ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Reconciling roles of the South China Sea summer monsoon and ENSO in prediction of the Indian Ocean dipole

To cite this article: Yazhou Zhang *et al* 2024 *Environ. Res. Lett.* **19** 014041

IOP Publishing Environ. Res. Lett. 19 (2024) 014041 https://doi.org/10.1088/1748-9326/ad122a

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
31 August 2023

REVISED
13 November 2023

ACCEPTED FOR PUBLICATION
4 December 2023

PUBLISHED
13 December 2023

LETTER

Reconciling roles of the South China Sea summer monsoon and ENSO in prediction of the Indian Ocean dipole

Yazhou Zhang[1,2], Jianping Li[1,*], Yina Diao[1,*], Qiuyun Wang[1], Renguang Wu[3], Ting Liu[2,4], Yishuai Jin[1], Zhaolu Hou[1] and Haili Wang[5]

[1] Frontiers Science Center for Deep Ocean Multispheres and Earth System (FDOMES)/Key Laboratory of Physical Oceanography/Academy of the Future Ocean/College of Oceanic and Atmospheric Science, Ocean University of China, Qingdao, People's Republic of China
[2] State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, People's Republic of China
[3] School of Earth Sciences, Zhejiang University, Hangzhou, People's Republic of China
[4] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, People's Republic of China
[5] State Key Laboratory of Tropical Oceanography (South China Sea Institute of Oceanology Chinese Academy of Sciences), Guangzhou, People's Republic of China
[*] Authors to whom any correspondence should be addressed.

E-mail: ljp@ouc.edu.cn and diaoyn@ouc.edu.cn

Keywords: Indian Ocean dipole predication, El Niño-Southern Oscillation, South China Sea summer monsoon, multi-factor empirical model, dynamical models

Supplementary material for this article is available online

Abstract

The Indian Ocean dipole (IOD) is a remarkable interannual variability in the tropical Indian Ocean. The improved prediction of IOD is of a great value because of its large socioeconomic impacts. Previous studies reported that both El Niño-Southern Oscillation (ENSO) and South China Sea summer monsoon (SM) play a dominant role in the western and eastern pole of the IOD, respectively. They can be used as predictors of the IOD at 3 month lead beyond self-persistence. Here, we develop an empirical model of multi-factors in which the western pole is predicted by ENSO and persistence and the eastern pole is predicted by SM and persistence. This new empirical model outperforms largely the average level of the dynamical models from the North American multi-model ensemble (NMME) project in predicting the peak IOD in boreal autumn, with a correlation coefficient of ∼0.86 and a root mean square error of ∼0.24 °C. Furthermore, the hit rate of positive culminated IOD in this new empirical model is equivalent to that in current NMME models (above 65%), much higher than that for negative culminated IOD. This improvement of skill using the empirical model suggests a perspective for better understanding and predicting the IOD.

1. Introduction

The Indian Ocean dipole (IOD) is a dominant interannual variability in the tropical Indian Ocean, characterized by a dipole pattern in sea surface temperature (SST) anomalies between the tropical western and eastern Indian Ocean (Saji *et al* 1999, Webster *et al* 1999). Numerous extreme weather and climate events with large socioeconomic impacts are attributed to the IOD, such as the extreme hot, droughts, and floods in Indian Ocean rim countries (Ashok *et al* 2003, 2004, Guan and Yamagata 2003, Cai *et al* 2009, 2011, Chen *et al* 2020, Duan *et al* 2020, Zhou *et al* 2021). Therefore, improving the prediction skill of the IOD would benefit the affected regions and thereby mitigate local socioeconomic losses.

Continuous effects have been made to improve the prediction of IOD using the state-of-the-art coupled models at seasonal time scales in past decades (Wajsowicz 2005, 2007, Luo *et al* 2007, 2008, Zhao and Hendon 2009, Shi *et al* 2012, Zhu *et al* 2015, Liu *et al* 2017, Zhao *et al* 2019, 2020, Song *et al* 2022). The dynamic forecasting skill of IOD events is typically limited to 3–4 months ahead (approximately one season) due to the strong boreal winter–spring 'predictability barrier' (Shi *et al* 2012, Liu *et al*

2017) though some individual strong IOD events (i.e. 2006 and 2019) were successfully predicted two seasons in advance (Luo *et al* 2008, Doi *et al* 2020). Some studies found that the statistical or dynamic climate models that capture well the IOD–El Niño-Southern Oscillation (ENSO) relationship exhibit a superior predictive skill of the IOD, while the ENSO-independent IOD events seem to have a lower predictability (Song *et al* 2008, Zhao and Hendon 2009, Shi *et al* 2012, Yang *et al* 2015, Zhao *et al* 2020).

Many empirical models have been implemented to predict the IOD events. For example, multiple linear regression and canonical correlation analysis have been used in previous studies to predict the Indian Ocean SST (Kug *et al* 2004, Dommenget and Jansen 2009, Chen *et al* 2022). A simple stochastic-dynamical model forced by forecasted ENSO conditions can predict the IOD well up to 6 month ahead (Zhao *et al* 2019, 2020). The above studies demonstrated that the predictability of IOD beyond persistence is largely influenced by ENSO. Recent studies suggest that the South China Sea summer monsoon (SM) is another climate factor affecting the IOD development (Zhang *et al* 2018, 2021). Furthermore, the SM and ENSO have strikingly different effects on the eastern and western pole of the IOD, with the dominant contributions of ENSO to the western pole and SM to the eastern pole, respectively (Zhang *et al* 2019). This leads us to ask whether the IOD prediction skill could be improved if considering both the preceding SM and ENSO signals, especially at its western and eastern poles. In this study, an empirical model of multi-factors (i.e. the SM, ENSO, and self-persistence) will be established using multiple linear regression, and their predictive skill of the IOD events is further assessed and compared with the North American multi-model ensemble (NMME) forecasting system.

2. Datasets and methodology

2.1. Datasets

The SST observations used here were obtained from the improved Extended Reconstructed SST version 5 on a 2° × 2° grid for the period 1948–2022 (Huang *et al* 2017). The NMME project is generally used to improve prediction skill through the error compensation and greater consistency and reliability between models (Hagedorn *et al* 2005, DelSole *et al* 2014). For comparison, the 12 models from the NMME project were utilized during 1982–2018 (Kirtman *et al* 2014), including the hindcasts (1982–2010) and real-time forecasts (2011–2018). Each model consists of 4–20 ensemble members, and the forecasts are provided at lead times from 1 month to 4 months. In addition to the ensemble mean forecast characteristics of each individual model, the grand multi-model ensemble (MME) forecasts are employed with equal weight given to each individual model. Table S1 summarized the shortened model names, time period, ensemble size, and lead months for these 12 models. Since the NMME forecast is initialized at the beginning of each month, the lead time is defined as the number of months between the latest available observed data and the center of the 3 month running hindcasting target period. For example, if the latest available observed data is January, the forecast for the January–February–March season has 1 month lead, for February–March–April season 2 month lead, and so on. Monthly anomalies are calculated with respect to climatology from January 1982 to December 2010 in both the observations and each individual NMME model, and the linear trends were removed from the datasets.

Previous studies (e.g. Shi *et al* 2012, Liu *et al* 2017, Doi *et al* 2020, Zhao *et al* 2020) had reported the predictive skill of the IOD based on the IOD mode index, which is defined as the area-averaged SST anomalies in the western Indian Ocean (50°–70° E, 10° S–10° N) minus those in the eastern Indian Ocean (90°–110° E, 10° S–0°). The Niño-3.4 (hereafter as N34) index, the area-averaged SST anomalies over 120°–170° W and 5° N–5° S, is used to describe ENSO.

The SM index was calculated as the area-averaged summer dynamical normalized seasonality for the 925 hPa wind field within the South China Sea monsoon domain (100°–125° E, 0°–25° N), which is available at http://lijianping.cn/dct/page/65578 (Li and Zeng 2002, 2003, Li *et al* 2010). The intensity of the SM index is given by:

$$\delta = \frac{\|\bar{V}1 - Vm,n\|}{\|\bar{V}\|} - 2,$$

where $||*||$ is a norm on the monsoon domain of integration, $\bar{V}1$ and $\bar{V}$ represent January climatological wind vector and the mean of January and July climatological wind vectors, respectively, and Vm,n denotes monthly wind vectors in the mth month of the nth year.

The significance of correlations between variables X and Y was tested using a two-tailed Student's t-test. The effective number of degrees of freedom (N_{eff}) is approximately estimated as follows (e.g. Pyper and Peterman 1998, Li *et al* 2013):

$$\frac{1}{N_{\mathrm{eff}}} \approx \frac{1}{N} + \frac{2}{N}\sum_{i=1}^{N}\frac{N-j}{N}\rho_{XX}(j)\rho_{YY}(j),$$

where N is the total length of the time series, $\rho_{XX}(j)$ and $\rho_{YY}(j)$ denote autocorrelations of two time series X and Y, respectively.

To compared with the NMME results, the same study period (1982–2018) for the observations is also chosen. The seasonal mean in this study is averaged for boreal JJA (June–July–August), JAS (July–August–September), ASO (August–September–October), SON (September–October–November), and OND (October–November–December).

2.2. Methodology

The empirical model used here is established through holdout method (Devroye and Wagner 1979). Holdout (simple) validation depends on a single partitioning of the data. The time series is divided into two parts: the training period is from 1948 to 1981 and the hindcasting period is from 1982 to 2018. As for a single predictor variable, the model is trained during the training period though the linear regression of the predictor variable x three months earlier on the dependent variable z. Though this training model, we can get the regression coefficient a and constant g. Thus, the forecast model can be expressed as:

$$z_i = f_i(x) = ax_{i-3} + g, \tag{1}$$

where is $f_i(x)$ a function fitted to the predictor variable x. This gives us the relation between the dependent variable z at the ith month and the predictor variable x at $(i-3)$th month (x_{i-3}) during the hindcasting period.

With two predictor variables, we can also construct a training model by using binary regression of the predictor variables x and y 3 months earlier on the dependent variable z during the training period. Thus, the forecast model is established through the regression coefficients a and b, and the constant g in this training model, which is shown as follows:

$$z_i = f_i(x,y) = ax_{i-3} + by_{i-3} + g, \tag{2}$$

where $f_i(x,y)$ is a function fitted to the predictor variables x and y. This gives us the dependent variable z at the ith month related to predictor variables x and y at the $(i-3)$th month (x_{i-3} and y_{i-3}) during the hindcasting period.

Similar to the model that includes two predictor variables, the forecast model based on three predictor variables is constructed as follows:

$$z_i = f_i(x,y,p) = ax_{i-3} + by_{i-3} + cp_{i-3} + g, \tag{3}$$

where $f_i(x,y,p)$ is a function fitted to the predictor variables x, y and p. This gives us the dependent variable z at the ith month related to predictor variables x, y and p at the $(i-3)$th month (x_{i-3}, y_{i-3} and p_{i-3}) during the hindcasting period.

The real predictor variables used here are the standardized SM, ENSO (represented by $N34$), and self-persistence (PS) of the IOD three months earlier, and the dependent variable is IOD. The SM + N34 + PS forecast models are built as follows:

$$y_i = aSM_{i-3} + bN34_{i-3} + cPS_{i-3} + g, \tag{4}$$

where y_i is the IOD at the ith month, SM_{i-3}, $N34_{i-3}$ and PS_{i-3} is the standardized SM, ENSO, and self-persistence of the IOD at the $(i-3)$th month, respectively. a, b, and c are the regression coefficients and g is a constant.

According to the previous studies, the SM is more important in the eastern pole of the IOD and ENSO is more important in the western pole of the IOD (Zhang *et al* 2018, 2019). A new empirical model (hereafter as RC-model) is reconstructed as follows:

Step 1. The western pole of the IOD is first predicted by N34 + PS three months earlier using binary regression;

Step 2. The eastern pole of the IOD is predicted by SM + PS three months earlier using binary regression;

Step 3. The predicted IOD is defined as the difference between the western pole predicted by N34 + PS model in Step 1 and eastern pole predicted by SM + PS model in Step 2.

The leaving-one-out cross-validation is applied to estimate the stability of the RC-model according to the previous studies (Grantz *et al* 2005, Regonda *et al* 2006, Wang *et al* 2019), which is briefly introduced as follows: (1) given a time series with length L, one time point is chosen to be a hindcasting point, and the rest of the time series (length: $L-1$) is used to build the prediction model using regression analysis. (2) The ensemble hindcast is produced until each time point of the given time series redoes the step (1).

Following Shi *et al* (2012) and Zhao *et al* (2020), we evaluate the ability of the models to predict three categories of IOD events: (i) positive IOD events in which the IOD amplitude in SON exceeds 1 standard deviation, (ii) negative IOD events in which the IOD amplitude in SON is less than −1 standard deviation, and (iii) neutral IOD events that fall in between. The contingency table (table S2) is made for the occurrence of observed and predicted IOD events using the SON DMI for each individual model. The hit rate (HR) for correctly forecasting the occurrence of a positive/negative IOD event is defined as:

$$\mathrm{HR_{posi}} = \frac{a}{a+b+c} \times 100\%, \tag{5}$$

$$\mathrm{HR_{nega}} = \frac{i}{g+h+i} \times 100\%. \tag{6}$$

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

$$\mathrm{FAR} = \frac{d+f}{d+e+f} \times 100\%. \tag{7}$$

See table S2 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 SON during 1982–2018.

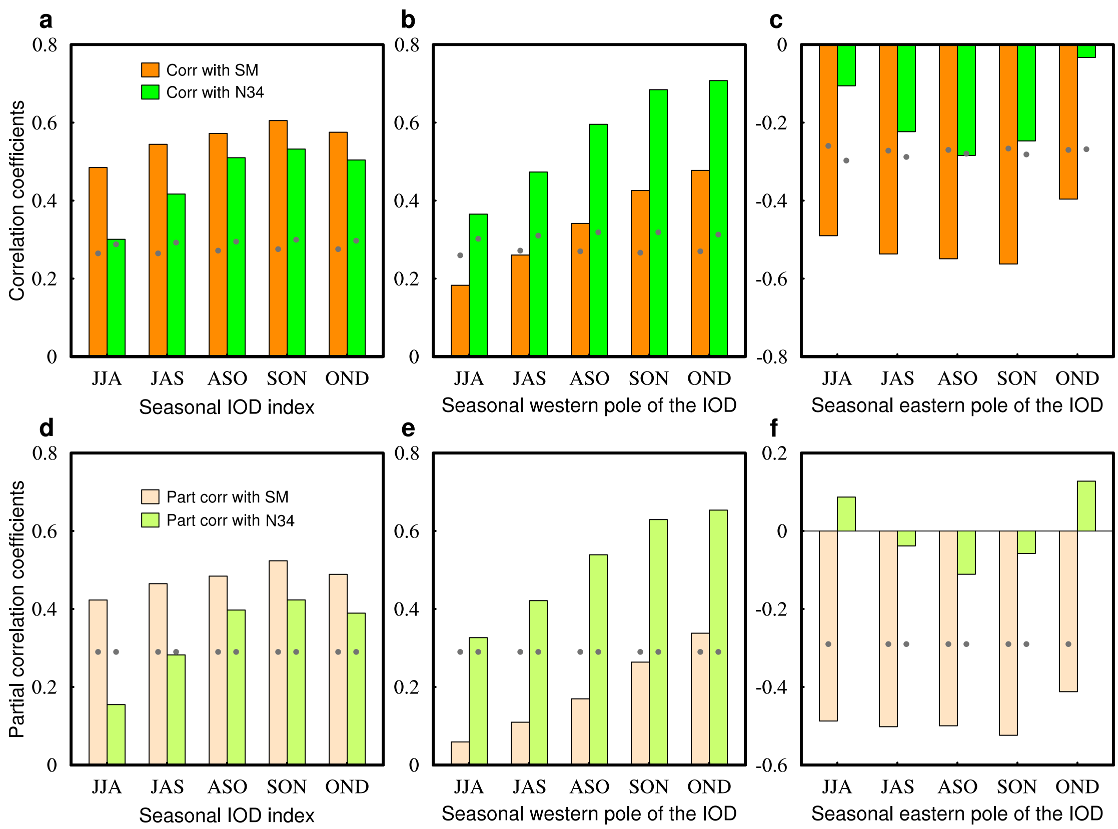

Figure 1. Relationships of the South China Sea summer monsoon (SM) and El Niño-Southern Oscillation (N34) during boreal summer (June–July–August, JJA) with the seasonal mean Indian Ocean dipole (IOD) index from JJA to OND (October–November–December). (a) For IOD index, (b) for eastern pole of the IOD, and (c) western pole of the IOD. (d) As (a), but for partial correlations between the SM (N34) and IOD after removing JJA N34 (SM) signals. (e)–(f) As (a), but for western and eastern poles of the IOD. Black dots in (a)–(f) indicates the correlation (partial correlation) coefficients beyond 99% confidence level, respectively.

3. Results

3.1. Connections of the SM and ENSO to the IOD

The relationships of the IOD with SM and ENSO are first investigated (figure 1). The SM and ENSO are significantly correlated with the IOD from JJA to OND during 1948–2018. However, there exist discrepancies in the relation of the SM and ENSO with the eastern and western poles of the IOD (figure 1(a)). The western pole of the IOD has a larger correlation coefficient with JJA ENSO than with the SM (figure 1(b)). In contrast, the correlation coefficient of the western pole of the IOD with SM is significant from JJA to OND, while that with JJA ENSO is insignificant except for ASO (figure 1(c)). After removing the SM (JJA ENSO) signal, the correlation coefficient of IOD with JJA ENSO (SM) is still significant from ASO (JJA) to OND, and the eastern (western) pole of IOD is only significantly correlated with the SM (JJA ENSO) (figures 1(d)–(f)). These results imply that there is a closer relation of eastern and western pole of the IOD to the SM and ENSO, respectively.

From the spatial pattern of the SST anomalies in the tropical Indian Ocean associated with the SM and JJA ENSO, we find that significant cold SST anomalies associated with the SM mainly exist in the tropical eastern Indian Ocean from JJA to SON based on the composite analyses of the ten positive IOD years (table S3), with weak warm SST anomalies in the tropical western Indian Ocean (figures S1(a) and (b)). Compared to the SM, the positive SST anomalies associated with JJA ENSO are clearly observed in the tropical western Indian Ocean, accompanied with weak negative SST anomalies in the tropical southeastern Indian Ocean, especially in SON (figures S1(c) and (d)). The tropical Indian Ocean SST anomalies in JJA and SON associated with both the SM and JJA ENSO are much stronger than those purely associated with either the SM or JJA ENSO (figures S1(e) and (f)), indicative of a synergistic effect of the SM and ENSO (Li *et al* 2019, Zhang *et al* 2019). The IOD in JJA intensified by the SM and JJA ENSO tends to contribute to the peak IOD in SON through strong self-persistence (figures S1(g) and (h)). These results support the findings that the SM and ENSO have remarkably different effects in the development and peak period of the IOD and its eastern and western poles (Zhang *et al* 2019).

The explained percentage is estimated to demonstrate the aforementioned results (section S1 and

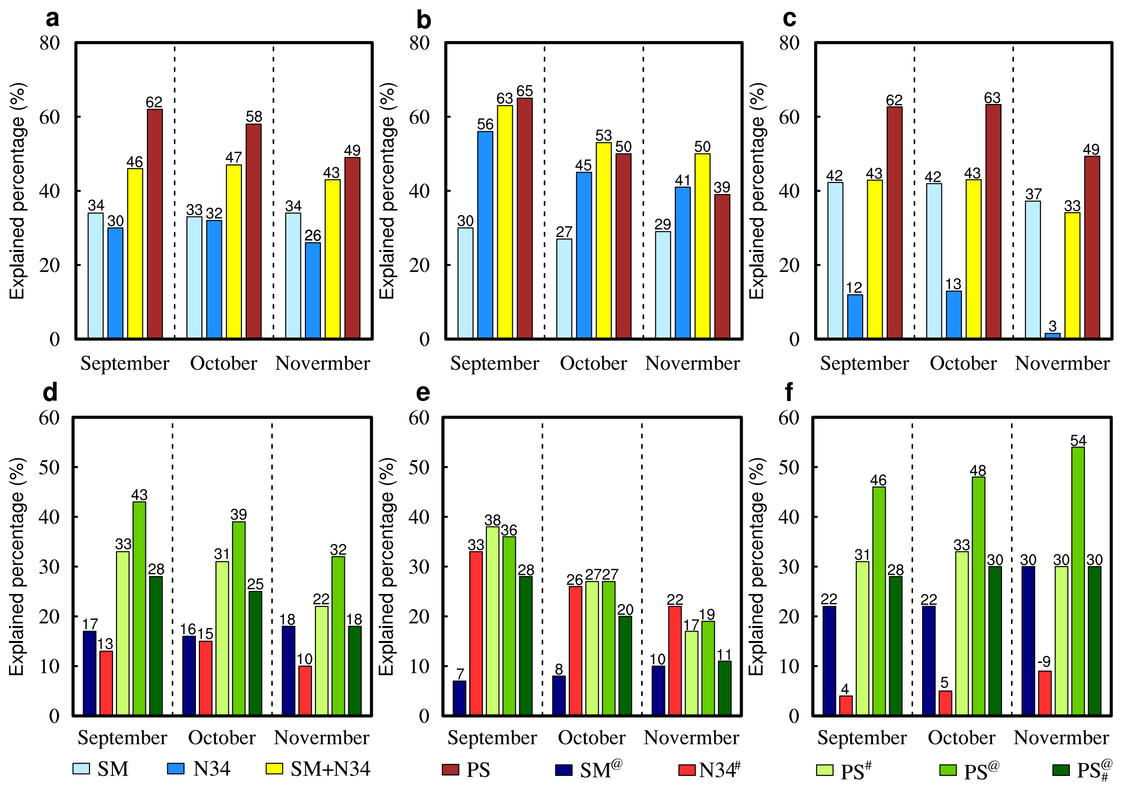

Figure 2. Explained percentage (%) of the monthly mean SST anomalies by the SM (light blue bars), ENSO (N34, blue bars), self-persistence (PS, orange bars), and both SM and ENSO (yellow bars) for the (a) IOD, (b) western pole and (c) eastern pole of the IOD in September, October, and November obtained using the preceding (3 months earlier) contributors' during positive years (table S3). (d)–(f) Same as (a)–(c), but for the partial explained percentage by the JJA SM-independent ENSO ($SM^{@}$, dark blue bars), ENSO-independent SM ($N34^{\#}$, red bars), self-persistence-independent SM ($PS^{\#}$, green bars) self-persistence-independent ENSO ($PS^{@}$, grass green bars), self-persistence-independent SM and ENSO ($PS^{@}_{\#}$, dark green bars).

Wang *et al* 2019). The mean percentage of SON IOD explained by the SM (∼33%) is relatively larger than that of JJA ENSO (∼29%), and this is also true for the SM-independent JJA ENSO (17%) and JJA ENSO-independent SM (∼13%) (figures 2(a) and (d)). Further analysis illustrates sharp distinctions in the eastern and western pole of the IOD. For the western pole of the IOD in SON, the contributions explained by the SM are nearly half of JJA ENSO, while for eastern pole of the IOD in SON, the contributions explained by the SM exceed three times of those by JJA ENSO, even reaching twelve times in November (figures 2(b) and (c)). Similar results can be obtained for the independent contributions of the SM and JJA ENSO (figures 2(e) and (f)).

The combined explained percentage of the SM and JJA ENSO are nearly ∼45%, larger than that of either factor alone (∼33% and ∼29%) but weaker than that of self-persistence (∼56%) of IOD in JJA (figure 2(d)). After removing the SM and JJA ENSO signals, the mean explained percentage of the self-persistence of JJA IOD has been reduced by more than half (figure 2(d)). The independent explained percentage of the self-persistence of JJA IOD (both removing SM and JJA ENSO) is equivalent to that of the JJA ENSO-independent SM (SM-independent JJA ENSO) in the western (eastern) pole of SON IOD (figures 2(e) and (f)). Therefore, the effects of the preceding SM, ENSO, and self-persistence in JJA on SON IOD are not negligible, and these three factors can be used as predictors for the IOD, especially at peak season in SON.

3.2. Prediction of the IOD based on empirical and NMME models

Previous studies have revealed the relative roles of the SM and ENSO on the IOD development, especially in its the eastern and western poles (Zhang *et al* 2018, 2019, 2021). The SM mainly affects the eastern pole of the IOD through the regional Hadley over the Western North Pacific and Maritime Continent, and the Walker circulation serves as the atmospheric bridge linking ENSO and western pole of the IOD (Zhang *et al* 2019). When the IOD events are purely linked to the SM, the anomalous regional Hadley circulation induced by the SM reinforces the eastern pole of IOD, but with little effect on the western pole of the IOD, which corresponds to the stronger eastern pole of the IOD. On the contrary, when the IOD events are only associated with ENSO, the significant Walker circulation is observed over the tropical Atlantic Ocean, but with casual signals over the

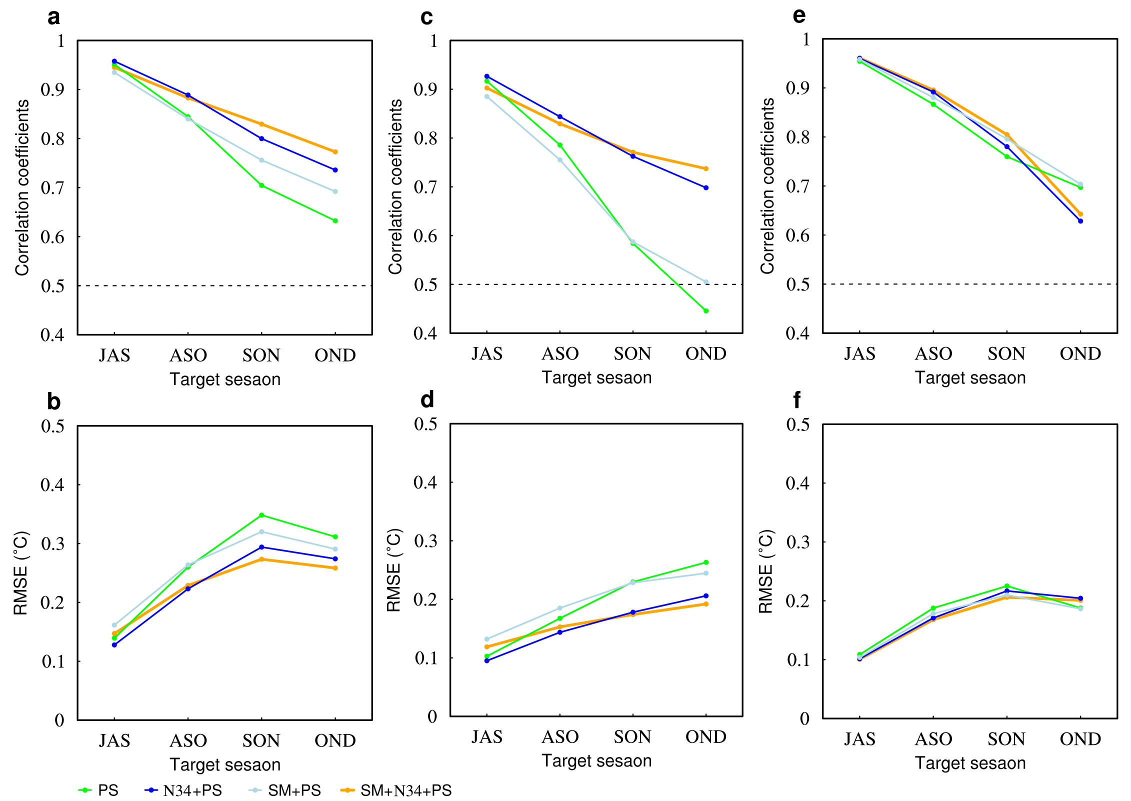

Figure 3. (a) Temporal correlation and (b) root mean square error (RMSE, °C) between predictions and observations during 1982–2018 for the IOD. The PS model (green line with dots), N34 + PS model (blue line with dots), SM + PS model (light blue line with dots), and SM + N34 + PS model (orange thick line with dots). (c)–(f) Same as (a), (b) but for but for western pole and eastern pole of the IOD.

tropical Pacific Ocean. This leads to the stronger western pole of the IOD. As the SM and ENSO co-occur, the eastern pole and western pole of the IOD are both intensified, leading to the stronger IOD events (Zhang *et al* 2019). Overall, the intensity of the eastern pole and western pole of the IOD are mainly affected by the SM and ENSO via atmospheric bridges, respectively.

Based on these possible physical connections with the preceding (3 months earlier) SM, ENSO (represented by N34), as well as the self-persistence (hereafter as PS) of the IOD, new physics-based empirical models (the SM + PS, N34 + PS, and SM + N34 + PS models) for the IOD are established. The predicted skill of the SM + N34 + PS model for the IOD is first assessed by comparison with single-factor model (the PS model) and double-factors models (the SM + PS and N34 + PS models).

These empirical models are constructed using a holdout method (section 2.2). In the development periods (JAS and ASO), the prediction skill of the SM + N34 + PS model is nearly equivalent to that from the single-factor model (PS) and double-factors models (the N34 + PS and SM + PS), and the former model outperforms largely the latter models at the peak (SON) and decay (OND) seasons (figure 3(a)). For the target season, the culminated IOD in SON predicted by the SM + N34 + PS model exhibits a highest correlation (~0.83) and lowest root mean square error (RMSE, ~0.27 °C), superior to the double-factors N34 + PS model (correlation coefficient of 0.8, RMSE of 0.29 °C), SM + PS model (correlation coefficient of 0.76, RMSE of 0.32 °C), and PS model (correlation coefficient of 0.7, RMSE of 0.35 °C) (figures 3(a) and (b)). Moreover, the declining rate of the SM + N34 + PS model is lowest with a value of ~12% at the target seasons compared to the single-factor model (PS) and double-factors models (the N34 + PS and SM + PS) (figure 3(a)). These results indicate that the SM + N34 + PS model has higher prediction skill for the IOD relative to single- or double-factors models.

For the western pole of the IOD, the N34 + PS and SM + N34 + PS models perform better with high correlation coefficient and low RMSE owing to involving ENSO signals (especially after ASO), whereas the PS and SM + PS models show a sharp drop in terms of correlation coefficient and RMSE due to relatively poor persistence of SST in the tropical western Indian Ocean (figures 3(c) and (d)). On contrary, these four models show equivalent prediction skill in predicting the eastern pole of the IOD due to the strong self-persistence of SST in the tropical eastern Indian Ocean (figures 3(e) and (f)). It is noted that the PS and SM + PS models show relative higher prediction skill than the N34 + PS and SM + N34 + PS models at the decay season in OND, suggesting that the ENSO

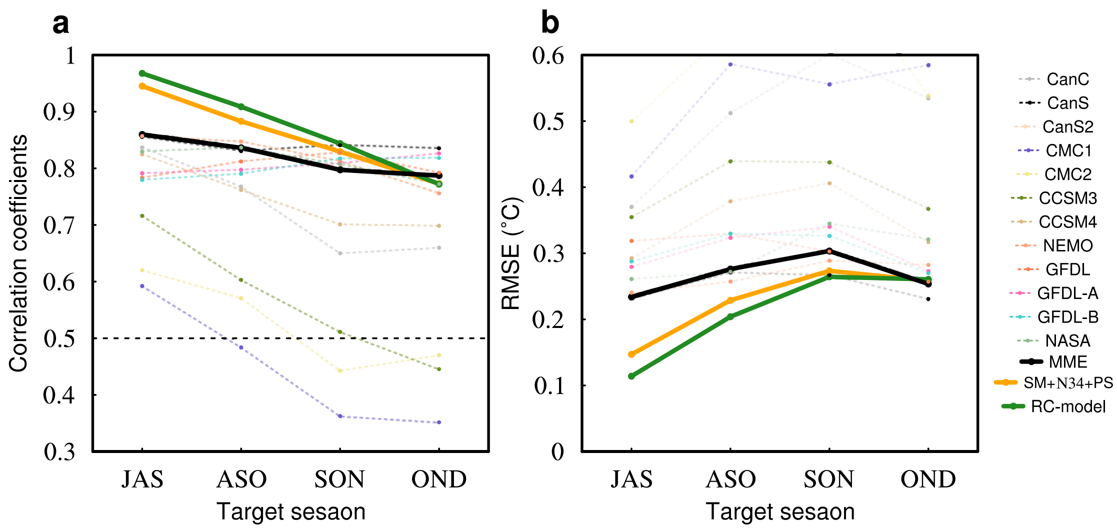

Figure 4. (a) Temporal correlation and (b) RMSE (°C) between predictions and observations for the peaking IOD in SON from 1982 to 2018. Each line represents one model. 12 hindcast models (thin dashed lines with circles), the SM + N34 + PS model (orange thick line with dots), the RC-model (reconstructed by the difference between the western pole predicted by N34 + PS model and eastern pole predicted by SM + PS), and the MME (black thick line with dots) averaged in 12 hindcast models are employed.

may be conducive to the decay of the eastern pole of the IOD (figures 3(e) and (f)).

The significance in the improvement between the pair models among the SM + N34 + PS model and PS, N34 + PS, and SM + PS models (section S2 and figure S2). Compared to the PS model, the improvement of the SM + PS, N34 + PS, and SM + N34 + PS models is significant for the IOD and its western pole during SON and OND (figure S2). For the eastern pole of the IOD, the two- and three-factors models are significant relative to the PS model besides the N34 + PS in SON, while the reversed results occur in OND (figure S2). This is because the persistence in the eastern pole of the IOD in SON is partially explained by the SM and ENSO may play a damping role in the IOD decay phase during OND. These results further verify the strikingly distinct roles of the SM and ENSO in predicting the IOD, especially in its eastern and western poles.

The NMME system has been widely used to assess the predictability of the IOD (Zhao *et al* 2019, 2020, Ling *et al* 2022, Lu *et al* 2022). Model predictable skills for the IOD are compared for the SM + N34 + PS model and 12 NMME models. The MME (averaged 12 NMME models) forecast exhibits relatively superior skill in predicting the IOD than most individual model in terms of both correlation coefficient and RMSE (figure 4). The IOD predicted by the SM + N34 + PS model is superior to that predicted by the MME in the development and peak phases of the IOD from JAS to SON (figure 4).

In fact, the SM is more important in the eastern pole of the IOD and ENSO is more important in the western pole of the IOD (Zhang *et al* 2019). Thus, we reconstruct a new empirical model (hereafter as RC-model) in which the IOD is obtained as the difference between the western pole predicted by N34 + PS model and eastern pole predicted by SM + PS model (section 2.2). The robust stability of this RC-model in predication performance is demonstrated by using leaving-one-out cross-validation (figure S3). This RC-model has the highest correlation coefficient and lowest RMSE, improving the prediction skill of the IOD at the development and peak seasons from JAS to SON (figure 4). The correlation coefficient (∼0.86) and RMSE (∼0.24 °C) between the RC-model and observations for the culminated IOD at a lead time of 3 months is nearly equivalent to that predicted by the current machine learning methods (Ratnam *et al* 2020, Liu *et al* 2021, Ling *et al* 2022), demonstrating the advantage of the RC-model relative to the SM + N34 + PS model and MME. However, the MME is slightly better than the empirical model in the performance of the decay phase in OND (figure 4). One of the reasons for the low prediction skill in OND for the SM + N34 + PS model is because the early winter is a transitional period for the most IOD events with the lowest signal-to-noise ratio (Zhao *et al* 2020).

To further document the skills for each individual IOD event, we use one standard deviation threshold to categorize the observations, RC-model, and dynamics model forecasts. The hit rate and the false alarm rate for the RC-model forecasts are shown in figure 5. The RC-model forecast performs better in terms of hit rate for positive culminated IOD in SON, much higher than that in MME and half of dynamical models (figure 5). By contrast, the hit

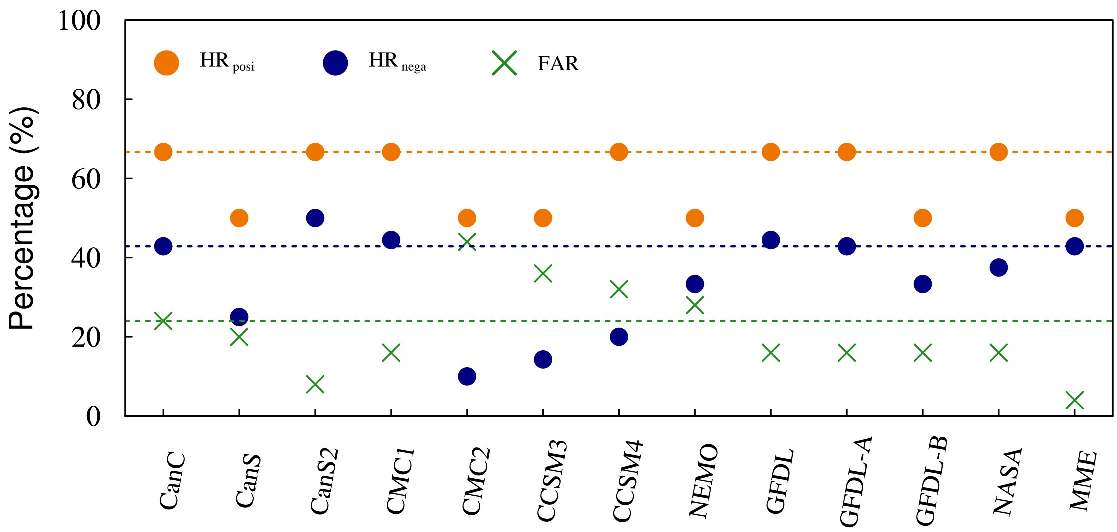

Figure 5. Hit rate for (dashed lines) RC-model, each dynamics models, and MME forecasts of the (orange dots) positive culminated IOD events, (blue dots) negative culminated IOD events, and (green crosses) false alarm rate for both positive and negative culminated IOD events in SON that exceed 1 standard deviation. Dashed orange, blue, and green lines in are the frequency that use RC-model forcing of occurrence of positive, negative, and neutral culminated events, respectively.

rate for negative culminated IOD of MME is equivalent to the RC-model though some dynamical models (CanSIPSv2, CMC1-CanCM3, and GFDL-CM2p1-aer04) outperform slightly (figure 5). Moreover, the hit rate for negative culminated IOD is lower than that for positive culminated IOD in both the RC-model and dynamical models; that is to say, there is asymmetric feature: hit rates for positive culminated IOD events being in the top-ranked group but for negative culminated IOD events hit rates being in the poor performance group. Compared to the hit rate, the false alarm rate of the MME performs best in all individual dynamical models, much better than that in the RC-model forecast (figure 5). Overall, these results suggest the advantage of the RC-model in predicting the IOD in SON at 3 month lead, which has potential value of operational applications for Climate Prediction Department.

4. Discussion and conclusions

Some extreme weather and climate events and relevant strong socioeconomic impacts are attributed to the IOD, and the prediction of the IOD is a challenging scientific issue since it was discovered. Previous studies reported that the culminated IOD is significantly correlated with both the SM and ENSO three months earlier, and the joint explained percentage of the preceding (3 months earlier) SM and JJA ENSO is nearly 20% higher than that of either factor alone. Moreover, the SM and ENSO mainly contribute to the eastern and western pole of the IOD, respectively. Therefore, the preceding (3 months earlier) SM and ENSO can be used as the crucial predictors for the culminated IOD beyond persistence. Here, we develop a multi-factor empirical model based on the preceding SM, ENSO as well as the self-persistence to predict the IOD events during the developing and peaking seasons.

The western pole of the IOD is mostly impacted by the ENSO and the eastern pole of the IOD is mainly dominated by SM (Zhang *et al* 2018, 2019, 2021). Thus, a new RC-model is defined as the western pole predicted by N34 + PS model minus the eastern pole predicted by SM + PS model. This RC-model exhibits a highest skill in predicting the culminated IOD with the high correlation coefficient (∼0.86) and low RMSE (∼0.24 °C) at a 3 month lead time, superior to the single- and multi-factor empirical models (PS, N34 + PS, SM + PS, and SM + N34 + PS). Moreover, this RC-model shows an advantage relative to the average skill level of dynamical NMME models, with a higher hit rate for the positive culminated IOD.

The regression method used in this study may be either simultaneous or at some fixed lag time due to the diversity of ENSO in the amplitudes, temporal evolution, and spatial patterns (Capotondi *et al* 2015, Zhao *et al* 2021), and this leads to the influence of preceding ENSO signals on the culminated IOD may be not completely removed. Actually, the spring ENSO signals have the bigger influence on the relationship between the JJA ENSO and SON IOD than the preceding winter ENSO signals (figures S4(a)–(h)). However, there remains a significant dipole structure associated with JJA ENSO after removal of the preceding spring ENSO (figures S4(i)). These results suggest the significant influence of preceding summer ENSO on the following autumn IOD, while it is not neglected the delayed influence of the preceding spring ENSO. Therefore, the contributions of the ENSO isolated with the preceding season signals to the culminated IOD should be investigated deeply, and the linear inverse model method may be a good tool to achieve this goal in the future (Zhao *et al* 2021).

Additionally, although the RC-model is demonstrated to be an efficient model for improved IOD prediction, the positive IOD events have been reported to be more predictable relative to the negative IOD events. During the prediction period of 1982–2018, the positive IOD events always coexist with El Niño (1982, 1985, 1994, 1997, 2006, 2015), while the four negative IOD events (1990, 1992, 1996, 2005) occurs without La Niña. The intensity of these pure negative IOD events predicted by the RC-model is closer to the observation than that predicted by NMME models (figure S5), indicating the superior skill of the empirical model in predicting the IOD events without ENSO. The IOD events display asymmetric variations in the positive and negative phases (Hong *et al* 2008, Cai *et al* 2009), and the asymmetry of ENSO may have certain contributions since the linear model transfers the asymmetry of the ENSO forcing to the IOD directly (Zhao *et al* 2020). Therefore, more effects should be made to study the asymmetry of the IOD so as to improve the prediction skill of IOD in both the statistical and dynamical models.

Data availability statement

The training and observational data sets used in this analysis are publicly available and were downloaded from the NOAA Extended Reconstructed Sea Surface Temperature V5 from their website at https://psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html. The NMME Forecasts of Monthly Climate Anomalies are available from the website at www.cpc.ncep.noaa.gov/products/NMME/.

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

Thanks for the helpful suggestions and comments from Dr X Qi, and Dr Z Chen. This work was jointly sponsored by the National Natural Science Foundation of China (NSFC) Project (42130607; 42105055), Scientific Research Fund of the Second Institute of Oceanography, MNR (QNYC2101), National Natural Science Foundation of China (NSFC) Project (42075025), Taishan Pandeng Scholar Project, State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences (LTO2309), Guangzhou Science and technology planning project (202201010524), and the open fund of State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources (QNHX2217). Thanks for the Center for High Performance Computing and System Simulation, Laoshan Laboratory (Qingdao) for providing computing resource.

Code availability

All the codes that contribute to the data preparation, simulation output analysis, and statistical analysis will be provided by the corresponding author upon reasonable request.

Conflict of interest

The authors declare no competing interests.

ORCID iDs

Yazhou Zhang https://orcid.org/0000-0002-6141-8684
Jianping Li https://orcid.org/0000-0003-0625-1575
Yina Diao https://orcid.org/0000-0002-9035-0439
Renguang Wu https://orcid.org/0000-0003-4712-2251

References

Ashok K, Guan Z, Saji N and Yamagata T 2004 Individual and combined influences of ENSO and the Indian Ocean dipole on the Indian summer monsoon *J. Clim.* **17** 3141–55

Ashok K, Guan Z and Yamagata T 2003 Influence of the Indian Ocean dipole on the Australian winter rainfall *Geophys. Res. Lett.* **30** 1329

Cai W J, Cowan T and Sullivan A 2009 Recent unprecedented skewness towards positive Indian Ocean dipole occurrences and its impact on Australian rainfall *Geophys. Res. Lett.* **36** L11705

Cai W J, Van Rensch P, Cowan T and Hendon H 2011 Teleconnection pathways of ENSO and the IOD and the mechanisms for impacts on Australian rainfall *J. Clim.* **24** 3910–23

Capotondi A *et al* 2015 Understanding ENSO diversity *Bull. Am. Meteorol. Soc.* **96** 921–38

Chen J, Yu J, Wang X and Lian T 2020 Different influences of southeastern Indian Ocean and western Indian Ocean SST anomalies on eastern China rainfall during the decaying summer of 2015/16 extreme El Niño *J. Clim.* **33** 5427–43

Chen P, Sun B, Wang H and Yang L 2022 Improving the CFSv2 prediction of the Indian Ocean dipole based on a physical-empirical model and a deep-learning approach *Int. J. Climatol.* **42** 9200–14

DelSole T, Nattala J and Tippett M 2014 Skill improvement from increased ensemble size and model diversity *Geophys. Res. Lett.* **41** 7331–42

Devroye L P and Wagner T J 1979 Distribution-free performance bounds for potential function rules *IEEE Trans. Inf. Theory* **25** 601–4

Doi T, Behera S and Yamagata T 2020 Predictability of the super IOD event in 2019 and its link with El Niño Modoki *Geophys. Res. Lett.* **47** e2019GL086713

Dommenget D and Jansen M 2009 Predictions of Indian ocean SST indices with a simple statistical model: a null hypothesis *J. Clim.* **22** 4930–8

Duan J, Li Y, Zhang L and Wang F 2020 Impacts of the Indian Ocean dipole on sea level and gyre circulation of the western tropical Pacific Ocean *J. Clim.* **33** 4207–28

Grantz K, Rajagopalan B, Clark M and Zagona E 2005 A technique for incorporating large-scale climate information

in basin-scale ensemble streamflow forecasts *Water Resour. Res.* **41**
Guan Z and Yamagata T 2003 The unusual summer of 1994 in East Asia: IOD teleconnections *Geophys. Res. Lett.* **30** 235–50
Hagedorn R, Doblas-Reyes F and Palmer T 2005 The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept *Tellus* A **57** 219–33
Hong C, Lu M and Kanamitsu M 2008 Temporal and spatial characteristics of positive and negative Indian Ocean dipole with and without ENSO *J. Geophys. Res. Atmos.* **113** D08107
Huang B, Thorne P, Banzon V, Boyer T, Chepurin G, Lawrimore J, Menne M, Smith T, Vose R and Zhang H 2017 Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons *J. Clim.* **30** 8179–205
Kirtman B *et al* 2014 The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction *Bull. Am. Meteorol. Soc.* **95** 585–601
Kug J, Kang I, Lee J and Jhun J 2004 A statistical approach to Indian Ocean sea surface temperature prediction using a dynamical ENSO prediction *Geophys. Res. Lett.* **31** L09212
Li J P, Sun C and Jin F 2013 NAO implicated as a predictor of Northern Hemisphere mean temperature multi-decadal variability *Geophys. Res. Lett.* **40** 5497–502
Li J P, Wu Z, Jiang Z and He J 2010 Can global warming strengthen the East Asian summer monsoon? *J. Clim.* **23** 6696–705
Li J P and Zeng Q C 2002 A unified monsoon index *Geophys. Res. Lett.* **29** 115-1–4
Li J P and Zeng Q C 2003 A new monsoon index and the geographical distribution of the global monsoons *Adv. Atmos. Sci.* **20** 299–302
Li J P, Zheng F, Sun C, Feng J and Wang J 2019 Pathways of influence of the northern hemisphere mid–high latitudes on East Asian climate: a review *Adv. Atmos. Sci.* **36** 902–21
Ling F, Luo J, Li Y, Tang T, Bai L, Ouyang W and Yamagata T 2022 Multi-task machine learning improves multi-seasonal prediction of the Indian Ocean dipole *Nat. Commun.* **13** 7681
Liu H, Tang Y, Chen D and Lian T 2017 Predictability of the Indian Ocean dipole in the coupled models *Clim. Dyn.* **48** 2005–24
Liu J, Tang Y, Wu Y, Li T, Wang Q and Chen D 2021 Forecasting the Indian Ocean dipole with deep learning techniques *Geophys. Res. Lett.* **48** e2021GL094407
Lu Z, Dong W, Lu B, Yuan N, Ma Z, Bogachev M and Kurths J 2022 Early warning of the Indian Ocean dipole using climate network analysis *Proc. Natl Acad. Sci. USA* **119** e2109089119
Luo J, Behera S, Masumoto Y, Sakuma H and Yamagata T 2008 Successful prediction of the consecutive IOD in 2006 and 2007 *Geophys. Res. Lett.* **35** L14S02
Luo J, Masson S, Behera S and Yamagata T 2007 Experimental forecasts of the Indian Ocean dipole using a coupled OAGCM *J. Clim.* **20** 2178–90
Pyper B and Peterman R 1998 Comparison of methods to account for autocorrelation in correlation analyses of fish data *Can. J. Fish. Aquat.* **55** 2127–40
Ratnam J, Dijkstra H and Behera S 2020 A machine learning based prediction system for the Indian Ocean dipole *Sci. Rep.* **10** 284
Regonda S Kumar, Rajagopalan B, Clark M and Zagona E 2006 A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin *Water Resour. Res.* **42**
Saji N, Goswami B, Vinayachandran P and Yamagata T 1999 A dipole mode in the tropical Indian Ocean *Nature* **401** 360–3
Shi L, Hendon H, Alves O, Luo J, Balmaseda M and Anderson D 2012 How predictable is the Indian Ocean dipole? *Mon. Weather Rev.* **140** 3867–84
Song Q, Vecchi G and Rosati A 2008 Predictability of the Indian Ocean sea surface temperature anomalies in the GFDL coupled model *Geophys. Res. Lett.* **35** L02701
Song X, Tang Y, Liu T and Li X 2022 Predictability of Indian Ocean dipole over 138 years using a CESM ensemble-prediction system *J. Geophys. Res. Oceans* **127** e2021JC018210
Wajsowicz R 2005 Potential predictability of tropical Indian Ocean SST anomalies *Geophys. Res. Lett.* **32** L24702
Wajsowicz R 2007 Seasonal-to-interannual forecasting of tropical Indian Ocean sea surface temperature anomalies: potential predictability and barriers *J. Clim.* **20** 3320–43
Wang Q *et al* 2019 Tropical cyclones act to intensify El Niño *Nat. Commun.* **10** 3793
Webster P, Moore A, Loschnigg J and Leben R 1999 Coupled ocean–atmosphere dynamics in the Indian Ocean during 1997–98 *Nature* **401** 356–60
Yang Y, Xie S, Wu L, Kosaka Y, Lau N and Vecchi G 2015 Seasonality and predictability of the Indian Ocean dipole mode: ENSO forcing and internal variability *J. Clim.* **28** 8021–36
Zhang Y, Li J, Xue J, Feng J, Wang Q, Xu Y, Wang Y and Zheng F 2018 Impact of the South China Sea summer monsoon on the Indian Ocean dipole *J. Clim.* **31** 6557–73
Zhang Y, Li J, Xue J, Zheng F, Wu R, Ha K and Feng J 2019 The relative roles of the South China Sea summer monsoon and ENSO in the Indian Ocean dipole development *Clim. Dyn.* **53** 6665–80
Zhang Y, Li J, Zheng F, Yu M, Feng J and Sun C 2021 Impact of the South China Sea summer monsoon on the Indian Ocean dipole in CMIP5 models *J. Clim.* **34** 1963–81
Zhao M and Hendon H 2009 Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model *Q. J. R. Meteorol. Soc.* **135** 337–52
Zhao S, Jin F and Stuecker M 2019 Improved predictability of the Indian Ocean dipole using seasonally modulated ENSO forcing forecasts *Geophys. Res. Lett.* **46** 9980–90
Zhao S, Stuecker M, Jin F, Feng J, Ren H, Zhang W and Li J 2020 Improved predictability of the Indian Ocean dipole using a stochastic dynamical model compared to the North American multimodel ensemble forecast *Weather Forecast.* **35** 379–99
Zhao Y, Newman M, Capotondi A, di Lorenzo E and Sun D 2021 Removing the effects of tropical dynamics from North Pacific climate variability *J. Clim.* **34** 9249–65
Zhou Z, Xie S and Zhang R 2021 Historic Yangtze fooding of 2020 tied to extreme Indian Ocean conditions *Proc. Natl Acad. Sci. USA* **118** e2022255118
Zhu J, Huang B, Kumar A and Kinter J III 2015 Seasonality in prediction skill and predictable pattern of tropical Indian Ocean SST *J. Clim.* **28** 7962–84