Prediction of Summer Precipitation in North China: Role of the Evolution of Sea Surface Temperature Anomalies from Boreal Winter to Spring

YU LIANG,a LEI FAN,b,c AND JIANLING YANGc

a College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao, China
b Key Laboratory of Physical Oceanography/Collaborative Innovation Center of Marine Science and Technology, Ocean University of China, Qingdao, China
c Key Laboratory of Meteorological Disaster Monitoring and Early Warning and Risk Management of Characteristic Agricultural in Dry Areas Regions, CMA, Yinchuan, China

(Manuscript received 25 July 2022, in final form 1 February 2023, accepted 8 February 2023)

ABSTRACT: Prediction of summer precipitation in north China (NCP) has long been a challenge partly because its low correlation with previous sea surface temperature (SST) anomalies (SSTA) limits the application of SST in NCP prediction. This study aims to extract optimal predictors of NCP from the SST field using an objective method—empirically optimal screening (EOS). It finds that the optimal precursory signal of NCP lies in the change of SSTA from winter to spring rather than the SSTA itself. This study identifies two optimal precursory signals predicting a positive (negative) NCP anomaly: the anomalous SST cooling (warming) from winter to spring in the coastal area of Somalia and Peru. Interestingly, these two presummer conditions have considerable independence, but they lead to a similar summer development of La Niña (El Niño). In summer, the tropical precipitation anomaly pattern associated with La Niña (El Niño) development excites a meridional wave train over the western Pacific and the circumglobal teleconnection in the Northern Hemisphere. Both of the anomalous wave trains show abnormal high (low) pressure over northeast Asia, which induces the south (north) wind anomalies over north China and produces abundant (deficient) precipitation there. These results highlight the importance of the SST evolution from winter to spring, break through the limitation of SST application in NCP prediction, and thus bring a prospect of improving NCP forecast skills.

SIGNIFICANCE STATEMENT: Sea surface temperature (SST) anomalies are most used as predictors in climate prediction. However, the forecast of summer precipitation in north China is limited by its low correlation with prior SST anomalies. In this paper, we find that the optimal precursory signal of north China precipitation (NCP) is not the SST anomaly itself, but the changes of SST anomalies from winter to spring in the coastal area of Somalia and Peru. These two precursory signals are almost independent yet indicate similar summer situations leading to NCP anomaly. These results highlight the importance of the dynamic evolution of sea surface temperature in improving the forecast skill of NCP.

KEYWORDS: Climate prediction; Seasonal forecasting; Statistical forecasting; ENSO; Monsoons; Sea surface temperature

1. Introduction

Prediction of midlatitude precipitation has always been a challenge due to its low signal-to-noise ratio, especially for land precipitation (Rowell et al. 1995; Fan et al. 2016; Wang et al. 2022). Some densely populated midlatitude areas are vulnerable to severe socioeconomic impacts of drought and flood disasters and thus have a high demand for climate prediction skills. North China is such a region. It is located at the northern boundary of the East Asian monsoon system (Wang 2001; Huang et al. 2007; Lei and Duan 2011; Zhou et al. 2009). Its rainy season is boreal summer (June–August, JJA), so the term “north China precipitation” (NCP) used in this paper refers to summer precipitation. One of the most important factors modulating the NCP variability is El Niño–Southern Oscillation (ENSO), with different phases of ENSO having different effects on the NCP (Huang and Wu 1989; Wu et al. 2003; Zhao et al. 2017; Wen and Hao 2021). During the post–El Niño summers, the Indo-Western Pacific Ocean capactor (IPOC; Xie et al. 2016) effects generate the Pacific–Japan (PJ) pattern (Nitta 1986) [also known as the East Asia–Pacific (EAP) pattern; Huang and Sun 1992] with an anticyclonic anomaly over the northwestern Pacific and a cyclonic anomaly to its north. This anomalous circulation pattern makes the rain belt concentrate in the Yangtze River valley in southern China and thus causes relatively little precipitation in north China. Although the well-known PJ/EAP pattern is a primary mode of the East Asian summer monsoon (EASM) variability, its accompanying precipitation anomalies mainly concentrate in the south of China instead of the north (Fan et al. 2013). Thus, the PJ/EAP is not the primary circulation pattern governing the NCP variability (Feng and Hu 2004). In fact, above-normal precipitation in north China is most correlated with an anomalous anticyclone on the east side of north China and comparatively low pressure on its west side, which allow stronger
southerly winds to move northward and bring above-normal NCP (Feng and Hu 2004; also see Fig. 1c). Such a circulation pattern usually occurs in the summer during the development of La Niña, especially for the eastern Pacific type of La Niña. The reverse is also true for El Niño cases (Wen and Hao 2021). In addition, the summer NCP is also influenced by the contemporaneous middle-to-high-latitude systems over land, such as the Eurasian (EU) teleconnection (Wang et al. 2017) and Silk Road (SR) teleconnection (Lu et al. 2002; Wang and He 2015).

Although studies on the contemporaneous factors of NCP have been extensive, the prediction of NCP using precursory predictors is still a challenge (Yim et al. 2016). In climate prediction, the slowly varying sea surface temperature (SST) field is widely used because SST has the advantage of a high signal-to-noise ratio. Nevertheless, although the ENSO-related SST in the tropical central-eastern Pacific is well known to affect the EASM, the correlation coefficient between ENSO and precipitation over China is considerably weak (Hu et al. 2017). This weak correlation is partly due to the large effects of mid-to-high-latitude systems on the NCP and partly due to the large sensitivity of precipitation anomaly patterns over China to the diversity of ENSO types (Feng et al. 2011; Wen et al. 2020, 2022). As for the SST anomaly (SSTA) fields in the previous winter and spring, their low correlations with the summer NCP have limited the application of SSTA in the NCP prediction. Thus, many researchers have resorted to quantities other than the tropical SSTs to predict the NCP, such as the sea level pressure, surface air temperature, geopotential height, and other atmospheric indices (Huang and Zhou 2006; Fan et al. 2009; Xing et al. 2016; Wang et al. 2017; Huang and Wang 2020; Huang et al. 2022). However, for the fundamental question of what the most predictive SST signal for the NCP is, previous studies have yet to reach a consistent conclusion. Considering that the contemporaneous summer SST can affect the summer NCP, the dynamic evolution of previous SST from winter to summer is probably of predictive significance to summer SST and then to summer NCP. This reasoning constitutes one of the motivations of this study, that is, to focus on the evolution of the SSTA field from winter to spring to extract the precursory signal of summer NCP.

The other motivation for this study comes from the perspective of the method. A traditional approach to search predictors is to calculate a correlation map between the forecast target and SST field and then subjectively select some “hotspots” with significant correlation values and verify their physical mechanisms. However, this approach may cause the multiple-comparison problem (also known as the look-elsewhere effect; Saville 1990) due to sampling error and thus may involve artificial skill (DelSole and Shukla 2009; Fan 2019). Although cross validation of independent samples can verify the validity of predictors, it is still difficult to ensure that these predictors are optimal for the forecast target. Another common approach is calculating the principal components (PCs) of SST and finding their optimal combination as predictors. However, our previous work (Fan 2019) has shown that the predictors selected by the above two approaches are not necessarily optimal. In fact, the empirically optimal predictors can be obtained by extracting PCs from SSTs in some selected regions where the correlation between SST and the forecast target exceeds a certain threshold. The correlation threshold is not necessarily the critical value at the significance level of 0.05 or 0.01 but is empirically derived from cross-validation experiments in which the predictor selection and validation are entirely independent. This approach is called empirically optimal screening (EOS). This method enables researchers to avoid subjectively and directly determining predictors from correlation maps, and it objectively gives refined and orthogonal predictors with mitigated artificial prediction skills. Given these advantages of the method, it is promising to obtain the optimal precursory SST signal of summer NCP, which is the objective of this study. The rest of the paper is organized as follows. Section 2 details the data and methodology, and section 3 presents the results. Finally, section 4 presents conclusions and a discussion.
2. Datasets and methods

a. Datasets

The National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed Sea Surface Temperature version 5 (ERSSTv5; Huang et al. 2017) is used for the construction of the EOS model and the analysis of the predictors. For the precipitation in China, we use the observational precipitation data of 160 stations provided by the National Climate Center of the China Meteorological Administration (CMA). The station precipitation data are used because the original observational data dates are from earlier than the satellite data and are the basis for other reconstructed data.

We also use the NOAA’s Precipitation Reconstruction (PREC) dataset (Chen et al. 2002) for a broader pattern of precipitation analysis (including the ocean). Other atmosphere data used are the monthly winds, SLP, and geopotential height, which are all from National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) Reanalysis project (Kalnay et al. 1996).

The spatial resolution of ERSSTv5, PREC, and NCEP–NCAR Reanalysis datasets are $2^\circ \times 2^\circ$, $2.5^\circ \times 2.5^\circ$, and $2.5^\circ \times 2.5^\circ$, respectively. To show the precipitation over north China in more detail, we interpolated the PREC data to the spatial resolution of $1^\circ \times 1^\circ$ by bilinear interpolation. In terms of temporal resolution, all the data used in this study are monthly.

All the data are truncated to a uniform period of 1951–2020, and they are preprocessed by deducting seasonal cycles and long-term trends. Specifically, the deduction of the seasonal cycles is achieved by subtracting the climatological means, and the long-term trends are fitted using linear regression.

b. Definition of the NCP index

As for the determination of the spatial range of the NCP index, the precipitation within this range should have highly consistent variations. In this regard, previous researchers have thoroughly studied the regionalization of precipitation in China (Min et al. 2016; Hao and Hou 2018). They suggested that the NCP index can be defined as the spatial average precipitation over nine stations (red dots in Fig. 1a): Chengde, Zhangjiakou, Beijing, Tianjin, Shijiazhuang, Xingtai, Changzhi, Taiyuan, and Linfen. This regionalization is also supported by the empirical orthogonal function (EOF) results of the north China precipitation in Fig. 1, which shows that the nine states indeed have highly consistent variations and can be taken as the core area of the NCP. Therefore, we follow previous studies and use the nine-station averaged precipitation series as the NCP index in this paper.

c. The EOS method

When we extract effective predictors from a factor field (e.g., SST), a good practice that leads to better forecasting skills is to look for orthogonal predictors that can provide nonoverlapping predictive information (Fan et al. 2011; Liu et al. 2012). A common way to achieve this is to perform the principal component analysis (PCA, also known as the EOF analysis) on the factor field and then choose the best combination of the PCs as the final predictors by using stepwise regression (or best-subset regression). However, our previous study (Fan 2019) has pointed out that simply extracting PCs of the entire factor field does not guarantee that the resulting factors are optimal for the predictand because the purpose of the PCA calculation itself is to decompose the total variance of the entire field, not specifically target the information of the predictand’s precursory signal. Therefore, preliminarily screening the factor field before PCA is necessary. The screen can be based on the correlation map between the predictand and the factor field. Data grids with a correlation coefficient higher than a threshold $R_t$ can be selected for the following PCA process.

Then the remaining question is how to determine the value of $R_t$. As to this question, Fan (2019) pointed out that too small an $R_t$ cannot achieve a good effect of purification, whereas too large an $R_t$ may cause the omission of essential factors and involve artificial factors due to sampling error. Furthermore, the optimal $R_t$ is not necessarily the critical correlation at the significance level of 0.05 or 0.01; the reason is that the type II error in statistical hypothesis testing tells us that areas not passing the significance test could also be effective for the prediction. Given the above considerations, the optimal $R_t$ can be identified empirically using historical data. We can try different values of $R_t$ by gradually increasing it from zero to a sufficiently high value, then record the corresponding hindcast skill for each value of $R_t$. Thereupon, the optimal $R_t$ can be identified from the curve of the hindcast skill as a function of $R_t$. The hindcast skill for each specific value of $R_t$ can be obtained in a cross-validation way, such as the leave-one-out or the tenfold cross validation. For example, in the leave-one-out cross validation, we use $n - 1$ samples to extract factors under an $R_t$ and then predict the remaining sample. The same step is repeated $n$ times under the same $R_t$ so that we can obtain the hindcast for the whole sequence and evaluate the hindcast skill under that $R_t$.

After the optimal $R_t$ has been obtained from the curve of prediction skill with $R_t$, grid points with a correlation higher than the optimal $R_t$ can be selected for the following PCA calculation. Finally, we can choose the optimal combination of these PCs using stepwise or best-subset regression. The resulting PCs provide nonoverlapping forecast information of the target (NCP).

The above procedure of empirically selecting optimal factors from a factor field is referred to as the EOS method, described in detail in the paper by Fan (2019). The EOS method prevents us from subjectively looking for key areas directly from the correlation map but objectively extracts optimal predictive signals from the factor field.

3. Results

a. Choosing the season of SST field

Before extracting factors from the SST field, we need to decide which season of the previous SST field to use. For this purpose, we examined the correlation maps of the summer NCP with the SSTA field in the prior seasons: early winter (December–January mean; hereafter SST DJ, Fig. 2a), later
winter and early spring (February–March, figure omitted), later spring (April–May mean; hereafter SST_AM, Fig. 2b), and the simultaneous summer (JJA, Fig. 2c). We can see that the summer NCP is closely correlated to the simultaneous summer SSTA (Fig. 2c) with a developing La Niña relating to a positive NCP, but the correlations of SST anomalies in the prior winter and spring (Figs. 2a,b) with NCP are so weak that few areas show significant correlation. Thus, the winter and spring SSTA fields are not suitable for the extraction of efficient factors. Fortunately, however, Fig. 2d shows that the SST_AM minus SST_DJ difference (ΔSST) is closer to the NCP than the winter or spring SSTA itself. This result indicates that the evolution of the SST field from winter to spring (ΔSST) provides a good source for the further extraction of factors using the EOS method. To get a more considerable and clear change from winter to spring, we use the SSTA differences of the April–May (AM) average minus the December–January (DJ) average rather than the differences of the March–May (MAM) minus the December–February (DJF). In fact, a comparison indicates that the results are not sensitive to these two schemes (figure not shown). Thus, the following shows the EOS results based on the ΔSST field of SST_AM minus SST_DJ.

b. EOS results

We extracted predictors using EOS (Fig. 3) based on the ΔSST field (60°S–60°N, 0°–360°) of the period 1951–2010, then tested the prediction skill for the period 2011–20. Figure 3a shows the cross-validated hindcast skill variation with the screening threshold Rt. Because those grid points with correlations higher than Rt are selected to perform PCA, the skill value 0.17 at Rt = 0 represents the hindcast skill without any preliminary screening (i.e., all the grid points in the ΔSST field are used to extract PCs). Such a low skill (0.17) indicates the necessity of the preliminary refinement of the ΔSST field before the PC extraction. The largest skill is achieved when Rt = 0.37.
FIG. 3. EOS results. (a) Changes of cross-validation skill with $R_t$ of SST anomalies for the training period 1951–2010, compared with the observed actual SST anomaly percentage (gray bars). (b) The leave-one-out hindcasted NCP anomaly percentage (blue line; unit: 100%) using the optimal $R_t$ (0.37) shown in (a) for the training period 1951–2010, compared with the observed actual NCP anomaly percentage (gray bars). (c) The predicted (blue) and observed (gray bars) NCP anomaly percentage for the test period (2011–20), with green dashed lines denoting the 95% confidence interval.

Then the naturally raised question is, which areas of $\Delta$SST have been screened out for the PC(s) extraction, and why can they predict the NCP? First, we checked the correlation map between the summer NCP series and $\Delta$SST field (Fig. 4a), which is the same as the correlation pattern in Fig. 2d but for a different data period (1951–2010). Figure 4a shows that the selected grids (area enclosed by purple lines in Fig. 4a, 88 grids in total) with correlation values exceeding $R_t$ (0.37) mainly concentrate in the coastal area of Somalia in the western Indian Ocean (WIO) and Peru in the tropical eastern Pacific (EP). After a stepwise regression analysis for the PCs of the selected $\Delta$SSTs (encircled by the purple dashed lines in Fig. 4a), the first two PCs are finally selected as the optimal combination of predictors for NCP prediction. Their correlation maps with the $\Delta$SST field mainly reflect the changes in the SSTA from

FIG. 4. (a) The correlation map of the NCP index with $\Delta$SST field (data period: 1951–2010; purple dashed line denoting correlation exceeding the optimal $R_t$). Correlation maps of $\Delta$SST with (b) PC2 and (c) PC1. The areas with a significance level above 0.05 are hatched, and the green boxes denote the spatial extent of $\Delta$WIO and $\Delta$EP defined in this paper. (d),(e) The two PCs (blue lines) in comparison with the observed NCP (gray bars). (f) The hindcasted NCP (red line) in a leave-one-out way using the two PC predictors in comparison with the observed standardized NCP series (bars).
boreal winter to spring along the Somali coast (PC2, Fig. 4b) and the Peruvian coast (PC1, Fig. 4c). Their correlations with the summer NCP index are 0.50 (PC2, Fig. 4d) and 0.40 (PC1, Fig. 4e), respectively. It is interesting that the ranking of PCs is not necessarily consistent with the ranking of correlations with NCP; this is because, according to the PCA algorithm, which of the WIO and EP is reflected in the PC1 depends on the combined effect of the number of grids (area) with synergistic variation and their magnitudes (variances). In this case, the SSTA changes in EP is reflected in PC1 just because it has a larger variance than the SSTA changes in WIO. Note that the PCs in the following paragraphs all refer to the EOS result (i.e., the PCs of selected ΔSSTs), not the EOF result in Fig. 1.

When these two predictors are used together for a leave-one-out cross validation of the entire data from 1951 to 2020 (Fig. 4f), the predictive correlation between the hindcasted and the actual observed NCP can be as high as 0.58. Note that this skill is higher than the skill (0.39) in Fig. 3b because, in Fig. 3b, the factors used in predicting the ith year are extracted from the data of other years; thus, it is more like a situation of the real forecast. However, in Fig. 4f, the prediction of each year uses the same factors selected based on the information for all years; thus, artificial skill is inevitably introduced (DelSole and Shukla 2009). Overall, the EOS method reveals that the decrease of SSTA in the two coastal areas from boreal winter to spring is almost independent of each other, and they both can predict a positive summer NCP anomaly, so their combination is selected as the optimal predictor for NCP.

c. Physical links between ΔWIO and NCP

To understand how the SST changes from winter to spring in the two coastal areas are linked to the summer NCP, we examined the evolution of SST and atmospheric anomaly patterns associated with the two ΔSST indices through composite analysis. In the following, to simplify the representation of the two predictors, we use an index of spatially averaged SSTA changes (ΔSST) as an alternative to each PC, and the spatial ranges of the two areas are, respectively, the coast of Somalia in the western Indian Ocean (10°S–10°N, 40°E–60°E, green box in Fig. 4b, hereafter ΔWIO) and the coast of South America area in the equatorial eastern Pacific (10°S–10°N, 90°W–80°W, green box in Fig. 4c, hereafter ΔEP). These two spatially averaged indices have high correlations (above 0.8) with corresponding PCs (Figs. 4b,c), and the correlation between the two indices is 0.25. The weak correlation between the ΔWIO and ΔEP indicates that their variations have large independence though covariation exists in some cases. Considering that the correlation of these two indices (ΔWIO and ΔEP) with NCP is negative, in the following, we use −ΔWIO and −ΔEP to represent the anomalous situation related to a positive NCP anomaly.

The composite analysis is performed separately for positive and negative NCP cases to examine the phase asymmetry. The case grouping is based on whether the standardized indices are greater than 0.5. Figure 5 shows the patterns of SST, SLP anomalies, and the 850-hPa wind anomaly associated with a cooling trend of WIO anomaly from winter to spring (Figs. 5a–c; −ΔWIO > 0.5) and a warming trend of WIO anomaly (Figs. 5d–f; −ΔWIO < −0.5) from winter to spring. We can see (Fig. 5) that the anomalous situations for positive and negative cases are basically symmetric, with −ΔWIO (ΔWIO) indicating the development of the La Niña (El Niño) event. The cooling (warming) trend of WIO from winter to spring indicates an increasing convergence (divergence) over the
eastern Indian Ocean and the Maritime Continent (Figs. 5b,e). The latent heat anomaly caused by the precipitation anomalies over the convergence (divergence) region excites baroclinic responses with stronger negative SLP anomalies (Figs. 5b,e) and positive 200-hPa geopotential height anomalies (Figs. 6b,e) over the Indian Ocean (Yang et al. 2007). It is worth noting that over the subtropical regions south and north of the tropical Indian Ocean, the strength of the baroclinic response is not symmetrical, with the SLP anomaly (Figs. 5b,e) and 200-hPa geopotential height (Figs. 6b,e) anomaly in the north much stronger than that in the south. In the positive NCP cases (Fig. 5b), this asymmetry causes the northward flow from the anomalous anticyclone over the southwest Indian Ocean to cross the equator along the Somali coast and reach South Asia. The convergence over the eastern Indian Ocean enhances the local rainfall and latent heat release, which gives positive feedback on the convergence and excites the Bjerknes feedback in the equatorial Pacific and leads to the summer ENSO development.

In summer, significant La Niña conditions develop with colder central-eastern Pacific SST (Fig. 5c) and positive precipitation anomalies over the central Pacific (Fig. 6c), and vice versa for the negative NCP cases (Figs. 5f and 6f). An interesting feature is that although the SST anomaly field is basically symmetrical (Figs. 6c,f), the precipitation anomaly field is not symmetrical (Figs. 6c,f). The precipitation anomaly over the central Pacific causes the tropical Matsuno–Gill response (Matsuno 1966; Gill 1980), with equatorial Kelvin waves spreading eastward throughout the tropics. Meanwhile, it also excites a meridional wave train with a barotropic anticyclone (cyclone) response over northeastern Asia (Figs. 6c,f). The tropic and midlatitude circulation responses jointly induce south (north) wind anomalies to their west, which is favorable (unfavorable) for the water vapor transport to the NCP area (Figs. 6c,f).

d. Physical links between ∆EP and NCP

Regarding the positive NCP cases (Figs. 7a–c), the prior cooling trend of EP from winter to spring indicates an El Niño that is rapidly declining and turning into an EP-type La Niña with abnormal SST anomalies originating in the coastal areas of Peru. The negative NCP cases (Figs. 7d–f) are symmetric to the positive NCP cases. As La Niña (El Niño) develops, the circumpolar positive (negative) geopotential height anomalies at 200 hPa in winter and spring weaken and even disappear in summer. Instead, the geopotential height at 200 hPa shows a zonal wave (Figs. 8c,f) train along the westerly waveguide in the Northern Hemisphere, which is known as the circumglobal teleconnection (CGT) pattern (Ding and Wang 2005; Ding et al. 2011). Ding et al. (2011) found that the summer CGT pattern often appears during the development of ENSO, and it is mainly stimulated by the Indian summer monsoon rainfall anomaly, especially for the south Asian sector of the CGT (Yang et al. 2009). Here, in the same line of reasoning, the precipitation intensity over the Indian Ocean is the main difference between the ∆WIO scenario (Figs. 6c,f) and the ∆EP scenario (Figs. 8c,f); thus, this precipitation difference probably
causes the difference in the South Asia geopotential height anomalies between the two scenarios. In addition to the Indian Ocean effect on the CGT, the negative rainfall anomalies over the central Pacific may also reinforce the northeast Asian sector of the CGT via the meridional wave train as in the $\Delta WIO$ scenario (Figs. 6c,f). The northeast Asian anticyclone anomaly and the relatively low pressure over the south and central Asia jointly induce the anomalous equatorial east winds carrying water vapor to turn northward and reach deep into north China, enhancing the NCP.

It is interesting and noteworthy that the anomaly patterns of the summer atmosphere associated with the two predictors (Figs. 6c,f and 8c,f) all resemble the contemporaneous atmosphere patterns associated with the summer NCP in Fig. 1c to a considerable extent, especially for the precipitation and 200-hPa height patterns in Fig. 8c. This feature indicates that the two predictors extracted from the $\Delta SST$ field in this study can explain the prevailing situation of NCP anomaly to a large extent; thus, it also implies that other factors may contribute little to the predictability of the NCP.

The prompting effect of phase transition from El Niño to La Niña on the NCP is also reported by Zhao et al. (2017), and the effect of summer EP ENSO on the NCP is in line with the results of Wen and Hao (2021). Nevertheless, the value of
this study lies in that from the perspective of prediction; it suggests that the evolution of the equatorial SSTA off the coast of Peru from winter to spring is beneficial to the NCP prediction.

The correlation between the two indices ($\Delta$WIO and $\Delta$EP) is very weak, so they have almost independent contributions to the prediction of NCP. In the $-\Delta$EP scenario (Fig. 7), the cooling (warming) of EP SSTA from winter to spring does not accompany a significant change of the contemporaneous WIO SSTA because the Indian Ocean is warmed (cooled) by the previous El Niño (La Niña) from winter to spring. Dramatic changes in the coastal EP SSTA from winter to spring are usually related to a previous ENSO event. In contrast, the change of WIO SSTA from winter to spring does not require such background conditions. That is the fundamental reason why the $\Delta$WIO and $\Delta$EP indices are nearly independent. The predictable information of the two predictors does not overlap; thus, their combination is optimal for NCP prediction.

Precipitation prediction in midlatitude north China has long been problematic, mainly because the low correlation coefficients between the NCP and the prior SST (Figs. 2a,b) have limited the application of SST in the NCP prediction. However, this study suggests that although the prior SST field does not contain effective signals of NCP, the SSTA changes from winter to spring can provide more predictive information. The effectiveness of these two coastal SST changes—the $\Delta$WIO and the $\Delta$EP—can be further verified by the weak partial correlation between the NCP and SSTA after removing the effects of $\Delta$WIO and $\Delta$EP (Fig. 9). Generally weak partial correlation coefficients are shown not only in the winter and spring SST, but also in the contemporaneous summer SST. Compared with the simple correlation maps (Figs. 2a–c), the weaker partial correlations (Figs. 9a–c) demonstrate that the two predictors ($\Delta$WIO and $\Delta$EP) well represent the critical information of SST for the NCP prediction.

**4. Summary and discussion**

In the above, two predictors for the NCP prediction from the SST field are extracted using the EOS method. The SSTA changes from boreal winter to spring in two coastal areas, namely, the coastal area of Somalia in the WIO and Peru in the EP, are selected as two orthogonal precursory signals reflecting the dynamic evolution of the abnormal situations over the western Indian Ocean and the eastern Pacific Ocean from winter to spring. These two presummer conditions are uncorrelated, but they both lead to a similar situation of ENSO development in summer. In summer, the anomalous SLP and precipitation patterns configure the Pacific east wind anomalies to converge over the Maritime Continent. Meanwhile, the tropical precipitation anomalies excite the abnormally high pressure over northeast Asia, which induces the moisture-carrying air that gathered in the western Pacific region to turn northward and reach deep into northeast Asia, producing abundant precipitation over north China.
This study highlights that the optimal precursory signal of the summer NCP from the SST field lies in the evolution of SSTA rather than the SSTA itself. There are two reasons for this conclusion. First, the boreal spring is a mixed season of two independent processes: prior winter ENSO attenuation and summer ENSO development. Thus, the spring SST has little predictive significance for the summer ENSO conditions. Second, the summer NCP is highly correlated with contemporaneous ENSO development. Thus, the SST evolution in prior seasons rather than the SSTA itself is essential to predict the summer situation of ENSO.

Finally, it should be noted that the present prediction model has limitations because the linearity of the method makes it impossible to consider the phase asymmetry of ENSO evolution. Future improvements of the prediction skill can be achieved by studying the predictability of NCP according to different prior backgrounds of ENSO with a focus on the summer ENSO forecast and exploring the synergistic effect of SST and other factors on NCP such as polar processes associated with sea ice.

Acknowledgments. In this study, we used the open-source Python packages NumPy (Harris et al. 2020), Seikit-learn (Pedregosa et al. 2011), xarray (Hoyer and Hamman 2017), eofs (Dawson 2016), SciPy (Virtanen et al. 2020), and Climate Data Operators (CDO; Schulzweida 2022) for data preprocessing, and ProPlot (https://github.com/proplot-dev/proplot) for data visualization. The authors thank Prof. Qinyu Liu and Prof. Xiao-Tong Zhong at Ocean University of China for constructive discussions. We also thank the three anonymous reviewers for their constructive comments. This work is supported by the Natural Science Foundation of China (41975089) and China Meteorological Administration Innovation and Development Project (CXFZ2021J024).

Data availability statement. The precipitation data of 160 stations from CMA can be downloaded at http://cmdp.nccma.net.cn/index.htm. The NOAA’s PREC data can be downloaded at https://psl.noaa.gov/data/gridded/data.prec.html provided by NOAA Physical Sciences Laboratory. The ERSSTv5 data and NCEP-NCAR Reanalysis data can be downloaded at https://psl.noaa.gov/data/gridded/. The EOS codes of this study are available upon request by email.

REFERENCES


