

• Original Paper •

# Sources of Subseasonal Prediction Skill for Heatwaves over the Yangtze River Basin Revealed from Three S2S Models

Jiehong XIE, Jinhua YU, Haishan CHEN, and Pang-Chi HSU\*

*Key Laboratory of Meteorological Disaster of Ministry of Education/Joint International Research Laboratory of Climate and Environment Change/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science & Technology, Nanjing 210044, China*

(Received 14 May 2020; revised 16 July 2020; accepted 12 August 2020)

## ABSTRACT

Based on the reforecast data (1999–2010) of three operational models [the China Meteorological Administration (CMA), the National Centers for Environmental Prediction of the U.S. (NCEP) and the European Centre for Medium-Range Weather Forecasts (ECMWF)] that participated in the Subseasonal to Seasonal Prediction (S2S) project, we identified the major sources of subseasonal prediction skill for heatwaves over the Yangtze River basin (YRB). The three models show limited prediction skills in terms of the fraction of correct predictions for heatwave days in summer; the Heidke Skill Score drops quickly after a 5-day forecast lead and falls down close to zero beyond the lead time of 15 days. The superior skill of the ECMWF model in predicting the intensity and duration of the YRB heatwave is attributable to its fidelity in capturing the phase evolution and amplitude of high-pressure anomalies associated with the intraseasonal oscillation and the dryness of soil moisture induced by less precipitation via the land–atmosphere coupling. The effects of 10–30-day and 30–90-day circulation prediction skills on heatwave predictions are comparable at shorter forecast leads (10 days), while the biases in 30–90-day circulation amplitude prediction show close connection with the degradation of heatwave prediction skill at longer forecast leads (> 15–20 days). The biases of intraseasonal circulation anomalies further affect precipitation anomalies and thus land conditions, causing difficulty in capturing extremely hot days and their persistence in the S2S models.

**Key words:** subseasonal prediction, heatwave, Yangtze River Basin, subseasonal-to-seasonal models

**Citation:** Xie, J. H., J. H. Yu, H. S. Chen, and P.-C. Hsu, 2020: Sources of subseasonal prediction skill for heatwaves over the Yangtze River basin revealed from three S2S models. *Adv. Atmos. Sci.*, **37**(12), 1435–1450, <https://doi.org/10.1007/s00376-020-0144-1>.

## Article Highlights:

- The major sources of subseasonal prediction skill for heatwaves over the Yangtze River Basin were identified based on S2S model data.
- Models with superior skill for heatwave prediction show higher fidelity in predicting both the 10–30-day and 30–90-day circulation anomalies.
- The land conditions modulated by circulation anomalies also contribute to the prediction skill for heatwave duration and intensity.

## 1. Introduction

Heatwaves, prolonged extremely high-temperature days, exert considerable impacts on not only human health and ecosystems but also infrastructure and social services. Several regions worldwide have frequently suffered from heatwave events in the context of global warming (Meehl and Tebaldi, 2004; IPCC, 2013). For example, Europe experi-

enced a severe heatwave in 2003 with more than 70 000 heat-related casualties (Robine et al., 2008). Western Russia was struck by the hottest summers in 2010, leading to the deaths of around 55 000 people (Hoag, 2014). During December 2013, a long-lasting hot period hit the southern sector of South America and caused a sharp rise in disease and a substantial surge in electricity consumption (Almeira et al., 2016). Record-breaking heatwaves have also been frequently reported in the last decade in densely populous countries of Asia, such as China (Yang and Li, 2005; Xia et al., 2016; You et al., 2017; Hsu et al., 2020), India

\* Corresponding author: Pang-Chi HSU  
Email: pangchi@nuist.edu.cn

(Nageswararao et al., 2020), Japan, and on the Korean Peninsula (Imada et al., 2019; Qian et al., 2020). For better risk management, an accurate prediction of heatwaves at a longer lead time beyond 10 days is crucial; however, this remains challenging for both scientific and operational communities.

To advance our understanding of the predictability at the subseasonal-to-seasonal timescale (i.e., beyond 10 days but less than a season) and improve the subseasonal-to-seasonal prediction skill for extreme events, the WWRP (World Weather Research Program) and WCRP (World Climate Research Program) under the World Meteorological Organization (WMO) jointly launched the Subseasonal to Seasonal Prediction (S2S) project (Vitart et al., 2017). One of the important tasks of the S2S project is to develop an extensive database of subseasonal (up to 60 days) forecasts and reforecasts for achieving the goals stated above. At present, 11 institutions provide their forecast models' outputs, including the China Meteorological Administration (CMA), the National Centers for Environmental Prediction (NCEP) of the U.S., the European Centre for Medium-Range Weather Forecasts (ECMWF), as well as others from Australia, Canada, France, Italy, Japan, Korea, Russia and the UK. The S2S model database provides a great opportunity for assessing the prediction skill for extreme events, such as mega-heatwaves in populous China, and more importantly for identifying the sources of prediction skill.

Some recent studies have evaluated the prediction skill of regional heatwave occurrence based on the S2S models. For example, Vitart and Robertson (2018) indicated that the ECMWF model shows capability in predicting high probabilities of extremely high temperature over Russia during the period of the 2010 Russian heatwave occurrence at a lead time of three weeks. For the strong heatwave that occurred in December 2013 over South America, the model from the Australian Bureau of Meteorology shows higher skill than the CMA model in predicting the temperature and circulation anomalies associated with this heatwave one to two weeks in advance (Osman and Alvarez, 2018). Based on the NCEP model's prediction outputs, Ford et al. (2018) documented that few heatwave events over the U.S. continent could be predicted at long lead times (> 15 days) owing to the misrepresentation of land-atmosphere feedback in the model. By assessing the prediction skill of the ECMWF model for a heatwave event over the Yangtze River basin (YRB) in 2012, Qi and Yang (2019) found that the model demonstrates limited skill at a two-week lead time because it poorly captures the circulation anomalies of intraseasonal oscillation. The importance of intraseasonal circulation anomalies to the prediction skill for heatwaves was also pointed out by Hsu et al. (2020), in which the real-time forecasts associated with the record-breaking heatwave over Northeast Asia in summer 2018 by two S2S models, from the CMA and Japan Meteorological Agency, were evaluated. The results of Hsu et al. (2020) suggested that when models correctly predict the phase evolution and amplitude of the trop-

ical western Pacific Madden-Julian Oscillation (MJO) convection, as well as the MJO-related teleconnection pattern, they show higher skill in terms of Northeast Asian heatwave prediction.

Previous studies using the S2S model data have tended to focus on heatwave case evaluations. The robustness of key factors pointed out to influence the heatwave prediction skill needs to be further verified by assessing a larger sample of heatwave events, which is one of the subjects of this study. Although the importance of the ability of models in predicting the intraseasonal circulation anomalies to the prediction of Asian heatwaves has been reported (Yang et al., 2018; Qi and Yang, 2019; Hsu et al., 2020), the relative effects of two distinct modes associated with Asian monsoon intraseasonal oscillation [i.e., the quasi-biweekly oscillation (QBWO) and the MJO] on the heatwave prediction skill have not been addressed and will be discussed in this study. In addition to the atmospheric condition, land surface processes have been identified as being crucial for subseasonal-to-seasonal prediction because of the longer memory of soil moisture than the atmosphere (Koster et al., 2004; Dirmeyer et al., 2009; Ford et al., 2018). The findings based on both observational analysis and model experiments suggest that the occurrence and development of heatwaves are generally induced by anomalous atmospheric circulations, while heatwave intensity and extent could be supported by the dryness of soil moisture (Fischer et al., 2007; García-Herrera et al., 2010; Zhang and Wu, 2011; Seo et al., 2018; Wang et al., 2019). The effect of land conditions on heatwave prediction at the subseasonal range is another issue worth investigating based on the S2S model data.

The YRB, with its large population, is one of the regions with the most frequent heatwaves in China during boreal summer (e.g., You et al., 2017). Although a few studies (Yang et al., 2018; Qi and Yang, 2019) have investigated the subseasonal prediction skill of YRB heatwaves based on limited case studies or on a single model assessment, the overall forecast skill for YRB heatwaves in the current subseasonal-to-seasonal prediction systems and the major sources of predictability need to be further disclosed by using larger prediction samples. These are the key steps for improving forecast accuracy and extending forecast leads of prolonged extreme hot events. Thus, the main aims of this study include a systematic evaluation of YRB heatwave prediction at the subseasonal timescale based on the reforecast (1999–2010) data from the CMA, NCEP and ECMWF S2S models, and a discussion of the key factors, including different intraseasonal modes and land processes that influence the prediction skill for YRB heatwaves.

The paper is organized as follows. In section 2, we introduce the data and methods. The heatwave prediction skills of the three S2S models are evaluated in section 3. The sources of prediction skill for the YRB mega-heatwave event in the summer of 2003 are analyzed and compared among the three models. In section 4, the relative effects of prediction biases with respect to the QBWO and MJO circula-

tions on heatwave prediction are discussed based on a larger sample of data (i.e., including all heatwave events over the reforecast period). The contributions of land surface conditions to the heatwave prediction skill are also examined. Section 5 summarizes the major findings of our study.

## 2. Data and methods

### 2.1. Reforecasts and verification data

The reforecast data from three models (CMA, NCEP and ECMWF) were collected from the server of the S2S database (<http://apps.ecmwf.int/datasets/data/s2s>). Basic descriptions of these models can be found in Table 1. The common period of reforecast for the model comparisons in this study is from 1999 to 2010. We focus on the summer season (June–August, or JJA) for heatwave-related assessments. Note that the CMA and NCEP models were initiated every day, while the ECMWF model produced the forecast twice a week (on Monday and Thursday). To more easily and conveniently compare the prediction skills among the three models, we applied a data processing method proposed by previous studies (Yang et al., 2018; Qi and Yang, 2019) to reprocess the twice-weekly data of the ECMWF model outputs to daily reforecast data (as in the NCEP and CMA models). Briefly, the forecasts from  $N - 2$  to  $N + 2$  days were used to represent the results of the  $N$ -day lead forecast. If there are two values representing a specific lead time forecast, an arithmetic average is applied. According to the results of sensitivity tests, we found that this data processing method will not change any of the main conclusions obtained from either the daily-based reforecast data or the original outputs, as reported by Yang et al. (2018).

The verification datasets include the gridded daily-mean near-surface air temperature (SAT) at 2 m from the CN05 database (Xu et al., 2009) provided by the National Climate Center of China with a horizontal resolution of  $0.5^\circ$ . Large-scale circulations associated with geopotential height at 500 hPa (H500) and zonal wind at 850 hPa (U850) are from the ERA-Interim reanalysis (Dee et al., 2011) at a  $1.5^\circ \times 1.5^\circ$  spatial resolution. The land surface data, including the surface sensible heat flux and soil moisture from ERA-Interim, are also utilized. The surface-layer soil moisture at 0–100 cm is obtained by summing the soil product in three levels (0–7 cm, 7–28 cm, 28–100 cm) with linear weights. Although the information on fluxes and soil depends largely on the reanalysis models, these fields have been shown to have reasonable amplitude and variation as compared with observation (Decker et al., 2012; Liu et al.,

2019). Daily outgoing longwave radiation (OLR) provided by NOAA at a spatial resolution of  $1^\circ$  is used to identify deep convections. For better comparison, we unified the spatial resolution of model outputs to be consistent with observations.

To eliminate the systematic biases of individual models, we removed the climatological mean state for each lead-time forecast. The long-term background components (with a period longer than 90 days) were also removed, to retain the components of subseasonal variation. Then, the high-frequency components ( $< 10$  days) associated with synoptic weather systems, which are not the target of subseasonal-to-seasonal prediction, were also subtracted, by applying a preceding five-day mean of each prediction time step (i.e., the average from day  $-5$  to day 0). The resultant component has roughly the time scale of 10–90 days, referred to as intraseasonal signals. To further separate the intraseasonal signals into the QBWO and MJO, we applied a preceding 15-day mean (average from day  $-15$  to day 0) in the last step to remove the high-frequency component with a period longer than 30 days, and thus obtained the MJO-related signals. The QBWO-related signals were derived from the difference between the intraseasonal (10–90 days) and MJO (30–90 days) components.

### 2.2. Definitions of heatwave and evaluation metrics

Given the fact that models generally produce errors when representing extreme amplitudes, an absolute (or a fixed) threshold for defining heatwave events as used in observation is not applicable in models. Thus, we adopted the threshold relative to climatological SAT distributions in the individual datasets to define the observed and modeled heatwaves. Specifically, a heatwave occurs when the SAT exceeds a relative criterion of the 90th percentile for at least three consecutive days. Considering the seasonal cycle of SAT, the criterion of the 90th percentile is defined for each day using a 15-day moving window (day  $-7$  to day 7) during JJA of 1999–2010.

The pattern correlation coefficient (PCC) and the temporal correlation coefficient (TCC) are used to respectively evaluate the similarity of the spatial distribution and temporal evolution between the prediction and observation. The root-mean-square error (RMSE) estimates the amplitude biases of prediction. The categorical verification score known as the Heidke Skill Score (HSS) (Heidke, 1926) serves as the metric to evaluate the hit rate of heatwave days. The HSS is related to two quantities—the proportion correct (PC) measurement and its maximum-likelihood estimate of probability ( $E$ ), which are defined as follows:

**Table 1.** Description of the CMA, NCEP and ECMWF models' reforecast data.

Model	Forecast range	Model resolutions	Type	Forecast frequency	Period	Ensemble size
CMA	0–60 d	T106, L40	Fixed	Daily	1994–2014	4
NCEP	0–44 d	T126, L64	Fixed	Daily	1999–2010	4
ECMWF	0–46 d	Tco639/Tco319, L91	On the fly	2/week	Past 20 years	11

$$PC = \frac{a+d}{a+b+c+d} = \frac{a+d}{n}, \quad (1)$$

where  $a$  represents the number of observed heatwave days that are correctly predicted,  $b$  is the number of false alarms,  $c$  represents the observed heatwave days that are not predicted, and  $d$  is the number of correct rejections. According to Eq. (1), the perfect value for PC is the unity and the reference value is the chance agreement. Therefore,  $E$  is written as

$$E = \left(\frac{a+c}{n}\right)\left(\frac{a+b}{n}\right) + \left(\frac{b+d}{n}\right)\left(\frac{c+d}{n}\right), \quad (2)$$

and HSS is then defined as

$$HSS = \frac{PC - E}{1 - E} = \frac{2(ad - bc)}{(a+c)(c+d) + (a+b)(b+d)}, \quad (3)$$

in which HSS ranges from  $-\infty$  to 1. A negative value indicates that the random forecast is better, while a value of zero means no skill. A perfect prediction obtains an HSS value of 1.

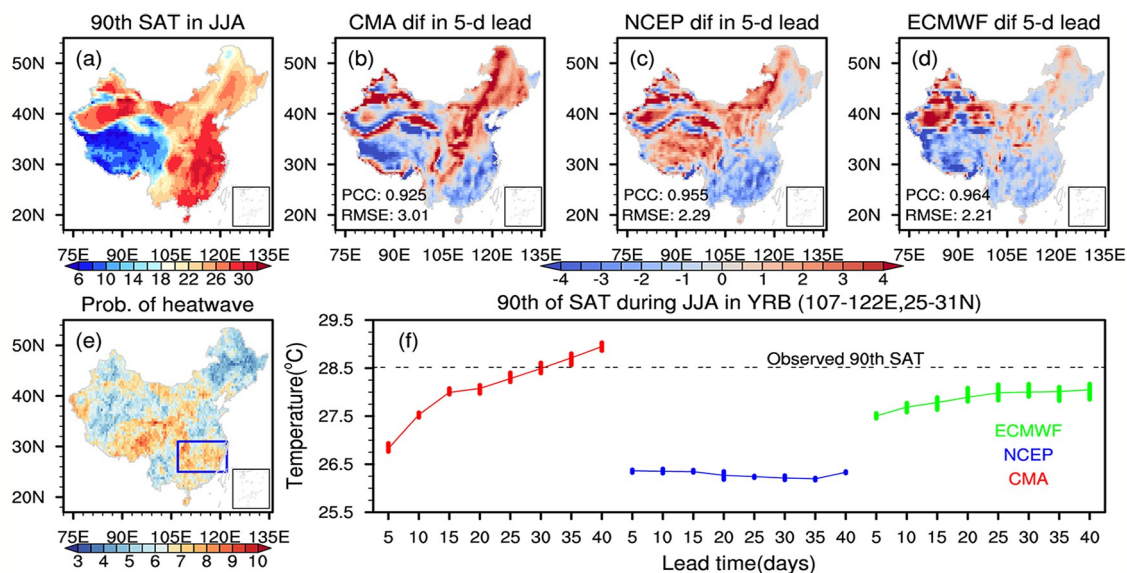
### 2.3. QBWO and MJO indices

The states of the QBWO and MJO over the Asian monsoon region were described by the real-time boreal summer intraseasonal oscillation indices proposed by Lee et al. (2013). The first two leading modes of the multivariate empirical orthogonal function (EOF) of deep convection (OLR)

and low-level circulation (U850) anomalies over the Asian summer monsoon region ( $10^{\circ}\text{S}$ – $40^{\circ}\text{N}$ ,  $40^{\circ}$ – $160^{\circ}\text{E}$ ) present the MJO-related signals, which propagate northwards/northeastwards from the tropical Indian Ocean towards East Asia. The third and fourth EOF modes capture the northwestward propagating QBWO from the tropical central-western Pacific towards East Asia. The respective life cycles of the MJO and QBWO can be divided into eight distinct phases by constructing phase diagrams of the corresponding principal components (PCs). The variability of MJO-related PCs  $[(PC1^2 + PC2^2)^{1/2}]$  and QBWO-related PCs  $[(PC3^2 + PC4^2)^{1/2}]$  is defined to describe their amplitude. Active MJO and QBWO events are selected based on an amplitude greater than 1. The predicted MJO and QBWO states were obtained by projecting the modeled OLR and U850 anomalies at different forecast lead times onto the observed EOF modes.

### 3. Assessments of subseasonal heatwave prediction skill

The threshold (90th percentile) of JJA SAT for the heatwave definition over China is shown in Fig. 1a. High SAT occurs over southeastern and northwestern China. The 90th-percentile values of JJA SAT maximize over the YRB with SAT higher than  $30^{\circ}\text{C}$  in some grids here. Thus, a heatwave will be recorded in these grids over the YRB when the high SAT (greater than the 90th percentile) lasts more than



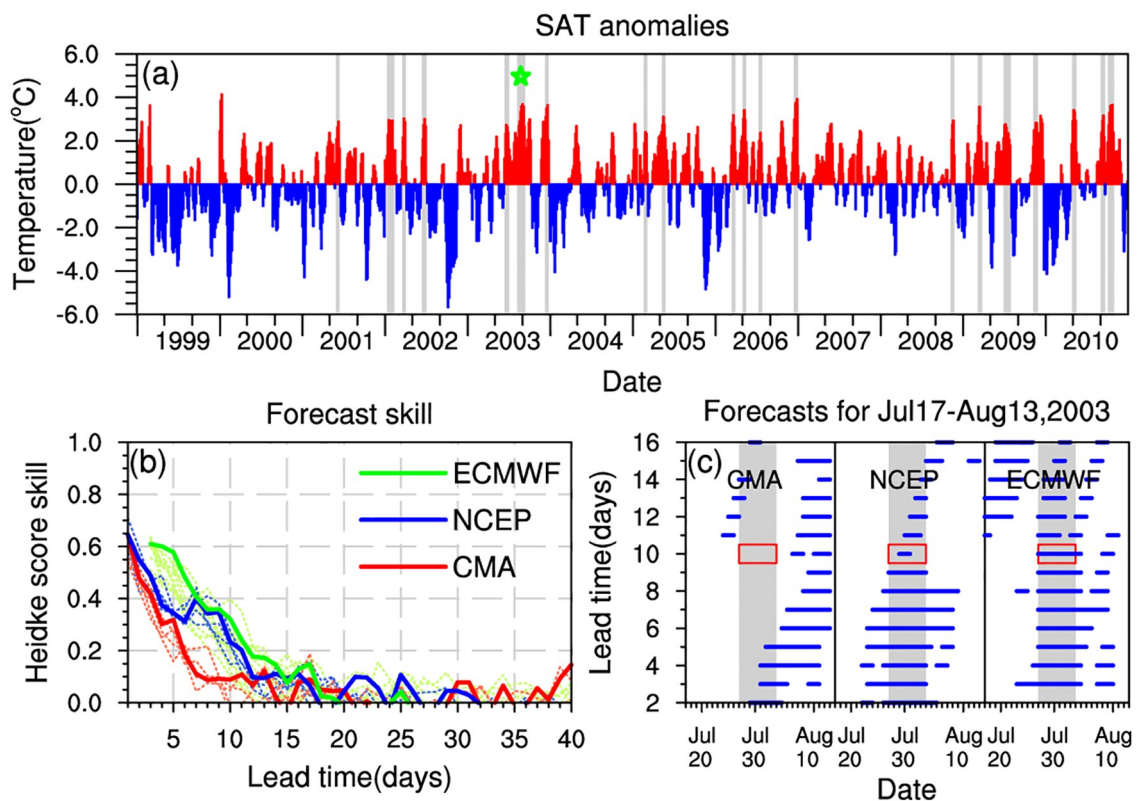
**Fig. 1.** Distributions of (a) 90th percentile values of observed SAT in JJA of 1999–2010 (units:  $^{\circ}\text{C}$ ) and the prediction biases (model minus observation) at a 5-day forecast lead time in the ensembles of (b) CMA, (c) NCEP and (d) ECMWF models, respectively. The PCCs and RMSEs between forecasted and observed values are given in the bottom-left corner of each panel. (e) The climatological probabilities (units: %) of observed heatwave occurrence in JJA of 1999–2010. The blue box delineates the key region of the YRB ( $25^{\circ}$ – $31^{\circ}\text{N}$ ,  $107^{\circ}$ – $122^{\circ}\text{E}$ ) for assessment. (f) The 90th percentile values of area (YRB)-averaged SAT (units:  $^{\circ}\text{C}$ ) at the forecast leads of 5, 10, 15, 20, 25, 30, 35, and 40 days, respectively. The red, blue and green curves represent the averages of ensemble members from the CMA, NCEP and ECMWF models, respectively, along with inter-member spreads shown by whiskers. Observed 90th percentile value is displayed by the horizontal dotted line for comparison.

three days. Based on the heatwave definition, we analyzed the climatological probability of heatwave occurrence over China (Fig. 1e) and found that heatwaves occur more frequently over the YRB (25°–31°N, 107°–122°E) and southwestern China (You et al., 2017). Because of the high summer-mean SAT (Fig. 1a) and the high probability of heatwave occurrence (Fig. 1e) detected in the YRB, the populated region of the YRB appears to be vulnerable to heatwaves. Hence, we selected the YRB as the key area for investigating the subseasonal prediction of heatwaves.

Ensemble-predicted (mean of ensemble members) 90th-percentile values of JJA SAT at the lead time of 5 days reveal high PCCs of ~0.93 (CMA) to ~0.96 (NCEP and ECMWF) with observation, while cold biases were predicted by all three models over southeastern China (Figs. 1b–d), including the YRB. Figure 1f compares the 90th-percentile values of regional (YRB) SAT predicted at lead times of 5 to 40 days from the ensemble mean (color curves) of the three S2S models, respectively. The cold biases of 90th percentile thresholds in the three models are shown at nearly all different forecast leads, as compared to the observation (the horizontal dashed line). The heatwave

thresholds for the ensemble members of the CMA model (red curve) increase with the forecast lead from 5 days (~27°C) to 40 days (~29°C). The NCEP model displays an evident underestimation of SAT. The thresholds of a predicted heatwave in the NCEP model lie between 25.9°C to 26.2°C for different leads (blue curve). As for the ECMWF model predictions, the SAT criterion for the heatwave definition is around 27.5°C–28°C (green curve).

Based on the relative threshold (90th percentile) of SAT and its duration (at least three days), we identified individual YRB heatwave events in summers (JJA) over the period of 1999–2010 (gray shading in Fig. 2a). In total, there were 20 heatwave cases observed during this period. Most of the heatwave events lasted for three days (nine cases) and four days (seven cases). The longest heatwave event, of 8 days, exceeding the 90th percentile of SAT occurred in the late July and early August of 2003 (marked by the green asterisk in Fig. 2a). Figure 2b compares the three S2S models' HSSs estimated by the fraction of correct predictions for heatwave days in the YRB during the summer seasons after eliminating the corrected predictions due to random chance. In both the individual members' and the



**Fig. 2.** (a) Observed area-averaged SAT anomalies (color bars; units: °C) and detected heatwave events (gray shading) over the YRB during the summers (JJA) of 1999–2010. The green asterisk marks the period of mega-heatwave in 2003. (b) Heidke Skill Score (y-axis) of heatwave predictions at different forecast leads from 1 to 40 days (x-axis) for the individual members (dashed curves) and the ensemble predictions (solid curves) of the CMA (red), NCEP (blue) and ECMWF (green) models. (c) Comparisons of observed and ensemble-forecasted heatwave occurrence periods for the mega-heatwave event in 2003. The gray shading marks the heatwave occurrence period in the observation (27 July to 3 August 2003). The blue horizontal lines indicate the predicted heatwave occurrence periods at different forecast leads (y-axis) by the CMA (left), NCEP (middle) and ECMWF (right) models, respectively. Red boxes mark the predicted heatwave at the lead time of 10 days.

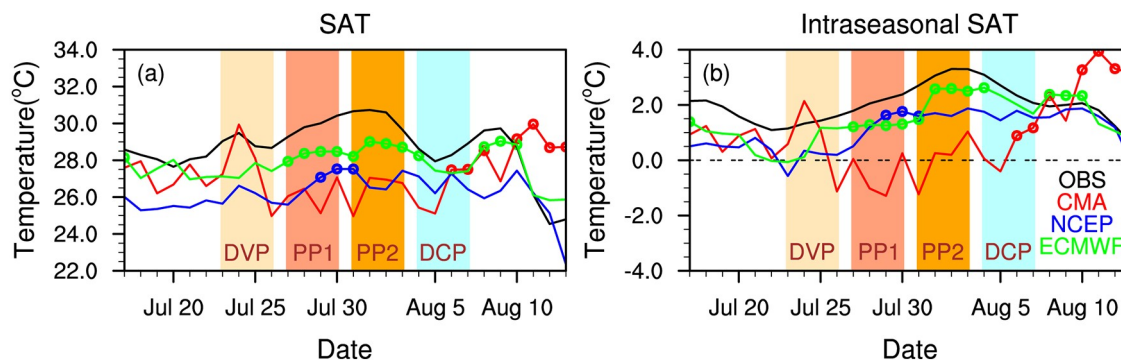
ensemble predictions, the HSS drops quickly as the forecast lead increases from 5 to 15 days in all three models. The model predictions show no skill (HSS of  $\sim 0$ ) beyond the 15-day lead. Compared to the predictions of individual members (dashed curves), the ensemble prediction (based on the mean of all members' SAT) shows higher skill (solid curves) at most of the forecast lead times. Among the three models, the ensemble prediction of ECMWF is superior at the lead time of 5–20 days. The HSS was reduced to 0.1 at the lead time of 12 days in all ensemble predictions from three models. The results suggest that the S2S models have some capability in predicting the extreme hot days associated with YRB heatwave events beyond the time scale of weather forecasts (i.e., 7–10 days in advance), while the skill is limited beyond the lead time of 12–15 days. Since the ensemble predictions have a higher and more stable predictive skill (Fig. 2b), only the ensemble prediction results are shown in the following analyses.

To further examine the sources of prediction skill for YRB heatwaves, we selected the prolonged mega-heatwave in the summer of 2003 (green asterisk in Fig. 2a), which caused tremendous losses to the economy but had a low predictability, as a case study. Figure 2c shows the model-predicted occurrence of the 2003 heatwave event at different forecast lead times (blue curves) against the observation (gray shading). Overall, the CMA model failed to predict the occurrence period of this heatwave event. The predicted heatwave appeared later than observed at all lead times (left-hand panel of Fig. 2c). In contrast, the NCEP model showed good skill in predicting this long-lasting heatwave at 9 days in advance (middle panel of Fig. 2c), although it tended to overestimate the duration of the heatwave at the lead times of 3–8 days. From the forecast lead of 10 days and beyond, the NCEP model underestimated the lifespan of this heatwave. Compared to the notable overestimation biases in the NCEP model, the ECMWF model revealed a more reasonable prediction of the heatwave period at the forecast leads of 5–10 days (right-hand panel of Fig. 2c). Given the consider-

able differences in prediction skill among the three models at the 10-day lead time forecast, in the following we examine the factors leading to the prediction biases at the forecast lead time of 10 days and identify the key sources of heatwave prediction skill.

The observed and predicted evolutions of SAT associated with the YRB mega-heatwave (27 July to 3 August) in the late July and early August of 2003 are shown in Fig. 3a. The SAT varied slowly with time and experienced different stages. To examine how well the models can predict the temporal variation of SAT anomalies associated with this mega-heatwave, we divided the heatwave period into four 4-day stages. The observed SAT (black curve in Fig. 3a) increased before the heatwave's initiation (23–26 July, referred to as the developing period) and reached its peak stages as the heatwave was detected (27–30 July for peak period I and 31 July–3 August for peak period II), followed by a decaying period during 4–7 August. The four stages are also clear in the 10–90-day intraseasonal component of observed SAT (black curve in Fig. 3b). The evolutions of total and intraseasonal SAT are highly consistent, suggesting that the intraseasonal oscillation might play a role in causing the prolonged high SAT anomalies.

Similar to the observation, the evolutions of intraseasonal SAT anomalies in the models' predictions are consistent with the total SAT (Figs. 3a and b). The predicted error of intraseasonal SAT anomalies leads to the failure of heatwave prediction in the CMA model at the forecast lead of 10 days. A positive SAT anomaly was predicted by the CMA model during the observed developing period but negative anomalies of SAT appeared in the observed heatwave peak stages (Fig. 3b). When the observed heatwave started to diminish in the decaying period, the SAT predicted by the CMA model showed a significant increase. As a result, the CMA model did not capture the timing of heatwave occurrence as observed (Figs. 2c and 3a). The NCEP and ECMWF models both correctly predicted the evolving features of the SAT anomaly, with an



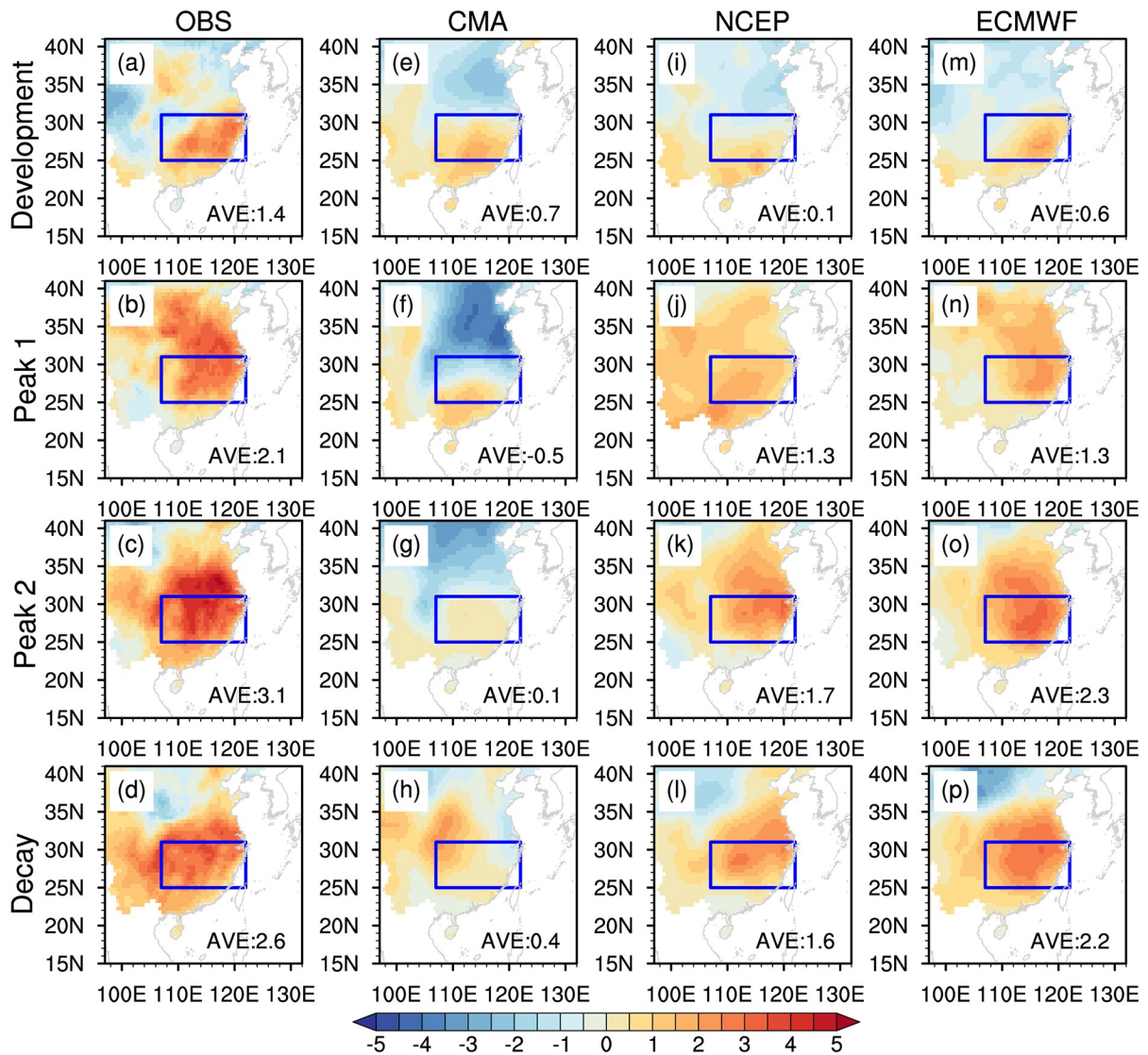
**Fig. 3.** Temporal evolution of (a) SAT and (b) its intraseasonal component over the YRB ( $25^{\circ}$ – $31^{\circ}$ N,  $107^{\circ}$ – $122^{\circ}$ E) for the 2003 mega-heatwave event (units:  $^{\circ}$ C). Black curves show the observations; red, blue, and green curves represent the 10-day lead predictions by the CMA, NCEP, and ECMWF models, respectively. The bars of color shading denote the developing period (DVP, 23–26 July), peak period I (PP1, 27–30 July), peak period II (PP2, 31 July–3 August) and decaying period (DCP, 4–7 August) of the observed mega-heatwave event. Circles on the curves indicate the model-predicted heatwave days.

increasing tendency of SAT anomalies in the developing period but a relatively flat curve of SAT maximum during the observed heatwave periods (peak periods I and II). For the prediction of amplitude, the ECMWF model shows better skill than the NCEP model. The former predicted the large increases in SAT anomaly, close to the observed values in peak period II, while the latter underestimated the SAT anomalies during the observed peak periods I and II (Fig. 3b). Thus, the heatwave predicted by the NCEP model revealed a shorter duration than observed (Figs. 2c and 3a). Overall, the ECMWF model shows higher skill in predicting the initiation, amplitude and duration of this YRB mega-heatwave in the summer of 2003.

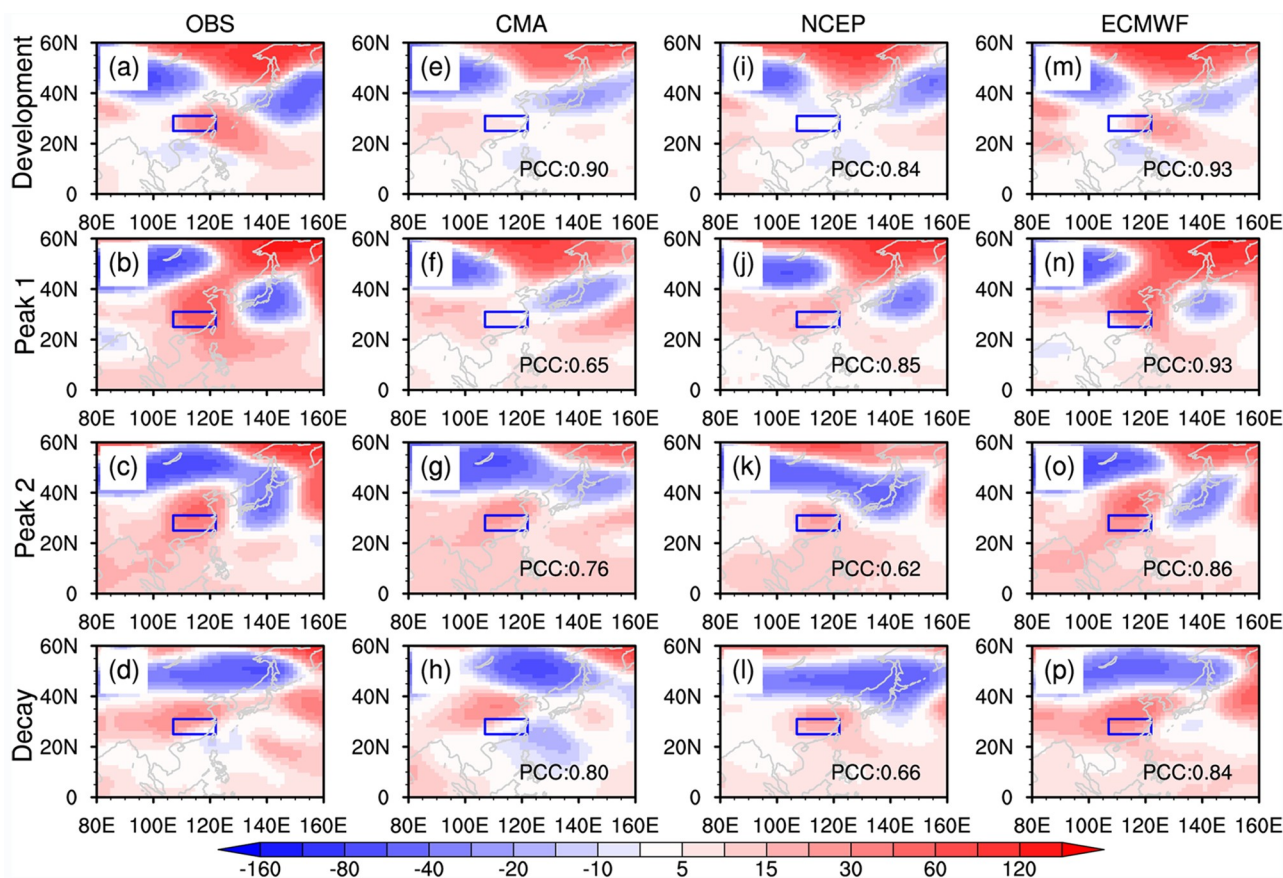
The results of Fig. 3 suggest a linkage between the evolution of intraseasonal SAT anomalies and the life cycle of the heatwave event. Once the model captures the evolution

of intraseasonal SAT anomalies, a higher skill can be achieved in predicting the occurrence period and the amplitude of the heatwave. The intraseasonal SAT anomalies and heatwave occurrence are evidently controlled by the background large-scale circulation anomalies (Hsu et al., 2017). Thus, the modulations of SAT anomalies by the intraseasonal circulation anomalies shown in the observation and the three models' predictions are compared in Figs. 4–6.

Observationally (Figs. 4a–d), the positive intraseasonal SAT anomaly appeared in the southeast coastal areas of China in the developing period, and then developed northwards and strengthened continually during the heatwave occurrence periods (peak periods I and II). In the developing period, the amplitude of SAT anomalies started to decrease over the YRB (Fig. 4d). All three models predicted the positive SAT anomalies in the YRB at the lead



**Fig. 4.** Distributions of intraseasonal SAT anomalies (units: °C) in the developing period (a, e, i, m), peak period I (b, f, j, n), peak period II (c, g, k, o) and decaying period (d, h, l, p) of the 2003 mega-heatwave event. Blue boxes delineate the YRB area: (a–d) observations, and (e–h) CMA, (i–l) NCEP and (m–p) ECMWF model prediction results at the lead time of 10 days. Area-averaged SAT anomalies over the YRB are given at the bottom of each panel.



**Fig. 5.** As in Fig. 4 but for the distributions of intraseasonal H500 anomalies (units: gpm). PCCs between observations and predictions over the Eurasian continent and western North Pacific ( $0^{\circ}$ – $60^{\circ}$ N,  $80^{\circ}$ – $160^{\circ}$ E) are given at the bottom of each panel.

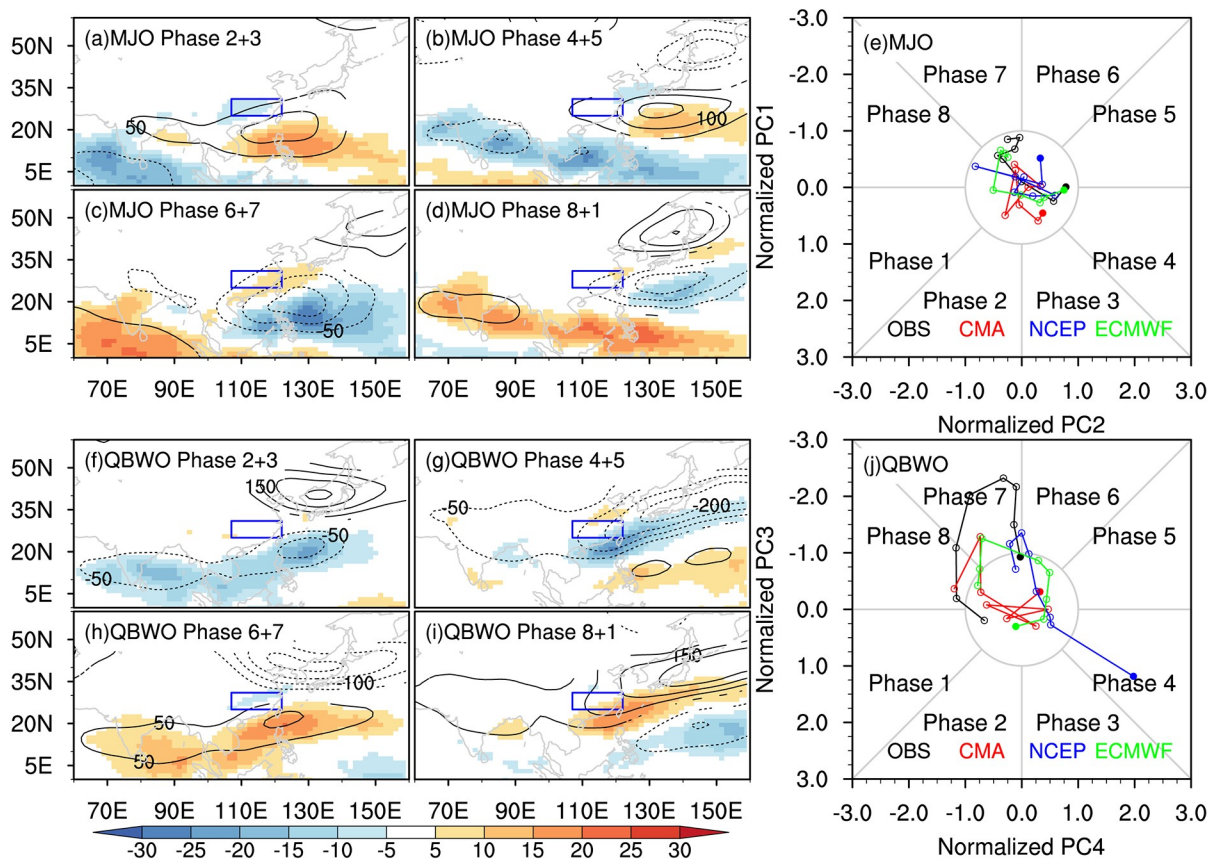
time of 10 days, while they underestimated the heatwave amplitude (Figs. 4e–p), due partly to the biased locations of the SAT maximum center. The CMA model did not predict the heatwave occurrence (Figs. 2c and 3) during the observed heatwave period because the SAT anomaly was too small in the peak periods I and II, and it was confined to southern China without moving northwards towards the YRB during the peak phases (Figs. 4f–g). The center of the SAT anomaly in the decaying period was also incorrectly predicted by the CMA model (Fig. 4h). The NCEP model predicted the development of positive SAT anomalies from South China towards the YRB, reaching their maximum in peak period II, although the amplitude of the SAT anomalies was around half that of the observed (Figs. 4i–l). Thus, the heatwave identified by the NCEP prediction is of weaker amplitude and shorter duration (Figs. 2c and 3). The ECMWF model shows good prediction skill in both the amplitude and evolution of the SAT anomalies (Figs. 4m–p), being similar to the observation (Figs. 4a–d). As a result, the ECMWF model successfully captures the duration and strength of this heatwave event at a 10-day lead (Figs. 2c and 3).

The variation of SAT anomalies is connected with large-scale circulations, in particular the high pressure anomaly-induced adiabatic heating effect (Black et al., 2004; Zhang et al., 2005; Zhang and Zhang, 2010). In Fig. 5, an

anomalous high-pressure anomaly in the mid-troposphere (a positive H500 anomaly) can be seen over the western North Pacific in the developing period (Fig. 5a), which propagates northwestwards to East Asia (including the YRB region) during the heatwave period (Figs. 5b–c), and then continually moves towards inland China in the decaying period (Fig. 5d). The location and amplitude of the H500 anomalies varies consistently with the SAT anomalies. In the initiation and developing stages before the heatwave occurrence, only the ECMWF model predicted the northwest-southeast tilted structure of the H500 anomaly over the western North Pacific at the lead time of 10 days (Figs. 5e, i and m). The ECMWF model also correctly represents the strengthening and northwestward-propagating features of intraseasonal circulation anomalies during the observed heatwave occurrence period (Figs. 5n and o). The PCCs of the intraseasonal H500 anomalies between the observation and ECMWF predictions are high (0.84–0.96) during the entire period associated with heatwave evolution. Lower PCCs of large-scale circulation anomalies were produced by the NCEP and CMA models (Figs. 5e–h and 5i–l). The less skillful predictions for the spatial and intensity evolutions of intraseasonal H500 anomalies will be the source of biases for anomalous SAT and heatwave predictions in the CMA and NCEP models.

To compare the relative contributions of the two dis-





**Fig. 6.** Composites of OLR (shading; units:  $W m^{-2}$ ) and H500 (contours; units: gpm) anomalies for (a) phases 2–3, (b) phases 4–5, (c) phases 6–7, and (d) phases 8 and 1 of the MJO relative to the weak phases (when the MJO amplitude is smaller than 1), based on the observational data. (f–i) As in (a–d) but for phase composites of the QBWO. Only the fields exceeding the 95% confidence level based on the Student’s *t*-test are shown. Panels (e) and (j) display the evolutions of MJO and QBWO phase indices during the heatwave occurrence period (27 July to 3 August 2003), respectively, in observation (black) and the 10-day lead predictions (red, CMA; blue, NCEP; green, ECMWF). The start date of 27 July is denoted by solid dots.

tinct modes of intraseasonal variability (30–90-day MJO and 10–30-day QBWO) to the formation and maintenance of this mega-heatwave event, we analyzed the typical evolutions of convection and circulation associated with the MJO (Figs. 6a–d) and QBWO (Figs. 6f–i) and their states during the heatwave period (Figs. 6e, j). Note that the evolution of intraseasonal H500 anomalies throughout the lifespan of the mega-heatwave (Fig. 5) resembles the northwestward-propagating QBWO (Figs. 6f–i). This is clear in the phase diagrams of the MJO and QBWO. The QBWO signals at phases 6–8 and 1 were enhanced during the heatwave period (Fig. 6j), while the MJO signals were relatively weak (Fig. 6e). The results of Figs. 5 and 6 suggest that the prediction skill of this mega-heatwave event is closely linked with the model capability in predicting the QBWO-related circulation anomalies. The three models all predicted weak MJO signals during the heatwave period, consistent with observation (Fig. 6e). In contrast, the skills of the three models in representing the activities of the QBWO were diverse. Among the three models, the performance with respect to the phase evolution and intensity was the best in the ECMWF model (green curve in Fig. 6j). The NCEP and CMA models pre-

dicted biased evolutions of QBWO phase and associated intensity (blue and red curves in Fig. 6j). Therefore, the prediction skill of this heatwave event seems related to the model prediction skill in terms of the QBWO activity, as the MJO effect was weak for this case. In other words, whether the models are able to correctly predict the occurrence of a heatwave with long-lasting high SAT anomalies depends largely on the predicted pattern and amplitude of intraseasonal circulation anomalies.

Through the atmosphere–land feedback, the clear sky induced by intraseasonal high-pressure anomalies will enhance the surface shortwave radiation flux, which heats the land surface and further favors increased upward sensible heat flux that strengthens and maintains the local SAT anomalies (Black et al., 2004; Zhang et al., 2005; Zhang and Zhang, 2010). The soil conditions beneath the surface also play a role in heatwave intensity through influencing the surface heat fluxes, particularly the evapotranspiration. Sufficient soil moisture may trigger an evaporative cooling effect; on the contrary, dried-out soils associated with a dry and hot spell tend to amplify the extreme high-temperature state (Fischer et al., 2007). Following the analysis of atmo-

spheric circulation anomalies (Figs. 5 and 6), we further diagnosed the anomalous sensible heat flux, standardized soil moisture anomaly, and precipitation anomaly for understanding the effects of land–atmosphere coupling on the prediction of this YRB mega-heatwave. As expected, positive anomalies of sensible heat flux were recorded over the YRB during the developing period and the occurrence periods (peak periods I and II) of the heatwave (Figs. 7a–c), while the degree of positive heat flux anomalies was decreased when the heatwave intensity started to decay in the decaying period (Fig. 7d). The increases in sensible heat flux and SAT were generated by the combined effects of atmospheric high-pressure anomalies (Figs. 5a–d) and the decreases in soil moisture due to deficient precipitation long before the heatwave’s occurrence from early July (black curves in the middle and bottom panels of Fig. 8). The observed SAT and soil moisture anomalies before and during the heatwave period (10 July to 5 August) show a high

TCC of  $-0.72$ , confirming the intense land–atmosphere coupling for heatwave occurrence and maintenance (black curves in the top two panels of Fig. 8). The assessments here of model predictions in terms of atmosphere–land interaction show that the heatwave prediction skill depends on the accuracy of the predicted circulation anomalies as well as the soil moisture related to precipitation anomalies. For example, the surface sensible fluxes predicted by the CMA model at a 10-day lead were mostly opposite to the observed (Figs. 7e–h), which could be related to the biased prediction of atmosphere–land coupling. We analyzed the prediction initiated on 17 July (blue curves in Fig. 8), which is about 10 days ahead of the heatwave’s occurrence, and found that the CMA model predicted small precipitation anomalies and near-normal soil moisture (lack of a dry land surface) during the observed heatwave period (Figs. 8b and c). Therefore, the predicted SAT anomalies showed small amplitude (Figs. 4e–h) that failed to meet the heatwave cri-

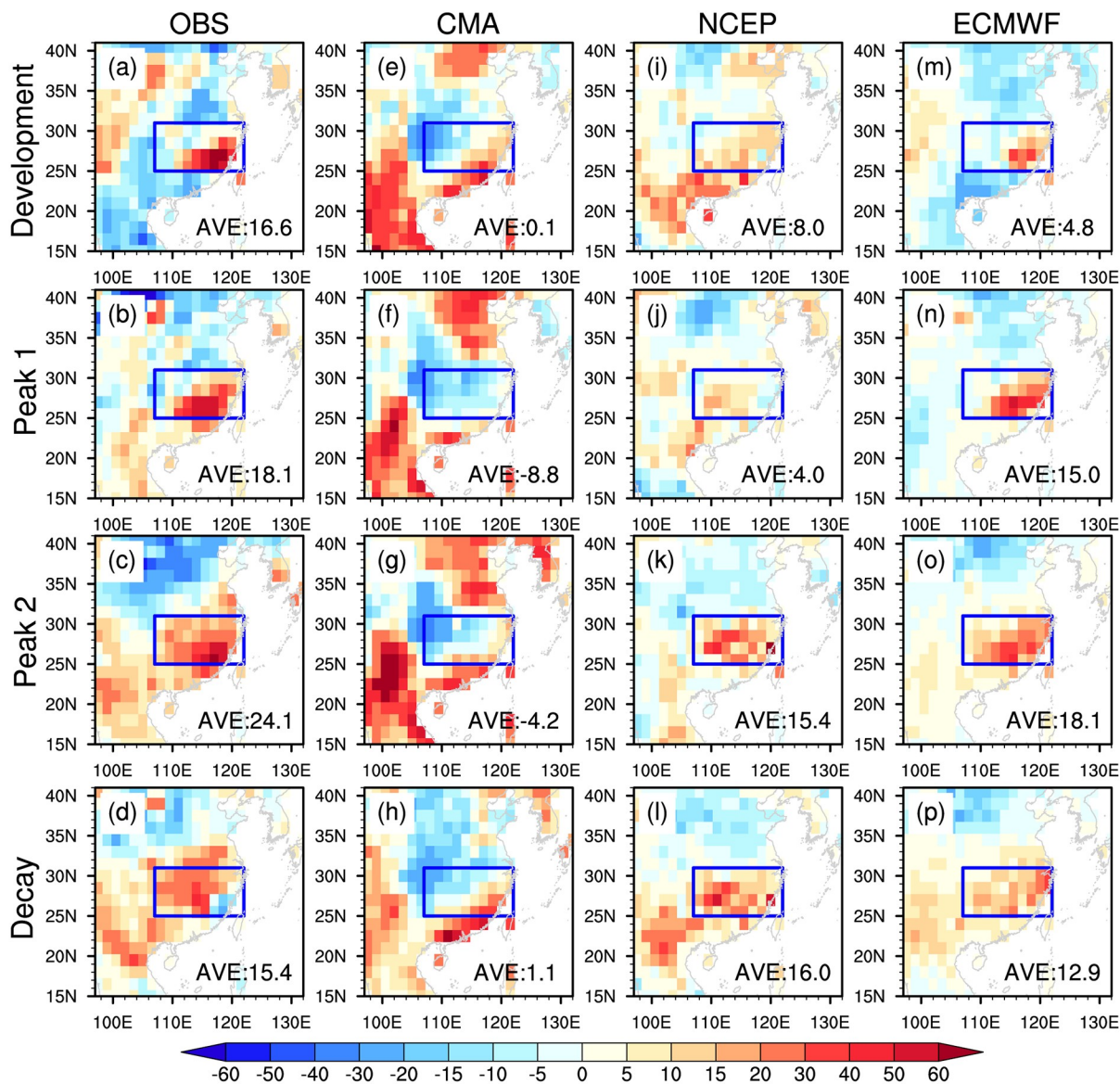


Fig. 7. As in Fig. 4 but for the surface sensible heat flux anomalies over the land area (units:  $W m^{-2}$ ).

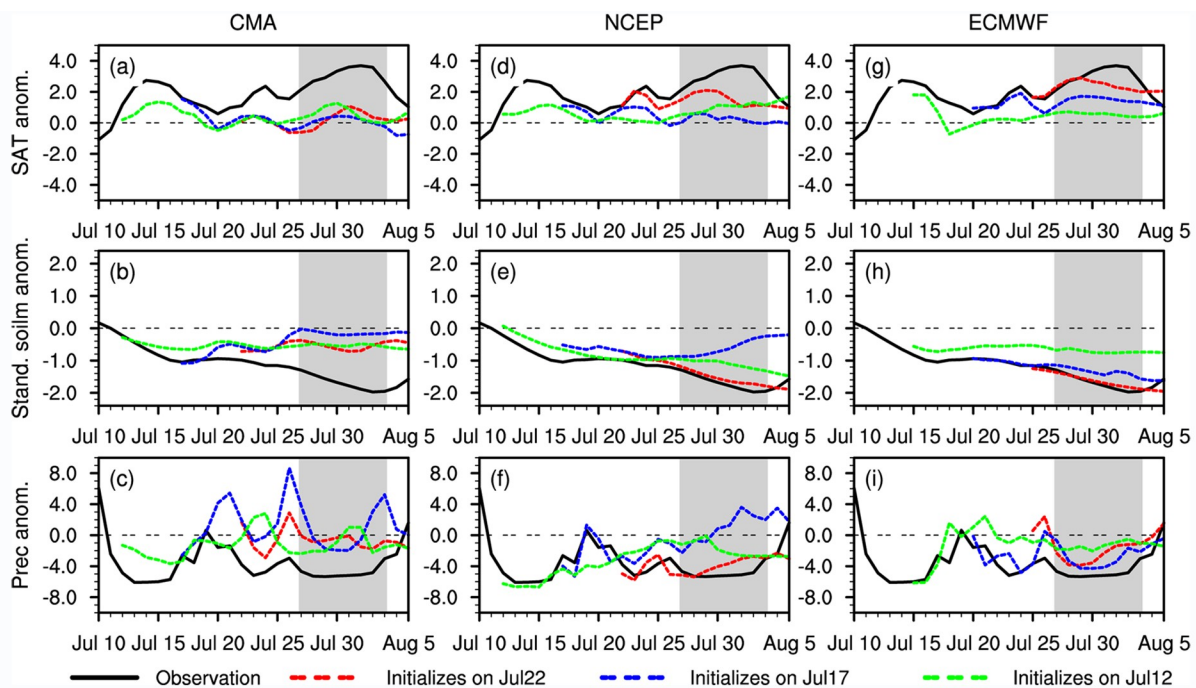
terion (Fig. 2c). The NCEP model also displayed a biased prediction of soil moisture after integration for a few days from 17 July (blue curve in Fig. 8e). Accompanied by an increasing trend in precipitation (Fig. 8f), soil moisture did not decrease continually as observed (Fig. 8e). Thus, a wetter land surface would have favored evaporative cooling and led to insignificant increases in SAT anomalies (Fig. 8d). On the other hand, the ECMWF model predicted a decreasing trend in soil moisture induced by reduced precipitation, consistent with the observation (blue and black curves in Figs. 8h–i). The drying of soil moisture (Fig. 8h) would have contributed to the maintenance of positive SAT anomalies (Fig. 8g). The superior capability of ECMWF model in predicting the soil moisture over China was also documented by Zhu et al. (2019).

By comparing the land surface processes related to this mega-heatwave case predicted at different initial times (different colors in Fig. 8), we note that the biases in SAT concurred with the anomalous land conditions. Thus, the prediction skill seems not to be sensitive to the model initial conditions a long time ahead of the heatwave's occurrence. As long as the model's integrations correctly reproduce the atmosphere–land feedback, the SAT anomalies (and thus the heatwave events) will be better predicted. For instance, the CMA model predicted near-normal precipitation and soil moisture (without decreasing trends) after model initiation at different initial dates (Figs. 8b–c), and the predicted SAT anomalies maintained a flat curve without rising up to form

a heatwave event (Fig. 8a). For the NCEP and ECMWF models, the dry soil associated with reduced precipitation was predicted correctly when the models were initiated 5 days ahead of the heatwave's occurrence (red curves in Figs. 8e–f and h–i), and thus the predicted SAT anomalies were close to the observation (Figs. 8d and g). For the prediction initiated 10 days ahead of the heatwave's occurrence (blue curves), the NCEP model started at a dry condition while it predicted an increasing tendency of precipitation and soil moisture (Figs. 8e–f), leading to small increases in SAT anomalies (Fig. 8d). Likewise, the underestimation of SAT anomalies was apparent in the ECMWF prediction initiated 15 days in advance of the heatwave (green curves) when the model did not capture the decreasing soil moisture and precipitation (Figs. 8 h–i). Note that these precipitation and soil moisture anomalies are highly linked to the evolution of intraseasonal circulation anomalies (Figs. 6–7), suggesting the leading role of intraseasonal oscillation in the predictability of YRB heatwaves (Qi and Yang, 2019).

#### 4. Influences of biases in intraseasonal circulation anomalies and air–land feedback on the heatwave prediction skill

In the previous section, we found that the superior prediction skill of the ECMWF model with respect to the 2003 YRB mega-heatwave prediction at the lead time of 10 days, at which the other two models predicted either none or a



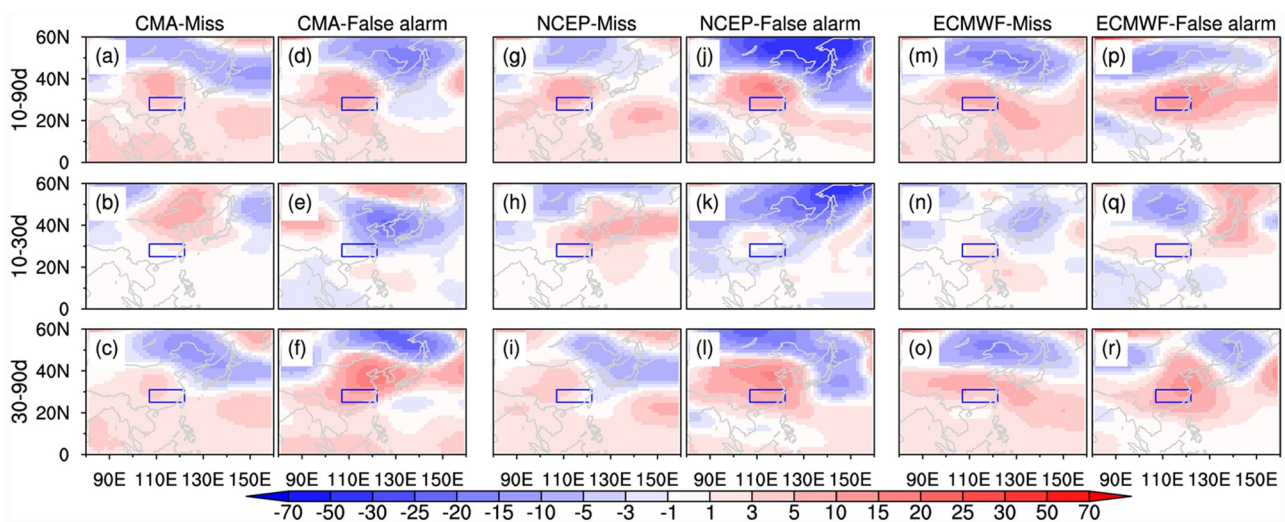
**Fig. 8.** Temporal evolutions of the area-averaged (YRB-averaged) (a, d, g) SAT anomaly (units:  $^{\circ}\text{C}$ ), (b, e, h) standardized soil moisture anomaly, and (c, f, i) precipitation anomaly (units:  $\text{kg m}^{-2}$ ) during the period of 10 July to 5 August 2003, predicted by the (a–c) CMA, (d–f) NCEP (middle), and (g–i) ECMWF models. Black solid curves represent the observations; red, blue and green dashed curves denote the predictions initialized on 12 (green), 17 (blue) and 22 (red) July, which are 5, 10, and 15 days ahead of heatwave initiation, respectively. The gray shading indicates the observed heatwave occurrence period.

weak heatwave, could be the result of its better prediction skill in capturing the evolution and intensity of QBWO-related high-pressure anomalies and the drying of soil moisture induced by reduced precipitation. To further verify whether the intraseasonal circulation anomalies (associated with the MJO and QBWO) and the land–atmosphere interaction play crucial roles in the skillful prediction at the sub-seasonal time scale for other YRB heatwave cases, we examined the biases in circulation anomalies (Figs. 9 and 10) and soil moisture anomalies (Fig. 11) against the biases in SAT anomalies associated with all YRB heatwaves during summers in 1999–2010 at different forecast leads.

The 2003 mega-heatwave case study shows that the failure of models to capture the timing and duration of heatwave occurrence results from the errors in predicted atmospheric circulation anomalies related to the intraseasonal oscillation (Figs. 3b, 5 and 6j). Considering all YRB heatwave events in 1999–2010, Figs. 9 and 10 compare the relative effects of misrepresentations of MJO- and QBWO-related circulation patterns and amplitudes on the errors in heatwave predictions at the forecast lead time of 15 days. When the models failed to predict heatwaves (but they did occur in observations), all three models were generally able to predict the prevalence of high-pressure anomalies over the YRB (Figs. 9a–c, g–i, and m–o) but tended to underestimate their intensities, with about half the amplitude of the observed H500 anomalies (not shown). As a result, the SAT anomalies were too small to reach the heatwave criterion. On the contrary, if the models predict strong high-pressure anomalies associated with intraseasonal oscillation (10–90-day signals), they are more likely to predict the occurrence of a heatwave, even if these heatwaves did not exist (Figs. 9d, j and p). Moreover, the causes of false alarms seem to be more relevant to the MJO-related H500 anomalies (Figs. 9f, l and r) than to the

QBWO signals (Figs. 9e, k and q).

In Fig. 10, we further quantitatively compare the biases in the strength of H500 anomalies associated with the QBWO (orange dots) and MJO (blue dots) against the biases in SAT anomalies for all heatwave days in the observation. Unsurprisingly, the biases in SAT anomalies of heatwaves increase as the forecast lead times become longer from a 10-day lead to 15-day and 20-day leads. Prediction errors in the amplitude of SAT anomalies are small and mostly underestimated (with mean biases of  $-0.71$  to  $-1.2$ ) at the 6–10-day forecast lead (Figs. 10a, d, g), and they rise up to the range of  $-1.15$  to  $-1.67$  at the forecast leads of 11–20 days (bottom two panels in Fig. 10). The prediction errors in the amplitude of H500 anomalies associated with the activities of both the QBWO and MJO are positively correlated with the biases in amplitude of SAT anomalies. These positive correlation coefficients are statistically significant, suggesting that the underestimations of SAT anomalies, which cause the underestimations of heatwave duration and intensity, come mostly from the biases in the strength of large-scale circulation anomalies related to intraseasonal oscillation. We further found that, for a shorter forecast lead time (Figs. 10a, d and g), the effects of QBWO- and MJO-related H500 biases on the fidelity in SAT predictions are comparable. As the lead times become longer (Figs. 10c, f and i), the MJO-related H500 biases show a tighter connection with SAT biases (with a higher correlation coefficient). The higher correlation coefficients of biases between MJO-related H500 and SAT anomalies than those between QBWO-related H500 and SAT anomalies were shown in the predictions of all three models at the forecast lead of 16–20 days, implying that the MJO activity provides a source of YRB heatwave prediction skill at the subseasonal timescale beyond a 10-day forecast lead (Yang et al., 2018;

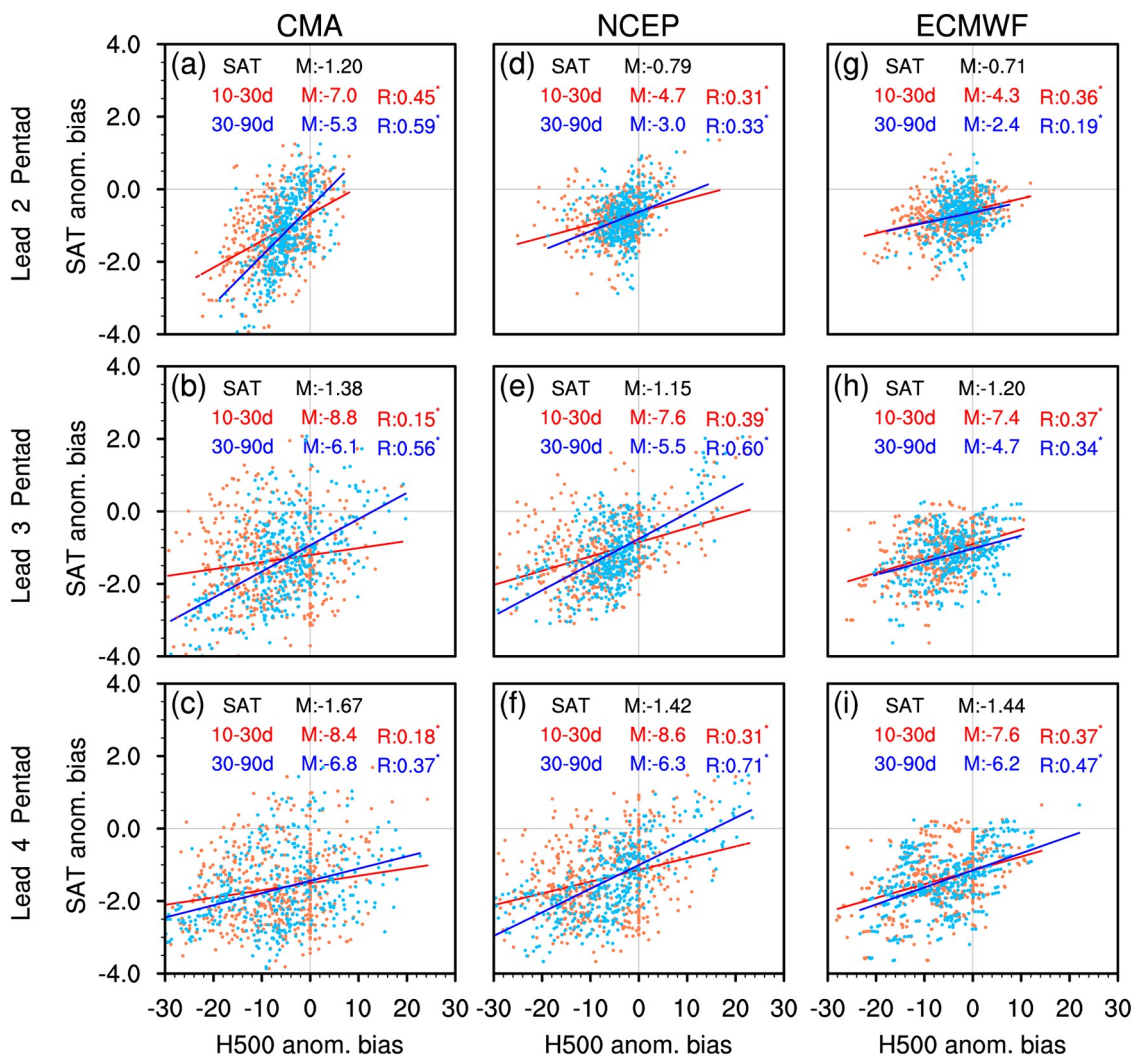


**Fig. 9.** Composites of intraseasonal circulation H500 anomalies (units: gpm) during the dates when models show errors in heatwave prediction at the 15-day forecast lead. (a–c) Composites of 10–90 d, 10–30 d, and 30–90 d H500 anomalies when the CMA model missed the heatwaves (i.e., heatwaves were not predicted to occur, but did occur), respectively. (d–f) As in (a–c) but for the composites of false alarm cases (i.e., heatwaves were predicted to occur but did not occur) in the CMA model. Panels (g–i, m–o) and (j–l, p–r) are the same as (a–c) and (d–f), except for the results of the NCEP and ECMWF models, respectively.

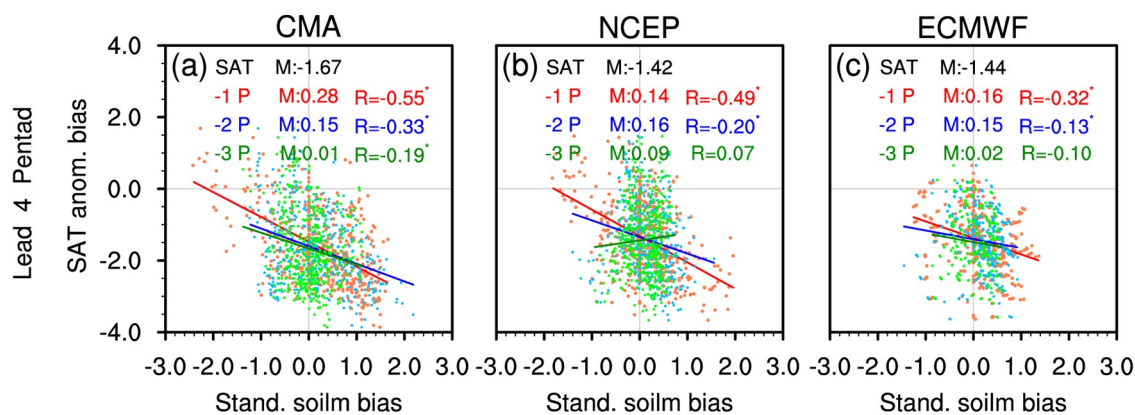
Hsu et al., 2020). For a shorter forecast lead time (e.g., 6–10 days), the model’s capability in representing the QBWO activity is also important for a skillful prediction of YRB heatwaves (Figs. 6 and 10).

A correct prediction of soil moisture anomalies affected by anomalous precipitation is another source of heatwave prediction skill at the subseasonal time scale, according to the prediction evaluations of the 2003 YRB mega-heatwave event (Fig. 8). When the ECMWF model predicted the continuous drying of soil moisture due to a lack of precipitation, positive SAT anomalies occurred and endured for more days. The CMA and NCEP models failed to predict the dryness in the land surface condition, and the SAT anomalies revealed little change without forming a heatwave event. Based on the assessments of all YRB heatwave events, Fig. 11 confirms the effect of preceding soil mois-

ture changes on the prediction bias of SAT anomalies during the observed heatwave periods. The SAT anomalies during the heatwave periods (y-axis) are all negatively correlated with the anomalous soil moistures in advance of heatwave occurrence (x-axis). In the case of predictions at a 20-day lead, the soil moisture shows a smaller bias at the beginning of prediction (green dots for prediction 11–15 days ahead of heatwave occurrence) and the biases in soil moisture have a relatively weak linkage with the SAT anomalies associated with heatwave occurrence (correlation coefficients are 0.19, 0.07 and 0.1 for the CMA, NCEP and ECMWF models, respectively). The correlation coefficients increase along with the increased time of model integration (blue and red colors for conditions 6–10 days and 1–5 days before heatwave occurrence, respectively), indicating that the errors in SAT prediction associated with heatwaves



**Fig. 10.** Scatterplots for (a–c) CMA, (d–f) NCEP and (g–i) ECMWF model prediction biases in SAT anomalies (units: °C, y-axis) against the intraseasonal H500 anomalies (units: gpm, x-axis) associated with the MJO (blue dots) and QBWO (orange dots) for all heatwave days over the YRB during 1999–2010: predictions at lead times of (a, d, g) 6–10 days, (b, e, h) 11–15 days, and (c, f, i) 16–20 days. The linear fit curves for blue and orange dots are presented in dark blue and red, respectively. The average of biases (M) for each variables and the correlation coefficient (R) between SAT biases and H500 biases are given in each panel. R with an asterisk indicates the correlation coefficients are significant at the 95% confidence level based on the Student’s *t*-test.



**Fig. 11.** As in Figs. 10c, f and i but showing scatterplots between the SAT biases during heatwave days against the soil moisture biases at 1–5 days (orange dots), 6–10 days (blue dots) and 11–15 days (green dots) in advance of YRB heatwave occurrence.

are closely related to the predicted biases in soil moisture in advance. In other words, once the biases in soil moisture prediction grow, they will cause a larger bias in the predictions of SAT anomalies through the land–atmosphere interaction, and thus the errors in heatwave prediction. Therefore, an accurate prediction of soil moisture evolution is important to a longer lead (such as the subseasonal time range) prediction for heatwaves over the YRB.

## 5. Summary

The YRB is one of the most densely populated regions in China, and has begun to suffer from frequent extreme high-temperature hazards during boreal summer (Fig. 1e). Understanding the prediction skill of heatwaves in operational forecast models and the key sources of prediction skill and/or biases is the basis to improve the forecast skill of heatwaves with longer lead times (Vitart et al., 2017; Vitart and Robertson, 2018; Yang et al., 2018). Here, we used the reforecast data of 1999–2010 from the CMA, NCEP and ECMWF models that participated in the WMO S2S project to assess the prediction skill of heatwaves over the YRB and discuss the effects of errors in predicted intraseasonal circulation anomalies and land surface conditions on the prediction of heatwaves at the subseasonal timescale beyond 10 days.

For the prediction of SAT in summer (JJA), all the three models show underestimated temperature (cold biases) in southeastern China, including the YRB (Fig. 1). Based on the relative criteria of high temperature at the 90th percentile in the summer season climatologically for the individual models, we defined predicted YRB heatwaves as when the area (YRB)-averaged SAT exceeded a threshold of no less than 3 days. Assessments of the three models' HSS, used to measure the proportion of correct predictions for heatwave days during the summers of 1999–2010, revealed that the three models lost their skill (with an HSS around zero) beyond the forecast lead time of 15 days (Fig. 2b). At the forecast lead of 10 days, the three models

showed diverse skill in their YRB heatwave predictions. The NCEP and ECMWF models performed better than the CMA model. To understand the sources of prediction skill in these three operational models, the 2003 mega-heatwave event with extreme hot days that lasted for 8 days over the YRB was first analyzed. The key factors causing the prediction biases of the heatwave were then confirmed and discussed in more depth by considering all heatwave predictions in the model reforecast period (1999–2010).

For the predictions of the 2003 mega-heatwave event at a 10-day forecast lead, the ECMWF model captured the intensity and duration of high SAT anomalies and thus the long-lasting heatwave event, while the NCEP model underestimated the persistence of high SAT anomalies and the lifespan of the heatwave. The CMA model unfortunately failed to predict the heatwave observed (Figs. 2c and 3). The better prediction skill of the ECMWF model could be attributable to its good skill in predicting the amplitude and evolution of high-pressure (H500) anomalies associated with intraseasonal oscillation (Figs. 4–6) and the relevant land surface process (Figs. 7 and 8). Observationally, the high-pressure anomalies induced by the northwestward-propagating QBWO favored the persistent high SAT anomalies over the YRB (Figs. 5a–d and 6j). Along with the reduced precipitation before the occurrence of the heatwave, the drying of soil moisture also contributed to the maintenance of high SAT anomalies (black curves in Fig. 8). The positive roles of intraseasonal oscillation and soil moisture have been documented in both observational and modeling studies (Fischer et al., 2007; Dirmeyer et al., 2009; Yang et al., 2018; Wang et al., 2019; Hsu et al., 2020). In this study, we emphasize their effects on heatwave prediction in the S2S models. When the NCEP and CMA models biasedly predicted the evolution and intensity of the QBWO (Fig. 6j), the occurrence timing and amplitude of high SAT anomalies showed errors (Fig. 3), leading to limited skill in heatwave prediction (Fig. 2c). The weak circulation anomalies in the NCEP and CMA models could further induce smaller decreases in precipitation and soil moisture, which res-

ult in near-normal SAT anomalies via the land–atmosphere coupling (Fig. 8). In contrast, the ECMWF model correctly predicted the propagation of enhanced QBWO-related high-pressure anomalies and the decreasing tendency of soil moisture (Figs. 6j and 8), meaning the amplitude and persistence of high SAT anomalies, as well as the heatwave duration, predicted by this model was closer to observation (Figs. 2c and 3).

The findings with respect to the major factors influencing the prediction skill as revealed by the evaluation of the 2003 YRB mega-heatwave event were confirmed by the analyses using the three models' predictions for all heatwave events in 1999–2010 (Figs. 9–11). A correct prediction of QBWO circulation anomalies plays a crucial role in the accuracy of the 2003 mega-heatwave prediction as the MJO shows little effect on this case; however, for most YRB heatwave events, both the MJO and QBWO circulation anomalies predicted by the models affect the prediction skill for YRB heatwaves in summer (Figs. 9 and 10). In this study, we further found that for a shorter forecast lead time (e.g., a 10-day lead), the effects of QBWO- and MJO-related H500 prediction skill on the fidelity in SAT predictions are comparable. As the lead times become longer (over 15–20 days), the biases in MJO circulation predictions cause a significant degrading of SAT predictions and thus heatwave prediction skill (Fig. 10). Thus, the MJO activity over the Asian monsoon regions could be the vital source of YRB heatwave prediction skill at the subseasonal timescale.

Through the land–atmosphere interaction, the prediction errors in soil moisture induced by precipitation anomalies are also a factor influencing the prediction skill of SAT anomalies (Fig. 11). If the models successfully predicted soil moisture anomalies and precipitation changes, which are largely influenced by the intraseasonal circulation anomalies, before the heatwave occurrence, they could also better predict the amplitude and duration of SAT anomalies associated with the heatwave. In other words, the skill of the models in representing the land–atmosphere coupling is closely connected with the YRB heatwave prediction skill at the subseasonal range, similar to the results of Ford et al. (2018) who focused on evaluating U.S. heatwave predictions.

The contributions of skillful predictions of intraseasonal circulation anomalies and the land surface condition to the prediction of YRB heatwaves found in this study call for further investigation into how to improve the model prediction skill for intraseasonal oscillation and land–atmosphere coupling. Identifying the factors and key processes affecting the phase evolution and strength of the QBWO and MJO in the S2S models (He et al., 2019), as well as their influences on the land surface processes, are the aims of our ongoing research. Besides, the findings of this study are based on assessment and analysis of S2S models' reforecast data. Recent studies found that moisture initialization is important for accurate prediction of the MJO and its teleconnection patterns (Ren et al., 2016; Wu et al., 2020). To what

extent the initial atmospheric and soil moisture anomalies exert impacts on the prediction skill of YRB heatwaves at the subseasonal time scale is worth further studying by conducting forecast experiments.

**Acknowledgements.** The authors would like to thank the anonymous reviewers for their comments, which helped improve the manuscript. This study was supported by the National Key R&D Program of China (Grant Nos. 2018YFC1505804 and 2018YFC1507704) and NSFC (Grant No. 41625019). We appreciate the operational centers for providing their model outputs through the S2S database.

## REFERENCES

- Almeira, G., M. M. Rusticucci, and M. Suaya, 2016: Relacion entre mortalidad y temperaturas extremas en buenos aires y rosario. *Meteorologica*, **41**, 65–79.
- Black, E., M. Blackburn, G. Harrison, B. Hoskins, and J. Methven, 2004: Factors contributing to the summer 2003 European heatwave. *Weather*, **59**, 217–223, <https://doi.org/10.1256/wea.74.04>.
- Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. B. Zeng, and M. G. Bosilovich, 2012: Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *J. Climate*, **25**, 1916–1944, <https://doi.org/10.1175/JCLI-D-11-00004.1>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Dirmeyer, P. A., C. A. Schlosser, and K. L. Brubaker, 2009: Precipitation, recycling, and land memory: An integrated analysis. *Journal of Hydrometeorology*, **10**, 278–288, <https://doi.org/10.1175/2008JHM1016.1>.
- Fischer, E. M., S. I. Seneviratne, P. L. Vidale, D. Lüthi, and C. Schär, 2007: Soil moisture-atmosphere interactions during the 2003 European summer heat wave. *J. Climate*, **20**, 5081–5099, <https://doi.org/10.1175/JCLI4288.1>.
- Ford, T. W., P. A. Dirmeyer, and D. O. Benson, 2018: Evaluation of heat wave forecasts seamlessly across subseasonal timescales. *npj Climate and Atmospheric Science*, **1**, 20, <https://doi.org/10.1038/s41612-018-0027-7>.
- García-Herrera, R., J. Díaz, R. M. Trigo, J. Luterbacher, and E. M. Fischer, 2010: A review of the European summer heat wave of 2003. *Critical Reviews in Environmental Science and Technology*, **40**, 267–306, <https://doi.org/10.1080/10643380802238137>.
- He, Z., P. C. Hsu, X. W. Liu, T. W. Wu, and Y. X. Gao, 2019: Factors limiting the forecast skill of the boreal summer intraseasonal oscillation in a subseasonal-to-seasonal model. *Adv. Atmos. Sci.*, **36**, 104–118, <https://doi.org/10.1007/s00376-018-7242-3>.
- Heidke, P., 1926: Berechnung Des Erfolges und der güte der windstärkevorhersagen im sturmwarnungsdienst. *Geografiska Annaler*, **8**, 301–349, <https://doi.org/10.1080/20014422.1926.11881138>.
- Hoag, H., 2014: Russian summer tops 'universal' heatwave index. *Nature*, <https://doi.org/10.1038/nature.2014.16250>.
- Hsu, P.-C., J.-Y. Lee, K.-J. Ha, and C.-H. Tsou, 2017: Influences of boreal summer intraseasonal oscillation on heat waves in

- monsoon Asia. *J. Climate*, **30**, 7191–7211, <https://doi.org/10.1175/JCLI-D-16-0505.1>.
- Hsu, P.-C., Y. T. Qian, Y. Liu, H. Murakami, and Y. X. Gao, 2020: Role of abnormally enhanced MJO over the Western Pacific in the formation and subseasonal predictability of the record-breaking Northeast Asian heatwave in the summer of 2018. *J. Climate*, **33**, 3333–3349, <https://doi.org/10.1175/JCLI-D-19-0337.1>.
- Imada, Y., M. Watanabe, H. Kawase, H. Shioyama, and M. Arai, 2019: The July 2018 high temperature event in Japan could not have happened without human-induced global warming. *SOLA*, **15A**, 8–12, <https://doi.org/10.2151/sola.15A-002>.
- IPCC, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 1535 pp.
- Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti, 2004: Realistic initialization of land surface states: Impacts on subseasonal forecast skill. *Journal of Hydrometeorology*, **5**, 1049–1063, <https://doi.org/10.1175/JHM-387.1>.
- Lee, J.-Y., B. Wang, M. C. Wheeler, X. H. Fu, D. E. Waliser, and I.-S. Kang, 2013: Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Climate Dyn.*, **40**, 493–509, <https://doi.org/10.1007/s00382-012-1544-4>.
- Liu, L. W., D. Wei, X. W. Wang, and S. J. Cheng, 2019: Multi-data intercomparison of soil moisture over China. *Journal of Arid Meteorology*, **37**, 40–47. (in Chinese with English abstract)
- Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, **305**, 994–997, <https://doi.org/10.1126/science.1098704>.
- Nageswararao, M. M., P. Sinha, U. C. Mohanty, and S. Mishra, 2020: Occurrence of more heat waves over the central east coast of India in the recent warming era. *Pure Appl. Geophys.*, **177**, 1143–1155, <https://doi.org/10.1007/s00024-019-02304-2>.
- Osman, M., and M. S. Alvarez, 2018: Subseasonal prediction of the heat wave of December 2013 in Southern South America by the POAMA and BCC-CPS models. *Climate Dyn.*, **50**, 67–81, <https://doi.org/10.1007/s00382-017-3582-4>.
- Qi, X., and J. Yang, 2019: Extended-range prediction of a heat wave event over the Yangtze River Valley: Role of intraseasonal signals. *Atmos. Ocean. Sci. Lett.*, **12**, 451–457, <https://doi.org/10.1080/16742834.2019.1669408>.
- Qian, Y. T., H. Murakami, P.-C. Hsu, and S. Kapnick, 2020: Effects of anthropogenic forcing and natural variability on the 2018 heatwave in Northeast Asia. *Bull. Amer. Meteor. Soc.*, **101**, S77–S82, <https://doi.org/10.1175/BAMS-D-19-0156.1>.
- Ren, H.-L., J. Wu, C.-B. Zhao, Y.-J. Cheng, and X.-W. Liu, 2016: MJO ensemble prediction in BCC-CSM1.1(m) using different initialization schemes. *Atmos. Ocean. Sci. Lett.*, **9**, 60–65, <https://doi.org/10.1080/16742834.2015.1116217>.
- Robine, J. M., S. L. K. Cheung, S. Le Roy, H. van Oyen, C. Griffiths, J. P. Michel, and F. R. Herrmann, 2008: Death toll exceeded 70,000 in Europe during the summer of 2003. *Comptes Rendus Biologies*, **331**, 171–178, <https://doi.org/10.1016/j.crv.2007.12.001>.
- Seo, E., and Coauthors, 2018: Impact of soil moisture initialization on boreal summer subseasonal forecasts: Mid-latitude surface air temperature and heat wave events. *Climate Dyn.*, **52**, 1695–1709, <https://doi.org/10.1007/s00382-018-4221-4>.
- Vitart, F., and A. W. Robertson, 2018: The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate and Atmospheric Science*, **1**, 3, <https://doi.org/10.1038/s41612-018-0013-0>.
- Vitart, F., and Coauthors, 2017: The subseasonal to seasonal (s2s) prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, <https://doi.org/10.1175/BAMS-D-16-0017.1>.
- Wang, P. Y., Q. Zhang, Y. Yang, and J. P. Tang, 2019: The sensitivity to initial soil moisture for three severe cases of heat waves over western China. *Front. Environ. Sci.*, **7**, 18, <https://doi.org/10.3389/fenvs.2019.00018>.
- Wu, J., H.-L. Ren, B. Lu, P.-Z. Zhang, C. B. Zhao, and X. W. Liu, 2020: Effects of moisture initialization on MJO and its teleconnection prediction in BCC subseasonal coupled model. *J. Geophys. Res.*, **125**, e2019JD031537, <https://doi.org/10.1029/2019JD031537>.
- Xia, J. J., K. Tu, Z. W. Yan, and Y. J. Qi, 2016: The super-heat wave in eastern China during July–August 2013: A perspective of climate change. *International Journal of Climatology*, **36**, 1291–1298, <https://doi.org/10.1002/joc.4424>.
- Xu, Y., X. J. Gao, Y. Shen, C. H. Xu, Y. Shi, and F. Giorgi, 2009: A daily temperature dataset over China and its application in validating a RCM simulation. *Adv. Atmos. Sci.*, **26**, 763–772, <https://doi.org/10.1007/s00376-009-9029-z>.
- Yang, H., and C. Y. Li, 2005: Diagnostic study of serious high temperature over south China in 2003 summer. *Climatic and Environmental Research*, **10**, 80–85, <https://doi.org/10.3969/j.issn.1006-9585.2005.01.008>. (in Chinese with English abstract)
- Yang, J., T. Zhu, M. N. Gao, H. Lin, B. Wang, and Q. Bao, 2018: Late - July barrier for Subseasonal forecast of summer daily maximum temperature over Yangtze river basin. *Geophys. Res. Lett.*, **45**, 12 610–12 615, <https://doi.org/10.1029/2018GL080963>.
- You, Q. L., Z. H. Jiang, L. Kong, Z. W. Wu, Y. T. Bao, S. C. Kang, and N. Pepin, 2017: A comparison of heat wave climatologies and trends in China based on multiple definitions. *Climate Dyn.*, **48**, 3975–3989, <https://doi.org/10.1007/s00382-016-3315-0>.
- Zhang, J. Y., and L. Y. Wu, 2011: Land-atmosphere coupling amplifies hot extremes over China. *Chinese Science Bulletin*, **56**, 3328, <https://doi.org/10.1007/s11434-011-4628-3>.
- Zhang, M., G. Fu, and J. T. Guo, 2005: Diagnostical analysis of an extreme high-temperature weather event in North China on 15 July, 2002. *Periodical of Ocean University of China*, **35**, 183–188, <https://doi.org/10.3969/j.issn.1672-5174.2005.02.003>. (in Chinese with English abstract)
- Zhang, Y. X., and S. B. Zhang, 2010: Causation analysis on a large-scale continuous high temperature process occurring in North China plain. *Meteorological Monthly*, **36**, 8–13. (in Chinese with English abstract)
- Zhu, H. C., H. S. Chen, Y. Zhou, and X. Dong, 2019: Evaluation of the subseasonal forecast skill of surface soil moisture in the S2S database. *Atmos. Ocean. Sci. Lett.*, **12**, 467–474, <https://doi.org/10.1080/16742834.2019.1663123>.