authors acknowledge the NOAA Research and Development High Performance Computing Program for providing computing and storage resources that have contributed to the research results reported within this article (http:// rdhpcs.noaa.gov), as well as the Texas Advanced Computing Center (TACC) at The University of Texas at Austin (http://www.tacc.utexas.edu). CFSv2 and CFSR data are obtained from the National Centers for Environmental Information. Drs. John Brown and George Kiladis are thanked for providing internal review, and Dr. Kathy Pegion provided helpful suggestions. Three anonymous reviewers provided extensive helpful feedback on an earlier version of this work.

APPENDIX A

The Ocean Model

Like HYCOM (Bleck 2002), iHYCOM is designed to capture the most relevant dynamic processes affecting the global SST on all time scales, such as wind-forced gyre circulations and boundary currents, sea ice formation, thermally and mechanically forced mixed layer entrainment/detrainment, small-scale diapycnal mixing, and thermohaline-forced meridional overturning. Excluded, for the time being, are dynamic ice spreading, tidal effects, and a wave submodel for surface roughness prediction.

a. Features common to HYCOM and iHYCOM

Column physics, including the grid generator that controls vertical placement of grid points, remain unchanged from HYCOM. The two mutually consistent assumptions inherent in the HYCOM design (Spiegel and Veronis 1960)—(i) use of potential density as the primary buoyancy variable and (ii) seawater incompressibility—also carry over into iHYCOM. Potential density is referenced to 1-km depth in the current version of iHYCOM.

b. Variable staggering, time step, and treatment of barotropic mode

For historical reasons, FIM uses an unstaggered horizontal mesh, generally referred to as the Arakawa A grid, and this design is carried over into iHYCOM.

The third-order Adams–Bashforth scheme successfully employed in FIM (Lee and MacDonald 2009) failed in iHYCOM because it does not permit rigorous enforcement of positive definitiveness in the layer thickness tendency equation. For this reason, iHYCOM has inherited from HYCOM the traditional leapfrog time differencing scheme. The reasons why multistep Adams–Bashforth time differencing works in FIM but not in iHYCOM may be threefold:

- 1) Oceanic orography is steeper than terrestrial orography.
- The ocean is less stratified and therefore subject to more "sloshing" by large-amplitude internal gravity waves.
- 3) Beyond the shelf break, iHYCOM does not use terrain-following coordinates as FIM does; hence, massless layers are a standard feature on the sea floor, where they pose a particular numerical challenge due to the steepness of the bottom slopes.

HYCOM, like many other ocean models, gains efficiency by separating barotropic gravity waves from other types of fluid motion and transmitting them using a numerically efficient two-dimensional shallow-water model (Bleck and Smith 1990). Such split-explicit schemes are borderline unstable (Morel et al. 2008). One particular disadvantage of the split-explicit scheme in HYCOM is that it does not allow changes in bottom pressure. This forces HYCOM users to express surface freshwater fluxes as virtual salinity fluxes (see appendix B). By shedding this restriction, iHYCOM is able to explicitly simulate the equatorward mass flux resulting from the poleward atmospheric moisture flux (Huang 1993).

Despite the absence of a mode splitting scheme, iHYCOM retains some split-explicit flavor. The 3D momentum and continuity equations are solved on a short "barotropic" time step. Mass fluxes from this integration are summed up and used to advect tracers (including temperature and salinity) over much longer "baroclinic" time intervals. Column physics routines, likewise, are called on the longer time steps.

c. Continuity equation

The flux-corrected transport algorithm from FIM (Bleck et al. 2015) is also used in iHYCOM. Coastlines are defined to always coincide with gridcell boundaries, providing clean conditions for global mass conservation. When designing the land–sea mask on the shared horizontal grid, special attention has been paid to the proper rendition of narrow passages and isthmuses affecting oceanic flow patterns.

Bolus fluxes, needed to account for the role of subgrid-scale ocean eddies, are based on the slope of layer interfaces as they are in HYCOM. Specifically, "pressure fluxes" proportional to the pressure differential between adjacent grid points are computed on each layer interface. The difference of pressure fluxes on the upper and lower interface of each layer then defines the layer bolus flux. Pressure fluxes are oriented such that were they to act in isolation, they would flatten the interface they reside on. A flux-limiting process assures that bolus fluxes do not cause interfaces to intertwine or intersect the sea floor.

d. Momentum equation

On unstructured grids, evaluation of the horizontal pressure gradient force

$$\alpha \nabla p + \nabla \phi = \nabla M - p \nabla \alpha$$

(where α , p, ϕ , and M stand for potential specific volume, pressure, geopotential, and Montgomery potential, respectively) requires line integrations along the perimeter of each grid cell. Values of the integrand (M and α in this case) on cell edges must be interpolated from cells in the vicinity of each edge segment. If a grid cell happens to abut the coastline, we assign to the inland "ghost" point the value of the integrand in the offshore cell. The same procedure is followed if the neighboring cell is a massless cell on the sea floor.

Sidewall drag is evaluated by viewing the ocean bottom as an assemblage of hexagonal basalt-like columns of various heights. Consider a single coordinate layer in two adjacent grid columns. If the layer in column 1 overlaps the "basalt" slab at the bottom of column 2, the velocity vector in column 2 used for sidewall drag calculation in column 1 is replaced by a linearly weighted average of the above-bottom velocity from column 2 and a mirror image of the velocity from column 1. (A sign change may be involved, depending on the choice of free- or no-slip boundary conditions.) Vertically prorating the effect of sidewalls in this fashion avoids temporal discontinuities in sidewall drag in layers subjected to gravity wave sloshing near steep bottom slopes.

e. Conservation

Conservation of mass and tracers is paramount in circulation models used in longer-term simulations. For this reason, it is essential for layer models to solve conservation equations in flux form. Even so, conservation of tracers is difficult to enforce in situations where a grid cell loses most (but not all) of its mass during a single time step.

The problem is caused by the need to divide tracer *amount* (e.g., salt in a grid cell), which is the predictand in the conservation equations, by layer thickness to recover tracer *concentration* (e.g., salinity) after the transport step. As layer thickness approaches zero, meaning that the operation moves into the vicinity of the zero-over-zero singularity, the resulting concentration value may lie outside the proper bounds, especially if mass export is large in relation to the mass remaining in the grid cell.

As discussed in Bleck et al. (2010), the usual remedy is to introduce upper and lower bounds on the result of the division and to compensate for the implied nonconservation by adding an appropriate global offset to the tracer field. Note that nonconservation ceases to be an issue if a grid cell loses 100% of its mass—rather than, say, 99%.

No such remedies are needed in the case of mass. Mass is conserved in the model to round off error limits.

APPENDIX B

The Sea Ice Model

The sea ice model presently employed in FIM–iHYCOM is a very basic, thermodynamic single-layer energy-loan model. In banking terms, the model "lends" energy to the ocean— accepting ice as collateral—to keep SST from dropping below the freezing level $T_{\rm frz}$. At any given location, the "loan" must be paid back in full during the warm season before SST at that location is allowed to rise above $T_{\rm frz}$.

Ice formed this way gets advected by surface currents. There is no ice rheology per se, but a shaving device prevents excessive ice buildup by laterally spreading any amount that exceeds a prescribed thickness.

The second task of this model, aside from serving as an energy bank, is to generate an ice surface temperature $(T_{\rm skin})$ for atmospheric heat flux calculations. During times when the ocean loses heat (atmospheric heat flux $F_{\rm atm} < 0$), this is done, as described in Bleck (2002), by assuming that the ice–internal vertical heat flux expressed as top-to-bottom temperature gradient times thermal conductivity matches $F_{\rm atm}$. The physical reasoning here is that during the freezing season, no latent heat sources or sinks are available at the (ostensibly) dry ice surface to justify a heat flux discontinuity.

The above algorithm causes $T_{\rm skin}$ to approach $T_{\rm frz}$ as $F_{\rm atm} \rightarrow 0$, and it would yield $T_{\rm skin} > T_{\rm frz}$ when $F_{\rm atm} > 0$. In the latter case, the model sets $T_{\rm skin} = T_{\rm frz}$. The resulting flux discontinuity (zero ice–internal heat flux vs nonzero $F_{\rm atm}$) causes melting.

The model, which has been dubbed ENLOAN in deference to the energy loan concept, does not carry a separate snow layer. This limits the fidelity of the internal heat flux estimate, given the large difference between the thermal conductivities of snow and ice. An effort is made to account for this difference by assuming that a fixed portion of the total frozen layer is made up of snow. If κ_{ice} , κ_{sno} are the thermal conductivities of ice

and snow, respectively, and h_{ice} , h_{sno} are the corresponding ice and snow depths, the effective conductivity of the total frozen layer can be shown to be

$$\kappa_{\rm eff} = \kappa_{\rm ice} \kappa_{\rm sno} \frac{h_{\rm ice} + h_{\rm sno}}{h_{\rm ice} \kappa_{\rm sno} + h_{\rm sno} \kappa_{\rm ice}}.$$

Reasonable values of the four parameters involved yield an effective conductivity that is 30%–50% lower than κ_{icc} . It should be noted that the version of FIM–iHYCOM used for this article (also used for real-time SubX forecasts) assumes the total frozen layer to be ice (i.e., $\kappa_{eff} = \kappa_{icc}$). This assumption is responsible for the current warm bias in skin temperature over sea ice regions, which motivated the development of the equation for κ_{eff} shown above.

For historical reasons having to do with restrictions imposed by the barotropic–baroclinic mode splitting scheme in HYCOM (see Bleck 2002), melting and freezing processes do not involve an exchange of actual mass between ice and ocean; instead, they spawn virtual salt fluxes, whose purpose is to account for seawater buoyancy changes during freezing and melting. With ice salinity being lower than typical seawater salinity, ice formation/melting leads to a salinity flux into/out of the ocean. This can be thought of as a "salt loan" accompanying the energy loan.

iHYCOM assumes precipitation falling onto ice to come down in the form of snow. Since ENLOAN does not keep track of separate snow and ice amounts, lumping the two together requires that snow be made as salty as sea ice. To avoid creating an internal salt source in the process of converting fresh to salty snow, ENLOAN extracts the required salt amount (which is usually miniscule) from the underlying water column.

APPENDIX C

Freshwater Budget Closure

When designing a coupled ocean–atmosphere model for extended-range prediction purposes, there is a fine line between essential and nonessential processes to be included in the model. We argue that closure of the global freshwater budget belongs to the former group.

Only part of the water evaporating from the sea surface is returned by clouds shedding water directly over the ocean; the deficit is made up by river and glacier runoff. Lack of a terrestrial runoff scheme in a coupled model, therefore, will cause salinity to build up near the sea surface, with potential consequences for buoyancydriven circulation systems, which in turn can affect horizontal heat transport.

This issue is addressed in FIM-iHYCOM through a rudimentary river runoff scheme, in which watersheds are defined during model initialization on the basis of terrain slope. During model integration, any runoff (the portion of precipitation not absorbed by the landsurface module) accumulating in a grid cell is handed off to a lower neighbor, thereby eventually making its way to the ocean. To keep the runoff scheme local on distributed memory computers, the transfer occurs at the rate of one grid cell per time step. No attempt is made to simulate details like river flow speed or retention by reservoirs, nor are all watersheds derived from a terrain height database at 60-km resolution expected to be realistic. Furthermore, landlocked water bodies like the Caspian Sea had to be artificially elevated to provide a pathway for excess river inflow to the ocean. These issues are ignored at present, given that emphasis is entirely on closing the global freshwater budget in longer-term simulations.

REFERENCES

- Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modell.*, 4, 55– 88, https://doi.org/10.1016/S1463-5003(01)00012-9.
- —, and L. Smith, 1990: A wind-driven isopycnic coordinate model of the north and equatorial Atlantic Ocean: 1. Model development and supporting experiments. J. Geophys. Res., 95, 3273–3285, https://doi.org/10.1029/JC095iC03p03273.
- —, S. Benjamin, J. Lee, and A. E. MacDonald, 2010: On the use of an adaptive, hybrid-isentropic vertical coordinate in global atmospheric modeling. *Mon. Wea. Rev.*, **138**, 2188–2210, https://doi.org/10.1175/2009MWR3103.1.
- —, and Coauthors, 2015: A vertically flow-following icosahedral grid model for medium-range and seasonal prediction. Part I: Model description. *Mon. Wea. Rev.*, **143**, 2386–2403, https:// doi.org/10.1175/MWR-D-14-00300.1.
- Brunet, G. N., and Coauthors, 2010: Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. *Bull. Amer. Meteor. Soc.*, **91**, 1397–1406, https:// doi.org/10.1175/2010BAMS3013.1.
- D'Andrea, F., and Coauthors, 1998: Northern Hemisphere atmospheric blocking as simulated by 15 atmospheric general circulation models in the period 1979–1988. *Climate Dyn.*, 14, 385–407, https://doi.org/10.1007/s003820050230.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/ 10.1002/qj.828.
- DeMott, C. A., N. P. Klingaman, and S. J. Woolnough, 2015: Atmosphere–ocean coupled processes in the Madden–Julian oscillation. *Rev. Geophys.*, 53, 1099–1154, https://doi.org/ 10.1002/2014RG000478.
- Fu, X., J.-Y. Lee, P.-C. Hsu, H. Taniguchi, B. Wang, W. Wang, and S. Weaver, 2013: Multi-model MJO forecasting during DYNAMO/CINDY period. *Climate Dyn.*, **41**, 1067–1081, https://doi.org/10.1007/s00382-013-1859-9.
- Green, B. W., S. Sun, R. Bleck, S. G. Benjamin, and G. A. Grell, 2017: Evaluation of MJO predictive skill in multiphysics and

multimodel global ensembles. *Mon. Wea. Rev.*, **145**, 2555–2574, https://doi.org/10.1175/MWR-D-16-0419.1.

- Grell, G. A., and S. R. Freitas, 2014: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, 14, 5233–5250, https:// doi.org/10.5194/acp-14-5233-2014.
- Han, J., and H.-L. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. *Wea. Forecasting*, 26, 520–533, https://doi.org/10.1175/WAF-D-10-05038.1.
- Huang, R. X., 1993: Real freshwater flux as a natural boundary condition for the salinity balance and thermohaline circulation forced by evaporation and precipitation. *J. Phys. Oceanogr.*, 23, 2428–2446, https://doi.org/10.1175/1520-0485(1993)023<2428: RFFAAN>2.0.CO;2.
- Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. J. Hydrometeor., 2, 36–50, https://doi.org/10.1175/ 1525-7541(2001)002<0036:GPAODD>2.0.CO;2.
- Kim, H.-M., C. D. Hoyos, P. J. Webster, and I.-S. Kang, 2008: Sensitivity of MJO simulation and predictability to sea surface temperature variability. *J. Climate*, 21, 5304–5317, https://doi.org/ 10.1175/2008JCLI2078.1.
- —, —, , and —, 2010: Ocean–atmosphere coupling and the boreal winter MJO. *Climate Dyn.*, **35**, 771–784, https://doi.org/ 10.1007/s00382-009-0612-x.
- Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, https://doi.org/10.1175/ BAMS-D-12-00050.1.
- Klingaman, N. P., and S. J. Woolnough, 2014: The role of air-sea coupling in the simulation of the Madden–Julian oscillation in the Hadley Centre model. *Quart. J. Roy. Meteor. Soc.*, 140, 2272–2286, https://doi.org/10.1002/qj.2295.
- Lee, J.-L., and A. E. MacDonald, 2009: A finite-volume icosahedral shallow-water model on a local coordinate. *Mon. Wea. Rev.*, **137**, 1422–1437, https://doi.org/10.1175/ 2008MWR2639.1.
- Lin, H., N. Gagnon, S. Beauregard, R. Muncaster, M. Markovic, B. Denis, and M. Charron, 2016: GEPS-based monthly prediction at the Canadian Meteorological Centre. *Mon. Wea. Rev.*, **144**, 4867–4883, https://doi.org/10.1175/MWR-D-16-0138.1.
- Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, https://doi.org/ 10.3402/tellusa.v21i3.10086.
- MacLachlan, C., and Coauthors, 2015: Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, 141, 1072–1084, https://doi.org/10.1002/qj.2396.
- Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, 28, 702–708, https://doi.org/10.1175/1520-0469(1971)028<0702: DOADOI>2.0.CO;2.
- —, and —, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. J. Atmos. Sci., 29, 1109–1123, https://doi.org/10.1175/1520-0469(1972)029<1109: DOGSCC>2.0.CO;2.
- Mastrangelo, D., P. Malguzzi, C. Rendina, O. Drofa, and A. Buzzi, 2012: First outcomes from the CNR-ISAC monthly forecasting system. *Adv. Sci. Res.*, 8, 77–82, https://doi.org/10.5194/ asr-8-77-2012.

- Matsueda, M., 2011: Predictability of Euro-Russian blocking in summer of 2010. *Geophys. Res. Lett.*, 38, L06801, https://doi.org/ 10.1029/2010GL046557.
- Morel, Y., R. Baraille, and A. Pichon, 2008: Time splitting and linear stability of the slow part of the barotropic component. *Ocean Modell.*, 23, 73–81, https://doi.org/10.1016/j.ocemod.2008.04.001.
- NOAA, 2017: S2S Prediction Task Force: Subseasonal to Seasonal (2016–2019). NOAA Climate Program Office, accessed 4 December 2017, https://cpo.noaa.gov/Meet-the-Divisions/ Earth-System-Science-and-Modeling/MAPP/MAPP-Task-Forces/S2S-Prediction-Task-Force.
- Pegion, K., 2017: SubX protocol. Center for Ocean–Land– Atmosphere Studies, George Mason University, accessed 4 December 2017, http://cola.gmu.edu/kpegion/subx/protocol.html.
- —, and B. P. Kirtman, 2008: The impact of air–sea interactions on the simulation of tropical intraseasonal variability. *J. Climate*, 21, 6616–6635, https://doi.org/10.1175/2008JCLI2180.1.
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. Bull. Amer. Meteor. Soc., 91, 1015–1058, https:// doi.org/10.1175/2010BAMS3001.1.
- —, and Coauthors, 2014: The NCEP Climate Forecast System version 2. J. Climate, 27, 2185–2208, https://doi.org/10.1175/ JCLI-D-12-00823.1.
- Seo, H., A. C. Subramanian, A. J. Miller, and N. R. Cavanaugh, 2014: Coupled impacts of the diurnal cycle of sea surface temperature on the Madden–Julian oscillation. J. Climate, 27, 8422–8443, https://doi.org/10.1175/JCLI-D-14-00141.1.
- Smirnova, T. G., J. M. Brown, S. G. Benjamin, and J. S. Kenyon, 2016: Modifications to the rapid update cycle land surface model (RUC LSM) available in the Weather Research and Forecasting (WRF) Model. *Mon. Wea. Rev.*, **144**, 1851–1865, https://doi.org/10.1175/MWR-D-15-0198.1.
- Spiegel, E. A., and G. Veronis, 1960: On the Boussinesq approximation for a compressible fluid. Astrophys. J., 131, 442–447, https://doi.org/10.1086/146849.

- Sun, S., B. W. Green, R. Bleck, and S. G. Benjamin, 2018: Subseasonal forecasting with an icosahedral, vertically quasi-Lagrangian coupled model. Part II: Probabilistic and deterministic forecast skill. *Mon. Wea. Rev.*, **146**, 1619– 1639, https://doi.org/10.1175/MWR-D-18-0007.1.
- Vecchi, G. A., and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. J. Climate, 27, 7994–8016, https://doi.org/10.1175/JCLI-D-14-00158.1.
- Vitart, F., 2004: Monthly forecasting at ECMWF. Mon. Wea. Rev., 132, 2761–2779, https://doi.org/10.1175/MWR2826.1.
- —, and Coauthors, 2017: The Subseasonal to Seasonal (S2S) Prediction project database. *Bull. Amer. Meteor. Soc.*, 98, 163– 173, https://doi.org/10.1175/BAMS-D-16-0017.1.
- Woolnough, S. J., F. Vitart, and M. A. Balmaseda, 2007: The role of the ocean in the Madden–Julian oscillation: Implications for MJO prediction. *Quart. J. Roy. Meteor. Soc.*, **133**, 117–128, https://doi.org/10.1002/qj.4.
- Zhang, C., 2005: Madden-Julian oscillation. *Rev. Geophys.*, 43, RG2003, https://doi.org/10.1029/2004RG000158.
- —, 2013: Madden–Julian oscillation: Bridging weather and climate. Bull. Amer. Meteor. Soc., 94, 1849–1870, https://doi.org/ 10.1175/BAMS-D-12-00026.1.
- Zhang, L., A. Kumar, and W. Wang, 2012: Influence of changes in observations on precipitation: A case study for the Climate Forecast System Reanalysis (CFSR). J. Geophys. Res., 117, D08105, https://doi.org/10.1029/2011JD017347.
- Zhu, H., M. C. Wheeler, A. H. Sobel, and D. Hudson, 2014: Seamless precipitation prediction skill in the tropics and extratropics from a global model. *Mon. Wea. Rev.*, 142, 1556– 1569, https://doi.org/10.1175/MWR-D-13-00222.1.
- Zhu, Y., X. Zhou, M. Peña, W. Li, C. Melhauser, and D. Hou, 2017: Impact of sea surface temperature forcing on weeks 3 and 4 forecast skill in the NCEP Global Ensemble Forecasting System. Wea. Forecasting, 32, 2159–2174, https://doi.org/ 10.1175/WAF-D-17-0093.1.