Prediction of Diverse Boreal Summer Intraseasonal Oscillation in the GFDL SPEAR Model

BAOQIANG XIANG, a,b BIN WANG, c GUOSEN CHEN, d AND THOMAS L. DELWORTHA

a NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey
b Cooperative Programs for the Advancement of Earth System Science, University Corporation for Atmospheric Research, Boulder, Colorado
c International Pacific Research Center, University of Hawai‘i at Mānoa, Honolulu, Hawaii
d Earth System Modeling Center, Key Laboratory of Meteorological Disaster of Ministry of Education, Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing, China

(Manuscript received 4 October 2023, in final form 18 January 2024, accepted 22 January 2024)

ABSTRACT: Boreal summer intraseasonal oscillation (BSISO) is a primary source of predictability for summertime weather and climate on the subseasonal-to-seasonal (S2S) time scale. Using the GFDL SPEAR S2S prediction system, we evaluate the BSISO prediction skills based on 20-yr (2000–19) hindcast experiments with initializations from May to October. It is revealed that the overall BSISO prediction skill using all hindcasts reaches out to 22 days as measured by BSISO indices before the bivariate anomalous correlation coefficient (ACC) drops below 0.5. Results also show that the northeastward-propagating canonical BSISO (CB) event has a higher prediction skill than the northward dipole BSISO (DB) event (28 vs 23 days). This is attributed to CB’s more periodic nature, resulting in its longer persistence, while DB events are more episodic accompanied by a rapid demise after reaching maximum enhanced convection over the equatorial Indian Ocean. From a forecaster’s perspective, a precursory strong Kelvin wave component in the equatorial western Pacific signifies the subsequent development of a CB event, which is likely more predictable. Investigation of individual CB events shows a large interevent spread in terms of their prediction skills. For CB, the events with weaker and fluctuating amplitude during their lifetime have relatively lower prediction skills likely linked to their weaker convection-circulation coupling. Interestingly, the prediction skills of individual CB events tend to be relatively higher and less scattered during late summer (August–October) than those in early summer (May–July), suggestive of the seasonal modulation on the evolution and predictability of BSISO.

SIGNIFICANCE STATEMENT: The advance of subseasonal-to-seasonal (S2S) prediction largely depends on dynamical models’ ability to predict some major intrinsic modes in the climate system, including the boreal summer intraseasonal oscillation (BSISO). Using a newly developed S2S prediction system, we thoroughly evaluated its performance in predicting BSISO, and revealed the skill dependence on the BSISO propagation diversity. Here we provide physical explanations of what influences the BSISO predictions and identify different precursory signals for two types of BSISO, which have important implications for operational forecasts.

KEYWORDS: Madden-Julian oscillation; Summer/warm season; Hindcasts; Coupled models; Intraseasonal variability; Tropical variability

1. Introduction

With the rapid growth of societal demands, there has been an increasing desire to improve the prediction skill of the subseasonal-to-seasonal (S2S) prediction, with the time scale ranging from 10 days to 1 season. Nowadays, S2S prediction is emerging as a frontier and one of the top priorities of the forecasting community. In the past decade, substantial progress on S2S prediction has been achieved primarily due to advances in model physics, improvement in initialization, and increased model resolutions (e.g., Vitart 2014). Many modeling and operational centers have launched their real-time, or near-real-time, subseasonal predictions (Pegion et al. 2019; Vitart et al. 2017). Meanwhile, although a notable gap still exists between the
current S2S predictions and the end-user needs, the community has been making efforts to enhance the application and usefulness of the S2S prediction products for various stakeholders, ranging from agriculture to water resource management to the public health (e.g., White et al. 2022).

Despite the great importance of S2S predictions, it has been recognized that the S2S predictions tend to be more challenging than the weather and seasonal predictions because of the lack of predictability sources and the rapid loss of initial memory in the Earth system (Robertson et al. 2020). Factors determining a dynamic model's S2S prediction skills are complex. First, the S2S prediction skill heavily relies on the fidelity of initial conditions (e.g., Meehl et al. 2021), especially for relatively short-term forecasts. Second, a major challenge for S2S predictions is the model's mean state drifting away from the initialized state to its own mean state (Mariotti et al. 2018). Hence, a model with more realistic mean states may minimize the initialization shock and model drift issues and have a better chance to capture the teleconnections excited from the tropics, facilitating a skillful S2S prediction. Third, one crucial element influencing a model's S2S prediction skill is its ability to simulate and predict some internally generated major modes of climate variability. Those include some major modes in the extratropics, such as the Pacific–North American (PNA) teleconnection, North Pacific Oscillation (NPO), North Atlantic Oscillation (NAO), Scandinavian pattern (SCA), east Atlantic–west Russia pattern (EA-WR), and the extratropical intraseasonal oscillations (EISO) (Zhu et al. 2023). Meanwhile, some modes over the tropics are also key S2S predictability sources, such as El Niño–Southern Oscillation (ENSO), Madden–Julian oscillation (MJO) (Madden and Julian 1972), and boreal summer intraseasonal oscillation (BSISO) (Wang and Xie 1997). Some previous studies have demonstrated the importance of those modes in predicting other meteorological fields at the S2S time scale (Black et al. 2017; Xiang et al. 2020, 2019; Yang et al. 2023).

The wintertime MJO is characterized by prominent eastward circumglobal propagation over the tropics. A suite of statistic diagnostics (Walisser et al. 2009) and dynamics-oriented diagnostics (Wang et al. 2018) for MJO were designed to assess the models’ intrinsic modes. The BSISO displays distinctive and more complicated evolution characteristics with the coexistence of eastward, northward, and northwestward propagations, mainly anchored in the Asian monsoon region (Kikuchi et al. 2012; Wang and Rui 1990; Wang and Xie 1997). In the literature, the BSISO is also known as the summertime MJO, monsoon intraseasonal oscillation (MISO) (Sharmila et al. 2013), and boreal summer intraseasonal variability (BSISV) (Annamalai and Sperber 2005). In this paper, we refer to it as BSISO by the authors’ preference. Several theoretical and modeling studies have sought to explain the physical processes governing the northward propagation of BSISO, including the interaction between tropical wave dynamics and the mean easterly vertical wind shear (Drbohlav and Wang 2005; Jiang et al. 2004; Wang and Xie 1996, 1997), meridional moisture advection (Jiang et al. 2004), air–sea interactions (Fu et al. 2003), and convective momentum transport (Kang et al. 2010). The literature illustrates that BSISO has pronounced impacts on synoptic-scale low pressure systems (Kikuchi 2021), precipitation extremes (Muhammad and Lubis 2023), monsoon onset (Li et al. 2016), tropical cyclones (Kikuchi and Wang 2010; Moon et al. 2018; Nakano et al. 2021, 2015; Yoshida et al. 2014), and midlatitude weather (Moon et al. 2013). Consequently, a skillful prediction of BSISO could allow for better S2S predictions of Asian climate and extremes (Chang et al. 2021; Lee et al. 2017; Shibuya et al. 2021).

A wealth of literature has been focused on predicting the wintertime MJO, and some state-of-the-art dynamical models have demonstrated their useful MJO prediction beyond four weeks (Kim et al. 2018; Xiang et al. 2021). Unlike the extensive studies on the wintertime MJO prediction, the BSISO prediction in current dynamical models has received relatively less attention (Jie et al. 2017; Lee et al. 2015). More importantly, what impacts the prediction skills of BSISO in a model remains largely unknown.

A new S2S prediction system has been developed recently (Xiang et al. 2021), using the Geophysical Fluid Dynamics Laboratory (GFDL) Seamless System for Prediction and Earth System Research (SPEAR) model (Delworth et al. 2020). Measured by the Wheeler and Hendon (2004) indices (WH04), the SPEAR model shows an encouraging prediction skill of 30 days in predicting the wintertime MJO (Xiang et al. 2021). The SPEAR model also displays distinctly different prediction skills for four types of MJO. For example, the fast-propagating MJO shows a considerably higher prediction skill than other types of MJO, especially the standing MJO. To obtain a complete picture of this model’s ability to predict the intrinsic intraseasonal modes, here we evaluate the predictions of the summertime counterpart of MJO (i.e., BSISO) in this system. Particularly, the focus is placed on assessing the predictions of two distinctive clusters of BSISO and exploring the factors influencing the BSISO prediction skills.

The paper is organized as follows. Section 2 introduces the model, hindcast experiments, as well as the analysis methodology. Section 3 describes the overall BSISO prediction skill and its skill dependence on the BSISO diversity. Section 4 presents the potential factors influencing the BSISO predictions. We end with a summary and discussion in section 5.

2. Model, experiments, and analysis methodology

2.1. Model and experiments

The model used in this study is the GFDL SPEAR model (Delworth et al. 2020). The SPEAR is a fully coupled climate model with atmospheric and land models identical to AM4.0/ LM4.0 (Zhao et al. 2018a,b) but with a dynamical vegetation model, and an ocean model (MOM6) and sea ice model (SIS2) (Adcroft et al. 2019). Three resolutions of atmospheric and land model configurations were developed for SPEAR and here we use the intermediate horizontal atmospheric and land model resolution (0.5°) that coupled to a 1° ocean and sea ice model. The atmospheric AM4.0 model contains 33 vertical levels with the top of the atmosphere at 1 hPa and the MOM6 ocean model has 75 vertical levels. The reader is referred to Delworth et al. (2020) for additional details about this model.

A nudging technique is used to achieve initialization to bring the model states into close correspondence with observations.
The atmospheric nudging fields include winds, temperature, and specific humidity using the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) analysis data (6-hourly interval) (Gelaro et al. 2017). The sea surface temperature (SST) is nudged to NOAA Optimum Interpolation 1/4 Degree Daily SST Analysis (OISST v2) (Reynolds et al. 2007). Hindcast experiments were performed every five days from 2000 to 2019, and ten ensemble members were generated by using perturbed nudging strengths for both the atmospheric fields and the ocean SST. The nudging of circulation is applied to the whole atmosphere. However, the nudging of the moisture field is confined to the free atmosphere with the lowest several model layers (roughly the boundary layer) unperturbed.

The studied period is from May to October 2000–19. There are a total of 740 hindcast cases with initial conditions starting in May–October, and each has 10 members. Each hindcast is integrated for 45 days. More information regarding the initialization can be found in Xiang et al. (2021).

b. Analysis and evaluation methodology

To achieve the observational anomalies, we remove the time mean and the first three harmonics of the observational climatological annual cycle. Further, the time-mean anomalies over the previous 120 days were subtracted to largely remove the interannual variations. The anomalies for model hindcasts were calculated by first removing the model hindcast climatology and then removing the previous 120 days’ time-mean anomalies from the combination of observations and hindcasts. For instance, for the third day from the initial date, we removed the 120 days’ mean of anomalies including the first 2 days from the hindcast and 118 days from observations right before the forecast. The model hindcast climatology is obtained by averaging all years of hindcasts for a particular start time and lead time and thus the hindcast climatology is a function of calendar day and lead time.

To assess the prediction of BSISO, we adopt two commonly used indices: the Lee et al. (2013, hereafter L13) indices and the Wheeler and Hendon (2004, hereafter WH04) indices. The observed and predicted L13 indices were obtained by projecting the anomalies of two-dimensional outgoing longwave radiation (OLR) and 850-hPa zonal winds onto the first two observed leading multivariate empirical orthogonal function (EOF) modes over the Asian monsoon domain (10°S–40°N, 40°–160°E) (Figs. 1a,b) from May to October. The first EOF mode (EOF1) yields a dipole convection pattern between the tropical Indian Ocean and the western North Pacific. A dynamically coherent lower-tropospheric circulation pattern is observed with cyclonic and anticyclonic circulation in the Indian Ocean and western North Pacific, respectively (Fig. 1a). The second mode (EOF2) has a tripole convection pattern, and the enhanced convection pattern…
has a tilted structure extending from the Indian continent south-eastward to the equatorial western Pacific together with strong lower-level wind convergence (Fig. 1b). Those two modes together show a propagation pattern as the time series of EOF1 mode (PC1) attains the highest positive correlation when leading the time series of EOF2 (PC2) by 13 days (Lee et al. 2013). However, EOF1 has a larger explained variance than EOF2 (8.7% vs 5.3%) so a substantial portion of EOF1 itself represents a stationary component (Zhu and Wang 1993). The BSISO magnitude is estimated by combining the projected PC1 and PC2 (\(\text{VPC}1^2 + \text{VPC}2^2\)). Similarly, the WH04 indices were achieved by projecting the latitudinal mean (15°S–15°N) anomalies of OLR, 850-hPa zonal winds, and 200-hPa zonal winds onto the first two observed leading multivariate EOF modes. Using the above L13 and WH04 indices as predictands, the so-called bivariate anomaly correlation coefficient (ACC) is used here to measure its forecast skill following Lin et al. (2008).

To verify the hindcasts, we use the NOAA daily mean interpolated OLR data (Liebmann and Smith 1996) and European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) data as observations (CS2 2017). All data are interpolated to 1° x 1° resolution for analysis.

3. Prediction of BSISO and its dependence on the BSISO diversity

a. Overall BSISO prediction

Using the L13 indices as a metric, Fig. 1c illustrates that the useful skill from the 10-member ensemble mean with all hindcasts reaches up to 22 days measured by the criterion before the bivariate ACC drops below 0.5. The SPEAR model outperforms most of the state-of-the-art S2S models and is also comparable with ECMWF model (Jie et al. 2017). We also examine the ACC using the WH04 indices and the ensemble-mean forecast skill reaches up to 30 days (Fig. 1d), the same as the wintertime MJO prediction skill (Xiang et al. 2021). The forecast within the weather time scale (around the first 10 days) is largely deterministic and the advantage of multimember ensemble mean becomes prominent beyond that (Figs. 1c,d). Interestingly, the skills measured by the L13 indices are generally lower than those using WH04 indices (He et al. 2019; Jie et al. 2017). Note that WH04 mainly represents the tropical eastward-propagation variations, while L13 explains more northward-propagating variance and the corresponding subtropical component is typically less predictable than that over the tropics (e.g., Judt 2020).

Is the forecast skill dependent on the amplitude and phases of BSISO? The ACC is systematically lower even from day 1 for initially weak cases than the initially strong cases (supplemental Fig. 1 in the online supplemental material). Here the initially strong (weak) cases are defined as the BSISO amplitude greater (less) than 1. The initially strong cases achieve a higher skill compared to the initially weak cases (25 vs 21 days) (supplemental Fig. 1a). We also evaluate the potential skill dependence on the phases and reveal that the initially strong cases starting at phases 2, 6, and 7 have relatively lower skills than the other phases (supplemental Fig. 1b). For the definition of those eight different BSISO phases, refer to L13.

b. Prediction of the canonical BSISO and northward dipole BSISO

The above analysis considers all initializations, including those in which BSISO is inactive. Individual BSISO events vary in magnitude and propagation characteristics as documented in the literature (e.g., Fu et al. 2018; Wang and Rui 1990). Similar to the wintertime MJO diversity (Wang et al. 2019; Wei et al. 2023), Chen and Wang (2021, hereafter CW21) revealed three clusters of BSISO with distinct propagation characteristic: the canonical BSISO (CB), the northward dipole BSISO (DB), and the eastward expansion BSISO. It is straightforward to ask whether the BSISO predictions rely on the propagation characteristics. Note that the occurrence of the eastward expansion BSISO is less frequent than CB (11 vs 29), and 8 out 11 of the eastward expansion events take place during the transition seasons (May and October), which are less representative of the classical BSISO events. Hence, the emphasis is placed on the first two clusters of BSISO (CB and DB). During the studied period (2000–2019), there were a total of 29 observed CB and 11 DB events following the classification of CW21 (Table 1). For each event, it is selected when the anomalous OLR in the equatorial Indian Ocean (5°S–10°N, 80°–100°E) is below its one standard deviation, and day 0 refers to the date when the anomalous OLR in the equatorial Indian Ocean reaches its minimum (CW21). For convenience, we refer to the period before (after) day 0 as the developing (decaying) phases, and the period around day 0 as the peak phase.

We present, in Fig. 2, the composite pentad mean (from pentad 4 to pentad +4) maps of the observational anomalous OLR and lower-tropospheric circulations associated with the CB
and DB events. During the developing phases (before day 0), both CB and DB are preceded by northeastward-propagating suppressed convection anomalies in the Indian Ocean despite that the DB phase leads CB by roughly one pentad. Meanwhile, the western Pacific is dominated by northwestward-propagating signals for both cases. Apart from the similarities, dramatic differences are noted between DB and CB. First, for DB, a considerable area of the western Pacific exhibits modestly enhanced convection
convection accompanied by prominent westerly anomalies at pentad 24 (Fig. 2j), which become more concentrated and then propagate northwestward during the ensuing pentads −3 to −1 (Figs. 2k–m). CB differs markedly with dominant suppressed convection in the western Pacific (Figs. 2a–d). Second, CB displays a more pronounced Kelvin wave component in the Maritime Continent and equatorial western Pacific than DB during the peak and decaying phases, indicated by the equatorially symmetric zonal wind component with its maximum near the equator (Figs. 2d–i vs Figs. 2m–r). Third, CB has largely balanced anomalies between the enhanced and suppressed convection, but DB has predominantly suppressed convection anomalies. Further, the suppressed convection anomalies in the western Pacific last much longer in DB than in CB. Fourth, CB tends to be more symmetric with the decaying phases almost a mirror image of the developing phases, but DB shows a more asymmetric feature between those two phases. Particularly, the decaying phases of CB gradually evolve into another robust suppressed convective phase of BSISO in the Indian Ocean (Figs. 2g–i), distinguished from the rapid demise of DB (Figs. 2p–r).

The results above suggest that CB is more likely to trigger a following suppressed convective phase starting from the equatorial Indian Ocean while DB does not. In other words, CB is more like an oscillatory mode while DB is likely an episodic mode. The contrasting features are also seen in the composite PC1–PC2 phase-space diagram (Fig. 3a). CB has a relatively weaker amplitude during the developing stage but retains a stronger amplitude in the following phases, in contrast with a rapid decay of DB. CB and DB generally have a comparable PC1 during their life cycle, while PC2 is different: in the developing phases CB has a weaker PC2 than DB, and the opposite is true during the decaying phases.

To assess the prediction skills for CB and DB, we choose the hindcast cases with their initial dates from day −230 (30 days before day 0) to day +15 (15 days after day 0) so that each event has about eight to nine hindcast cases. The purpose of considering the cases with shifted initial dates to an earlier time is to cover the whole lifetime of BSISO in prediction, including the developing and decaying phases. Eventually, 252 and 102 hindcast cases are selected for CB and DB, respectively. Measured by the L13 indices, CB shows a useful skill of 28 days, while the prediction of DB appears to be limited to 23 days (Fig. 3b). A consistent skill dependence is also found using WH04 indices (33 days vs 24 days; see supplemental Fig. 2), confirming that the CB is intrinsically more predictable than DB.

The question then arises as to why the prediction of DB is more challenging than CB. To address this, we evaluate the model prediction skills at different initial phases (Fig. 4). For the cases with initialization at day −25 and −20, CB has comparable but slightly lower prediction skills than DB (Fig. 4a). It suggests that the model is equally skillful in predicting the developing phases for both CB and DB. However, the contrast is prominent for the cases initialized at day −5 and around day 0 (Fig. 4b). The skills for CB remain similar and even slightly higher than those initialized at day −25 and day −20, while the skill for DB drops suddenly after around 10 days, leading to a much lower skill than CB (17 days vs 29 days) (Fig. 4b). The contrast in predicting their decaying phases between CB and DB is largely responsible for the different prediction skills as shown in Fig. 3b. The reason behind this is related to the contrasting evolutions: CB and DB have similar evolution with northeastward propagation during the developing phases while they differ dramatically during the decaying phases (Fig. 2).
and at around day 0.

For the cases with initialization conditions at 5 days before day 0, 22 cases are used for CB and DB, respectively. (b) As in (a), but for the cases with initialization conditions at 25 and 20 days before day 0, so that a total of 58 and 20 forecast lead days. The cases are selected with initialization conditions as shown in the cases initialized at day 0 and at around day 0.

The previous section examines the overall skill for all hindcast cases as well as two groups of BSISO. A study of predictions of individual BSISO events may provide more insight into the physical processes affecting BSISO predictions. Since the observational DB events are too few to perform statistical analysis, the following analysis is centered on the more frequently occurring CB events. For each CB event, we evaluate its skill by considering all forecast cases from day −30 to day +15 so that each event has approximately eight to nine hindcast cases. Note that for some events with their day 0 occurring in May and late October, we also include some hindcast cases initialized during April and November for a fair comparison among events.

A salient interevent spread is found in terms of the bivariate ACC (supplemental Fig. 4). For some events, the ACC drops to below zero with a lead time of 30 days, while the ACC still remains above 0.7 for some other events. Why do the skills differ markedly among individual events? Addressing this issue has profound implications for identifying the “time window of opportunity” during which better predictions are achieved than in other periods.

Inspection of Fig. 7 reveals a strong skill dependence on the seasonality and also the observed BSISO amplitude. Given the considerable fluctuation for the ACC with the lead times (supplemental Fig. 4), here we use the time-averaged ACC during the first four weeks to represent the forecast skill for individual events. It is shown that the multievent mean skills tend to be lower, but not statistically significant, during May–July than in August–October (Fig. 7a). Meanwhile, the ACC is also more scattered (larger interevent spread) in May–July than in August–October, with the standard deviation of ACC in May–July about 1.7 times larger than that in August–October (Fig. 7a). During August–October, 11 out of 13 events have skills higher than 0.72, while 10 out of 16 May–July events have skills lower than 0.72. Further, 5 out of 6 with the lowest skills among all the CB events occur in May–July. It suggests that the prediction of CB tends to be more challenging in the early summer season than in the late summer season.

4. Understanding factors influencing the CB predictions

Figure 5 shows the model’s ability to predict the peak phase (around day 0) at different lead times. For both CB and DB, the model reasonably well predicts the observed convection–circulation anomalous patterns over the equatorial Indo-Pacific sector for almost all forecasts 5–20 days in advance. Additionally, their time evolution is also well captured for both CB and DB, as shown in the cases initialized at day −20 (supplemental Fig. 3). This affirms that the contrasting prediction skills between CB and DB (Fig. 3b) are not related to the prediction of their corresponding developing phases. One noteworthy point is that for DB the enhanced convection anomalies in the Indian Ocean are substantially underpredicted even with a lead time of 5 days for DB (Fig. 5g), while the model is skillful in predicting the suppressed convection anomalies and the associated anticyclonic circulations in the western North Pacific for all leads time from 5 to 20 days (Figs. 5g–j).

Then, why is the prediction skill of the decaying phases so different between CB and DB (Fig. 4b)? Figure 6 displays the composite time evolution of the observed and predicted convection–circulation patterns starting from around day 0. For the cases initialized at around day 0, the model well predicts the evolution pattern of CB (with a decaying amplitude). For DB, the model is unable to predict the observed easterly and suppressed convection anomalies in the equatorial Indian Ocean at week 3 (Fig. 6o vs Fig. 6k), together with significantly but erroneously predicted suppressed convection anomalies over the Arabian Sea at weeks 3 and 4 (Figs. 6o,p). The relatively lower prediction skill for DB is related to the weaker persistence of the enhanced convection signal over the Indian Ocean. In contrast, CB has longer persistence by triggering a successive suppressed convective phase of BSISO initiated from the equatorial Indian Ocean.

Figure 4. (a) the bivariate ACC for CB and DB as a function of forecast lead days. The cases are selected with initialization conditions at 25 and 20 days before day 0, so that a total of 58 and 22 cases are used for CB and DB, respectively. (b) As in (a), but for the cases with initialization conditions at 5 days before day 0 and at around day 0.

Comparison of Skills at Different Lead Times

(a) Pred of developing phase (Init on day -25 & -20)

(b) Pred of decaying phase (Init on day -5 & 0)
We further present the scatter diagram between the ACC and the observed amplitude for individual CB events (Fig. 7b). The amplitude is the observed time-averaged amplitude over the period between day −21 and day +25. A significant correlation is identified (r = 0.66). It sheds light on the importance of the magnitude as another factor influencing the BSISO predictions, with stronger magnitude having higher prediction skills.

To further understand the potential factors governing the skills, it is enlightening to compare two groups of CB events with the time-averaged ACC (in the first 4 weeks) greater or less than 0.8 standard deviations of the multi-event mean ACC. Eventually, seven and six events are identified as the good and poor CB prediction groups. Figure 8 depicts the observed and predicted time evolution of the PC1–PC2 diagram for those two groups. Some observational differences can be identified, providing hints for understanding the factors influencing the predictive ability of those two groups. First, the events with good skills have a relatively larger amplitude than the poor group, in agreement with Fig. 7b. Second, the good skill group retains its strong magnitude after day 0, while the poor skill group displays considerable magnitude fluctuation with a rapid decay after day +5 and then a reintensification after that. This is mainly related to the rapid decay of PC2 (Fig. 8b). For the poor prediction group, the suppressed convection anomalies in the Indian Ocean largely decouple from circulation anomalies as well as the enhanced

![Observed and Predicted Peak Phase](image-url)

**Fig. 5.** The model’s skill in predicting the target phase with the strongest anomalous convection in the equatorial Indian Ocean (at around day 0) with different lead times for two types of BSISO. (a),(f) The observed composite anomalies of OLR (shading; W m^{-2}) and 850-hPa winds (m s^{-1}; not shown when wind speed is less than 0.5 m s^{-1}) averaged over days 1–5 for the CB and DB. (b)–(e),(g)–(j) The composite results from model predictions with a lead time of 5–20 days. The green stippling denotes the regions at the 5% significance level for composite OLR anomalies.
convection in the equatorial western Pacific. This is in accordance with the weak PC2.

The model predictions for those two groups have very distinctive features. For the hindcasts initialized from day −20, the good prediction group faithfully captures the amplitude and propagation speed in spite of a fast evolution after day 15 (blue curve in Fig. 8a). Yet, for the poor prediction group the predicted evolution tends to be slower than the observations (Fig. 8b). For predictions initialized at around day 0, the good skill group displays a high coherence with observations regarding both the magnitude and propagation speed within the first 15 days (light blue in Fig. 8a). In contrast, for the poor prediction group, the model prediction can only properly capture the observed rapid decay at around day +10, but fails to replicate the reintensification after that (light blue in Fig. 8b). Again, the major difference seems to appear in the decaying phases between the good and the poor prediction groups.

FIG. 6. The time evolution of observed and predicted composite anomalies of OLR (shading; W m⁻²) and 850-hPa winds (m s⁻¹; not shown when wind speed is less than 0.5 m s⁻¹) from week 1 to week 4 starting from around day 0, for (left) CB and (right) DB. The green stippling denotes the regions at the 5% significance level for composite OLR anomalies.
We further examine the spatial patterns of convection–circulation anomalies from the peak to the decaying phases (Fig. 9). The good prediction group has remarkably similar patterns with observations even during weeks 3 and 4 (Figs. 9g,h vs Figs. 9c,d). For the poor prediction group, the observed suppressed convection in the central Indian Ocean and Maritime Continent from weeks 2 to 4 is nearly completely absent in the predictions, and the enhanced convection in the western Pacific is largely underpredicted (Figs. 9j–l vs Figs. 9n–p). Note that the convection anomalies are usually more difficult to predict than the atmospheric circulations in an S2S forecast system (Xiang et al. 2015). The model struggles to predict the anomalous convection in the Indian Ocean and Maritime Continent with the absence of strong coupling between convection and circulations. Hence, the relatively lower prediction skill for the poor prediction group is arguably due to the decoupling between convection and circulation.

5. Summary and discussion

Predicting summertime weather and climate is usually thornier than that in the wintertime. A better understanding of the
predictability sources for various weather and climate phenomena is vital for improving the summertime S2S predictions. As the dominant intraseasonal signal in the summertime, skillful predictions of BSISO contribute to the predictions of other climate and extremes in both regional and remote areas. Using a newly developed GFDL SPEAR S2S prediction system, we evaluate the BSISO prediction based on 20-yr hindcast experiments (2000–19) initialized from May to October. It is revealed that the overall BSISO prediction skill is up to 22 days measured before the bivariate ACC of L13 indices drops to 0.5. The propagation of BSISO displays large diversity among individual events. We explore the predictions of two clusters of BSISO, and results show that the canonical BSISO (CB) tends to be more predictable than the northward dipole BSISO (DB) (28 vs 23 days). This is due to the fact that DB tends to decay rapidly after reaching its strongest convection in the Indian Ocean, while CB has longer persistence by triggering

FIG. 9. As in Fig. 6, but for composite results of (left) the good prediction group and (right) the poor prediction group started from around day 0.

Unauthenticated | Downloaded 04/06/24 03:29 PM UTC
a successive suppressed convection phase of BSISO. This is linked to the fact that CB exhibits a stronger Rossby wave component as well as the westerly wind anomalies prevailing in the equatorial Indian Ocean than DB (Fig. 6b vs Fig. 6j), raising the odds for a rapid phase transition to the suppressed convection phase owing to the horizontal dry advection. From a forecaster’s perspective, a BSISO event with a weak Kelvin wave component is more likely to further develop into a DB and, therefore, less predictable. By contrast, a strong Kelvin wave component signifies the subsequent development of CB, which is likely to be more predictable.

The potential predictability in a model represents its intrinsic upper limit of prediction skills. Here we evaluate the potential predictability of CB and DB by treating one ensemble as the truth with a perfect model assumption. It is revealed that the potential predictability of BSISO is about 35 days, indicating that there is room for 13 days to improve this model. Meanwhile, CB achieves higher potential predictability than DB (35 vs 32 days), implying the actual lower prediction skill for DB is also ascribable, at least in part, to its relatively lower intrinsic predictability.

The prediction of individual CB events displays a large interevent spread, particularly during May–July. However, the events in August–October have relatively higher prediction skills accompanied by less interevent spread. For CB, the events with weaker magnitude and larger magnitude fluctuations during their lifetime (weaker persistence) are usually less predictable. We argue that in late summer the BSISO propagation domain expands into the western North Pacific, increasing its periodicity and predictability. We also investigate the potential impact of the background interannual variation, while we do not find evident physical explanations for the contrasting evolution among individual events.

We compare our composite results with Fig. 1 of CW21. The evolution for CB remains very similar while some noticeable differences are found for DB between this study and CW21. For example, there is a clear tilted mode at pentad 0 in this study (Fig. 2m), which is nearly absent in Fig. 1h of CW21. Also, CW21 shows that the magnitude of the suppressed convection anomalies in the western Pacific is comparable with that of enhanced convection anomalies in the Indian Ocean during the decaying phases (Figs. 1k,n,q of CW21), in contrast to the dominant suppressed convection in the western Pacific revealed from this study. There are several possible reasons accounting for this disparity. One is due to the different studied periods so the samples for the composite analysis are different. The other is that CW21 made the composite analysis centered around day 0 exactly following the observations. Here, the date of day 0 could be slightly shifted given that the hindcast experiments were carried out every 5 days, so a slight mismatch may also contribute to the differences between those two studies.

There are several questions that remain to be addressed. For instance, what causes the distinct evolution between CB and DB? As mentioned in CW21, the Kelvin wave component, indicated by the equatorially symmetric zonal winds with its maximum near the equator, is weaker in DB than CB so it does not show an evident eastward propagation feature. For DB, the physical processes responsible for the northward movement in the Indian Ocean are due to the horizontal moisture advection, while that in the western Pacific is primarily related to the SST change. Nevertheless, the underlying root causes driving their contrasting behaviors have not been established. In addition, why does the prediction skill of individual CB events tend to be lower with a larger interevent spread in May–July than August–October? What controls the contrasting evolutions between the well-predicted and poorly predicted CB events? Future thorough investigations are necessary to understand the underlying physics of the presented results here, and to establish the robustness of the results by examining other S2S prediction systems. In addition, since the convection–circulation anomalies are largely distinguished in the western Pacific between CB and DB (Fig. 2), how those two clusters of BSISO modulate tropical cyclone activities is another interesting topic worthy of investigation.

Acknowledgments. We thank the review comments from Drs. Jiale Lou, Donghyuck Yoon, and three anonymous reviewers. This research from the Geophysical Fluid Dynamics Laboratory is supported by NOAA’s Science Collaboration Program and administered by UCAR’s Cooperative Programs for the Advancement of Earth System Science (CPAESS) under award NA21OAR4310383. We are grateful for support through a grant from the NOAA Climate Test Bed program (NA22OAR4590217). BW acknowledges support from NSF/Climate Dynamics Award 2025057.


REFERENCES


