Improved Predictability of the Indian Ocean Dipole Using a Stochastic Dynamical Model Compared to the North American Multimodel Ensemble Forecast

SEN ZHAO,[a,b] MALTE F. STUECKER,[c,d,e] FEI-FEI JIN,[b] JUAN FENG,[f] HONG-LI REN,[g,h]
WENJUN ZHANG,[a] AND JIANPING LI[i]

[a] CIC-FEMD/ILCEC, Key Laboratory of Meteorological Disaster of Ministry of Education, and
College of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, China
[b] Department of Atmospheric Sciences, University of Hawai'i at Mānoa, Honolulu, Hawaii
[c] Department of Oceanography, and International Pacific Research Center, School of Ocean and
Earth Science and Technology, University of Hawai'i at Mānoa, Honolulu, Hawaii
[d] Center for Climate Physics, Institute for Basic Science, Busan, South Korea
[e] Pusan National University, Busan, South Korea
[f] College of Global Change and Earth System Science, Beijing Normal University, Beijing, China
[g] Laboratory for Climate Studies, and CMA–NJU Joint Laboratory for Climate Prediction Studies,
National Climate Center, China Meteorological Administration, Beijing, China
[h] Department of Atmospheric Sciences, School of Environment Studies, University of Geoscience, Wuhan, China
[i] Key Laboratory of Physical Oceanography/Institute for Advanced Ocean Studies/Frontiers Science Center
for Deep Ocean Multispheres and Earth System, Ocean University of China, and Qingdao National
Laboratory for Marine Science and Technology, Qingdao, China

ABSTRACT

This study assesses the predictive skill of eight North American Multimodel Ensemble (NMME) models in predicting the Indian Ocean dipole (IOD). We find that the forecasted ensemble-mean IOD–El Niño–Southern Oscillation (ENSO) relationship deteriorates away from the observed relationship with increasing lead time, which might be one reason that limits the IOD predictive skill in coupled models. We are able to improve the IOD predictive skill using a recently developed stochastic dynamical model (SDM) forced by forecasted ENSO conditions. The results are consistent with the previous result that operational IOD predictability beyond persistence at lead times beyond one season is mostly controlled by ENSO predictability and the signal-to-noise ratio of the Indo-Pacific climate system. The multimodel ensemble (MME) investigated here is found to be of superior skill compared to each individual model at most lead times. Importantly, the skill of the SDM IOD predictions forced with forecasted ENSO conditions were either similar or better than those of the MME IOD forecasts. Moreover, the SDM forced with observed ENSO conditions exhibits significantly higher IOD prediction skill than the MME at longer lead times, suggesting the large potential skill increase that could be achieved by improving operational ENSO forecasts. We find that both cold and warm biases of the predicted Niño-3.4 index may cause false alarms of negative and positive IOD events, respectively, in NMME models. Many false alarms for IOD forecasts at lead times longer than one season in the original forecasts disappear or are significantly reduced in the SDM forced by forecasted ENSO conditions.

1. Introduction

The Indian Ocean dipole (IOD) phenomenon is a prominent climate pattern in the tropical Indian Ocean, characterized by year-to-year fluctuations of a dipole structure in sea surface temperature (SST) anomalies between the southeastern equatorial Indian Ocean (SEIO) and the western equatorial Indian Ocean (WIO) (Saji et al. 1999; Webster et al. 1999). Like El Niño–Southern Oscillation (ENSO) in the Pacific, the changes in the zonal SST gradient and coherent thermocline anomalies across the Indian Ocean is coupled with the atmospheric circulation (Yamagata et al. 2004; Schott et al. 2009; McPhaden and Nagura 2014). Importantly, the IOD affects weather and climate in many areas of

ⓐ Denotes content that is immediately available upon publication as open access.

Corresponding author: Malte F. Stuecker, stuecker@hawaii.edu

the world, especially in Indian Ocean rim areas such as Australia, India, East Africa, and East Asia (Ashok et al. 2001, 2003; Guan and Yamagata 2003; Saji and Yamagata 2003; Yamagata et al. 2004; Yuan et al. 2008; Cai et al. 2011; Qiu et al. 2014; Lu et al. 2018a). Therefore, skillful IOD predictions allow for the implementation of mitigation measures to climate variability and thereby can provide societal benefits in various areas such as agriculture, fisheries, marine ecosystems, human health, as well as potential increased resilience to natural disasters (e.g., Abram et al. 2003; Cai et al. 2009; Hashizume et al. 2012; Takaya et al. 2014; Yuan and Yamagata 2015).

The predictability of Indian Ocean SST anomalies associated with the IOD has previously been assessed using a range of coupled climate models and statistical models (e.g., Wajsowicz 2005, 2007; Luo et al. 2005, 2007; Song et al. 2008; Zhao and Hendon 2009; Dommenget and Jansen 2009; Shi et al. 2012). For instance, Shi et al. (2012) assessed the predictive skill of the SST anomalies associated with the IOD for the period of 1982–2006 by using ensemble seasonal forecasts from six coupled models developed by the Australian Bureau of Meteorology, National Centers for Environmental Prediction (NCEP), European Centre for Medium-Range Weather Forecasts (ECMWF), and Frontier Research Centre for Global Change. They found that the maximum lead time for skillful prediction of SSTs in the WIO is about 5–6 months compared to only 3–4 months in the SEIO (when all start calendar months are considered). Other studies found that skillful prediction of IOD (i.e., the anomalous zonal SST gradient) events is limited to a lead time of approximately one season (Shi et al. 2012; Liu et al. 2017), with slightly higher skill seen only for some individual strong IOD events, perhaps up to about two seasons (Luo et al. 2008; Shi et al. 2012). The prediction failure of IOD events at longer lead times was mostly attributed to a strong boreal winter "predictability barrier" (Wajsowicz 2005, 2007; Feng et al. 2014) (i.e., forecast skill drops rapidly for the target boreal winter season regardless of the forecast start time).

Some studies (Song et al. 2008; Zhao and Hendon 2009; Yang et al. 2015) showed that IOD events that co-occur with ENSO events are more predictable, while the remaining events appear to be initiated by weather noise and exhibit a lower predictability. These results indicate that a poorly simulated IOD–ENSO relationship might be one reason that limits the predictive skill of the IOD in operational forecasts (Shi et al. 2012). In fact, there is a considerable debate regarding the IOD–ENSO relationship within the scientific community. On one hand, some modeling studies (Iizuka et al. 2000; Behera et al.

2006) argued that the IOD is in fact an intrinsic climate mode that is largely independent from ENSO. For instance, Behera et al. (2006) found that only about 42% of IOD events were affected by the ENSO. On the other hand, other studies hypothesized that the IOD mode is not independent of the tropical Pacific and ENSO (Annamalai et al. 2003; Loschnigg et al. 2003; Zhang et al. 2015; Yang et al. 2015; Kajtar et al. 2017; Stuecker et al. 2017). By using a partially coupled model experiment with decoupled SST over the tropical Pacific, Crétat et al. (2018) and Wang et al. (2019) showed that the IOD still exists without ENSO, but with weaker amplitude and reduced Bjerknes feedback in the Indian Ocean. Furthermore, several studies demonstrated evidence that only about one-third of IOD events occur independently of ENSO events (Loschnigg et al. 2003; Stuecker et al. 2017). Recently, Stuecker et al. (2017) developed a new null hypothesis framework for the IOD and showed that most of the observed IOD variability can be explained by deterministic interactions between the annual cycle and ENSO [ENSO combination mode (C-mode)] (Stuecker et al. 2013, 2015). Zhao et al. (2019) further demonstrated improved IOD predictions using seasonally modulated ENSO forcing and provided evidence that IOD predictability beyond persistence is largely controlled by ENSO predictability and signal-to-noise ratio.

In operational seasonal forecasting, the use of multimodel ensemble prediction generally results in improved skill due to error compensation and greater consistency and reliability between models (Hagedorn et al. 2005; DelSole et al. 2014). The North American Multimodel Ensemble (NMME) system (Kirtman et al. 2014) was recently developed to harness this idea. The NMME system is used for seasonal predictions since 2011 and was made an operational forecast system in 2016. Many studies have shown that the NMME system has advanced the forecasting skill of ENSO and relevant climate variables (Barnston et al. 2015, 2019; Chen et al. 2017).

Given this recent improvement of ENSO prediction in the NMME system, one might wonder if a similar skill improvement also exists for IOD prediction or if the enhanced ENSO skill can be translated into a better IOD prediction skill using the simple model framework developed by Stuecker et al. (2017) and Zhao et al. (2019). Furthermore, we ask the question whether we are near the intrinsic predictability limit associated with the chaotic nature of the coupled ocean–atmosphere system. For instance, Newman and Sardeshmukh (2017) argued that the Indian Ocean SST forecast skill of the NMME system is close to the predictability limit estimated using signal-to-noise ratios

TABLE 1. Basic information for the eight NMME climate models.

Model	Name used here	Period	Ensemble size	Lead times (months)
CMC1-CanCM3	CMC1	Jan 1982–Jul 2019	10	0.5–11.5
CMC2-CanCM4	CMC2	Jan 1982–Jul 2019	10	0.5–11.5
COLA-RSMAS-CCSM4	CCSM4	Jan 1982–Jul 2019	10	0.5–11.5
NCEP-CFSv2	CFSv2	Jan 1982–Jul 2019	24	0.5–9.5
GFDL-CM2p1-aer04	GFDL	Jan 1982–Dec 2017	10	0.5–11.5
GFDL-CM2p5-FLOR-A06	GFDL-A	Jan 1982–Jul 2019	12	0.5–11.5
GFDL-CM2p5-FLOR-B01	GFDL-B	Jan 1982–Jul 2019	12	0.5–11.5
NASA-GMAO-062012	NASA	Jan 1982–Jan 2018	12	0.5–8.5

from a simplified NMME linear inverse model (LIM) forecast. Furthermore, Liu et al. (2017) suggested that the SST forecast at each pole of the IOD has little room for improvement, while there is a large potential to improve the gradient forecasts of the two poles, at least by 0.2–0.3 correlation skill, based on potential predictability estimates using multimodel forecasts from the Ensemble-Based Predictions of Climate Changes and Their Impacts (ENSEMBLES). Such potential predictability estimates are of course model dependent; therefore, it is interesting to compare IOD potential predictability using the NMME models with the newly developed stochastic dynamical IOD model (Stuecker et al. 2017; Zhao et al. 2019).

The remainder of this paper is organized as follows. Section 2 presents the data and methodology. Section 3 evaluates the forecast skill of the NMME systems. Section 4 discusses the further improvements using the SDM. Section 5 includes the discussion and summarizes the main conclusions.

2. Data and methodology

a. Data

We utilize the hindcasts (1982–2010) and real-time forecasts (2011–19) of eight models from the NMME project, which are CMC1-CanCM3, CMC2-CanCM4, COLA-RSMAS-CCSM4, NCEP-CFSv2, GFDL-CM2p1-aer04, GFDL-CM2p5-FLOR-A06, GFDL-CM2p5-FLOR-B01, and NASA-GMAO-062012. For simplicity, these model names are shortened as CMC1, CMC2, CCSM4, CFSv2, GFDL, GFDL-A, GFDL-B, and NASA, respectively. Table 1 summarizes the time period, ensemble size, and lead months for these eight models used here. The number of ensemble size ranges from 10 to 24, and the maximum lead time varies from 8.5 to 11.5 months. The NMME forecasts were initialized on or near the first day of each month. The lead time is defined as the number of months between forecast start time and the center of the month being predicted. For example, for a forecast starting at the beginning of January, the forecast for January has 0.5-month lead, for February a

1.5-month lead, and so on. Besides looking at the ensemble mean forecast characteristics of each individual model, the grand multimodel ensemble (MME) forecasts are studied with equal weight given to each individual model. All gridded SST forecast data on a global 1° grid analyzed here are publicly available in the International Research Institute for Climate and Society (IRI) Data Library (http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME). The SST observations used here are the NOAA Optimum Interpolation SST data, version 2 (Reynolds et al. 2002; http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html).

Monthly anomalies are calculated with respect to the climatology from January 1982 to December 2010 in both the observations and most of the NMME models except for CCSM4 and CFSv2. Several studies have noted a discontinuity in the forecast bias of the CFSv2 SST hindcasts for the central and eastern tropical Pacific occurring around 1999, which has been related to a discontinuity in the data assimilation and initialization procedure (Xue et al. 2011; Kumar et al. 2012; Barnston and Tippett 2013; Barnston et al. 2019). Both CFSv2 and CCSM4 model share the same initial conditions (Kirtman et al. 2014), which come from the Climate Forecast System Reanalysis (Saha et al. 2010). Therefore, following the method of Barnston et al. (2019), we eliminate the discontinuous forecast biases by calculating the forecast anomalies using two different climatological periods of 1982–98 and 1999–2010, respectively, for these two models. In calculating the anomalies for the NMME models, the dependence on both season and forecast lead time are considered following Kumar et al. (2017).

The IOD mode index (DMI) is defined as the area-averaged SST anomalies in the WIO (10°S–10°N, 50°–70°E) minus those in the SEIO (10°S–0°, 90°–110°E) (Saji et al. 1999). The predictive skill of the IOD has been studied by measuring the predictive skill of the DMI in many previous studies (e.g., Luo et al. 2007; Liu et al. 2017; Doi et al. 2017). The Niño-3.4 index (hereafter N3.4) is defined as the SST anomalies averaged over the region 5°S–5°N and 120°–170°W.

The N3.4 is used by many operational centers as a key oceanic variable that describes the ENSO state.

b. A stochastic dynamical model in predicting the DMI and hindcast experiments

We use a recently developed stochastic dynamical model (SDM) for the DMI by Stuecker et al. (2017) and Zhao et al. (2019). In this framework, the continuous time evolution of the DMI is described by the following equation:

$$\frac{dT}{dt} = -\lambda_0[1 + D\cos(\omega_a t + \varphi_D)]$$
$$+ \alpha_0[1 + A\cos(\omega_a t + \varphi_A)]\text{N3.4}(t) + \sigma_0\xi(t), \quad (1)$$

where T is the monthly DMI. The three terms on the right-hand side of Eq. (1) represent seasonally modulated damping that arises from the Indian Ocean coupled ocean–atmosphere feedback, the seasonally modulated ENSO forcing, and the stochastic forcing, respectively. The parameters of growth rate (λ_0, D, and φ_D), coupling strength (α_0, A, and φ_A), and stochastic forcing strength (σ_0) can be estimated via multivariate linear regression using the observed DMI and N3.4 indices following Zhao et al. (2019). The term ω_a denotes the annual frequency of the annual cycle [$=\pi/(6 \text{ months})$]. We used the training period of 1982–98 for estimating all the parameters as in Zhao et al. (2019). We find a very similar performance of the SDM when we train the parameters using the period after 1999 instead. This suggests the robustness of the SDM parameters despite observed decadal changes in IOD predictability (Han et al. 2014; Liu et al. 2017). We will also neglect the stochastic forcing term—which can be a crucial factor in the initiation of observed IOD events—as we compare our IOD model skill to ensemble mean forecasts following Zhao et al. (2019).

We conducted three types of experiments using the SDM. In the first experiment (SDM-P-P), we integrated the SDM initialized from the monthly observed (thus "perfect") DMI conditions by prescribing the observed/perfect ENSO forcing. This experiment is a measure of the upper IOD predictability limit provided by ENSO. In the second experiment (SDM-F-P), the observed ENSO forcing was replaced by the forecasted N3.4 index in each individual NMME model. Finally, in the third experiment (SDM-F-F), the ENSO forcing was the same as SDM-F-P but the initial DMI conditions were replaced with the DMI index at 0.5-month lead in each individual NMME model. The SDM-F-F is the SDM version that can be used in an operational forecast setting. Importantly,

our approach is linear when we use fixed parameters and neglect the stochastic forcing term. Therefore, we would obtain the same result if instead our method would be applied to all the members independently and then the ensemble average of the resulting DMI forecasts would be calculated.

c. Forecast verification metrics

To quantify the deterministic skill of the IOD predictions in different models and approaches with respect to the observations we use the anomaly correlation coefficient (ACC) and the root-mean-square error (RMSE) metrics. The ACC and RMSE metrics are two common measures of forecast accuracy quantifying errors in sign and amplitude. To assess seasonal performance, the RMSE is standardized for each season individually and referred to as normalized RMSE (NRMSE), so that the climatology forecasts (zero anomaly) result in the same RMSE-based skill (of zero) for all seasons and that all season's RMSE contribute equally to a seasonally combined RMSE (Barnston et al. 2012). For skill comparisons between the predictions from models of the International Research Institute for Climate and Society (IRI)/Climate Prediction Center (CPC) IOD prediction plume and other seasonal forecast products, which are for 3-month-averaged SST data, 3-month running means are used prior to calculating the deterministic verification metrics. The Fisher z transformation was used to test statistical significance of the ACC differences.

Large-amplitude IOD events exhibit the most pronounced climate impacts and thus are most important to predict. Following Stanski et al. (1989) and Shi et al. (2012), we further assess the ability of the models to predict three categories of IOD events: (i) positive IOD events that have a DMI amplitude exceeding 0.5 standard deviation, (ii) negative IOD events that have a DMI amplitude less than 0.5 standard deviation, and (iii) neutral IOD events that fall in between. The contingency table (Table 2) is made for the occurrence of observed and predicted IOD events using the DMI in September–November (SON) for each individual model. We use the observed 0.5 standard deviation threshold to categorize the model forecasts following Shi et al. (2012) for better comparison with their results. The hit rate (HR) for correctly forecasting the occurrence of a positive/negative IOD event is defined as

$$\text{HR}_{\text{positive}} = \frac{a}{a + b + c} \times 100\% \quad \text{or}$$
$$\text{HR}_{\text{negative}} = \frac{i}{g + h + i} \times 100\%. \quad (2)$$

TABLE 2. Contingency table of forecasting positive, neutral, and negative IOD events.

	Observed positive IOD	Observed neutral	Observed negative IOD	Total
Forecasted positive IOD	a	d	g	$M \, (a + d + g)$
Forecasted neutral	b	e	h	$N \, (b + e + h)$
Forecasted negative IOD	c	f	i	$O \, (c + f + i)$
Total	$J \, (a + b + c)$	$K \, (d + e + f)$	$L \, (g + h + i)$	$T \, (a + b + c + d + e + f + g + h + i)$

The false alarm rate (FAR), which is a measure of incorrectly forecasting an event when in reality a neutral event occurred, is defined as

$$\text{FAR} = \frac{d + f}{d + e + f} \times 100\% . \tag{3}$$

See Table 2 for the definitions of the letters a–i. Based on the threshold, 11 $(a + b + c)$ positive IOD events, 10 $(g + h + i)$ negative IOD events, and 16 $(d + e + f)$ neutral IOD events occur in the SON season during 1982–2018.

3. Prediction skill of IOD and biases of IOD–ENSO relationships in NMME models

Figures 1a and 1b shows the all-months stratified ACC and RMSE skill scores between forecasted and observed DMI as function of lead time. For initialization skill at 0.5-month lead, the MME (black lines in Figs. 1a,b) demonstrates the best ACC and RMSE skill scores. This result indicates a relatively better initialization skill for the zonal gradient of Indian Ocean SST when we use a multimodel ensemble mean. The CFSv2, NASA, and CCSM4 models (red, purple, and blue lines in Figs. 1a,b, respectively) are among the best performers in terms of initialization skill.

Each individual model and the MME are characterized by an ACC above persistence and an RMSE lower than persistence. For short lead times of 1.5–4.5 months, each individual model does not show statistically significant ACC and RMSE differences with the other models, while the MME is superior to each individual model, characterized by a higher ACC (about 0.1) and lower RMSE (about 0.05 K). For longer than 5.5-month lead times, the ACC drops below 0.5 and the RMSE increases to values as large as the climatological standard deviation of the DMI for all models, suggesting very limited predictive skill of the IOD for current operational models at longer lead times. The GFDL-A, GFDL-B (long dashed and solid orange lines in Figs. 1a,b), and CCSM4 models are among the top-ranked models for the ACC at leads longer than 5.5 months. The MME also shows competitive ACC skill and better RMSE skill than the top-ranked models at

leads longer than 5.5 months. These results indicate the superior performance of the MME for the IOD predictions compared to individual models, consistent with Wu and Tang (2019).

Given the important role of the IOD–ENSO relationship in IOD predictions (Song et al. 2008; Zhao and Hendon 2009; Luo et al. 2016; Stuecker et al. 2017; Zhao et al. 2019), we next evaluate the performance of the NMME models in predicting ENSO and the IOD–ENSO relationship. Figures 1c and 1d shows the all-months stratified ACC and RMSE skill scores between forecasted and observed N3.4 index as function of lead time. While the MME exhibits the highest ACC and RMSE skill scores for the DMI among the NMME models at nearly all lead times (Figs. 1a,b), the MME shows the highest ACC skill for the N3.4 index only at short lead times (0.5–3.5 months). At longer lead times, CMC2 has the highest skill (at 4.5 and 5.5 months) and CFSv2 exhibits good skills at even longer lead times (9.5 months; Figs. 1c,d). This outcome is consistent with the findings of Kirtman et al. (2014) and Barnston et al. (2019), which showed that some individual models may be superior to the MME at certain lead times, however the MME is always close to being top ranked.

As an important proxy for the IOD–ENSO relationships, Figs. 2a and 2b shows the lead–lag cross correlations between monthly N3.4 and DMI for the observations and the forecasts at lead times of 4.5 and 7.5 months for the NMME models. The IOD–ENSO relationship deteriorates away from the observed correlation with increasing lead time in some models (that consist of multiple ensemble members each; see Table 1) such as CFSv2 and CMC2. Although the predictive skill of ENSO has significantly improved from CMC1 to CMC2 (green dashed and solid lines in Figs. 1c,d), an improvement of predictive skill for the IOD is not clearly evident (Figs. 1a,b). It is important to note that the CMC2 model (solid green lines in Figs. 1a,b) is an outlier as the DMI RMSE in this model is as large as the DMI RMSE of the persistence forecast and also larger than the DMI RMSE in the CMC1 model (long dashed green lines in Figs. 1a,b). This may be related to a poor representation of the IOD–ENSO relationship in both the CMC1 and CMC2 models with a
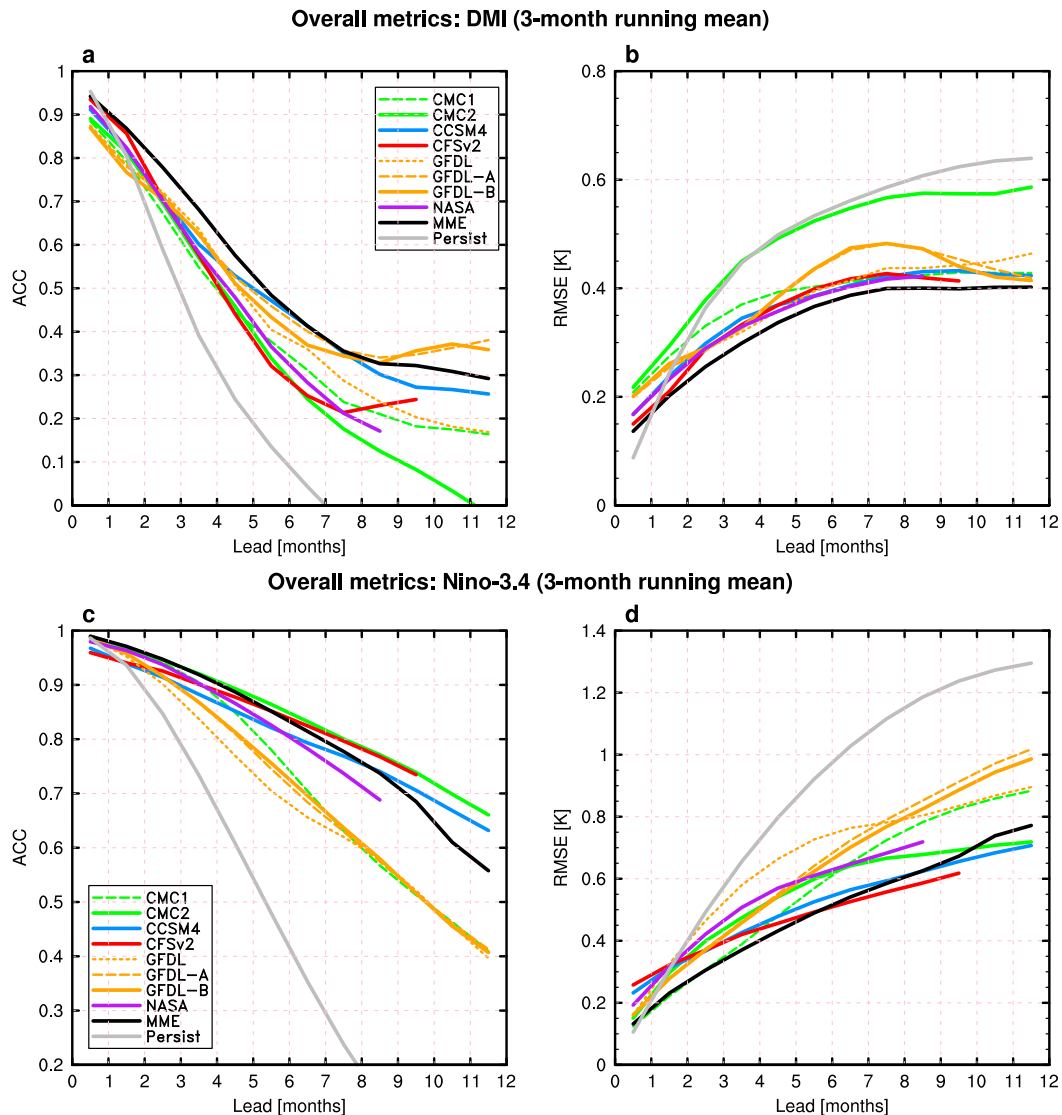
FIG. 1. (a) Anomaly correlation coefficient (ACC) and (b) root-mean-square error (RMSE; K) skill between the all-months observed and forecasted DMI, as a function of lead time for the individual models. (c),(d) As in (a) and (b), but for the Niño-3.4 index. Both forecasts and verification were smoothed with a 3-month running mean prior to computing the metrics.

positive lead–lag correlation coefficient for ENSO leading IOD months, which is opposite to what is seen in the observations and most other models (Figs. 2a,b). In this sense, the relatively poor predictive skill for the IOD in CFSv2 (Figs. 1a,b) might also be related to a poor representation of the IOD–ENSO relationship in CFSv2, with a much higher negative correlation for ENSO leading IOD months in CFSv2 compared to the observations (Figs. 2a,b). Overall, CCSM4 performs best simulating the IOD–ENSO relationship in terms of a cross-correlation relationship that looks closest to the observations among the NMME models (Figs. 2a,b).

Importantly, these biased lead–lag cross correlations between monthly N3.4 and DMI for the forecasts at lead times of 4.5 and 7.5 months can be corrected using the SDM (Figs. 2c,d) since it utilizes the observed IOD–ENSO relationship (Stuecker et al. 2017; Zhao et al. 2019). It should be noted that the SDM overestimates the positive correlations of DMI leading ENSO by the peak time around 2 months. This is because the stochastic forcing as well as some intrinsic Indian Ocean processes are not included in the current SDM. In the following section, we provide evidence that utilizing the observed IOD–ENSO relationship in the SDM can improve the IOD predictive skill in the NMME models.
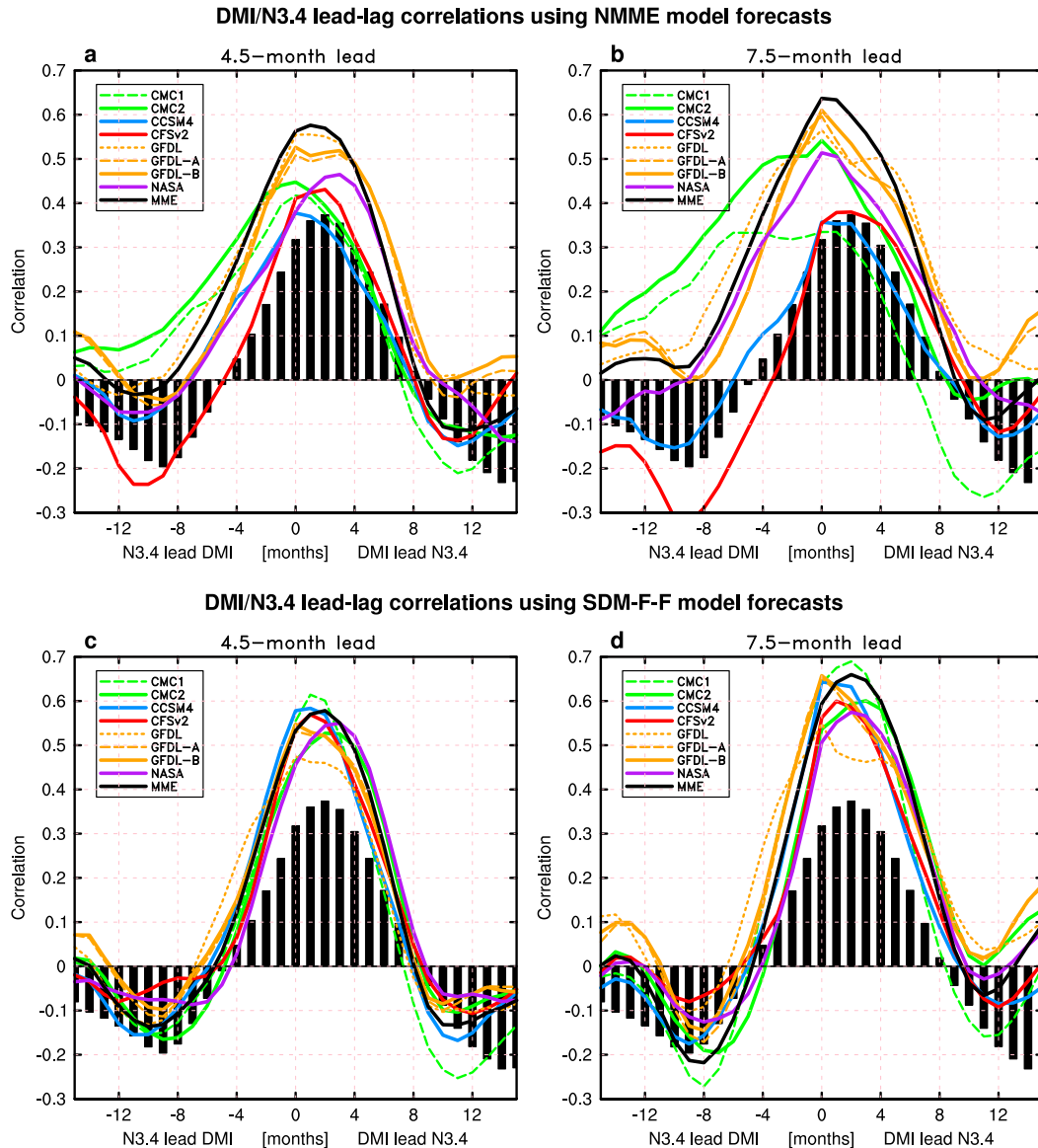
FIG. 2. Lead/lag cross correlations between monthly Niño-3.4 index and DMI for the observations (black bars) and the forecasts (curves) at lead times of (a),(c) 4.5 and (b),(d) 7.5 months for (top) the NMME models and (bottom) the SDM-F-F that uses forecasted ENSO forcing and DMI initial values from the NMME models.

4. Improved performance of SDM in predicting the IOD compared to NMME models

a. Deterministic all-months stratified skill of IOD prediction

The SDM forecasts driven by forecasted ENSO forcing from the CMC1, CMC2, and CFSv2 models exhibit significantly better predictive skill of the IOD in terms of both ACC and RMSE than the original IOD forecasts from each individual model (Figs. 3 and 4). For example, compared with the original CMC2 forecast, the corresponding (forced with forecasted CMC2 ENSO

conditions) SDM-F-F exhibits an improved ACC value of 0.15 and RMSE value of 0.15 K averaged over lead times from 4.5 to 9.5 months. Similarly, an improvement using the SDM is also evident for the NASA model at lead times longer than 5.5 months. The ACC scores of our SDM-F-F forecasts using forecasted ENSO forcing from the GFDL, GFDL-A, and GFDL-B models are not statistically different from those in the corresponding original models (Figs. 2e–g), partly due to relatively low ENSO prediction skills of these models (sharp drop of ACC and RMSE skill for N3.4 with increasing lead times; Figs. 1c,d). In general, the RMSE scores are
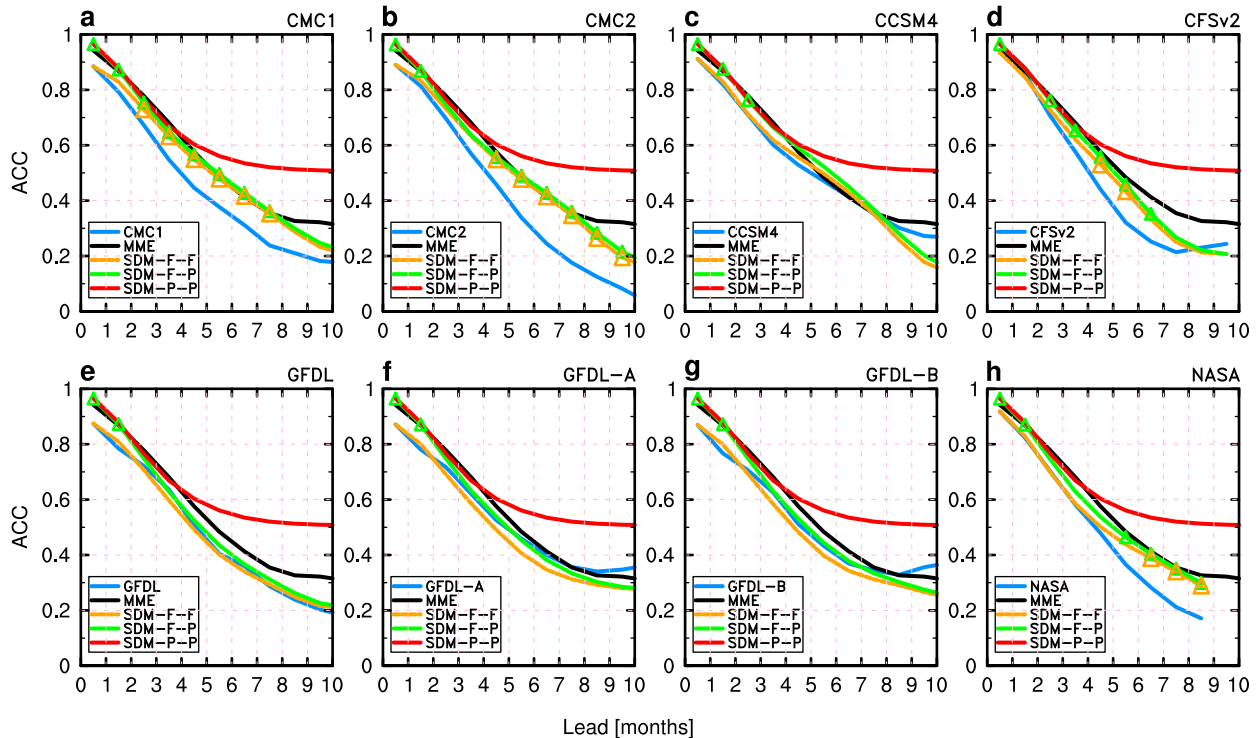
FIG. 3. (a) ACC skill between the all-months observed and forecasted DMI as a function of lead time for CMC1 (blue), MME (black), SDM-P-P (red), SDM-F-P (green), and SDM-F-F (orange). The SDM-F-P and SDM-F-F in (a) are SDM experiments using CMC1 forecasted ENSO forcing with observed and CMC1 forecasted DMI initial conditions, respectively. (b)–(h) As in (a), but for CMC2, CCSM4, CFSv2, GFDL, GFDL-A, GFDL-B, and NASA, respectively. The triangles denote that the ACC differences between SDMs (SDM-F-P in green and SDM-F-F in orange) and original model (blue) are statistically significant above the 90% significance level based on a two-sided test of the Fisher z transformation.

improved in the SDM compared to the original models, especially at longer lead times, approximately at the same level or better than the MME (Fig. 4). Importantly, the SDM-P-P forecasts that utilize the observed ENSO forcing and DMI initial conditions demonstrate superior performance of their IOD predictions compared to the MME at lead times from 4.5 to 9.5 months in terms of both their ACC and RMSE skill scores (red lines in Figs. 3 and 4). This strongly suggests that IOD predictions can be still further improved by improving the ENSO predictions in these models.

b. Seasonal variation in the IOD prediction skill

To explore the seasonality of IOD forecast skill, Figs. 5 and 6 show the ACC and NRMSE of each individual model, the MME, and the persistence, respectively, as a function of target month and lead time. Figures 5 and 6 show mostly consistent patterns among individual models with ACC values that peak and NRMSE values that reach their minimum at target months in boreal fall, which is the peak IOD season that exhibits the largest signal-to-noise ratio (Kumar and Hoerling 2000; Liu et al. 2017) and therefore has

the highest potential predictability (Luo et al. 2007). The MME exhibits relatively superior skill in boreal fall than each individual model in terms of both ACC and NRMSE scores. Still, the MME skill in boreal fall is still significantly lower than that from our SDM-P-P forecasts at longer lead times (Figs. 5i,l and 6i,l), again indicating the potential room for IOD prediction improvement by improving ENSO prediction.

Unlike the spring predictability barrier for the ENSO predictions, all models show a sharp drop of ACC and NRMSE skill at target months in boreal winter (December and January) regardless of the lead time (Figs. 5 and 6), indicating the existence of the winter predictability barrier for the IOD predictions (Wang et al. 2009; Shi et al. 2012; Feng et al. 2014). One of the reasons that the winter barrier is said to exist is because winter is a transitional time of the year for most IOD events that exhibits the lowest signal-to-noise ratio. The underlying mechanism might be due to the annual reversal of the monsoon winds (Li et al. 2003; Schott et al. 2009; Luo et al. 2016). The northwesterly surface wind is weak during boreal winter and spring, the thermocline is flat, and there is little or no upwelling in the
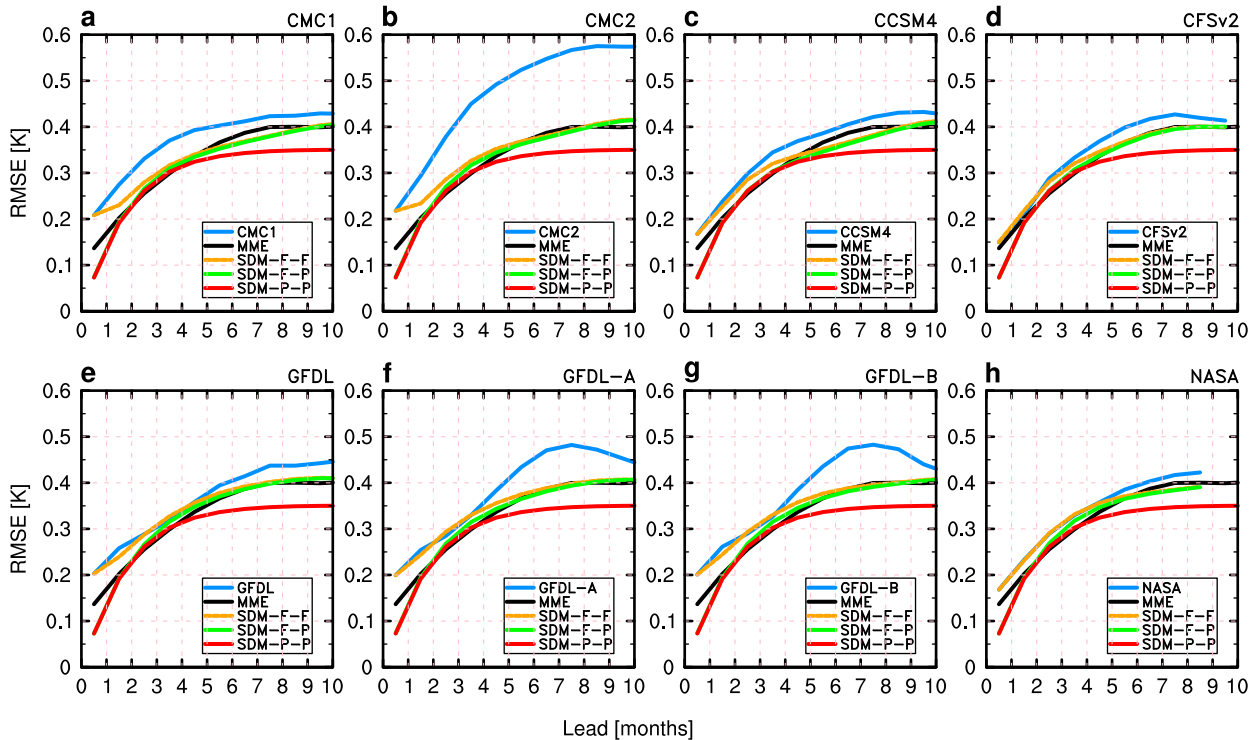
FIG. 4. As in Fig. 3, but for RMSE (K) skill.

eastern equatorial Indian Ocean, suggesting only a weak or absent Bjerknes feedback during this season (Schott et al. 2009). In contrast, the strong reversal of monsoon winds to southeasterly during boreal summer and fall is favorable for the Indian Ocean Bjerknes feedback and thus favors the development of IOD events. Furthermore, a negative thermodynamic air–sea feedback in boreal winter arises from the interaction between an anomalous atmospheric anticyclone and a cold SST anomaly off Sumatra (Li et al. 2003). Both SDM-F-F and SDM-P-P also exhibit the sharp drop of ACC and NRMSE during boreal winter (Figs. 5k,l and 6k,l). This suggests the winter predictability barrier for the IOD predictions cannot be overcome with the SDM approach.

An interesting feature is that unlike the skill seasonality in the persistence forecasts, many models' forecasts (CCSM4, GFDL, GFDL-A, GFDL-B, NASA, and the MME) illustrate a slight recovery of ACC and NRMSE skill at target months in late winter/early spring (February–April) for most lead times (Figs. 5 and 6). However, this rebound is not evident in the CMC1 and CMC2 models, and only weakly represented in CFSv2. By studying the persistence of observed SEIO and WIO SST anomalies, Ding and Li (2012) suggested that the winter predictability barrier for SST in SEIO is more strongly influenced by ENSO. Furthermore, this skill

rebound appears in the SDM-F-F forecasts that use forecasted ENSO forcing and DMI initial conditions from CMC1, CMC2, and CFSv2 models' forecasts (see example in Fig. 5k for CMC2). This further indicates that a poor representation of the IOD–ENSO relationship limits IOD predictability in these three models. The superior performance of the MME is also evident for this rebound (Figs. 5 and 6), which might be explained by both better ENSO prediction skill (Figs. 1c,d) and a more realistic IOD–ENSO relationship (Fig. 2).

c. Prediction skill for the IOD in peak season

Concentrating just on the SON season when the IOD tends to peak, Fig. 7a shows the skillful DMI lead time (defined by an ACC value of 0.6) ranging from 4.5-month lead (CMC1 and CMC2) to 6-month lead (most of other NMME models and the MME), which is significantly improved compared to a skillful 4-month lead reported by Shi et al. (2012) using older prediction systems. The superior performance of the MME is evident in terms of the NRMSE metric (Fig. 7b). Such a MME benefit was also found in other multimodel studies for ENSO (e.g., Barnston et al. 2019) and IOD predictions (Liu et al. 2017). If a skillful prediction is defined as ACC above 0.5 and NRMSE less than 1, the MME provides skillful predictions of DMI in SON at 6.5-month lead (Figs. 7a,b).
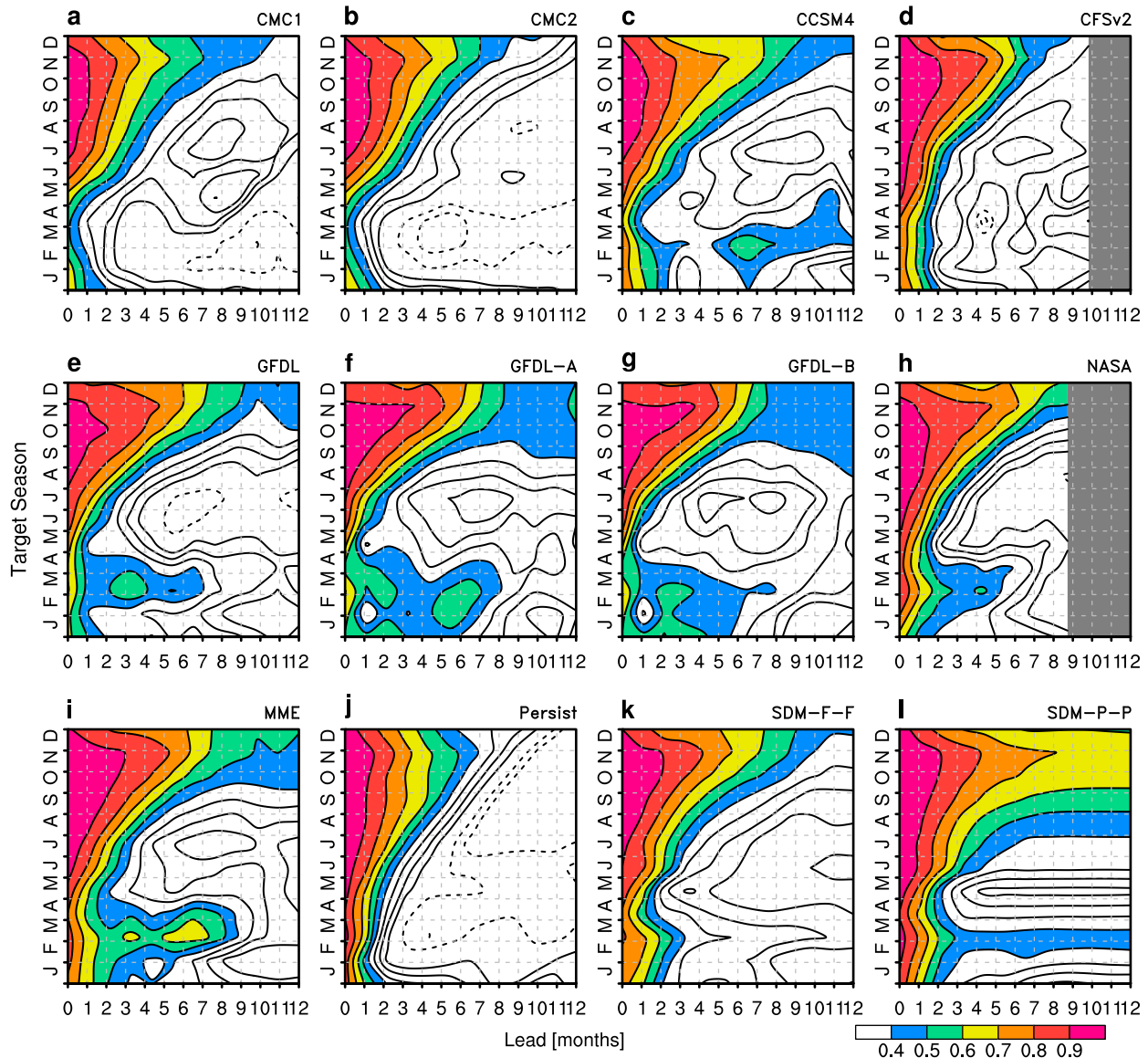
ACC seasonality: DMI (3-month running mean)

FIG. 5. ACC between model forecasts and observations as a function of lead time and target season. Each panel highlights individual models, the MME, persistence, SDM-F-F, and SDM-P-P. The SDM-F-F forecasts use forecasted ENSO forcing and DMI initial conditions from the ensemble mean of CMC2 and CFSv2. The model names are indicated at the top right of each panel. The contour interval is 0.1.

The SON stratified metrics for SDM-F-F forecasts are shown in Figs. 7c and 7d. A slightly improved ACC and considerable improved RMSE skill is seen for SDM-F-F forecasts compared to the original forecasts from CMC1, CMC2, CFSv2, and CCSM4 at most lead times, and for NASA at longer lead times. Importantly, the SDM-F-F provides a slightly better forecast than any of the original forecasts of the NMME models. Furthermore, the SDM-P-P forecast provide skillful IOD predictions up to 11 months ahead, which is strongly superior to the MME.

This implies there is ample scope to improve the NMME models in terms of IOD prediction skill and that the upper predictability limit at longer lead times has probably not yet been achieved because none of the NMME models are fully capturing the observed IOD–ENSO relationship (Fig. 2) and because both ENSO physics and ENSO prediction skill could likely be further improved upon (Kumar et al. 2017).

Figures 8a–c shows the hit rate for positive and negative IOD events in SON and the false alarm rate as a

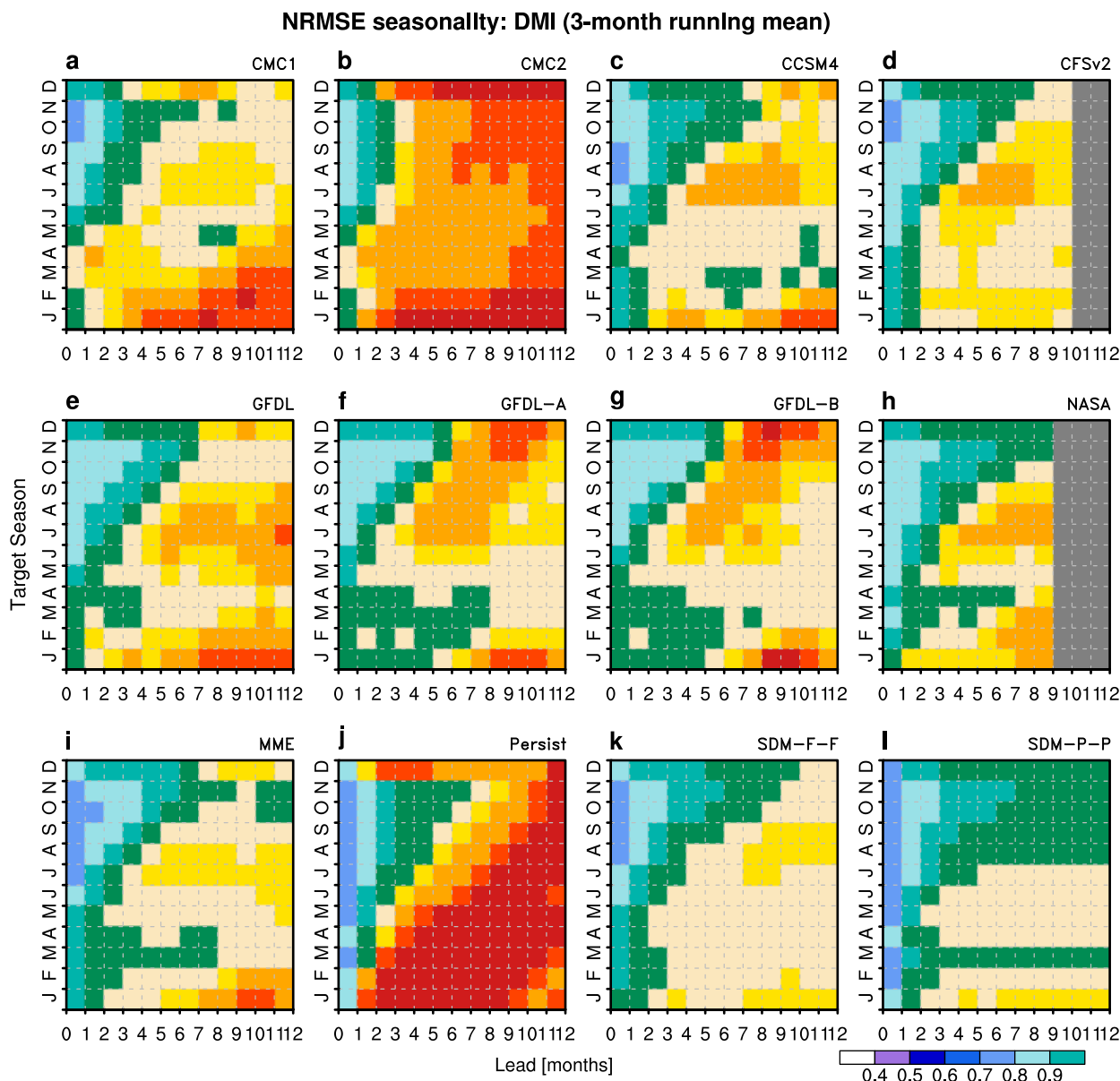## NRMSE seasonality: DMI (3-month running mean)

FIG. 6. As in Fig. 5, but for the normalized root-mean-square error (NRMSE).

function of lead time for the original forecasts of the NMME models. The observed frequency of occurrence of positive IOD events, negative IOD events, and neutral IOD events are, respectively, 11/37 ($=J/T$), 10/37 ($=L/T$), and 16/37 ($=K/T$) (see Table 2 for definitions of capital letters J–T) for the period of 1982–2018. As seen in Figs. 8a–c, the hit rate for positive IOD events and false alarm rate from original NMME forecasts is larger than the observed frequency of occurrence for IOD events and exhibit large model diversity. We find hit rates exceeding 50% ranges from 3.5 (CFSv2 and NASA) to 8.5 months (CMC2, GFDL-A, and GFDL-B)

and the MME in between, with false alarms exceeding 50% from 1.5 (CMC2) to 7.5 months (NASA) and the MME in between. The hit rate for negative IOD events exhibits relatively smaller model diversity than that for positive IOD events, with hit rate exceeding 50% ranges from 3.5 (NASA) to 6.5 months (CCSM4). Although some model original forecasts (such as CMC2, GFDL-A, and GFDL-B) usually correctly predict the occurrence of IOD events when an event actually occurs, they also often wrongly predict an event when none occurs; so that there is reduced confidence of an event occurring when one is forecasted. Nevertheless, these rate skills from the NMME

SON Metrics: NMME original forecasts
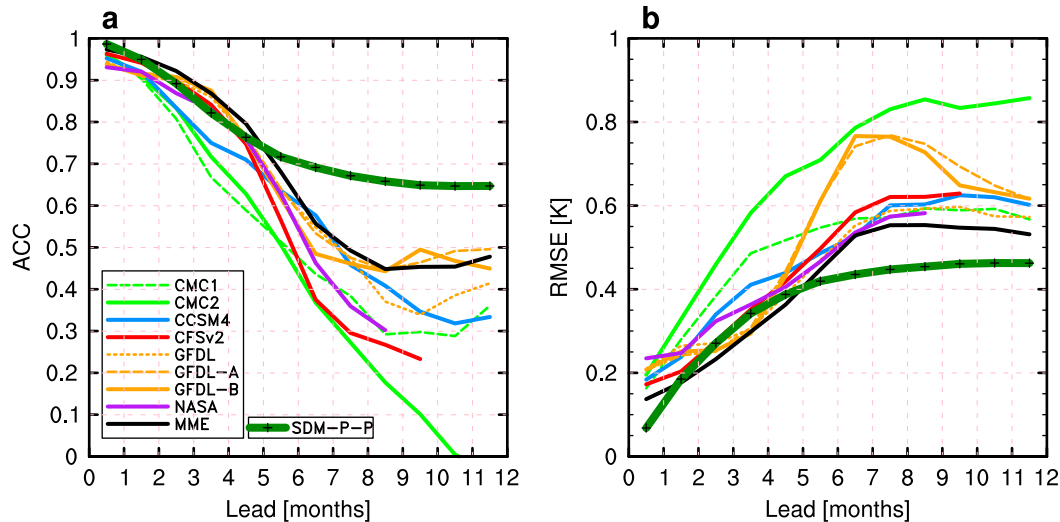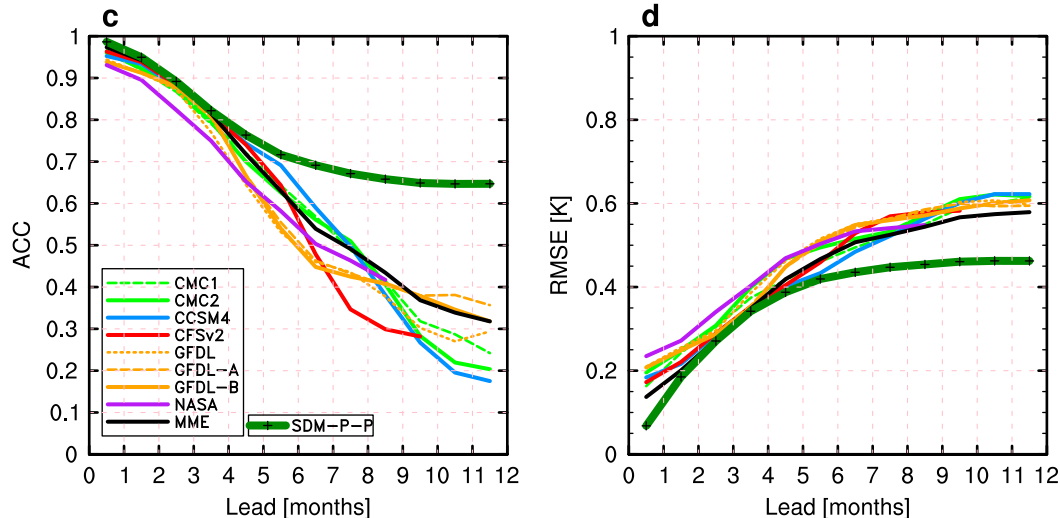

SON Metrics: SDM-F-F forecasts

FIG. 7. (a) ACC and (b) RMSE (K) skill between the observed and forecasted DMI at target season SON, as a function of lead time for the individual models. (c),(d) As in (a) and (b), but for SDM-F-F forecasts using forecasted ENSO forcing and DMI initial conditions from NMME models and MME. The metrics for SDM-P-P forecasts are also indicated.

original forecasts are higher than those from older prediction systems reported by Shi et al. (2012), indicating a marked improvement is clearly achieved through NMME systems.

The hit rate and the false alarm rate for the SDM-F-F forecasts are shown in Figs. 8d–f. We see reduced false alarm rates at longer lead times for all SDM-F-F forecasts compared with their corresponding original forecasts although the hit rate for negative IOD events is slightly decreased. The SDM-P-P forecasts at longer lead times are the best performers in terms of false alarm rate. Another interesting aspect shown in Fig. 8 is that the SDM-P-P forecasts exhibit asymmetric characteristics with

hit rates for positive IOD events being in the middle-ranked group while for negative IOD events hit rates being the worst performers. We hypothesize that this asymmetric characteristic is related to the asymmetry of ENSO since the linear SDM transfers the asymmetry of the ENSO forcing to the IOD directly. Any potential asymmetry in the statistical ENSO–IOD relationship is not included in the current SDM.

d. Individual IOD events

Figure 9 shows the DMI time series comparing the forecasts from individual models and the MME with

NMME original forecasts

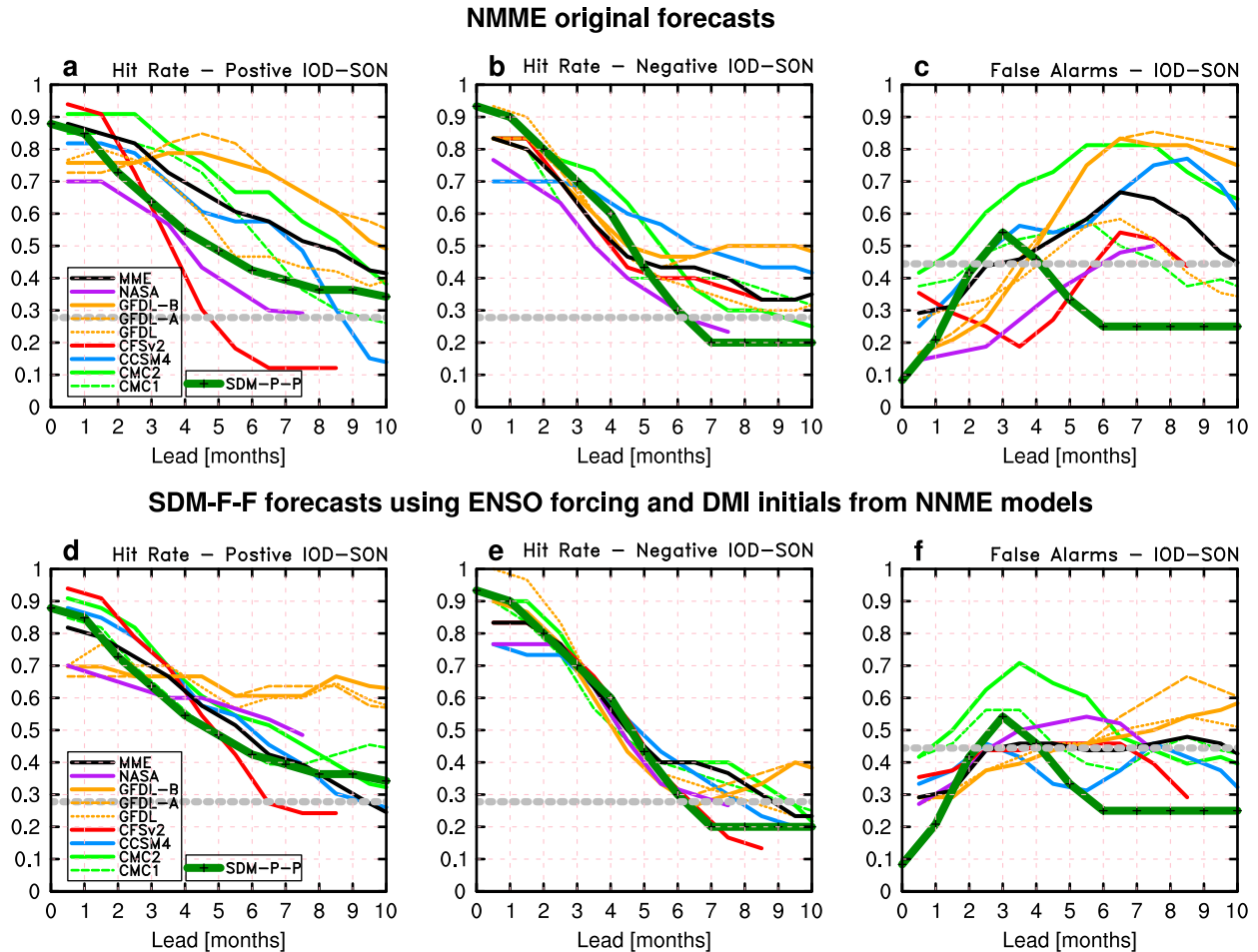## SDM-F-F forecasts using ENSO forcing and DMI initials from NNME models

FIG. 8. Hit rate for NMME original forecasts of (a) positive IOD events, (b) negative IOD events, and (c) false alarm rate for both positive and negative events in SON that exceed 0.5 observed standard deviation (0.3 K). Abscissa is lead time in months and ordinate is the percentage. Dashed gray lines in (a)–(c) are observed frequency of occurrence of positive, negative, and neutral events, respectively. A 1–2–1 filter across lead time was applied to the hit rate and false alarm rate prior to plotting. (d)–(f) As in (a)–(c), respectively, but for the SDM-F-F forecasts that use forecasted ENSO forcing and DMI initial conditions from NMME models and MME.

observations throughout the 1982–2019 period. The DMI time series for SDM-F-F and SDM-P-P forecasts are shown in Fig. 10. The forecasts shown at 0.5-, 2.5-, 4.5-, 6.5-, and 8.5-month lead times are generally matching the major patterns seen in the observations successfully, but their agreement weakens as expected with increasing lead times.

There is a large event-by-event forecast skill diversity evident for the IOD predictions among the NMME models (Fig. 9). This diversity arises from different contributions of ocean–atmosphere coupled processes that contribute to the development of the Indian Ocean dipole (Tanizaki et al. 2017). The strong positive IOD events of 1997 and 2015, which co-occurred with the super El Niño events of 1997/98 and 2015/16 in the Pacific (see observed N3.4 anomalies in Fig. 11), respectively, were well predicted by most of individual

models and by the MME even at lead times longer than two seasons, in terms of magnitude, development phase timing, and decay phase timing. Skillful predictions up two seasons in advance by most of individual models and by the MME hold also true for the 1998 and 2010 negative IOD events, which co-occurred with strong La Niña events (Fig. 11). Consistent with Zhao et al. (2019), CFSv2 failed to predict the occurrence of the 2015 IOD event one season ahead while the SDM successfully predict the event two seasons in advance.

The 2010 negative IOD event was well predicted in CCSM4, CFSv2, GFDL, GFDL-A, and GFDL-B models two seasons ahead, but was not successfully predicted by CMC1, CMC2, NASA, and the MME (Fig. 9). In contrast, Fig. 10 shows that the 2010 event was successfully predicted two seasons ahead by the SDM-F-F with forecasted ENSO forcing from CMC1
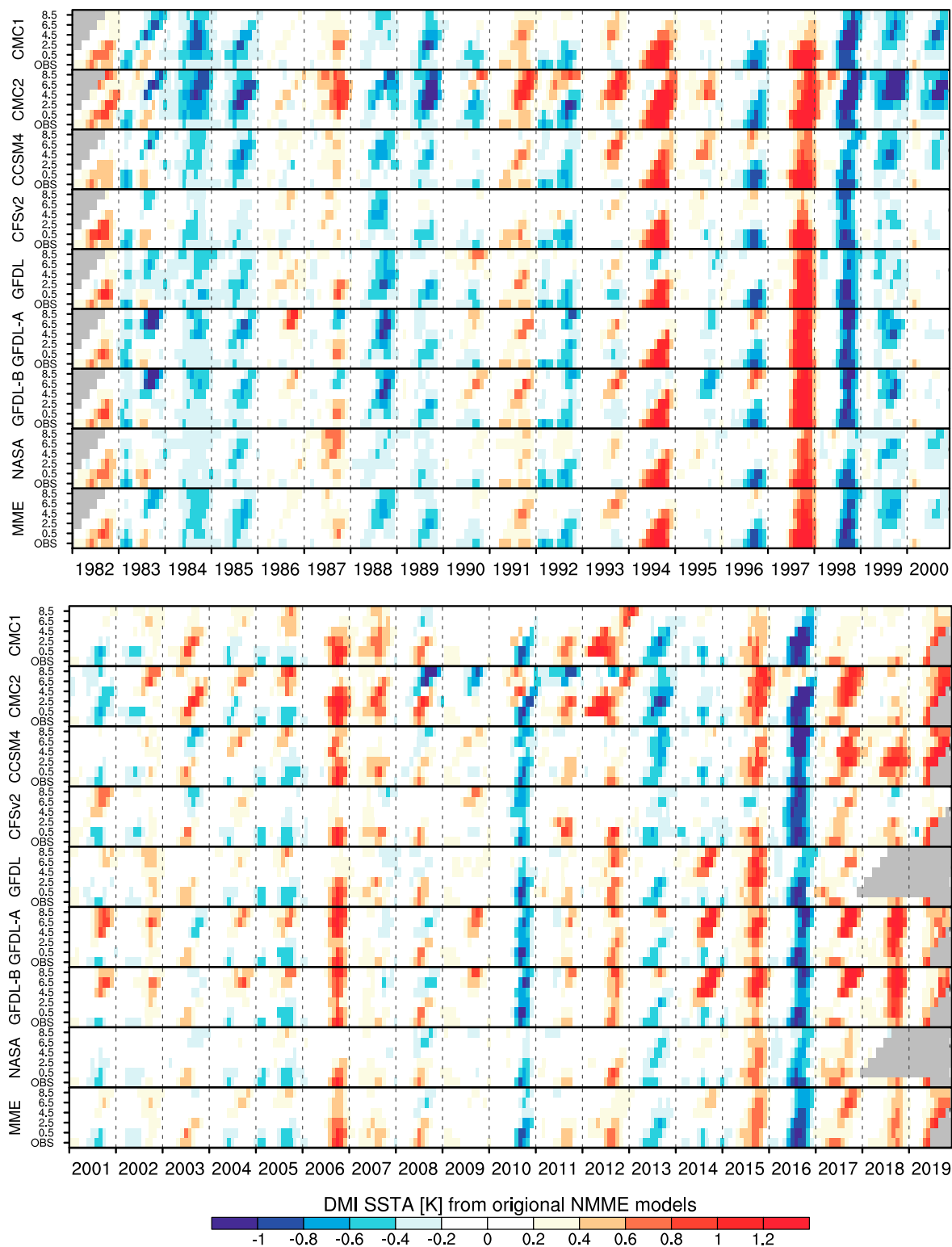
FIG. 9. Time series of running 3-month mean DMI SST anomaly observations and corresponding model forecasts (individual models and MME) for the same period from start times at 0.5-, 2.5-, 4.5-, 6.5-, and 8.5-month leads. The bottom row shows the observations, while the nine rows above show the forecasts at the five increasing lead times (months). Gray color indicates not available data (depending on model and lead time).
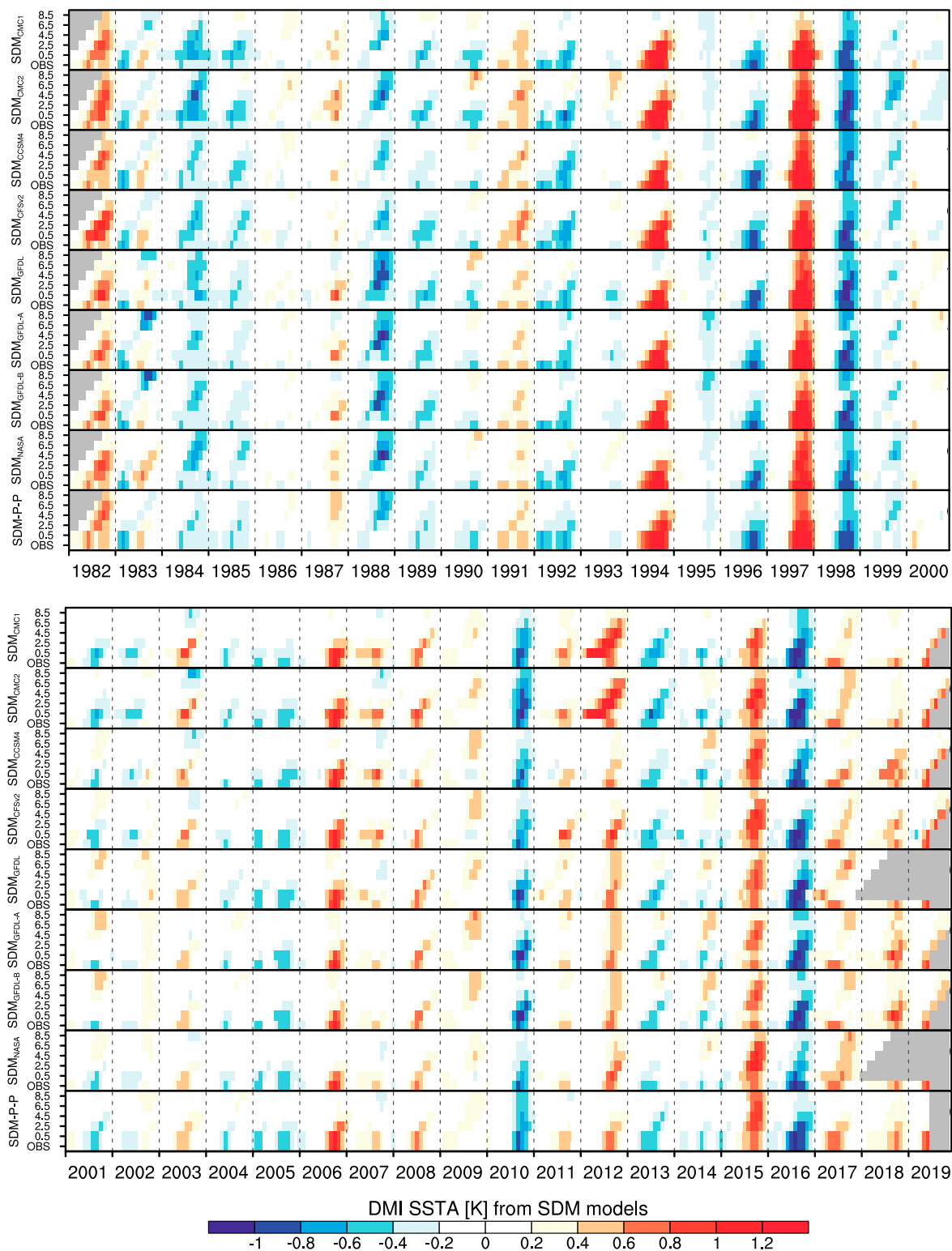
FIG. 10. As in Fig. 9, but for the SDM-F-F forecasts using forecasted ENSO forcing and DMI initial conditions from each individual model and the SDM-P-P forecasts using observed ENSO forcing and DMI initial conditions. The gray color indicates not available data (depending on model and lead time).

and CMC2, but was not successfully predicted two seasons ahead by the SDM-F-F with forecasted ENSO forcing from CCSM4, CFSv2, GFDL, GFDL-A, GFDL-B, and NASA due to the strong warm biases in the forecasted ENSO conditions at lead times of up to two seasons in these models (Fig. 11). Importantly, the 2010 event was well predicted in the SDM-P-P two seasons ahead, suggesting the dominate role of ENSO forcing in this event. It also suggests that the successful prediction of the 2010 IOD event in the original CCSM4, CFSv2, GFDL, GFDL-A, and GFDL-B forecasts are potentially due to error compensation between ENSO forcing and Indian Ocean intrinsic processes.

The strongest negative IOD event co-occurred in 2016 with a weak La Niña condition. Figure 10 shows that SDM-P-P failed to predict the development phase of the 2016 IOD event during June–August at a lead time of 2.5 months. The mature phase of the 2016 IOD event was well predicted 4.5 months ahead by SDM-P-P but up two seasons ahead by SDM-F-F forecasts using forcings from CMC1, GFDL, GFDL-A, GFDL-B, and NASA. The better performance of the SDM-F-F at longer lead times may be related to the cold biases of predicted N3.4, that is, the NMME models predicted stronger La Niña conditions than what actually occurred (Fig. 11). This supports the finding by Lim and Hendon (2017) that Indian Ocean surface and subsurface conditions may have played a dominant role in the 2016 negative IOD event based on an analysis of forecast sensitivity experiments using the Australian Bureau of Meteorology's dynamical seasonal forecast system. Lu et al. (2018b) also demonstrated that skillful predictions of the 2016 IOD event in two operational models was due to realistic representations of observed air–sea interactions and the precursor signal of early subsurface warming in the eastern Indian Ocean.

The 1994 and 2006 positive IOD conditions are two important examples of events that occurred during a neutral ENSO phase. The amplitudes and impacts of these events are comparable to the strongest 1997 IOD that co-occurred with El Niño conditions in the Pacific (Guan and Yamagata 2003; Luo et al. 2008). None of the original NMME model forecasts (including the MME) are able to predict the development phase of the 1994 IOD event during April–June (2 months in advance). Since there are only seasonally modulated damping processes controlling the evolution of the DMI in the SDM during ENSO neutral conditions, it is expected that the SDM forecasts fail to predict the development phase of ENSO-independent IOD events. Once the IOD starts gaining amplitude in JJA 1994, both the NMME models and the SDM can predict the event occurrence and

decay phase timing during October–December (OND) one season ahead (Figs. 9 and 10). This highlights that the development phase timing of ENSO-independent IOD events is very challenging to predict.

In contrast, the ENSO-independent 2006 positive IOD event was well predicted two seasons ahead by some of the NMME models (GFDL-A, GFDL-B, and CCSM4) in terms of magnitude, development phase timing, and decay phase timing. The ENSO-independent 2012 positive IOD event was predicted best by the GFDL, GFDL-A, and GFDL-B models. This suggests that the GFDL-A and GFDL-B models exhibit superior performance in predicting IOD event during a neutral ENSO state compared to the other NMME models. These events may serve as important examples that might help identify potential root causes of the low predictability in some models and higher predictability in others, thereby contributing to potential future skill improvement of ENSO-independent IOD event predictions.

A main reason for the limited IOD predictive skill in the NMME models is the considerable false alarm rate of negative/positive IOD events during neutral IOD phases (Fig. 9). Some false alarms occur ubiquitously among most of NMME models at longer lead times, such as the negative IOD events predicted for 1983, 1988, and 1999 that did not occur in reality. The same holds true for the predicted 1993 positive IOD event that did not occur. Some other false alarms are more model dependent. For instance, the false alarm of a predicted 2014 positive IOD event in GFDL, GFDL-A, and GFDL-B did not occur in other models, and was thus only weakly represented in the MME. The false alarm of the predicted 2000 negative IOD event at longer lead times in CMC1, CMC2, CCSM4, and NASA did not occur in CFSv2, GFDL, GFDL-A, and GFDL-B, and was also only weakly represented in the MME. Additionally, the observed 2017 positive IOD event reached its mature phase from May to July. However, its mature phase was wrongly predicted to occur between August and November by most of the NMME models and the MME.

The improvement of the SDM in predicting IOD events compared to the original NMME model forecasts is shown by the fewer amount of false alarms in the SDM (Fig. 10). Some false alarms (such as 1983, 1993, 2001, and 2009) in the original NMME model forecasts at longer lead times (Fig. 9) are not evident in the SDM-P-P forecast. Also they are not evident or only weakly represented in the SDM-F-F predictions that use forecasted ENSO forcing from the corresponding NMME model (Fig. 10). For example, the false alarm of the predicted 1983 negative IOD event disappears in the SDM-F-F forecasts for CMC2, CCSM4, and CFSv2. In addition, it is only weakly represented in the SDM-F-F
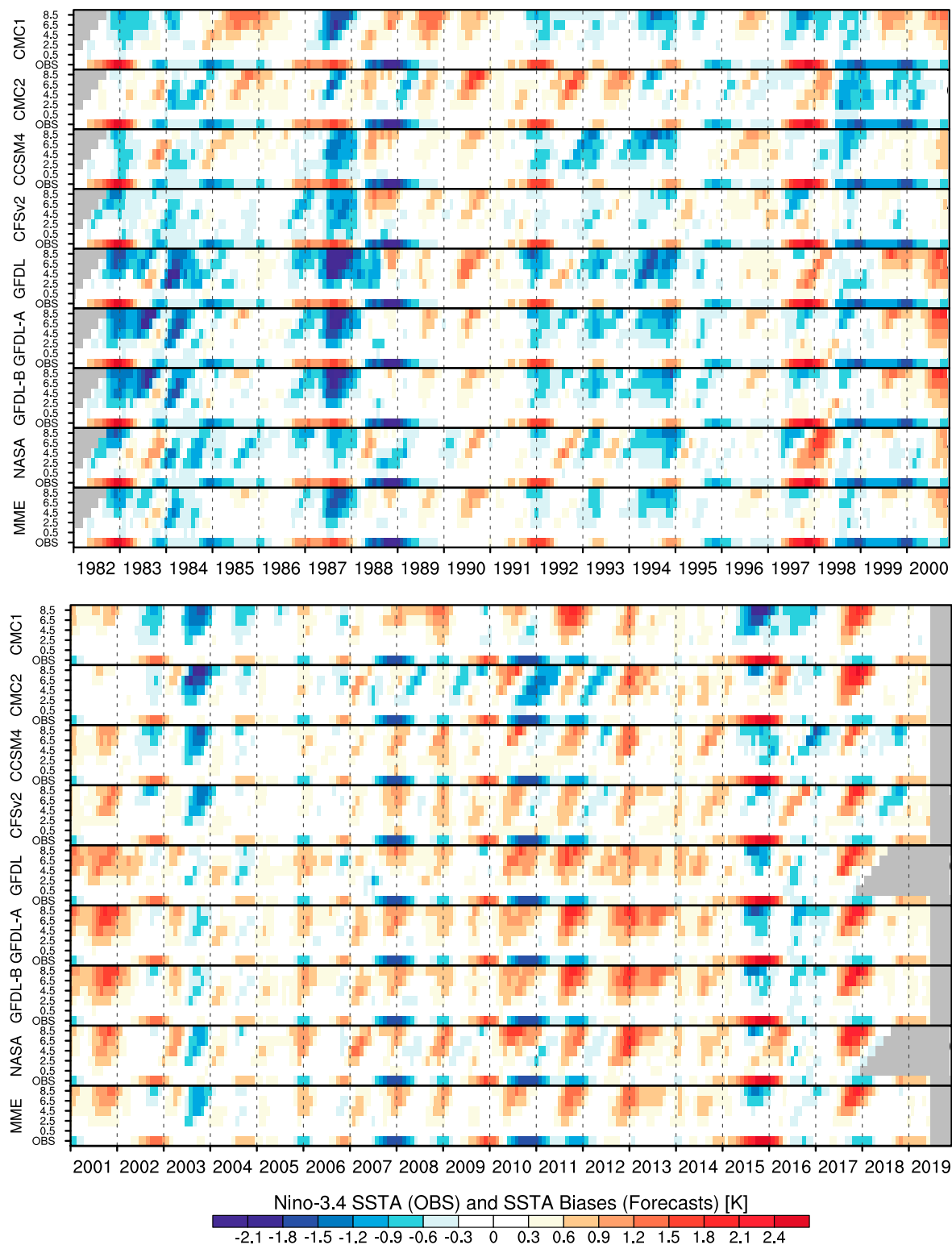
FIG. 11. Time series of running 3-month mean Niño-3.4 SST anomaly observations and biases of corresponding model forecasts (individual models and MME) for the same period from start times at 0.5-, 2.5-, 4.5-, 6.5-, and 8.5-month leads. The bottom row shows the observations, while the nine rows above show the biases at the five increasing lead times (months). The gray color indicates not available data (depending on model and lead time).

forecasts for CMC1, GFDL, GFDL-A, and GFDL-B (Fig. 10), which show considerable cold biases of the predicted N3.4 (Fig. 11). For another example, the false alarm of the predicted 2001 positive IOD event in the original forecasts weakens in the corresponding SDM-F-F predictions of CFSv2, GFDL-A, and GFDL-B (comparing Figs. 9 and 10), in which there are considerable warm biases of the predicted N3.4 (Fig. 11). These results suggest that cold and warm biases of the predicted N3.4 may cause false alarms of negative and positive IOD events, respectively, in the coupled models. Recently, Tompkins et al. (2017) demonstrated that the "overconfidence problem" in ENSO prediction is a common deficiency in most dynamical seasonal prediction systems including the NMME models. Therefore, reducing the false alarm rate in ENSO prediction should also lead to a reduction of the false alarm rate in IOD prediction.

5. Conclusions and discussion

In this study, predictability of the IOD (measured by the DMI) was studied by analyzing the hindcasts and real-time forecasts from eight NMME models with the help of a simple recently developed SDM (Stuecker et al. 2017; Zhao et al. 2019). As for the overall IOD predictive skill in original forecasts from NMME models, the MME forecast is found to be superior to the forecast of each individual model at short lead times (1.5–4.5 months). The three best performing individual models are CCSM4, GFDL-A, and GFDL-B (Fig. 1). If an ACC value of 0.5 is used as a standard of skillful predictions, we find that the MME IOD forecast is skillful up to about 4–5-month lead time, which is much longer than the skillful lead time of 2–3 months seen in ENSEMBLES (Liu et al. 2017). This indicates a gradual improvement of IOD predictions in current seasonal forecast systems.

Although CFSv2 and CMC2 are top-ranked models in predicting ENSO, they exhibit poor predictive skill for the IOD in terms of both ACC and RMSE (Fig. 1). The poor IOD prediction skills seen in CFSv2, CMC2, as well as CMC1, are likely related to a poor representation of the observed statistical and physical IOD–ENSO relationship in these models (Fig. 2). This attribution statement is further supported by significantly improved skills of SDM-F-F DMI forecasts that use forecasted ENSO forcing from these three models, in which the observed IOD–ENSO relationship is well reproduced (Figs. 3 and 4). In general, the skills for SDM-F-F DMI forecasts that use forecasted ENSO forcing from other NMME models were better than those for the NMME original DMI forecasts. Importantly, the SDM-P-P DMI forecasts demonstrate superior performance of IOD

predictions than the MME at lead-times of 4.5–9.5 months in terms of both ACC and RMSE skill scores (Figs. 3 and 4), shedding light on the potential room for improvement of IOD prediction skill by improving ENSO predictions.

An analysis on the effects of seasonality verifies the existence of the winter predictability barrier for the IOD predictions in NMME models. This is consistent with the low predictability limit of monthly SSTs over southeastern tropical Indian Ocean discussed by Li and Ding (2013). Comparing SDM-F-F and SDM-P-P forecasts confirms that the winter predictability barrier may not be overcome using the SDM approach. Most of models and the MME exhibit a slight recovery of ACC and NRMSE skills at target months in late boreal winter and early spring. This skill rebound does not exist in the original IOD forecasts from CMC1, CMC2, and CFSv2, but is seen in the corresponding SDM-F-F forecasts for these three models, suggesting that the winter predictability barrier for IOD predictions is strongly influenced by ENSO, consistent with Ding and Li (2012).

There is large event-by-event skill diversity for the IOD predictions among NMME models. The superior performance of the SDM is evident for most of the IOD events, especially IOD events that co-occurred with strong El Niño/La Niña events. Moreover, many false alarms at longer lead times in the original forecasts of NMME models and the MME forecast are much reduced in the SDM-F-F forecasts for the corresponding individual model. Our results also suggest that cold/warm biases of the predicted N3.4 may cause false alarms of negative/positive IOD events in the coupled models.

Our results have important implications for future model development. The physical basis for the IOD–ENSO relationship in the SDM is that the anomalous surface wind stress and heat fluxes induced by the seasonally modulated atmospheric ENSO (C-mode) circulation in the Indian Ocean are represented by the right-hand side ENSO forcing term in Eq. (1). Therefore, we suspect that the biases in the IOD–ENSO relationship in some CGCMs mostly arise from biases in the ENSO atmospheric teleconnection to the Indian Ocean, involving processes (and parameterizations in coupled models) of convection, clouds, and radiation. However, here we did not eliminate other potential predictability sources that might arise from Indian Ocean intrinsic dynamics via recharge oscillator dynamics (Feng and Meyers 2003; McPhaden and Nagura 2014; Wang et al. 2016; Lim and Hendon 2017; Lu et al. 2018b). Additionally, previous studies reported that the ENSO–IOD relationship varies depending on different ENSO types (Zhang et al. 2015;

Fan et al. 2017). Our SDM could potentially be further improved in the future by including Indian Ocean subsurface heat content as an additional resolved process and by considering different ENSO flavors.

REFERENCES

Abram, N. J., M. K. Gagan, M. T. McCulloch, J. Chappell, and W. S. Hantoro, 2003: Coral reef death during the 1997 Indian Ocean Dipole linked to Indonesian wildfires. *Science*, **301**, 952–955, https://doi.org/10.1126/science.1083841.

Annamalai, H., R. Murtugudde, J. Potemra, S. P. Xie, P. Liu, and B. Wang, 2003: Coupled dynamics over the Indian Ocean: Spring initiation of the zonal mode. *Deep-Sea Res. II*, **50**, 2305–2330, https://doi.org/10.1016/S0967-0645(03)00058-4.

Ashok, K., Z. Guan, and T. Yamagata, 2001: Impact of the Indian Ocean Dipole on the relationship between the Indian monsoon rainfall and ENSO. *Geophys. Res. Lett.*, **28**, 4499–4502, https://doi.org/10.1029/2001GL013294.

——, ——, and ——, 2003: Influence of the Indian Ocean Dipole on the Australian winter rainfall. *Geophys. Res. Lett.*, **30**, 1821, https://doi.org/10.1029/2003GL017926.

Barnston, A. G., and M. K. Tippett, 2013: Predictions of Niño-3.4 SST in CFSv1 and CFSv2: A diagnostic comparison. *Climate Dyn.*, **41**, 1615–1633, https://doi.org/10.1007/s00382-013-1845-2.

——, ——, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, https://doi.org/10.1175/BAMS-D-11-00111.1.

——, ——, H. M. van den Dool, and D. A. Unger, 2015: Toward an improved multimodel ENSO prediction. *J. Appl. Meteor. Climatol.*, **54**, 1579–1595, https://doi.org/10.1175/JAMC-D-14-0188.1.

——, ——, M. Ranganathan, and M. L. L'Heureux, 2019: Deterministic skill of ENSO predictions from the North American multimodel ensemble. *Climate Dyn.*, **53**, 7215–7234, https://doi.org/10.1007/S00382-017-3603-3.

Behera, S. K., J.-J. Luo, S. Masson, S. A. Rao, H. Sakuma, and T. Yamagata, 2006: A CGCM study on the interaction between IOD and ENSO. *J. Climate*, **19**, 1688–1705, https://doi.org/10.1175/JCLI3797.1.

Cai, W., T. Cowan, and M. Raupach, 2009: Positive Indian Ocean Dipole events precondition southeast Australia bushfires. *Geophys. Res. Lett.*, **36**, L19710, https://doi.org/10.1029/2009GL039902.

——, P. van Rensch, T. Cowan, and H. H. Hendon, 2011: Teleconnection pathways of ENSO and the IOD and the mechanisms for impacts on Australian rainfall. *J. Climate*, **24**, 3910–3923, https://doi.org/10.1175/2011JCLI4129.1.

Chen, L.-C., H. van den Dool, E. Becker, and Q. Zhang, 2017: ENSO precipitation and temperature forecasts in the North American multimodel ensemble: Composite analysis and validation. *J. Climate*, **30**, 1103–1125, https://doi.org/10.1175/JCLI-D-15-0903.1.

Crétat, J., P. Terray, S. Masson, and K. P. Sooraj, 2018: Intrinsic precursors and timescale of the tropical Indian Ocean Dipole: Insights from partially decoupled numerical experiment. *Climate Dyn.*, **51**, 1311–1332, https://doi.org/10.1007/S00382-017-3956-7.

DelSole, T., J. Nattala, and M. K. Tippett, 2014: Skill improvement from increased ensemble size and model diversity. *Geophys. Res. Lett.*, **41**, 7331–7342, https://doi.org/10.1002/2014GL060133.

Ding, R., and J. Li, 2012: Influences of ENSO teleconnection on the persistence of sea surface temperature in the tropical Indian ocean. *J. Climate*, **25**, 8177–8195, https://doi.org/10.1175/JCLI-D-11-00739.1.

Doi, T., A. Storto, S. K. Behera, A. Navarra, and T. Yamagata, 2017: Improved prediction of the Indian Ocean dipole mode by use of subsurface ocean observations. *J. Climate*, **30**, 7953–7970, https://doi.org/10.1175/JCLI-D-16-0915.1.

Dommenget, D., and M. Jansen, 2009: Predictions of Indian ocean SST indices with a simple statistical model: A null hypothesis. *J. Climate*, **22**, 4930–4938, https://doi.org/10.1175/2009JCLI2846.1.

Fan, L., Q. Liu, C. Wang, and F. Guo, 2017: Indian Ocean dipole modes associated with different types of ENSO development. *J. Climate*, **30**, 2233–2249, https://doi.org/10.1175/JCLI-D-16-0426.1.

Feng, M., and G. Meyers, 2003: Interannual variability in the tropical Indian Ocean: A two-year time-scale of Indian Ocean Dipole. *Deep-Sea Res. II*, **50**, 2263–2284, https://doi.org/10.1016/S0967-0645(03)00056-0.

Feng, R., W. Duan, and M. Mu, 2014: The "winter predictability barrier" for IOD events and its error growth dynamics: Results from a fully coupled GCM. *J. Geophys. Res. Oceans*, **119**, 8688–8708, https://doi.org/10.1002/2014JC010473.

Guan, Z. Y., and T. Yamagata, 2003: The unusual summer of 1994 in East Asia: IOD teleconnections. *Geophys. Res. Lett.*, **30**, 1544, https://doi.org/10.1029/2002GL016831.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus*, **57A**, 219–233, https://doi.org/10.3402/TELLUSA.V57I3.14657.

Han, W., and Coauthors, 2014: Intensification of decadal and multi-decadal sea level variability in the western tropical Pacific during recent decades. *Climate Dyn.*, **43**, 1357–1379, https://doi.org/10.1007/s00382-013-1951-1.

Hashizume, M., L. F. Chaves, and N. Minakawa, 2012: Indian Ocean Dipole drives malaria resurgence in East African highlands. *Sci. Rep.*, **2**, 269, https://doi.org/10.1038/srep00269.

Iizuka, S., T. Matsuura, and T. Yamagata, 2000: The Indian Ocean SST dipole simulated in a coupled general circulation model. *Geophys. Res. Lett.*, **27**, 3369–3372, https://doi.org/10.1029/2000GL011484.

Kajtar, J. B., A. Santoso, M. H. England, and W. Cai, 2017: Tropical climate variability: Interactions across the Pacific, Indian, and Atlantic Oceans. *Climate Dyn.*, **48**, 2173–2190, https://doi.org/10.1007/s00382-016-3199-z.

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.

Kumar, A., and M. P. Hoerling, 2000: Analysis of a conceptual model of seasonal climate variability and implications for seasonal prediction. *Bull. Amer. Meteor. Soc.*, **81**, 255–264, https://doi.org/10.1175/1520-0477(2000)081<0255:AOACMO>2.3.CO;2.

——, M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang, 2012: An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Mon. Wea. Rev.*, **140**, 3003–3016, https://doi.org/10.1175/MWR-D-11-00335.1.

——, Z.-Z. Hu, B. Jha, and P. Peng, 2017: Estimating ENSO predictability based on multi-model hindcasts. *Climate Dyn.*, **48**, 39–51, https://doi.org/10.1007/s00382-016-3060-4.

Li, J., and R. Ding, 2013: Temporal–spatial distribution of the predictability limit of monthly sea surface temperature in the global oceans. *Int. J. Climatol.*, **33**, 1936–1947, https://doi.org/10.1002/joc.3562.

Li, T., B. Wang, C.-P. Chang, and Y. Zhang, 2003: A theory for the Indian Ocean dipole–zonal mode. *J. Atmos. Sci.*, **60**, 2119–2135, https://doi.org/10.1175/1520-0469(2003)060<2119:ATFTIO>2.0.CO;2.

Lim, E.-P., and H. H. Hendon, 2017: Causes and predictability of the negative Indian Ocean Dipole and its impact on La Niña during 2016. *Sci. Rep.*, **7**, 12619, https://doi.org/10.1038/s41598-017-12674-z.

Liu, H., Y. Tang, D. Chen, and T. Lian, 2017: Predictability of the Indian Ocean Dipole in the coupled models. *Climate Dyn.*, **48**, 2005–2024, https://doi.org/10.1007/s00382-016-3187-3.

Loschnigg, J., G. A. Meehl, P. J. Webster, J. M. Arblaster, and G. P. Compo, 2003: The Asian monsoon, the tropospheric biennial oscillation, and the Indian ocean zonal mode in the NCAR CSM. *J. Climate*, **16**, 1617–1642, https://doi.org/10.1175/1520-0442(2003)016<1617:TAMTTB>2.0.CO;2.

Lu, B., H.-L. Ren, R. Eade, and M. Andrews, 2018a: Indian Ocean SST modes and their impacts as simulated in BCC_CSM1.1(m) and HadGEM3. *Adv. Atmos. Sci.*, **35**, 1035–1048, https://doi.org/10.1007/s00376-018-7279-3.

——, and Coauthors, 2018b: An extreme negative Indian Ocean Dipole event in 2016: Dynamics and predictability. *Climate Dyn.*, **51**, 89–100, https://doi.org/10.1007/s00382-017-3908-2.

Luo, J.-J., S. Masson, S. Behera, S. Shingu, and T. Yamagata, 2005: Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *J. Climate*, **18**, 4474–4497, https://doi.org/10.1175/JCLI3526.1.

——, ——, ——, and T. Yamagata, 2007: Experimental forecasts of the Indian Ocean dipole using a coupled OAGCM. *J. Climate*, **20**, 2178–2190, https://doi.org/10.1175/JCLI4132.1.

——, S. Behera, Y. Masumoto, H. Sakuma, and T. Yamagata, 2008: Successful prediction of the consecutive IOD in 2006 and 2007. *Geophys. Res. Lett.*, **35**, L14S02, https://doi.org/10.1029/2007GL032793.

——, C. Yuan, W. Sasaki, S. K. Behera, Y. Masumoto, T. Yamagata, J.-Y. Lee, and S. Masson, 2016: Current status of intraseasonal–seasonal-to-interannual prediction of the Indo-Pacific climate. *Indo-Pacific Climate Variability and Predictability*, S. K. Behera and T. Yamagata, Eds., World Scientific, 63–107.

McPhaden, M. J., and M. Nagura, 2014: Indian Ocean dipole interpreted in terms of recharge oscillator theory. *Climate Dyn.*, **42**, 1569–1586, https://doi.org/10.1007/s00382-013-1765-1.

Newman, M., and P. D. Sardeshmukh, 2017: Are we near the predictability limit of tropical Indo-Pacific sea surface temperatures? *Geophys. Res. Lett.*, **44**, 8520–8529, https://doi.org/10.1002/2017GL074088.

Qiu, Y., W. Cai X. Guo, and B. Ng, 2014: The asymmetric influence of the positive and negative IOD events on China's rainfall. *Sci. Rep.*, **4**, 4943, https://doi.org/10.1038/srep04943.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625, https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

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

Saji, N. H., and T. Yamagata, 2003: Structure of SST and surface wind variability during Indian Ocean dipole mode events: COADS observations. *J. Climate*, **16**, 2735–2751, https://doi.org/10.1175/1520-0442(2003)016<2735:SOSASW>2.0.CO;2.

——, B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360–363, https://doi.org/10.1038/43854.

Schott, F. A., S.-P. Xie, and J. P. McCreary Jr., 2009: Indian Ocean circulation and climate variability. *Rev. Geophys.*, **47**, RG1002, https://doi.org/10.1029/2007RG000245.

Shi, L., H. H. Hendon, O. Alves, J.-J. Luo, M. Balmaseda, and D. Anderson, 2012: How predictable is the Indian Ocean dipole? *Mon. Wea. Rev.*, **140**, 3867–3884, https://doi.org/10.1175/MWR-D-12-00001.1.

Song, Q., G. A. Vecchi, and A. J. Rosati, 2008: Predictability of the Indian Ocean sea surface temperature anomalies in the GFDL coupled model. *Geophys. Res. Lett.*, **35**, L02701, https://doi.org/10.1029/2007GL031966.

Stanski, H. R., L. J. Wilson, and W. R. Burrows, 1989: Survey of common verification methods in meteorology. 2nd ed. Research Rep. MSRB 89-5, WWW Tech. Rep. 8, WMO/TD 358, World Meteorological Organization, http://www.cawcr.gov.au/projects/verification/Stanski_et_al/Stanski_et_al.html.

Stuecker, M. F., A. Timmermann, F.-F. Jin, S. McGregor, and H.-L. Ren, 2013: A combination mode of the annual cycle and the El Niño/Southern Oscillation. *Nat. Geosci.*, **6**, 540–544, https://doi.org/10.1038/ngeo1826.

——, F.-F. Jin, A. Timmermann, and S. McGregor, 2015: Combination mode dynamics of the anomalous northwest Pacific anticyclone. *J. Climate*, **28**, 1093–1111, https://doi.org/10.1175/JCLI-D-14-00225.1.

——, A. Timmermann, F.-F. Jin, Y. Chikamoto, W. Zhang, A. T. Wittenberg, E. Widiasih, and S. Zhao, 2017: Revisiting ENSO/Indian Ocean Dipole phase relationships. *Geophys. Res. Lett.*, **44**, 2481–2492, https://doi.org/10.1002/2016GL072308.

Takaya, A., Y. Morioka, and S. K. Behera, 2014: Role of climate variability in the heatstroke death rates of Kanto region in Japan. *Sci. Rep.*, **4**, 5655, https://doi.org/10.1038/SREP05655.

Tanizaki, C., T. Tozuka, T. Doi, and T. Yamagata, 2017: Relative importance of the processes contributing to the development of SST anomalies in the eastern pole of the Indian Ocean Dipole and its implication for predictability. *Climate Dyn.*, **49**, 1289–1304, https://doi.org/10.1007/s00382-016-3382-2.

Tompkins, A. M., and Coauthors, 2017: The climate-system historical forecast project: Providing open access to seasonal forecast ensembles from centers around the globe. *Bull. Amer. Meteor. Soc.*, **98**, 2293–2301, https://doi.org/10.1175/BAMS-D-16-0209.1.

Wajsowicz, R. C., 2005: Potential predictability of tropical Indian Ocean SST anomalies. *Geophys. Res. Lett.*, **32**, L24702, https://doi.org/10.1029/2005GL024169.

——, 2007: Seasonal-to-interannual forecasting of tropical Indian Ocean sea surface temperature anomalies: Potential predictability

and barriers. *J. Climate*, **20**, 3320–3343, https://doi.org/10.1175/JCLI4162.1.

Wang, B., and Coauthors, 2009: Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Climate Dyn.*, **33**, 93–117, https://doi.org/10.1007/s00382-008-0460-0.

Wang, H., R. Murtugudde, and A. Kumar, 2016: Evolution of Indian Ocean dipole and its forcing mechanisms in the absence of ENSO. *Climate Dyn.*, **47**, 2481–2500, https://doi.org/10.1007/s00382-016-2977-y.

——, A. Kumar, R. Murtugudde, B. Narapusetty, and K. L. Seip, 2019: Covariations between the Indian Ocean Dipole and ENSO: A modeling study. *Climate Dyn.*, **53**, 5743–5761, https://doi.org/10.1007/S00382-019-04895-X.

Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. R. Leben, 1999: Coupled ocean–atmosphere dynamics in the Indian Ocean during 1997–98. *Nature*, **401**, 356–360, https://doi.org/10.1038/43848.

Wu, Y., and Y. Tang, 2019: Seasonal predictability of the tropical Indian Ocean SST in the North American multimodel ensemble. *Climate Dyn.*, **53**, 3361–3372, https://doi.org/10.1007/s00382-019-04709-0.

Xue, Y., B. Huang, Z.-Z. Hu, A. Kumar, C. Wen, D. Behringer, and S. Nadiga, 2011: An assessment of oceanic variability in the NCEP climate forecast system reanalysis. *Climate Dyn.*, **37**, 2511–2539, https://doi.org/10.1007/s00382-010-0954-4.

Yamagata, T., S. K. Behera, J.-J. Luo, S. Masson, M. R. Jury, and S. A. Rao, 2004: Coupled ocean-atmosphere variability in the Tropical Indian Ocean. *Earth's Climate: The Ocean-Atmosphere Interaction*, *Geophys. Monogr.*, Vol. 147, Amer. Geophys. Union, 189–212.

Yang, Y., S.-P. Xie, L. Wu, Y. Kosaka, N.-C. Lau, and G. A. Vecchi, 2015: Seasonality and predictability of the Indian Ocean dipole mode: ENSO forcing and internal variability. *J. Climate*, **28**, 8021–8036, https://doi.org/10.1175/JCLI-D-15-0078.1.

Yuan, C., and T. Yamagata, 2015: Impacts of IOD, ENSO and ENSO Modoki on the Australian winter wheat yields in recent decades. *Sci. Rep.*, **5**, 17252, https://doi.org/10.1038/srep17252.

Yuan, Y., H. Yang, W. Zhou, and C. Y. Li, 2008: Influences of the Indian Ocean dipole on the Asian summer monsoon in the following year. *Int. J. Climatol.*, **28**, 1849–1859, https://doi.org/10.1002/joc.1678.

Zhang, W., Y. Wang, F.-F. Jin, M. F. Stuecker, and A. G. Turner, 2015: Impact of different El Niño types on the El Niño/IOD relationship. *Geophys. Res. Lett.*, **42**, 8570–8576, https://doi.org/10.1002/2015GL065703.

Zhao, M., and H. H. Hendon, 2009: Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model. *Quart. J. Roy. Meteor. Soc.*, **135**, 337–352, https://doi.org/10.1002/qj.370.

Zhao, S., F.-F. Jin, and M. F. Stuecker, 2019: Improved predictability of the Indian Ocean Dipole using seasonally modulated ENSO forcing forecasts. *Geophys. Res. Lett.*, **46**, 9980–9990, https://doi.org/10.1029/2019GL084196.