Atmospheric Research 246 (2020) 105110

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Seasonal prediction of surface O_3-related meteorological conditions in summer in North China

Zhicong Yin[a,b,c,*], Yuyan Li[a], Bufan Cao[a]

[a] Key Laboratory of Meteorological Disaster, Ministry of Education/Joint International Research Laboratory of Climate and Environment Change (ILCEC)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing 210044, China
[b] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China
[c] Nansen-Zhu International Research Centre, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:
Ozone pollution
Prediction
Sea ice
Climate Forecast System
Antarctic oscillation

ABSTRACT

Ground-level ozone concentrations have evidently increased, and surface O_3 pollution has become one of the major air pollutions in the summer in North China. Climate factors could modulate the O_3 concentrations in summer. The surface O_3-related meteorological conditions in the summer in North China were predicted in this study. Based on the annual increment approach, five observed preceding predictors were used to establish the perdiction models whose performance significantly exceeded the Climate Forecast System. After adding the signals from the climate prediction model, the percentage of the same sign increased to 93.9%, and the bias of the independent tests in 2017 and 2018 were negligible for the climate anomalies. The linear correlation coefficient between the observed and simulated values was 0.84 ($P < .01$). It is notable that the hybrid prediction models visibly performed better in the recent decade than in previous decades, which played important roles and provided potentials to execute real-time seasonal predictions. This prediction model could allow the government to forecast summer O_3 pollution conditions in advance and consequently determine whether extra emission reductions are required to counteract the climate effects.

1. Introduction

Closely related to anthropogenic activities, the concentrations of volatile organic compounds and nitrogen oxides have seriously exceeded their natural levels in China (Wang et al., 2017). Volatile organic compounds mainly come from gasoline-burning cars and chemical solvents, and nitrogen oxides are produced by fossil fuels. In addition to their direct influences on air quality, the volatile organic compounds and nitrogen oxides interacted to form ground-level ozone (O_3, abbreviations were summarized in Table S1) in the presence of sunlight (Jin and Holloway, 2015). Surface O_3 pollution is harmful to humans, animals and crops (Tai et al., 2014). In terms of human health, surface O_3 can irritate the respiratory system and even result in permanent lung damage. In the past decade, the O_3 concentrations in the urban areas of China have increased (Wang et al., 2017). This increased trend in North China and the Yangtze River Delta has been more significant than that in the Pearl River Delta and the Sichuan Basin (Li et al., 2018). In the most severely O_3-polluted regions in China (i.e., North China), the summer-mean O_3 concentration was above 150 μg/m^3, and the daily O_3 concentration frequently exceeded 200 μg/m^3 during 2015–2018 (Yin et al., 2019a). In the future, the surface ozone concentration and the possibility of severe ozone pollution may both increase over eastern China (Wang et al., 2013). Thus, it is imperative to prevent and manage surface O_3 pollution in North China. Inspired by haze abatement, the fine forecasting and seasonal prediction of pollution-related meteorological conditions play important roles. Currently, O_3 forecasts for 24 h to 10 days have begun to appear in some big cities in China, but the seasonal prediction of O_3-related meteorological conditions is still lacking. However, this kind of prediction offers great value for the government in terms of long-term pollution control measures and economical productions (Wang, 2018).

The O_3-related meteorological conditions can be summarized as drought (less precipitation), intense sunlight (direct ultraviolet radiation) and high temperature (Yin et al., 2019c). Associated large-scale and local atmospheric circulations, which have mainly manifested as anti-cyclonic anomalies over North China, modulate the meteorological conditions to influence the formation, transportation and dispersion of surface O_3 (Yin et al., 2019d; Hu et al., 2019). The East Asian summer monsoon significantly impacted the interannual variations of summertime ground-level O_3 concentrations (Yang et al., 2014). The

---

[*] Corresponding author at: 219 Ningliu Road, Pukou District, Nanjing University of Information Science & Technology, Nanjing, China, 210044
E-mail address: yinzhc@163.com (Z. Yin).

https://doi.org/10.1016/j.atmosres.2020.105110
Received 24 January 2020; Received in revised form 27 May 2020; Accepted 14 June 2020

positive sea surface temperature anomalies in the tropical Atlantic enhanced the anti-cyclonic anomalies over the western north Pacific and then led to more precipitation (correspondingly more cloud and weaker ultraviolet radiations) in North China (Zhao et al., 2019). In addition to the summer monsoon from the south, the Eurasia pattern transported the impacts of Arctic sea ice on surface O_3 pollution in North China (Yin et al., 2019c, 2019d). The preceding May sea ice to the north of Eurasia might explain approximately 60% of the interannual variability in O_3-related weather conditions. Furthermore, Wang and He (2015) pointed out the severe summer drought in 2014 in North China and partially attributed it to the air-sea ice interactions in the polar region via the Eurasia pattern (Wei et al., 2004), and Lin and Li (2018) found the spring sea-ice anomalies in the Barents Sea favored for the northward migration of the East Asian summer subtropical rainy belt. The April–May Antarctic Oscillation positively correlated with the summer rainfall in North China (Qin et al., 2005). Meanwhile, Fan and Wang (2010) used the negative relationship between the Antarctic Oscillation and summer temperature to build seasonal prediction models. The wet-cool climate conditions might, in turn, caused weak surface O_3 pollution. Furthermore, the spring Eurasian snow cover had significant impacts on the summer rainfall (Zhang et al., 2017). Although not all of these aforementioned studies directly focused on the surface O_3 pollution, they analyzed the climate conditions closely related to O_3 concentrations, and established a scientific basis for seasonal prediction of the summer O_3-related meteorological conditions in North China.

The year-to-year increment prediction was proposed by Fan et al., (2008) based on Wang et al., (2000). The year-to-year increment is defined as the difference of a variable between the current and previous year (DY), which differed from the frequently used climate anomalies and highlighted the feature of tropospheric biennial oscillation in East Asia (Wang et al., 2000). Fan et al. (2012) examined the prediction performances of an anomalies scheme and a DY scheme and found that the later approach improved the prediction of the East Asian summer monsoon. Additionally, the DY approach was used to predict the July–August precipitation and performed well (Fan et al., 2008a, 2008b; Wang and Fan, 2005). Recently, the DY approach has begun to be applied in seasonal predictions of variables that were cross-influenced by socioeconomic and climate factors (Yin et al., 2016), such as the discoloration day for *Cotinus coggygria* leaves in Beijing (Yin et al., 2014) and dust frequency (Ji and Fan, 2019; Fan and Wang, 2004). Considering the need for haze prevention and management, statistical prediction models of haze pollution were carried out in North China (Yin and Wang, 2017) and the Yangtze River Delta (Yin et al., 2019b), respectively. Using the DY of sea ice and sea surface temperature, the long-term trend and turning points of winter haze days in North China were simulated well, and the percentage of the same mathematical sign was 91.7%. In this study, we tried to establish seasonal prediction models of O_3-related meteorological conditions in North China with the DY approach, which has the potential to serve air pollution management.

2. Datasets and methods

The observed hourly O_3 concentrations during 2015 to 2018 were available on the website of http://beijingair.sinaapp.com/#messy. The maximum daily average 8 h concentration of ozone (MDA8) was calculated to represent the general O_3 pollution level. Monthly sea ice concentrations and sea surface temperature (1° × 1°) were downloaded from the Met Office Hadley Center (Rayner et al., 2003). The Antarctic Oscillation index was computed by the NOAA climate prediction center (Barnston and Livezey, 1987). The monthly soil moisture values (0.5° × 0.5°) were downloaded from NOAA's Climate Prediction Centre (Huug et al., 2003). The snow depth was downloaded from the ERA-Interim dataset (Dee et al., 2011).

The 1° × 1° geopotential heights at 500 hPa and 200 hPa, zonal and meridional winds at 850 hPa and 10 m, precipitation, boundary layer height, air temperatures at 200 hPa and near the surface and low and medium cloud covers were also from the ERA-Interim dataset (Dee et al., 2011). The Climate Forecast System Version 2 (CFSv2) is a fully coupled model including atmosphere, ocean, land and sea ice modules (Saha et al., 2014) that is updated in real-time operations. Generally, the CFSv2 could provide the summer predictions in May and involved 24 forecast members. Here, the ensemble mean summer geopotential heights at 500 hPa, 10 m meridional winds, precipitation, boundary layer height, air temperatures at 200 hPa and near the surface were included.

The procedures to build the seasonal prediction models are organized as follows. First, the DY values (i.e., Y_t–$Y_{t\text{-}1}$) of variables, including the predictand and potential predictors, were calculated. Second, the predictors were selected based on the physical mechanisms. Finally, the multiple linear regression was trained and verified. Two schemes were designed to train the prediction models. The first prediction model was trained based on the 5 preceding factors and denoted as PM_{5F}. The other prediction model trained by the 5 external factors and the CFS predicted atmospheric circulations was named PM_{CFS}. The prediction models were verified both by the leave-one-out validation and the independent forecasts. When modeling, the data from 2017 and 2018 were unused as independent test samples, and the forecast biases for 2017 and 2018 were analyzed. Furthermore, due to a lack of independent tests, the recycling independent tests were designed. The prediction model was built using the data from 1986 to a certain year, and the OWI anomalies from the next year to 2018 were independently predicted. The expiration year of training data moved forward from 2007 to 2017, so that there were 11 prediction models with the same predictors and 66 independently predicted results. For example, the model trained by the data from 1986 to 2012 could produce independent predictions from 2013 to 2018.

3. OWI and CFSv2 predictions

The anomalous southerlies, local boundary layer height, local precipitation, and temperature difference between the surface and the 200 hPa layer significantly influenced the surface O_3 pollution in North China (Yin et al., 2019c, 2019d). The associated southerlies prevented the cold-dry air from the north to disperse the surface ozone and its precursors. Meanwhile, increased solar radiation and high surface air temperatures enhanced the photochemical reactions to produce more surface ozone. According to Yin et al. (2019c), the $V_{10m}I$ was the area-averaged meridional wind at 10 m (35–50°N, 110–122.5°N). The PI was the area-averaged precipitation (37.5–42.5°N, 112–127.5°N). The DTI was the area-averaged difference between the temperature at the surface and at 200 hPa (surface air temperature minus T200, 37.5–47.5°N, 110–122.5°N). The BI was the area-averaged boundary layer height (37.5–47.5°N, 112.5–120°N). The ozone weather index (OWI), i.e., the O_3-related meteorological condition, was defined as OWI = normalized V_{10m} I + normalized BI – normalized PI + normalized DTI (Yin et al., 2019d). In this study, the OWI were calculated basing on the ERA-Interim reanalysis data and the CFSv2 prediction results, respectively.

The observed OWI (calculated from ERA-Interim reanalysis) agreed well with the results of synoptic analysis and could represent the O_3-related meteorological conditions (Yin et al., 2019c, 2019d). In summer, the MDA8 showed an obvious monthly variation, i.e., decreasing from June to August (Fig. 1a). It is evident that the OWI well reproduced this monthly variation in O_3 concentrations in North China. Furthermore, the summer mean MDA8 increased from 2015 to 2017 and maintained a high level in 2018, indicating that the surface O_3 pollution in North China was aggravated in the past four years. The summer mean OWI exhibited consistent features and showed a good performance (Fig. 1a). Basing on the ozone observations during 2015–2018, the observed summer MDA8 DY from 2016 to 2018 can be calculated. As to the MDA8 DY, the OWI DY during 2016–2018 might simulate the variations. The OWI DY not only showed obvious features

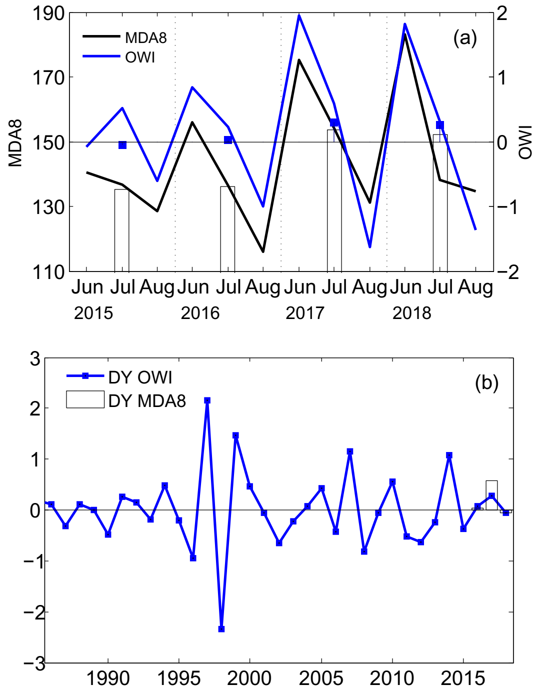

Fig. 1. (a) Monthly variations (June, July and August) in MDA8 (black line) and OWI (blue line) from 2015 to 2018. The summer mean MDA8 (black bars) and OWI (blue rectangles) are also shown. (b) Variations in summer OWI DY (blue line) from 1986 to 2018 and MDA8 DY (black bar) from 2016 to 2018. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of tropospheric biennial oscillation but also had a 135% larger variance than the original OWI, indicating a better predictability than the original OWI (Fig. 1b). The OWI DY dramatically fluctuated from 1996 to 2000 and then maintained a large amplitude, indicating a better predictability after mid-1990s (Fig. 1b).

The CFSv2 is a fully coupled model that routinely output the meridional wind at 10 m, precipitation, temperature at the surface and at 200 hPa and boundary layer height (Saha et al., 2014). Thus, the OWI can be directly calculated from the results of CFSv2 (Fig. 2). The correlation coefficient between observed OWI and CFS-forecast OWI was 0.02 (insignificant). The CFS-forecast OWI presented declining trend after 2008, which were not features of the observed OWI (Fig. 2a). These poor prediction abilities of OWI possibly came from the weak simulations of 4 sub-index (Fig. 2b). The largest correlation coefficient of 4 sub-indexes between observations and CFSv2 was only 0.2 ($V_{10m}I$), which was even insignificant at the 90% confidence level. The DY approach could take advantage of the observations from the previous year, and thus improve the prediction performance (Fan and Wang 2008). We corrected the CFS-forecast OWI by the DY approach, i.e., adding the CFS-forecast OWI DY to the observed OWI in the previous year (Fig. 2a). It is obvious that the trend, particularly the wrong decreasing trend after 2008, was corrected. However, the correlation coefficient between observed OWI and CFS-corrected OWI was still insignificant (0.05). To build statistical models might be a practicable route to obtain improved seasonal prediction of OWI.

4. Predictors and associated physical mechanisms

In the DY atmospheric circulations associated with the OWI DY, the Eurasia pattern and another Rossby-wave-like pattern from the Chukchi Sea to North China could be recognized both in the high- and mid-troposphere (Fig. 3a). The joint activity center of these two teleconnection patterns, i.e., the anomalous anti-cyclonic circulations, was located above North China (Yin et al., 2019d). The year to year increments of the area-averaged geopotential heights at 500 hPa (38–56°N, 95–128°E) were defined as the I_{AC} index. The correlation coefficient between OWI DY and I_{AC} was 0.44 (above the 99% confidence level) during 1986–2018. Associated sinking motions decreased the low and medium cloud cover and resulted in intense sunlight and high air temperatures near the surface (Fig. 3b). Thus, the photochemical reactions of volatile organic compounds and nitrogen oxides were enhanced to produce more ground level O_3 in North China. The climate link between the May sea ice and OWI in North China revealed by Yin et al. (2019c, 2019d) might provide the possibility of potential predictors, and thus, their DY relationships were analyzed and are shown in Fig. 4. The May sea ice DY in the Gakkel Ridge (82–85°N, 20–120°E) had a significant positive correlation with the OWI DY (Fig. 4a) and was averaged as the I_{SI} index, of which the correlation coefficient with OWI DY was 0.57 (above the 99% confidence level). More May sea ice in the current year would lead to more favorable meteorological conditions in the summer for the occurrence of O_3 pollution in this year than in the previous year. Two Rossby wave-like patterns transported the impacts of sea ice to the atmospheric circulations in East Asia and particularly enhanced the anticyclonic circulations over North China (Fig. 4b). Under such influences, the cloud cover was less, and the air near the surface was hotter than those in the previous year (Fig. 4c).

The spring Eurasian snow cover might induce atmospheric responses and then remotely influence the summer climate in China (Wu et al., 2009). We examined the relationship between the spring snow DY and the OWI DY and found that the May snow cover in West Siberia

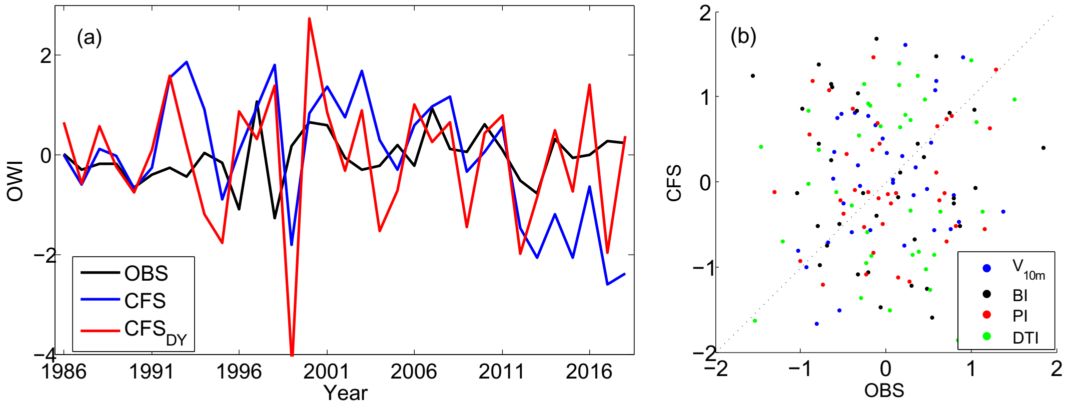

Fig. 2. (a) The summer mean OWI from ERA-Interim reanalysis (black), CFS prediction dataset (blue) and CFS data but corrected by the DY approach (red) from 1986 to 2018. (b) Scatter plots of the $V_{10m}I$ (blue), BI (black), PI (red) and DTI (green) from the ERA-Interim reanalysis and CFS prediction dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

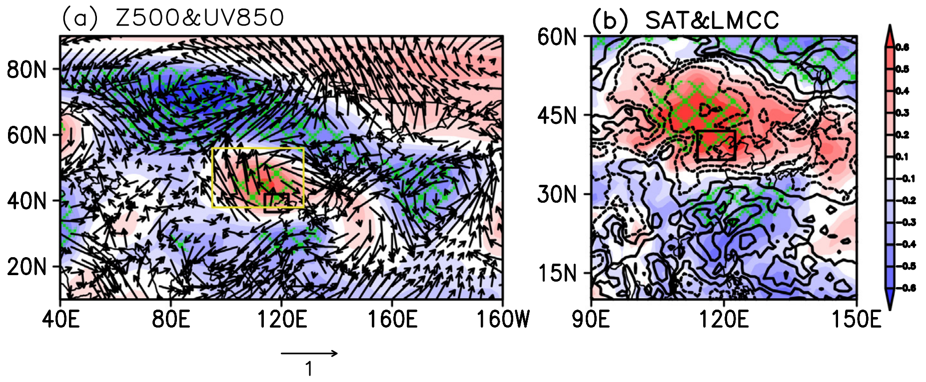

Fig. 3. DY correlation coefficients between OWI and (a) summer mean Z500 (shading), wind at 850 hPa (arrows), (b) SAT (shading), sum of medium and low cloud cover (contour) from 1986 to 2018. The black boxes indicated the location of North China. The Z500 in the yellow box of panel (a) is averaged as the I_{AC}. The green crosses indicate that the DYCCs with shading are above the 95% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

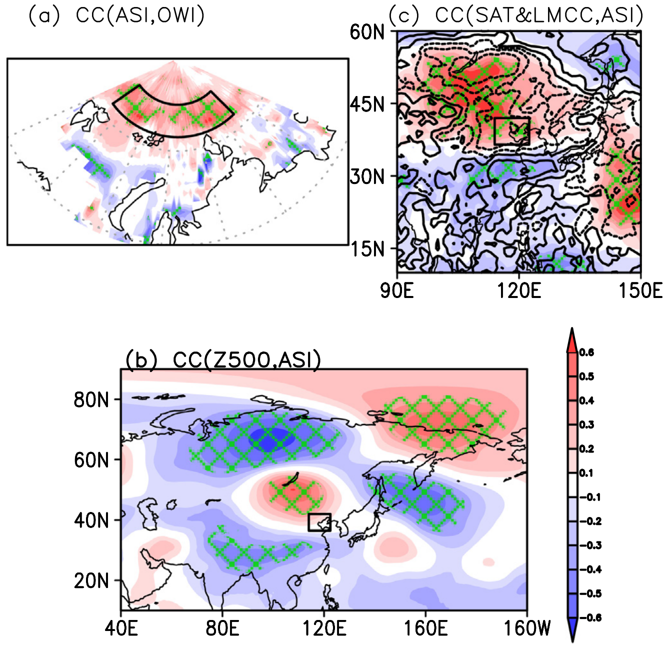

Fig. 4. DY correlation coefficients between (a) OWI and May sea ice concentrations (shading). The sea ice area in the black box is averaged as the I_{SI}. DY correlation coefficients between I_{SI} index and (b) summer Z500 (shading), (c) SAT (shading), sum of medium and low cloud cover (contour) from 1986 to 2018. The black boxes in panel b–c indicated the location of North China. The green crosses indicate that the CCs with shading are above the 95% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

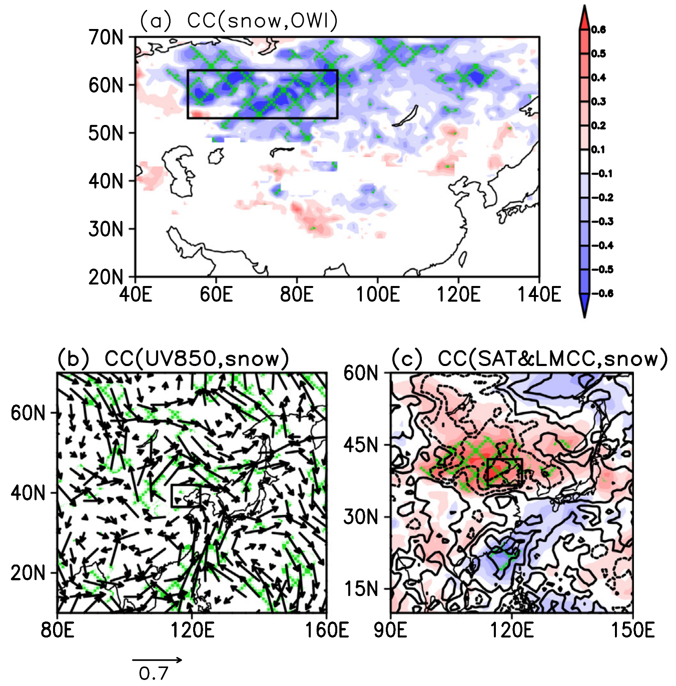

Fig. 5. DY correlation coefficients between (a) OWI and May snow concentrations (shading). The snow cover in the black box is averaged as the I_{SW}. DY correlation coefficients between I_{SW} and (b) summer mean wind at 850 hPa (arrows), (c) SAT (shading), sum of medium and low cloud cover (contour) from 1986 to 2018. The black boxes in panel b–c indicated the location of North China. The green crosses indicate that the CCs with shading in panels a-c and the wind in panel b are above the 95% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(53–63°N, 53–90°E) negatively correlated with the variation in the OWI DY (Fig. 5a). The I_{SW} index was calculated as the DY of the area-average snow depth and had a significant negative correlation (R = –0.63) with the OWI DY. A lower snow depth could stimulate atmospheric responses such as anticyclonic anomalies at 850 hPa over North China (Fig. 5b). The associated divergence of water vapor decreased the relative humidity and limited the precipitation, resulting in a dry environment near the surface. Furthermore, the anomalous anticyclonic circulations mostly indicated less cloud cover (Fig. 5c) and permitted more sunlight and ultraviolet radiation to reach the ground, which consequently led to high temperatures near the surface (Fig. 5c). Jointly, the dry-hot weather always increased the pace of photochemical reactions. At the mid and high latitudes, tracing back to early-spring (Jan-Feb-Mar), the soil moisture DY over the central Siberian Plateau (I_{SoM}) had a significant negative relationship with the OWI DY (Fig. 6a), and the correlation coefficient between I_{SoM} and the OWI DY was −0.49 (above the 99% confidence level). This drier soil moisture could enhance the Western Pacific subtropical high and induce significant anticyclonic activity over the east of China (Fig. 6b). As revealed by Yin et al. (2019a), the stronger and northward subtropical high was favorable for the production of ground O_3 levels. The North China was located in the center of the positive anomalies of surface air temperature and suffered high temperatures and intense sunlight (Fig. 6c). Thus, the OWI DY physically linked with the late-winter I_{SoM}.

In addition to the climate factors at the northern mid and high latitudes, the potential predictors in the tropics and in the Southern Hemisphere were also considered. The preceding tropical Atlantic sea surface temperature had positive relationship with the summer precipitation in North China (Yu, 2019). The warming sea surface in tropical Atlantic will enhanced the west subtropical Pacific High (Chen et al., 2015) that were beneficial to the rainfall in summer in North China (Qu et al., 2012). More precipitations indicated the ultraviolet radiations were prevented by more cloud cover and then the

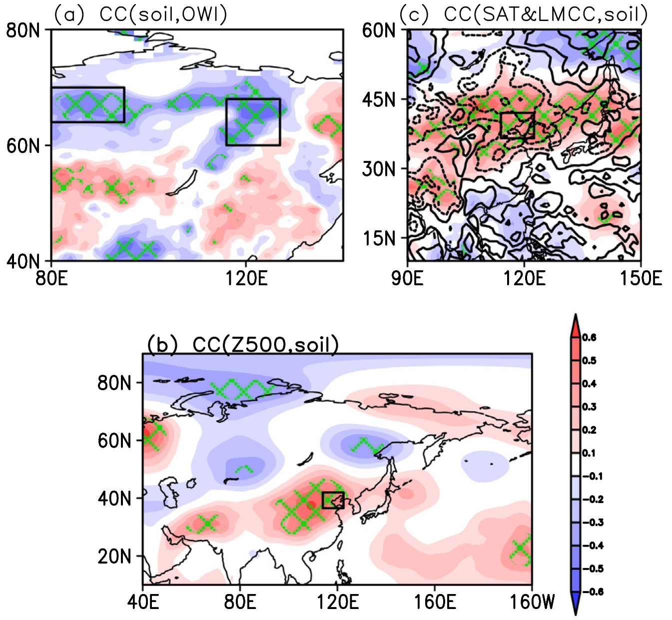

Fig. 6. DY correlation coefficients between (a) OWI and spring (Jan-Feb-Mar) soil moisture (shading). The spring soil moisture values in the black boxes were averaged as the I_{SoM}. DY correlation coefficients between I_{SoM} and (b) summer Z500 (shading), (c) SAT (shading), sum of medium and low cloud cover (contour) from 1986 to 2018. The black boxes in panel b–c indicated the location of North China. The green crosses indicate that the CCs with shading are above the 95% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

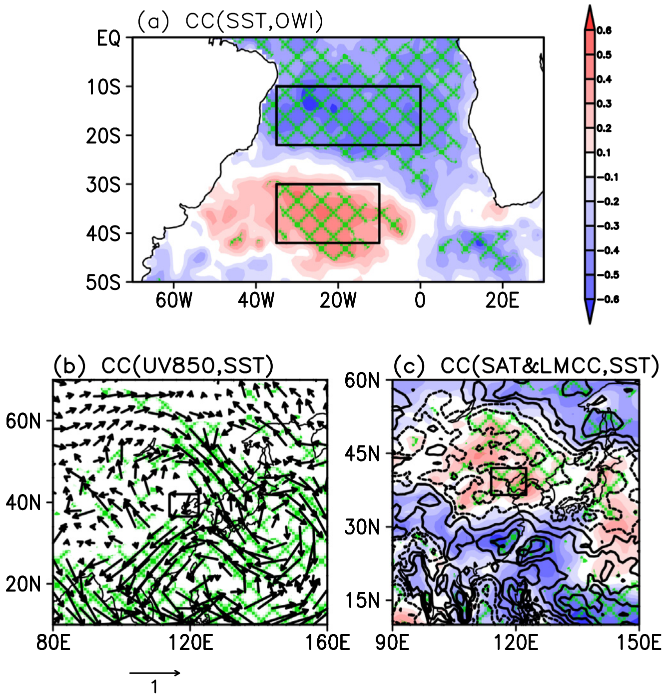

Fig. 7. DY correlation coefficients between (a) OWI and April–May sea surface temperature (shading). The differences of SST between the two black boxes are calculated as the I_{SST}. DY correlation coefficients between I_{SST} and (b) summer mean wind at 850 hPa (arrows), (c) SAT (shading), sum of medium and low cloud cover (contour) from 1986 to 2018. The black boxes in panel b–c indicated the location of North China. The green crosses indicate that the CCs with SST (a), wind (b) and cloud cover (c) are above the 95% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

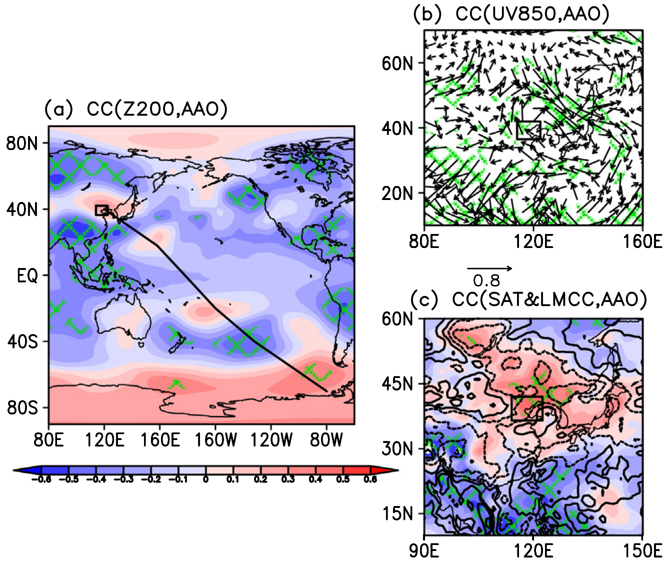

Fig. 8. DY correlation coefficients between April Antarctic Oscillation and (a) Z200 (shading), (b) wind at 850 hPa, (c) SAT (shading), sum of medium and low cloud cover (contour) from 1980 to 2018. The black boxes in panel a–c indicated the location of North China. The green crosses indicate that the CCs with shading are above the 95% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

photochemical reactions would be weakened (Fig. 7, Yin et al., 2019b). There were significantly correlated dipoles of the sea surface temperature in the Atlantic (Fig. 7a), which were located in the Brazil Basin (negatively correlated center) and the Tristan da Cunha islands (positively correlated center). The sea surface temperature difference DY between these two areas was defined as the I_{SST} index, and the correlation coefficient between the I_{SST} and the OWI DY was -0.6. At 850 hPa, a significant anticyclone was excited over North China and the transportation of water vapor was weakened (Fig. 7b). Both the higher surface air temperature and the lower cloud cover accelerated the photochemical reactions and converted more volatile organic compounds and nitrogen oxides into surface O_3 (Fig. 7c). The late-spring Antarctic Oscillation had a close relationship with the precipitation and temperature in North China (Qin et al., 2005; Fan and Wang, 2010), and thus, the DY correlation coefficient between the April Antarctic Oscillation and the OWI was calculated to be -0.43, which was above the 95% confidence level. In the upper troposphere, several alternate atmospheric centers combined as a Rossby-wave-like train to link the southern annular mode over the Antarctic region with anomalous atmospheric circulations over North China (Fig. 8a). The anticyclonic circulations over North China could also be recognized in the lower troposphere (Fig. 8b). Similar with the former analyses, the meteorological conditions associated with the lower I_{AAO} index became favorable for O_3 pollution occurrence (Fig. 8c).

5. Prediction models and validations

Based on the physical mechanisms, five preceding external forcing drivers, including I_{SI}, I_{SW}, I_{SoM}, I_{SST} and I_{AAO}, were selected as the potential predictors to forecast the O_3-related meteorological conditions in the summer in North China (Fig. 9). These predictors were located in varied latitudes to involve efficient and abundant signals as much as possible. It is evident there were several points located far away from the regression line in each panel in Fig. 9. Thus, multiple-factor modeling was used to introduce the synergistic effects and decrease the influence of the lack of nonlinear relationships (Yin and Wang, 2016). The cross correlation coefficients among the predictors were calculated,

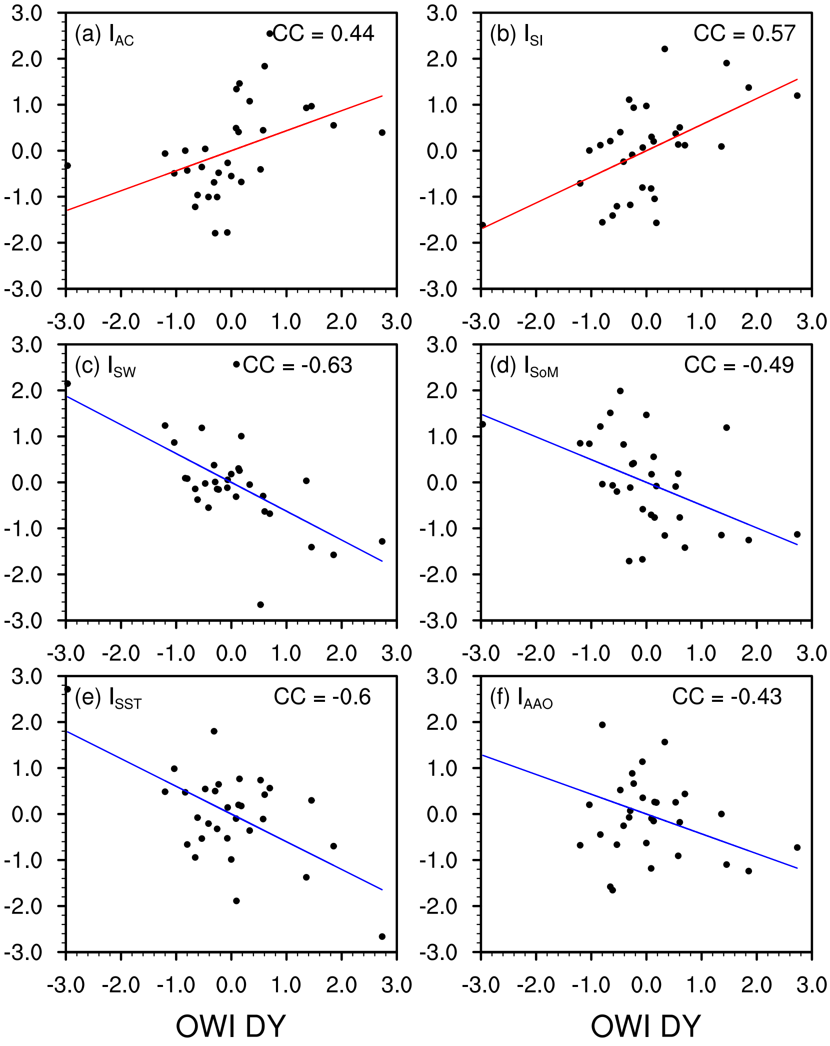

Fig. 9. Scatter plot of the OWI DY and the predictors, including (a) I_{AC}, (b) I_{SI}, (c) I_{SW}, (d) I_{SoM}, (e) I_{SST} and (f) I_{AAO}. The straight lines are the fit lines, and the CCs indicate the correlation coefficients between the OWI DY and the predictor.

most of them were insignificant (Table S2). The multicollinearity problem could be ignored when modeling with the multiple linear regression. The PM_{5F} was trained only with the preceding external forcing factors and was as follows: OWI DY = 0.29 × I_{SI} −0.32 × I_{SW} −0.34 × I_{SoM} −0.38 × I_{SST} −0.14 × I_{AAO}. The small variance inflation factors (Table S3) also indicate that the collinearity problem was insignificant.

It was foreseeable that the direct fitting performance would be fine, and thus, we only present the verifications of the leave-one-out cross validation and independent tests (Table 1). The linear correlation coefficient between the observed and fitted OWI DY values was 0.84 and was significant at the 99.99% confidence level, indicating that the PM_{5F} could reproduce the variation in the OWI DY. Furthermore, the large amplitudes during 1995–2000 were well simulated (Fig. 10a). The root-mean-square-error of the PM_{5F} was 0.43, and the predicted biases for 2017 and 2018 were − 0.14 and 0.26, respectively. In general, the goal of seasonal prediction was to obtain the climatic anomalies. Thus, the observed OWI in the previous year was added to the predicted OWI DY to achieve the final results (Fig. 10b). The percentage of the same sign (i.e., the mathematic sign of the fitted and observed anomalies was the same) was 87.9% during 1986–2018, which showed a better performance in the period after the mid-1990s compared to the prior periods. Furthermore, the long-term trend, which was removed by the DY approach, was accurately reintroduced. In the recycling independent tests of PM_{5F}, the percentage of the same sign was 83.3%, indicating that even the data were not updated in a timely way, and the ability of the prediction model could maintain a high level (Fig. 11). It is notable that the predicted anomalies in 2017 were close to the measurement and showed the correct mathematic sign; however, the bias in 2018 was larger and only approximately 50% of the predicted mathematic signs were correct.

Table 1
The leave-one-cross validated root-mean square error (RMSE), mean absolute error (MAE), linear CC (LCC), linear CC after detrending (DCC), explained variance (EV), percentage of same sign (PSS) and predicted bias for 2017 ($Bias_{17}$) and 2018 ($Bias_{18}$). The subscript "ano" indicates that the index was evaluated using the anomalies. All of the parameters were calculated for the PM_{5F}, PM_{CFS} and PM_{AC}.

	PM_{5F}	PM_{CFS}	PM_{AC}
RMSE	0.43	0.43	0.39
MAE	0.35	0.33	0.30
LCC	0.84	0.84	0.87
EV	70.6%	70.6%	73.1%
$Bias_{17}$	−0.14	− 0.03	0.08
$Bias_{18}$	0.26	0.10	0.05
PSS_{ano}	**87.9%**	**93.9%**	**97%**
LCC_{ano}	0.65	0.65	0.72
DCC_{ano}	0.63	0.62	0.70

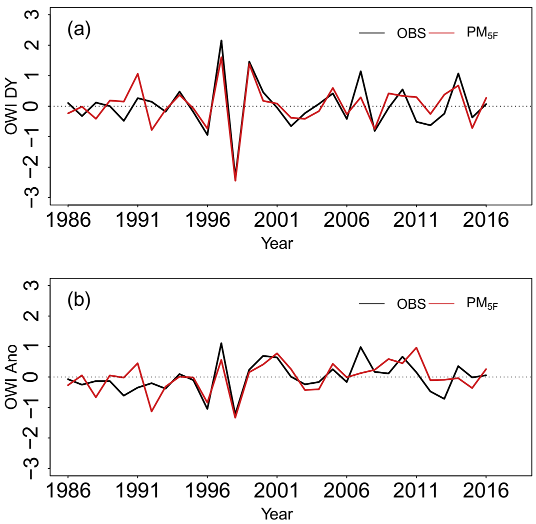

Fig. 10. Variation in the observed (black) and PM_{5F} (leave-one-out validation) simulated (red) OWI DY (a) and OWI anomalies (b) from 1986 to 2016. The results for 2017 and 2018 represent the observed (black circle) and independently predicted (red circle) OWI values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

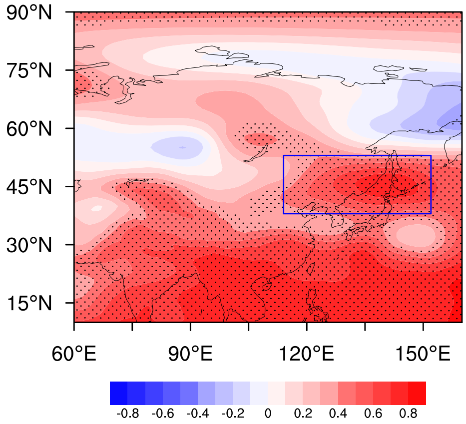

Fig. 12. The DY correlation coefficients of summer-mean Z500 between the ERA-interim reanalysis and CFS data. The black dots indicate that the CCs with shading are above the 95% confidence level. The blue box indicates the region for calculating the I_{CFS} index. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In addition to the observed preceding external forcings, the real-time climate numerical model could also provide some useful information for the seasonal predictions. Therefore, the summer atmospheric circulations predicted by the CFS were treated as the sixth predictor to improve the seasonal prediction performance. The correlation coefficients between the observed and CFS-predicted summer geopotential heights at 500 hPa were calculated. It is evident that the CFS model performed better in the subtropical and tropical latitudes than in the mid and high latitudes in East Asia (Fig. 12). We chose the well-simulated summer geopotential heights at 500 hPa value at 38–53°N, 114–152°E (I_{CFS}), which was near the I_{AC}, to represent the impacts of the anticyclonic circulations over North China. The prediction model trained here was denoted as PM_{CFS} and was as follows: OWI DY = 0.32 × I_{SI} −0.29 × I_{SW} −0.32 × I_{SoM} −0.4 × I_{SST} −0.15 × I_{AAO} + 0.14 × I_{CFS}. The PM_{CFS} showed seemingly similar performances with the PM_{5F}. The correlation coefficient and root-mean-square-error were the same to those of the PM_{5F} (Table 1). Nonetheless, the bias of the independent tests in 2017 and 2018 decreased to −0.03 and 0.10 (Fig. 13a), which were obviously smaller than those from the PM_{5F}. As to the climate anomalies, the percentage of the same sign (93.9%) was improved (Fig. 13b). The percentage of the same sign in the recycling independent tests of the PM_{CFS} was 86.4%, which was also larger than that of the former prediction model. The predicted anomalies in 2017 became close to the observations, and those in 2018 were significantly improved compared to the PM_{5F} (Fig. 14). In particular, the prediction models, which were built with the data after 2009, visibly performed better. The final target of the prediction model was to execute real-time seasonal predictions, and thus, the improvement of performances in the latest years played important roles and provided confidences for further routine operations.

To explore the reasons for limited improvements by adding CFS

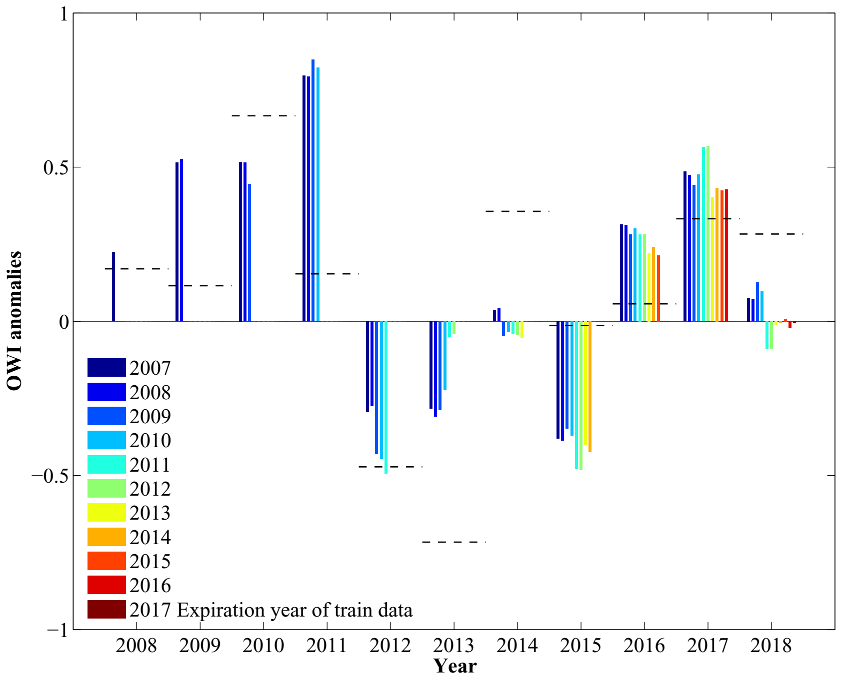

Fig. 11. The OWI anomalies from PM_{5F} recycling independent tests (bar) and observations (dashed line). The expiration year of the training data meant that the PM_{5F} was trained by the datasets from 1986 to this year, and the OWI anomalies from the next year to 2018 were independently predicted. For example, the results plotted with yellow bars (existing from 2015) were independently predicted by the PM_{5F} that had been trained by the data from 1986 to 2014, and the OWI in 2015–2018 were predicted by this model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

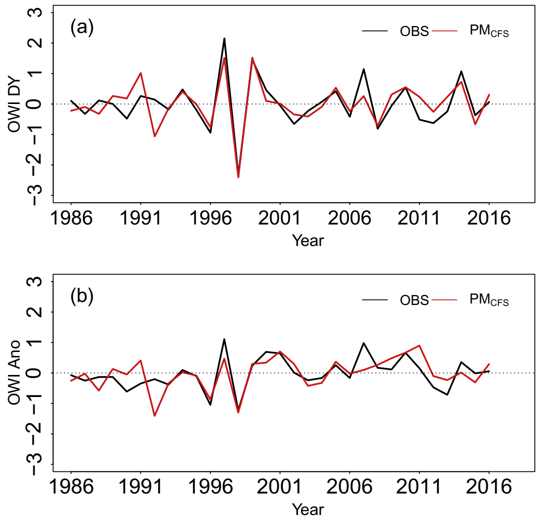

Fig. 13. Variation in the observed (black) and PM_{CFS} (leave-one-out validation) simulated (red) OWI DY (a) and OWI anomalies (b) from 1986 to 2016. The results for 2017 and 2018 represents the observed (black circle) and independently predicted (red circle) OWI values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

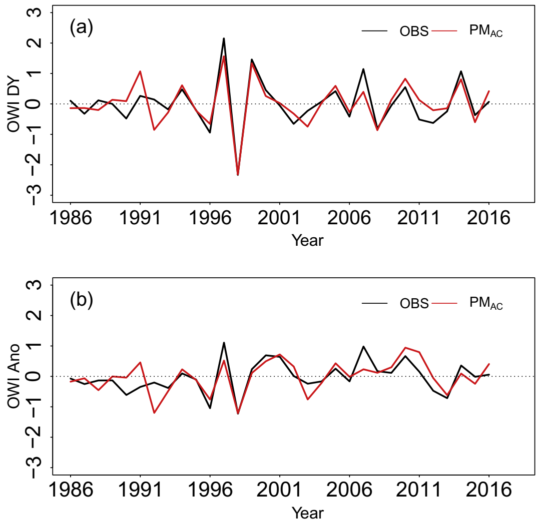

Fig. 15. Variation in the observed (black) and PM_{AC} (leave-one-out validation) simulated (red) OWI DY (a) and OWI anomalies (b) from 1986 to 2016. The results for 2017 and 2018 represent the observed (black circle) and independently predicted (red circle) OWI values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

information, another experiment was designed by replacing the CFS forecast with the observed I_{AC}. This prediction model was denoted as PM_{AC} and was trained with five preceding climate factors and the simultaneous I_{AC}. The correlation coefficient between the observed and fitted OWI DY values was 0.87 and was higher than the values for both the PM_{5F} and PM_{CFS}. The root-mean-square-error, MAE and biases for the independent tests in 2017 and 2018 were all decreased (Table 1). The PM_{AC} could explain 73.1% of the variance of the OWI DY and successfully reproduced the variation in the OWI DY and its extreme (Fig. 15a). After adding the observations in the previous year to the predicted OWI DY, the percentage of the same sign was 97% and was better than that of both the PM_{5F} and PM_{CFS}. As to the climate anomalies, the correlation coefficient between the observed and fitted OWI values were 0.72 and 0.7 before and after detrending, respectively. Thus, the long-term trend and the interannual variation of the OWI were well simulated (Fig. 15b). Furthermore, in the recycling independent tests, the percentage of the same sign was 100%, and the extreme values in 2010 (maximum) and 2013 (minimum) were closely obtained. More importantly, the predicted results in the near two years (i.e., 2017 and 2018) were significantly improved and were close to the measurements (Fig. 16). Comparing the results of the PM_{CFS} and PM_{AC}, we speculated that the predicted errors might partially come from the biased CFS signals. Thus, improving the CFS model or correcting the CFS results were potential and effects ways to gain a better prediction performance of the surface O_3-related meteorological conditions in the summer in North China.

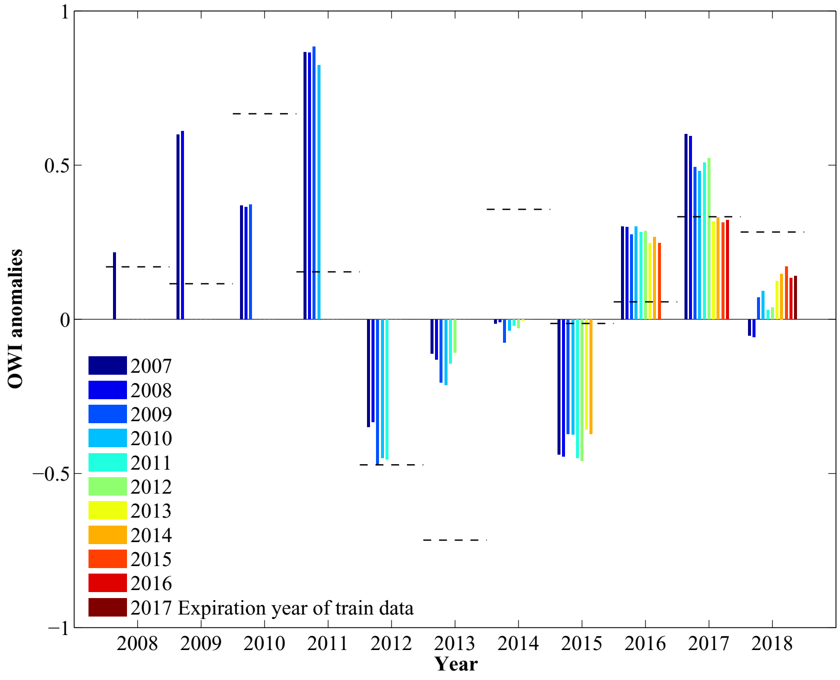

Fig. 14. The OWI anomalies from PM_{CFS} recycling independent tests (bar) and observations (dashed line). The expiration year of the training data meant that the PM_{CFS} was trained by the datasets from 1986 to this year, and the OWI anomalies from the next year to 2018 were independently predicted. For example, the results plotted with yellow bars (existing from 2015) were independently predicted by the PM_{CFS} that had been trained by the data from 1986 to 2014, and the OWI in 2015–2018 were predicted by this model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

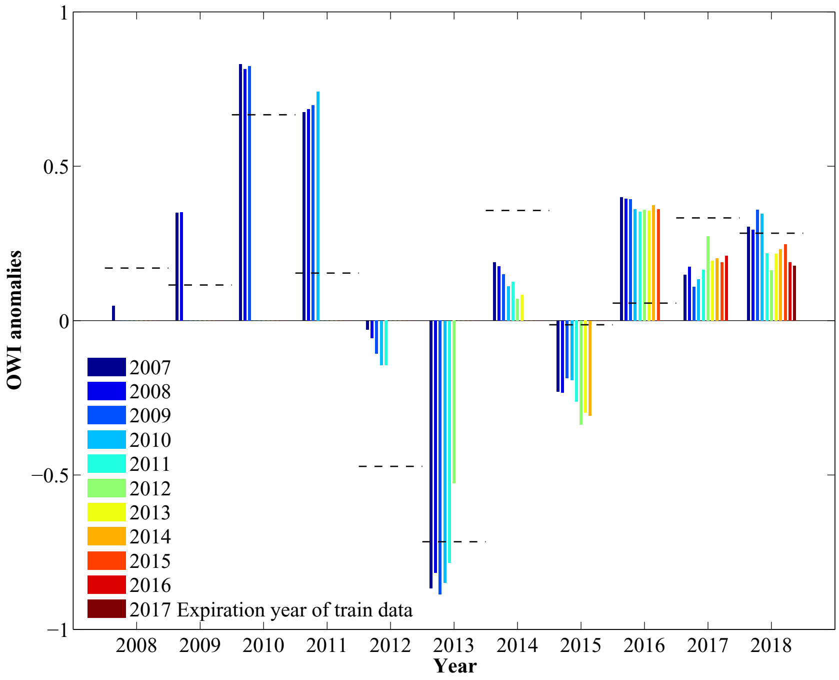

Fig. 16. The OWI anomalies from PM_{AC} recycling independent tests (bar) and observations (dashed line). The expiration year of the training data meant that the PM_{AC} was trained by the datasets from 1986 to this year, and the OWI anomalies from the next year to 2018 were independently predicted. For example, the results plotted with yellow bars (existing from 2015) were independently predicted by the PM_{AC} that had been trained by the data from 1986 to 2014, and the OWI in 2015–2018 were predicted by this model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

6. Conclusions and discussion

The ozone-related meteorological conditions (i.e., OWI) closely contributed to the O_3 concentrations in the summer in North China and were the final predictand in this study. Using the annual increment approach, the year-to-year difference of the OWI values was first predicted and then added to the observations in the previous year. In the first scheme, five leading climate drivers were used to build the prediction model, and the independent predicted values for 2017 and 2018 were close to the measurements. The percentage of the same sign was 87.9% in the cross validations during 1986–2018, and it was 83.3% in the recycling independent tests. Furthermore, the long-term trend and the interannual variations can both be reproduced. In the second scheme, the useful signals of the CFS outputs were superposed with the aforementioned predictors. As expected, the bias of the independent tests decreased, while the percentage of the same sign increased to 93.9%, indicating evident improvements.

To find the error sources, the observed atmospheric circulations replaced the CFS information. The rebuilt prediction model could explain 73.1% of the variance of the OWI DY values and successfully reproduced the variations in OWI and its extremes. The percentage of the same sign was 97% in the leave-one-out cross validation, and it was 100% in the recycling independent tests. Thus, the predicted errors of the CFS might partially contribute to the predicted biases. In addition, even considering the observed anticyclonic circulation in North China, the OWI in some years, e.g., 2007, 2011 and 2012, still was not successfully predicted. The possible reasons might be that the nonlinear relationships were not considered and some useful signals were not included.

In this study, we focused on the seasonal predictions and did not aim to fully explain the physical mechanisms of the predictors, which still required further dynamic and thermodynamic researches. These predictors were located in varied latitudes and their synergistic effects were introduced by the multiple-factor modeling. We also test the prediction performances when one predictor was excluded and found that the PM_{5F} and PM_{CFS} were the optimal choices (Table S4, Fig. S1). The predicted results in this study were the O_3-related meteorological conditions, which could influence the photochemical reactions to modulate the O_3 concentrations in summer in North China. Thus, when high OWI values are seasonally predicted, the government could make long-term pollution control decisions far in advance, and extra emission reductions might be required to counteract climate effects.

Acknowledgements

The National Key Research and Development Plan (2016YFA0600703), National Natural Science Foundation of China (91744311, 41705058 and 41991283), and the funding of Jiangsu innovation & entrepreneurship team supported this research.

Authors' contribution

Yin Z. C. designed and performed research. Yin Z. C. prepared the manuscript with contributions from Yuyan Li and Bufan Cao.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosres.2020.105110.

References

Barnston, A.G., Livezey, R.E., 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon. Wea. Rev. 115, 1083–1126.

Chen, W., Lee, J.Y., Lu, R., Dong, B., Ha, K.J., 2015. Intensified impact of tropical Atlantic SST on the western North Pacific summer climate under a weakened Atlantic thermohaline circulation. Clim. Dyn. 45, 2033–2046. https://doi.org/10.1007/s00382-014-2454-4.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., 2011. The ERAInterim reanalysis: configuration and performance of the data assimilation system. Q J R Meteor. Soc. 137, 553–597. https://doi.org/10.1002/qj.828.

Fan, K., Wang, H.J., 2004. Antarctic oscillation and the dust weather frequency in North China. Geophys. Res. Lett. 31, L10201. https://doi.org/10.1029/2004GL019465.

Fan, K., Wang, H.J., 2010. Seasonal prediction of summer temperature over Northeast China using a year-to-year incremental approach. J. Geophys. Res. 24 (3), 269–275.

Fan, K., Wang, H.J., Choi, Y.J., 2008a. A physically-based statistical forecast model for the middle-lower reaches of Yangtze River Valley summer rainfall. Chin. Sci. Bull. 54 (4), 602–609.

Fan, K., Lin, M.J., Gao, Y.Z., 2008b. Predicting precipitation in North China during the flood season by using inter-annual incremental method (in Chinese). Sci. China 38, 1452–1459.

Fan, K., Liu, Y., Chen, H.P., 2012. Improving the prediction of the East Asian summer monsoon: new approaches. Weather Forecast. 27, 1017–1030. https://doi.org/10.1175/WAF-d-11-000921.

Hu, B., Liu, H., Liu, Z., 2019. Spatiotemporal characteristics of ultraviolet solar radiation in China. Atmos. Ocean. Sci. Lett. 12 (4), 302–304. https://doi.org/10.1080/16742834.2019.1617627.

Huug, D., Huang, J., Fan, Y., 2003. Performance and analysis of the constructed analogue method applied to US soil moisture applied over 1981–2001. J. Geophys. Res. 108, 1–16.

Ji, L., Fan, K., 2019. Climate prediction of dust weather frequency over northern China based on sea-ice cover and vegetation variability. Clim. Dyn. 53, 687–705.

Jin, X.M., Holloway, T., 2015. Spatial and temporal variability of ozone sensitivity over China observed from the ozone monitoring Instrument. J. Geophys. Res. Atmos. 120, 7229–7246. https://doi.org/10.1002/2015JD023250.

Li, K., Jacob, D.J., Liao, H., Shen, L., Zhang, Q., Bates, K.H., 2018. Anthropogenic drivers of 2013–2017 trends in summer surface ozone in China. Proc. Natl. Acad. Sci. U. S. A. 116 (2), 422–427. https://doi.org/10.1073/pnas.1812168116.

Lin, Z.-D., Li, F., 2018. Impact of interannual variations of spring sea ice in the barents sea on east asian rainfall in june. Atmos. Oceanic Sci. Lett. 11 (3), 275–281. https://doi.org/10.1080/16742834.2018.1454249.

Qin, J., Wang, P.X., Gong, Y., 2005. Impacts of Antarctic oscillation on summer moisture transport and precipitation in eastern China. Chin. Geogr. Sci. 15 (1), 22–28.

Qu, Y.J., Li, Y.H., Shi, C.H., 2012. Temporal and Spatial Characteristics of August Precipitation in North China and its Causes in last 50 years. Plateau Meteorol. 32, 787–797.

Rayner, N.A., Parker, D.E., Horton, E.B., Folland, C.K., Alexander, L.V., Rowell, D.P., Kent, E.C., Kaplan, A., 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J. Geophys. Res. 108, 4407. https://doi.org/10.1029/2002JD002670.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., Becker, E., 2014. The NCEP climate forecast system Version 2. J. Clim. 27, 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.

Tai, A.P.K., Martin, M.V., Heald, C.L., 2014. Threat to future global food security from climate change and ozone air pollution. Nat. Clim. Chang. 4, 817–821.

Wang, H.J., 2018. On assessing haze attribution and control measures in China. Atmos. Ocean. Sci. Lett. 11 (2), 20–22.

Wang, H.J., Fan, K., 2005. Central-north China precipitation as reconstructed from the Qing dynasty: Signal of the Antarctic Atmospheric Oscillation. Geophys. Res. Lett. 32, L24705. https://doi.org/10.1029/2005GL024562.

Wang, H., He, S., 2015. The North China/Northeastern Asia severe summer drought in 2014. J. Clim. 68, 6667–6681.

Wang, H.J., Zhou, G.Q., Zhao, Y., 2000. An effective method for correcting the seasonal-interannual prediction of summer climate anomaly. Adv. Atmos. Sci. 17, 234–240.

Wang, Y., Shen, L., Wu, S., Mickley, L., He, J., Hao, J., 2013. Sensitivity of surface ozone over China to 2000-2050 global changes of climate and emissions. Atmos. Environ. 75 (4), 374–382. https://doi.org/10.1016/j.atmosenv.2013.04.045.

Wang, T., Xue, L.K., Brimblecombe, P., Lam, Y.F., Li, L., Zhang, L., 2017. Ozone pollution in China: a review of concentrations, meteorological influences, chemical precursors, and effects. Sci. Total Environ. 575, 1582–1596. https://doi.org/10.1016/j.scitotenv.2016.10.081.

Wei, J., Zhang, Q.Y., Tao, S.Y., 2004. Physical causes of the 1999 and 2000 summer severe drought in North China (in Chinese) Chin. J. Atmos. Sci. 28, 125–137.

Wu, B.Y., Yang, K., Zhang, R.H., 2009. Eurasian snow cover variability and its association with summer rainfall in China. Adv. Atmos. Sci. 26 (1), 31–44. https://doi.org/10.1007/s00376-009-0031-2.

Yang, Y., Liao, H., Li, J., 2014. Impacts of the East Asian summer monsoon on interannual variations of summertime surface-layer ozone concentrations over China. Atmos. Chem. Phys. 14, 6867–6879.

Yin, Z.C., Wang, H.J., 2016. Seasonal prediction of winter haze days in the north central North China Plain. Atmos. Chem. Phys. 60 (15), 1395–1400.

Yin, Z.C., Wang, H.J., 2017. Statistical prediction of winter haze days in the North China Plain using the generalized additive model. J. Appl. Meteorol. Climatol. 56, 2411–2419.

Yin, Z.C., Yuan, D.M., Ding, D.P., Xie, Z., 2014. Statistical prediction based on meteorology of Cotinus Coggygria leaves discoloration-day in the Fragrant hill (in Chinese). Meteor. Mon. 40, 229–233.

Yin, Z.C., Cao, B.F., Wang, H.J., 2019a. Dominant patterns of Summer ozone Pollution in Eastern China and Associated Atmospheric Circulations. Atmos. Chem. Phys. Discuss. https://doi.org/10.5194/acp-2019-430.

Yin, Z.C., Wang, H.J., Duan, M.K., 2019b. Outline of the realtime seasonal haze pollution prediction in China in recent years (in Chinese). Trans. Atmos. Sci. 42 (1), 2–13.

Yin, Z.C., Wang, H.J., Li, Y.Y., Ma, X.H., Zhang, X.U., 2019c. Links of climate variability among Arctic Sea ice, Eurasia teleconnection pattern and summer surface ozone pollution in North China. Atmos. Chem. Phys. 19, 3857–3871. https://doi.org/10.5194/acp-19-3857-2019.

Yin, Z.C., Yuan, D.M., Zhang, X.Y., Yang, Q., Xia, S.W., 2019d. Different contributions of Arctic Sea ice anomalies from different regions to North China summer ozone pollution. Int. J. Climatol. 40, 559–571.

Yu, X.C., 2019. Interannual and Interdecadal Variabilities in the Rainy Season Characteristics over North China and their Influence Factors (in Chinese). Lanzhou university, Lanzhou, pp. 33–35.

Zhang, R.N., Zhang, R.H., Zuo, Z.Y., 2017. Impact of Eurasian spring snow decrement on east Asian summer precipitation. J. Clim. 30, 3421–3437.

Zhao, W., Chen, W., Chen, S., Yao, S.L., Nath, D., 2019. Interannual variations of precipitation over the monsoon transitional zone in China during August–September: Role of sea surface temperature anomalies over the tropical Pacific and North Atlantic. Atmos. Sci. Lett. 20, e872.