Design and Construction of Daily-updating Objective Climate Prediction System Based on the Real-time Forecast of CFSv2

Hao Ma[1a*], Fei Yu[2b], Ming Yang[3c], Jingwen Ge[1d], Gaofeng Fan[1e], Ying Liu[1f], Zheyong Xu[4g], Jianjiang Wang[5h], and Hangyuan Sun[5i]

[a*]mahao20032003@aliyun.com, [b]yuf@cma.gov.cn, [c]yangmingstudy@126.com

[d]gejingwen@zju.edu.cn, [e]fangaofengcn@163.com, [f]yuhoubohui@sina.com,

[g]94674775@qq.com, [h]jjwwork@126.com, [i]15210891864@163.com

[1]Zhejiang Climate Center, Zhejiang Meteorological Bureau, Hangzhou 310000, China

[2]China Meteorological Administration Earth System Modeling and Prediction Center, Beijing 100081, China

[3]Zhejiang Meteorological Information and Network Center, Zhejiang Meteorological Bureau, Hangzhou 310000, China

[4]Zhoushan Meteorological Service Centre, Zhoushan Meteorological Bureau, Zhoushan 316021, China

[5]Quzhou Meteorological Observatory, Quzhou Meteorological Bureau, Quzhou 324000, China

Abstract—The CFSv2 forecast products have been widely used in climate prediction operation all over the world. Although the real-time forecast is able to basically capture large pattern of climate anomaly, there still exists obvious bias, which may have enormous impacts on predicted result and thus cannot be neglected. Presently, how to smartly use the massive modeling outputs to improve forecast skill is very important for objective prediction. In this paper, a statistical downscaling strategy for correcting systematic bias through recovering modeling-climatology to its observational counterpart is introduced, and with such methodology, an operational platform conducting real-time 1-30d and 10-30d temperature and precipitation objective prediction is constructed for Zhejiang province. Various verification schemes of the Ps score, Pc score, ACC, SCC, RMSE, the absolute bias, relative bias, and sign coherence are applied on long-term temperature and rainfall assessment. Given the behavior of 335 independent forecast ensembles from January 1st to November 30th in 2019, predictive ability of the downscaling model is generally satisfying. The performance of the Ps score, Pc score, and ACC are similar, and the skillful ensembles are much more than the skillless ones, especially for the Ps score and ACC. Comparatively, SCC can achieve even higher values. RMSE of the 10-30d temperature and precipitation as well as 1-30d rainfall prediction exhibits distinct seasonality. Result of the 1-30d forecast is moderately superior to that of the 10-30d range. While the capability of 1-30d temperature prediction is close to current operational level, the result of monthly-scale precipitation prediction is statistically better than subjective

ICIDC 2022, June 17-19, Qingdao, People's Republic of China

DOI 10.4108/eai.17-6-2022.2322676

forecast. In general, forecast presentation demonstrates this system is practically useful and valuable.

Keywords—CFSv2; systemic bias correction; extended-range; monthly scale; verification; operation platform

1 INTRODUCTION

With the fast development of economy and society, industrial and living need for climate prediction gradually increased in the past several decades. Although precise weather forecast within a week has been available, people still want to know the general climate trend for the future days, which is of great importance for scientifically preparing natural disasters defense and properly arranging substantial activities in advance. Besides, long-term climate prediction is valuable for agricultural production and business analysis.

In China, climate prediction, whose prescription is usually shorter than two seasons, is also referred to as "short-term climate prediction"[1]. Currently, the operational framework covers a broad range of time scales, i. e. from extended-range forecast to seasonal prediction. The main predictive technologies contain mathematical-physical statistics[2-4], numerical simulation[5-9], statistical-dynamical downscaling[10-14], machine learning and deep learning[15-17]... and so on. As global and regional climate models evolve rapidly, nowadays modeling result is able to reasonably predict large-scale anomalies[18], and objective prediction, instead of subjective decision, gradually becomes the primary strategy for modern multi-scale climate forecast, which will be further reinforced in the future[19].

In 2004, National Center for Environmental Prediction (NCEP) released the Climate Forecast System (CFS), which could be looked upon as an extension of common numerical weather prediction[20], and in 2011, the improved adaptation (version 2) of CFS (hereafter CFSv2) became operational[21]. Compared to the original version, CFSv2 upgraded nearly all aspects of the data assimilation and forecast model components[22]. The current version contains an atmospheric model with horizontal resolution of T126 and 64 vertical layers, a classical oceanic model (MOM3), a modified four-level soil model and an interactive three-layer sea ice model. Comprehensive evaluations have shown that CFSv2 exhibits beneficial skills in hindcasting temperature and precipitation in East Asia and China at extended-range and seasonal time-scale[23-27], and such skills can be further enhanced after useful statistical downscaling[23, 28-31]. Now, CFSv2 forecast results have been widely used in operational climate prediction in China[32-36]; however, it has also been demonstrated that there exist distinct systemic biases, which will significantly influence modeling output in the Indo-Pacific region[37], how to deal with these errors remains a thorny problem in operational climate prediction.

Currently, most climate prediction products are commonly issued every 10 days, and this slow frequency cannot fulfill the total need of government and society, especially for some sudden and urgent cases. Enormous public numerical predictions make climate trend outlook with high-frequency update possible[18]. Considering direct forecast from models usually appears apparent bias, it is necessary to conduct systemic bias correction to generate reasonable predictive result.

Zhejiang province, localized in Southeast China, is a small region with the area of only 105, 500km^2, but frequently attacked by various natural disasters such as Meiyu and typhoon. Generally speaking, provincial climate variability is complex and consequently the prediction skill is limited. In this paper, technical route for high-frequency update (daily-updating) objective prediction is investigated in Zhejiang province based on the CFSv2 products, and meanwhile, to demonstrate whether the scheme is effective, a careful evaluation for the modified forecast is also provided.

The paper is organized as follows. A brief description of the data and technical methods for forecast and verification are provided in section 2. Section 3 describes main performance of our downscaling model and central pages of the on-line operational platform. Concluding remarks and further discussion are given in section 4.

2 DATA AND METHODS

2.1 Data

Stational observation: In this research, observational data used are monthly 2-m temperature (hereafter temperature for short) and precipitation at 90 meteorological stations (Banshan, Hangzhou, Hanggao, Sanliting, Sandun, Bin'jiang, Xiaoshan, Qiaosi, Fuyang, Lin'an, Tonglu, Chun'an, Jiande, Donghu, Keqiao, Shangyu, Zhuji, Xinchang, Shengzhou, Gongchengxueyuan, Ertonggongyuan, Beilun, Zhenhai, Yinzhou, Cicheng, Cixi, Yuyao, Fenghua, Shipu, Ninghai, Jiaxing, Xiuzhouqu, Jiashan, Haining, Tongxiang, Haiyan, Pinghu, Quzhouxueyuan, Quzhou, Kaihua, Longyou, Changshan, Jiangshan, Dachendao, Huangyan, Taizhou, Tiantai, Sanmen, Xianju, Linhai, Wenling, Yuhuan, Jinhua, Jinyizhong, Pujiang, Lanxi, Yiwu, Dongyang, Pan'an, Wuyi, Yongkang, Huzhou, Nanxun, Changxing, Anji, Deqing, Dinghai, Putuo, Shengsi, Daishan, Lishui, Suichang, Longquan, Jinyun, Qingtian, Yunhe, Qingyun, Songyang, Jingning, Wenzhou, Yongqiang, Louqiao, Yueqing, Yongjia, Taishun, Wencheng, Pingyang, Ruian, Dongtou, and Cangnan) from 1982-2020. All these stations are regional representative stations at county-level, whose observational quality are relatively reliable[38].

Modeling output: For CFSv2, there are four control runs per day from the 0000, 0600, 1200, and 1800 UTC cycles of the real-time data assimilation system, out to 9 months (red lines in Fig. 1). In addition to the control runs, there are three additional perturbed runs at 0000 UTC, out to one season (blue lines in Fig. 1), and at the 0600, 1200, and 1800 UTC cycles, there are also three additional perturbed runs, out to 45 days (green lines in Fig. 1). Hence, there are totally 16 CFSv2 runs every day, of which four runs go out to 9 months, three runs go out to 1 season, and nine runs go out to 45 days[21]. Currently, only the climatology of the specific runs with forecast period of 45 days is available, and the climatological field is essential for subsequent downscaling, so in daily output, just nine members with 45 days forecast-range are adopted for further treatment.

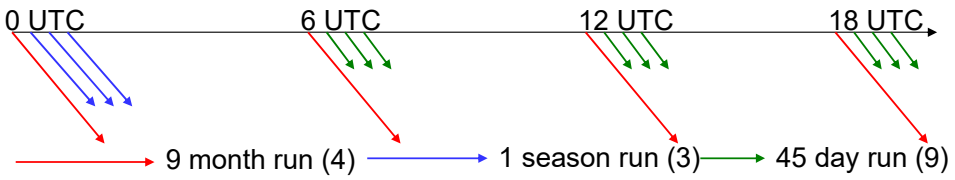

Fig. 1 Operational configuration of the CFSv2 (from Saha et al., 2014).

2.2 Forecast strategy

Traditional climate trend prediction (CTP) usually refers to monthly and seasonal time-scale forecast[39], however, with the rapid growth of governmental and public need, CTP for the extended-range becomes more and more important[40]. Considering the limited predictable scale of the selected CFSv2 forecast members, two time-span of the monthly scale (1-30 day, hereafter 1-30d) and extended-range scale (10-30 day, hereafter 10-30d) are chosen as the target forecast period. Next, we will introduce the detailed process of 10-30d forecast, and the procedure of 1-30d forecast is similar.

Basic logic: To develop a prediction scheme based on the CFSv2 result, modeling forecast ability should be understood first. Two examination indices, i. e. Spatial Correlation Coefficient (SCC, the detailed definition will be introduced in section 2.3.1) and Root Mean Squared Error (RMSE, the detailed definition will be given in section 2.3.1), are chosen to statistically reflect accumulated effect of stational biases for the reforecast of CFSv2 (Fig. 2).

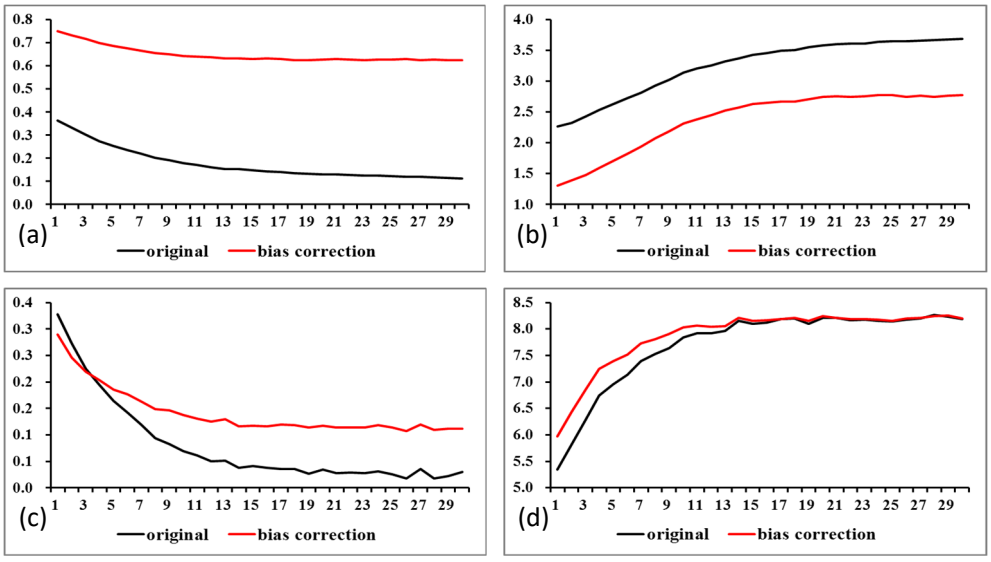

Fig. 2 Comparison of mean daily SCC (a, c) and SRMSE (b, d) for temperature (a, b) and precipitation (c, d) over Zhejiang province for the long-term 1-30d reforecast during 1991-2009 by CFSv2 before and after model-bias correction; unit of (b) and (d) is ℃ and mm respectively.

Generally speaking, the capability of CFSv2 on hindcasting temperature and precipitation is limited. For temperature, initial SCC in the short-term forecast range (STR, 1-10d) is smaller than 0.4, and slowly decreases in the extended-range (ER, 11-30d) (Fig. 2a, black line). Moreover, RMSE is larger than 2.0 ℃ even at the early stage, and gradually rises in the following days (Fig. 2b, black line), which demonstrates the existence of stable systematic bias between model-prediction and observation. For precipitation, the situation is similar, but the bias grows more rapidly for both SCC and RMSE (Fig. 2c and 2d, black lines). Considering the distinct systematic bias is clear, it can be speculated that general performance of CFSv2 may improve after correcting modeling bias, and thus we recover modeling climatology to the observed 30-year mean (1981-2010) climatology in order to reduce forecast bias (the detailed methodology will be introduced in the following part).

Forecast skill of CFSv2 is significantly ameliorated especially for temperature after amending climatology. SCC of temperature uplifts to larger than 0.6 (Fig. 2a, red line), and RMSE is declined by about 1 ℃ (Fig. 2b, red line). For rainfall, although the improvement of RMSE is unapparent (Fig. 2d, red line), SCC is evidently enhanced at the ER (Fig. 2c, red line). Overall, climatology-recovery strategy is effective in correcting model-bias for the hindcast result, and subsequently we will apply this solution on rectifying real-time forecast bias of CFSv2.

Suppose A is certain meteorological variable (temperature / precipitation), and then we have

$$A_O = A_{OC} + A'_O \tag{1}$$

And

$$A_M = A_{MC} + A'_M \tag{2}$$

here A_O and A_M represent observational A and originally forecasted A by CFSv2 respectively, while A_{OC} and A_{MC} stand for climatological A for observation and CFSv2 prediction, and A'_O and A'_M are observational and forecasted anomaly of A. Compared to the actual value of A, the bias of prediction for A' is much smaller, so we approximately have

$$A'_O \approx A'_M \tag{3}$$

Thus, (1) can be written as

$$A_O = A_{OC} + A'_O \approx A_{OC} + (A_M - A_{MC}) = A_M + (A_{OC} - A_{MC}) \quad (4)$$

In this meaning, the modified predicted A can be gotten by correcting modeling climatology. Due to the initial forecast of CFSv2 (A_M) can be gained timely, formation of observational and forecasted climatology is necessary for completing the forecast workflow.

Generation of observational climatology: In normal observation, temperature and precipitation at each station are observed four times (0200, 0800, 1400, and 2000 BJT) per day according to the requirement of China Meteorological Administration (CMA). Therefore, daily temperature / precipitation can be acquired by averaging / summing up four observed values in one day. It should be noted that there exists a lag of 8 hours between UTC and BJT, so a conversion from UTC to BJT is needed. Subsequently, for any day in a year, the climatological daily temperature / precipitation can be calculated as the mean value for the specific day in 30 years from 1981 to 2010. As for the special case of the leap year, which naturally occurs every four years, the climatology of February 29th is obtained by averaging values of all leap years in the total 30 years, i. e. 1984, 1988, 1992, 1996, 2000, 2004, and 2008. Furthermore, for certain day in a year, the climatological temperature / precipitation for the future 10-30d is available by averaging / adding the climatology of each day for the corresponding time-span, i. e. from the 10th day to the 30th day. In Zhejiang province, some automatic stations are lack of long-term observation, and their climatological values can be replaced by that of the nearest normal stations. Also, the observational series of some stations are not complete, and the sparse missing-values can be filled up using reasonable spatial interpolation[41].

Extraction of forecasted climatology: For CFSv2, each member at different time in every day has a separate climatology[21]. That is to say, in a specific day, forecast at 0600, 1200, and 1800 UTC have divergent climatology, and hence there are totally 1098 (3×366) climatology corresponding to all the forecast members. Similarly, for any ensemble, in the 45d time-span, daily forecasted climatology for per member can be formatted, and thus climatology for the 10-30d range (mean value / sum of the 10-30d climatological temperature / precipitation series) can be easily calculated.

Formation of operational forecast: Considering each ensemble has its own climatology in the CFSv2 forecast framework, to form a uniform forecast result, we just take the predicted anomaly as the mean value of various anomaly forecasted by nine members with 45d forecast length. In this meaning, forecasted anomaly is actually “average anomaly”, and (4) can be modified as

$$A_O \approx A_{OC} + (A_M - A_{MC}) = A_{OC} + \overline{\sum_i \sum_j (A_{Mij} - A_{MCi})}$$

$$= A_{OC} + \frac{1}{9} \times \{[(A_{M061} - A_{MC06}) + (A_{M062} - A_{MC06}) + (A_{M063} - A_{MC06})]$$
$$+ [(A_{M121} - A_{MC12}) + (A_{M122} - A_{MC12}) + (A_{M123} - A_{MC12})]$$
$$+ [(A_{M181} - A_{MC18}) + (A_{M182} - A_{MC18}) + (A_{M183} - A_{MC18})]\}$$
$$= A_{OC} + \frac{1}{9} \times [(A_{M061} + A_{M062} + A_{M063} - 3 \times A_{MC06}) + (A_{M121} + A_{M122} + A_{M123} - 3 \times A_{MC12})$$
$$+ (A_{M181} + A_{M182} + A_{M183} - 3 \times A_{MC18})] \quad (5)$$

2.3 Verification scheme

Spatial evaluation: To assess the result of forecasted spatial distribution, five testing solutions are provided, i. e. the Ps score, the Pc score, Anomaly Correlation Coefficient (ACC), SCC, and RMSE. The detailed algorithms will be introduced as follows.

The Ps score: The Ps score, formulated by CMA, is the current operational grading criterion in China, which can be calculated as

$$Ps = \frac{N0 \times 2 + N1 \times 2 + N2 \times 4}{(N - N0) + N0 \times 2 + N1 \times 2 + N2 \times 4 + M} \times 100 \quad (6)$$

where $N0$ represents the number of forecast stations with the same sign as that of the observational result, while $N1$ and $N2$ stand for the number of forecast stations successfully capture the first-grade ($-2.0 < T' \le -1.0$ or $1.0 \le T' < 2.0$; $-50 < R' \le -20$ or $20 \le R' < 50$; T' and R' denote temperature anomaly (TA) and the ratio of precipitation anomaly (RPA) respectively) and second-grade ($T' \le -2.0$ or $T' \ge 2.0$; $R' \le -50$ or $R' \ge 50$) anomalies respectively, and M reflects the number of missing stations (the forecasted temperature / rainfall grade lower than the second grads but in fact $T' \ge 3$ ℃ or $T' \le -3$ ℃, or $R' \ge 100\%$ or $R' = -100\%$).

The Pc score: The Pc score once prompted by CMA as a traditional examining method before the Ps score[42], whose calculation is relatively laconic and can be described as follows:

$$Pc = \frac{N0}{N} \times 100 \quad (7)$$

The meaning of $N0$ is the same as that in (6), and N represents the number of total stations used for scoring.

ACC: ACC is a widely used verifying scheme recommended by World Meteorological Organization (WMO)[43], and the definition is as follows:

$$ACC = \frac{\sum_{i=1}^{N}(\Delta X_{fi} - \overline{\Delta X_{fi}})(\Delta X_{oi} - \overline{\Delta X_{oi}})}{\sqrt{\sum_{i=1}^{N}(\Delta X_{fi} - \overline{\Delta X_{fi}})^2 \sum_{i=1}^{N}(\Delta X_{oi} - \overline{\Delta X_{oi}})^2}} \tag{8}$$

where ΔX_{fi} and ΔX_{oi} stand for predicted and observational stational TA / the ratio of rainfall anomaly, while $\overline{\Delta X_{fi}}$ and $\overline{\Delta X_{oi}}$ represent the corresponding climatological values during 1981-2010. The signification of N is the same as that in (7). In reality, $\overline{\Delta X_{fi}} = 0$ and $\overline{\Delta X_{oi}} = 0$, so (8) can be further simplified to

$$ACC = \frac{\sum_{i=1}^{N}\Delta X_{fi} \times \Delta X_{oi}}{\sqrt{\sum_{i=1}^{N}\Delta X_{fi}^2 \sum_{i=1}^{N}\Delta X_{oi}^2}} \tag{9}$$

SCC: SCC is similar to classical correlation coefficient, but the meteorological element (temperature / precipitation) varies in the spatial field rather than time series[44], which is capable to reflect the similarity between the forecasted and observational field. Specifically, the formula is as follows:

$$SCC = \frac{\sum_{i=1}^{N}(X_{fi} - \overline{X_{fi}})(X_{oi} - \overline{X_{oi}})}{\sqrt{\sum_{i=1}^{N}(X_{fi} - \overline{X_{fi}})^2 \sum_{i=1}^{N}(X_{oi} - \overline{X_{oi}})^2}} \tag{10}$$

Similarly, here X_{fi} and X_{oi} denote predicted and observational stational temperature / precipitation (or their anomalies), and $\overline{X_{fi}}$ and $\overline{X_{oi}}$ are the corresponding spatial-mean values. The meaning of N is the same as that in (7).

RMSE: RMSE is one of the most mature algorithms measuring the forecast bias, and can be calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(F_i - O_i)^2}{N}} \tag{11}$$

Here F_i and O_i refer to forecasted and observational stational temperature / precipitation respectively, and N also has the same meaning as that in (7).

Stational examination: For single station, three examining methods are provided, i. e. the absolute bias (AB), relative bias (RB), and sign coherence (SC). AB can be represented as the difference between prediction and observation ($X_f - X_o$), and RB can be easily obtained by dividing the observational value into AB ($\frac{X_f - X_o}{X_o}$). SC has two reverse results, i. e. both forecasted and observational values share the same signs or the opposite signs.

3 RESULTS

3.1 Fundamental construction of the operational system

Generally speaking, the daily-updating objective climate prediction system consists of two parts, the forecast module and the verification section. In the former part, current prediction of temperature and precipitation tendency for the future 1-30d and 10-30d and corresponding observational result (when observation can be gotten) are provided, and thus the predicted and observational temperature, TA, rainfall, and the ratio of rainfall anomaly are exhibited. With the real-time date moves, the forecast changes accordingly. The start date is adjustable so as to make the historical prediction information search convenient.

In the later part, AB, RB, and SC for single station and statistical assessment including the Ps score, the Pc score, ACC, SCC, and RMSE for forecast field are furnished for both 1-30d and 10-30d cases. Also, the start date can be changed freely according to forecasters' requirement. It should be noted that all the calculation results including forecast and evaluation can be exported backstage to the Excel documents.

3.2 General interface of the on-line platform

The prediction interface: In order to roundly display the objective prediction and improve sensory taste, all the forecasted results are presented in two forms: spatial-distribution figure and exhaustive table. The plot mainly displays predicted and observational temperature, TA, precipitation, and RPA anomaly (Fig. 3). All these results can be shown at province-level (Zhejiang province) and city-level (11 cities in Zhejiang province, i. e. Huzhou, Jiaxing, Hangzhou, Shaoxing, Ningbo, Zhoushan, Quzhou, Jinhua, Taizhou, Lishui, Wenzhou, not shown). Also detailed forecast for each station and corresponding observation are presented in a blow-by-blow table (not shown).

The verification interface: Similar to the exhibition manner of forecast products, the

verification interface also contains both figures and tables. Figures are mainly used to present spatial distribution of stational AB, RB, and SC, and for SC, red and blue dots stand for the same and opposite signs respectively (Fig. 4). Other statistical results including the Ps score, the Pc score, ACC, SCC, and RMSE, combined with stational biases, are merged into an integral table to show forecast skill in diverse aspects at province, city, and country levels (not shown).

3.3 Performance of large-ensemble prediction experiment

To objectively evaluate the performance of CFSv2 real-time downscaling forecast, a large-ensemble (335 independent days, from January 1st 2020 to September 30th 2020) prediction is carried out. Generally, after correcting systematic bias, downscaling result can give basically satisfying forecasts at both 10-30d and 1-30d range, and the detailed skill for TA and RPA (hereafter "RPA" for short) is analyzed as below.

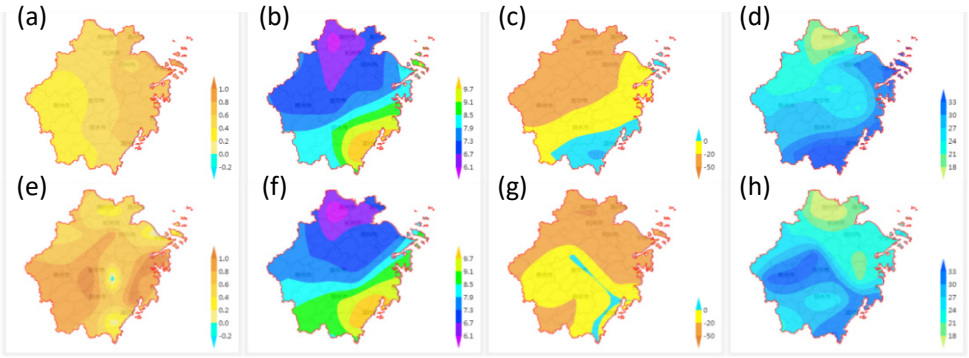

Fig. 3 The main prediction interface for future 10-30d TA (a), temperature (b), RPA (c), and precipitation (d) as well as comparison with observation (e-h) over Zhejiang province (start from November 30th, 2020).

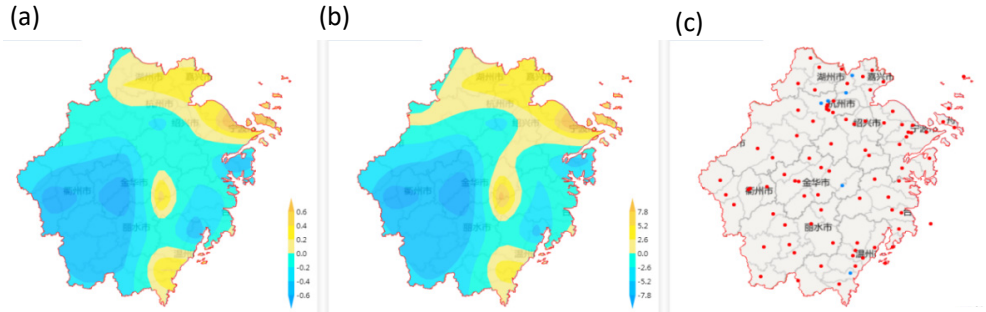

Fig. 4 The main verification interface of spatial distribution of AB (a), RB (b), and SC (c) for future 10-30d temperature prediction over Zhejiang province (start from November 30th, 2020); red and blue dots in (c) represent predicted TA and actual observation share the same and opposite signs respectively.

For 10-30d TA at the ER, mean values of the Ps score and Pc score are 76.5 and 71.8 respectively, and average ACC is 0.2 (Tab. 1). There are 252d (about 75.2%) and 227d (about

67.6%) the Ps and Pc score appearing positive skills (≥60), and more than half period both indices suggesting high skills (215d and 190d respectively). Similarly, there are 219d with ACC larger than 0.3, which even exceeds 0.5 for more than 200d (208d) (Tab. 2). Time series of the Ps score, Pc score, and ACC are broadly consistent, and the high-skill periods primarily occur over three periods: April 11th to June 3rd, July 10th to August 21st, and November 1st to November 23rd. It should be noted that all the three indices experience several dramatic fluctuations, which indicates the forecast skill may not be stable over some specific ranges (Fig. 5a-c). Unlike above indices, SCC almost keeps positive except on June 24th (Fig. 5d), and the averaged value is as large as 0.8 (Tab. 1), signifying the spatial distribution of forecasted 10-30d TA is well in accordance with observation. Temporal evolution of RMSE presents distinct season-dependence with small bias in spring and summer (from April to October) but large bias in autumn and winter (from November to March) (Fig. 5e), and the mean value of 1.3 ℃ is acceptable in the viewpoint of climate prediction (Tab. 1).

Compared to TA, forecast skill for RPA at the ER scale (10-30d) is relatively weak in all the verification indices, and mean Ps score, Pc score, and ACC are 71.3, 57.2, and 0.1 respectively (Tab. 1). Although the number of ensembles with positive skill are still more than that with negative skill for the Ps score (245d versus 90d) and ACC (197d versus 138d), in the sight of the Pc score, ensembles with positive and negative skill are very close (178d versus 176d). Also, for the Ps score and ACC, ensembles with high skill remain abundant (163d and 156d), whereas which are insufficient for the Pc score (98d) (Tab. 2). Due to temporal and spatial changeability of rainfall, stable high-skill is hard to gain, and comparatively persistent high-skill periods mainly distribute over the following spans: March 28th to April 7th, May 29th to July 4th, August 2nd to August 16th, and October 5th to October 23rd (Fig. 6a-c). Although SCC appears positive for most ensembles (262d), there do exist some periods with long-term negative skill (January 1st to January 16th, August 18th to September 7th, September 28th to October 10th, and November 17th to November 24th), which causes average SCC is only 0.3 (Tab. 1), far lower than that of 10-30d TA. It is interesting that when SCC is at low-level, the Ps score, Pc sore, and ACC sometimes situate at high-level, which implies that SCC is able to reflect a different aspect of forecast-level, in contrast to other indices (Fig. 6d). Similar to that of TA, RMSE of RPA also displays prominent seasonality with large and small bias occurring during May-August and September-April respectively, suggesting the amplitude of forecast bias may be associated with climatology (Fig. 6e).

TAB. 1 MEAN VALUES OF THE PS SCORE, PC SCORE, ACC, SCC, AND RMSE OF SBC FOR THE 10-30D AND 1-30D PREDICTION FORECASTED FROM JANUARY 1ST, 2020 TO NOVEMBER 30TH, 2020.

	Ps score	Pc score	ACC	SCC	RMSE
10-30d temperature	76.5	71.8	0.2	0.8	1.3
10-30d precipitation	71.3	57.2	0.1	0.3	60.1
1-30d temperature	81.6	76.4	0.2	0.8	1.1
1-30d precipitation	78.1	64.4	0.2	0.4	68.6

TAB. 2 DETAILED VERIFICATION OF THE NUMBER OF DAYS WITH DIFFERENT SKILLS CORRESPONDING TO THE FOUR INDICES OF THE PS SCORE, PC SCORE, ACC, AND SCC FOR THE 10-30D AND 1-30D PREDICTION FORECASTED FROM JANUARY 1ST, 2020 TO NOVEMBER 30TH, 2020 RESPECTIVELY (UNIT: D).

verification indices		10-30d temperature	10-30d precipitation	1-30d temperature	1-30d precipitation
Ps score	positive skill (>=60)	252	245	263	279
	high skill (>=80)	215	163	232	206
	excellent skill (>=90)	156	93	198	150
	negative skill (<60)	83	90	72	66
Pc score	positive skill (>=60)	227	178	241	205
	high skill (>=80)	190	98	215	135
	excellent skill (>=90)	175	50	195	82
	negative skill (<60)	108	166	94	130
ACC	positive skill (>0)	237	197	250	239
	high skill (>=0.3)	219	156	238	209
	excellent skill (>=0.5)	208	133	229	178
	negative skill (<=0)	98	138	85	96
SCC	positive skill (>0)	334	262	335	285
	high skill (>=0.3)	323	182	333	223
	excellent skill (>=0.5)	318	108	322	142
	negative skill (<=0)	1	73	0	50

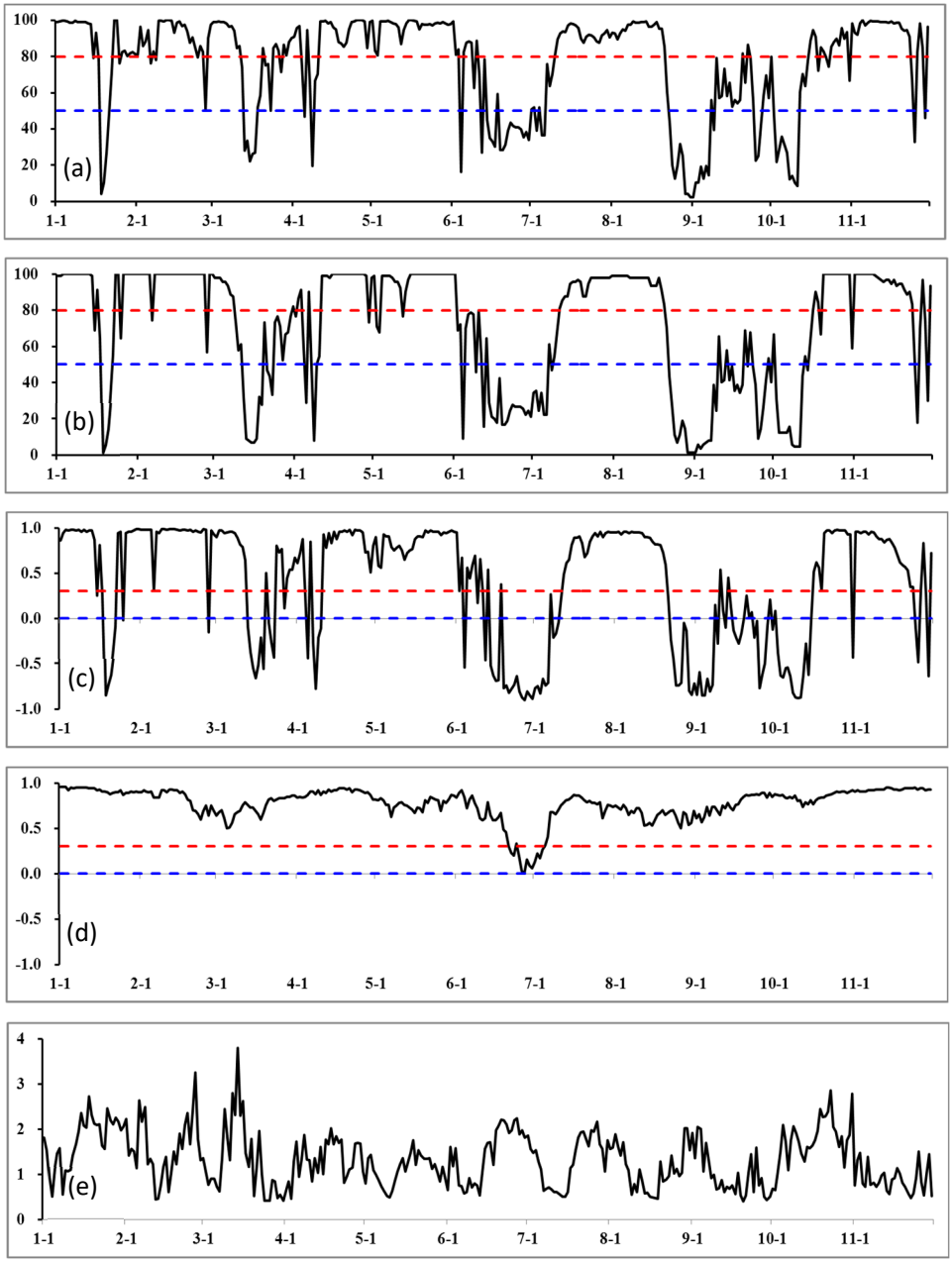

Fig. 5 Time series of the Ps score (a), Pc score (b), ACC (c), SCC (d), and RMSE (e) for the 10-30d temperature prediction by SBC forecasted from January 1st, 2020 to November 30th, 2020; the red and blue dashed lines in (a) and (b) indicate the score of 80 and 50, and in (c) and (d) stand for correlation coefficient of 0.3 and 0.0, respectively; unit of (e) is ℃.

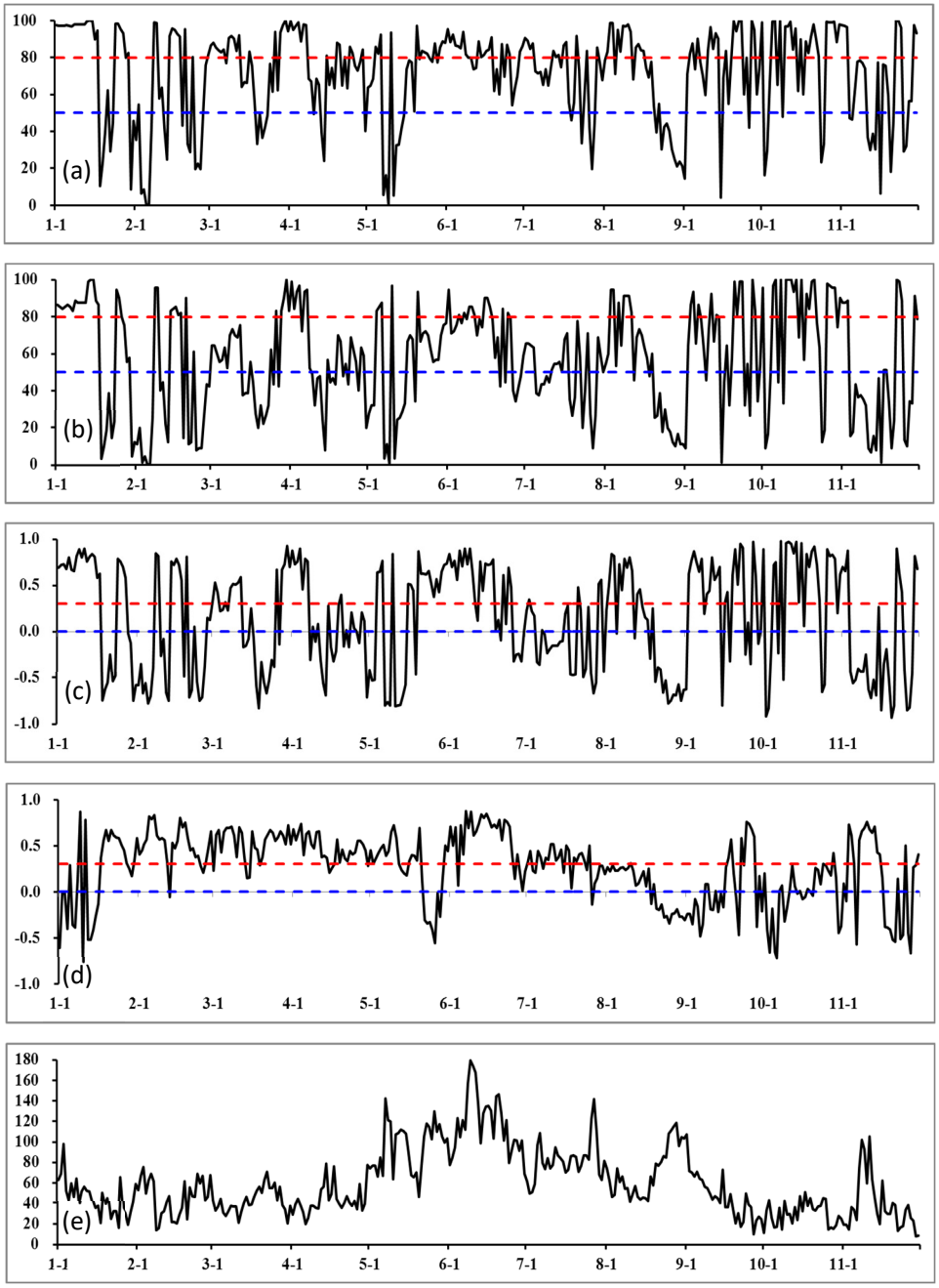

Fig. 6 Time series of the Ps score (a), Pc score (b), ACC (c), SCC (d), and RMSE (e) for the 10-30d precipitation prediction by SBC forecasted from January 1st, 2020 to November 30th, 2020; the red and blue dashed lines in (a) and (b) indicate the score of 80 and 50, and in (c) and (d) stand for correlation coefficient of 0.3 and 0.0, respectively; unit of (e) is mm.

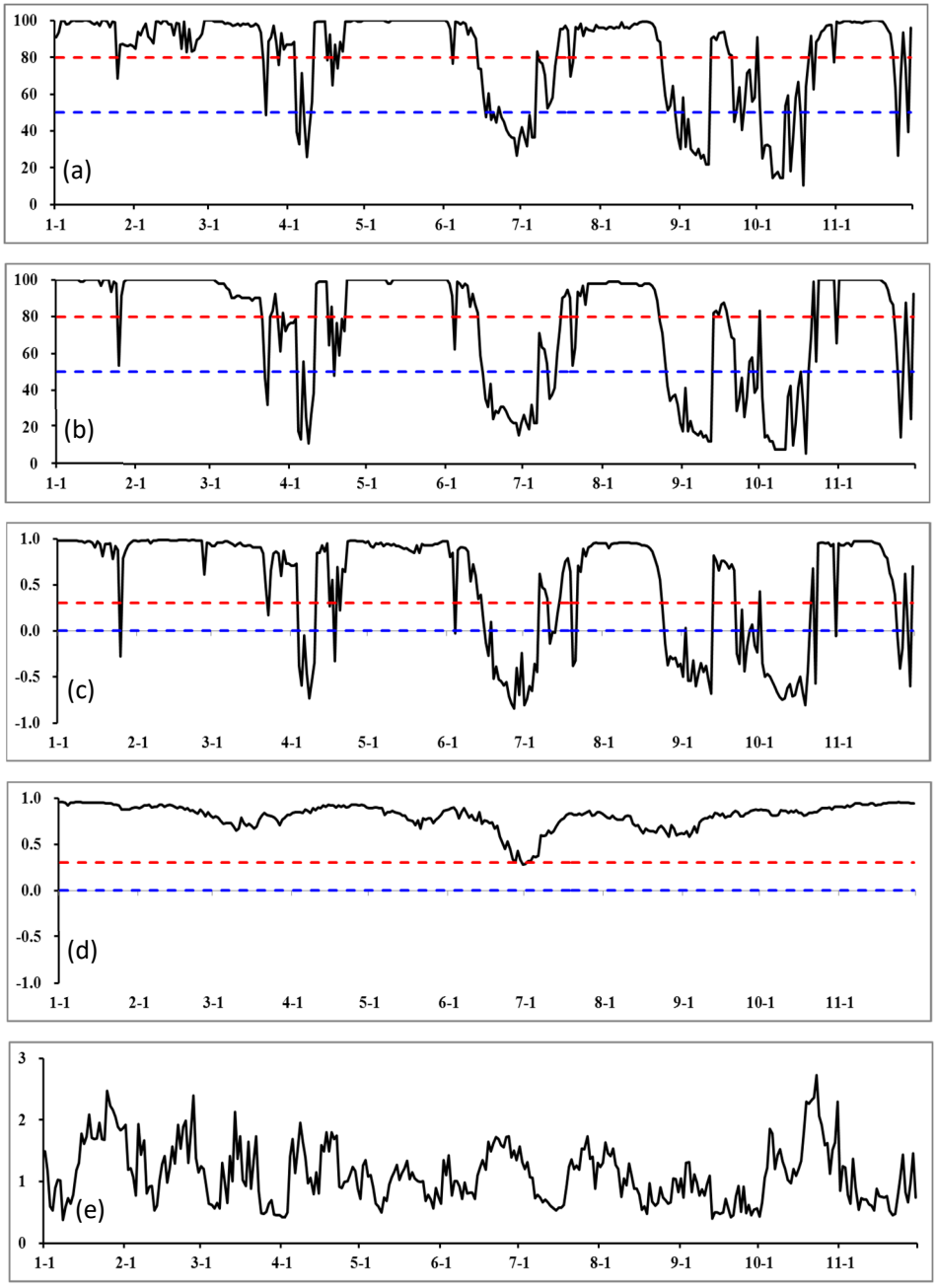

Fig. 7 Time series of the Ps score (a), Pc score (b), ACC (c), SCC (d), and RMSE (e) for the 1-30d temperature prediction by SBC forecasted from January 1st, 2020 to November 30th, 2020; the red and blue dashed lines in (a) and (b) indicate the score of 80 and 50, and in (c) and (d) stand for correlation coefficient of 0.3 and 0.0, respectively; unit of (e) is °C.

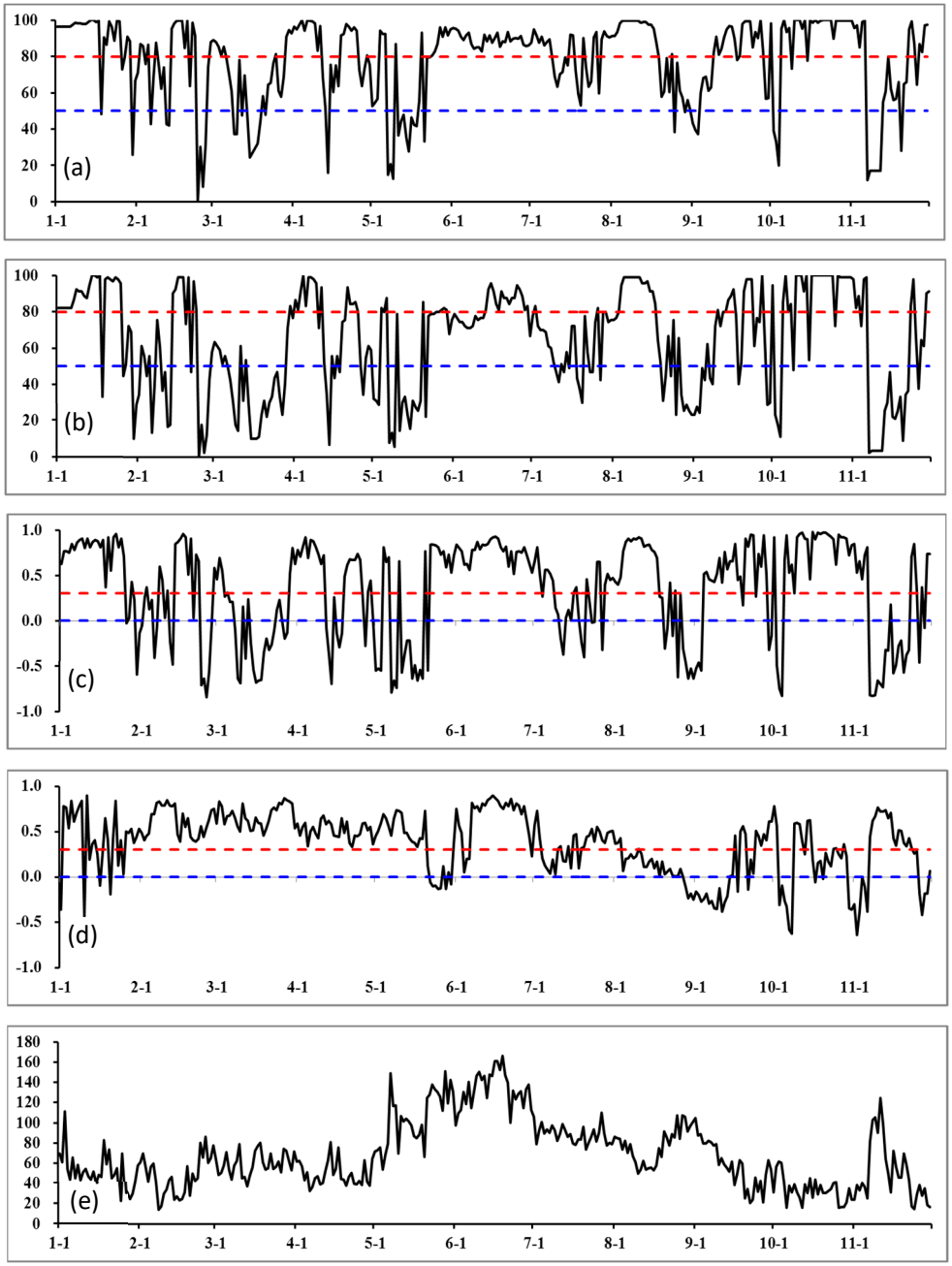

Fig. 8 Time series of the Ps score (a), Pc score (b), ACC (c), SCC (d), and RMSE (e) for the 1-30d precipitation prediction by SBC forecasted from January 1st, 2020 to November 30th, 2020; the red and blue dashed lines in (a) and (b) indicate the score of 80 and 50, and in (c) and (d) stand for correlation coefficient of 0.3 and 0.0, respectively; unit of (e) is mm.

Generally speaking, the forecast ability for the 1-30d range is slightly better than that for the 10-30d span, which demonstrates the prediction model may have better skill for long period. For 1-30d TA forecast, mean values of the Ps score, Pc score, and ACC lift to 81.6, 76.4, and 0.2 respectively (Tab. 1), and the Ps score is statistically close to that of subjective forecast; even for some special range, objective prediction exhibits obvious higher-skill[45]. Also, the number of ensembles with positive skill for the Ps score (263d), Pc score (241d), and ACC (250d) are moderately more than that in the 10-30d forecast case, and more importantly, ensembles with excellent skill go up noticeably (198d versus 156d for the Ps score; 195d versus 175d for the Pc score; 229d versus 208d for ACC). Comparatively speaking, the number of ensembles with high-skill is far more than that with negative skill (232d versus 72d for the Ps score; 215d versus 94d for the Pc score; 238d versus 85d for ACC), which further confirms the operational usefulness of this objective result (Tab. 2). Specifically, the ranges with high-skill primarily reside over the following spans: January 1st to January 16th, January 23rd to March 11th, April 11th to June 10th, July 10th to August 21st, and October 16th to November 22nd (Fig. 7a-c). As for SCC, all the ensembles show positive skill, and more than 90% ensembles present excellent skill (Tab. 2). It should be noted that for the 1-30d TA forecast, seasonality of RMSE is markedly decayed. Although the amplitude of RMSE during May-September is slightly weaker than that during October-April, the difference is subtle (Fig. 7e). Mean RMSE for 1-30d forecast (1.1 ℃) is moderately smaller than that for 10-30d forecast (1.3 ℃), which also certifies the superiority of long-term prediction.

Statistically, general performance for 1-30d precipitation forecast is skillful and the mean Ps score, Pc score, and ACC reach 78.1, 64.4, and 0.2 respectively (Tab. 1). Long-term evaluation has attested that mean ACC of operational monthly rainfall prediction is smaller than 0.1, so the objective result is distinctly superior to subjective prediction[45]. There are more ensembles with positive skill for the Ps score (279d), Pc score (205d), and ACC (239d) than that in the 10-30d forecast case (245d, 178d, and 197d, respectively). In addition, more than half ensembles appear high-skill (206d for the Ps score, 135d for the Pc score, and 209d for ACC), which remarkably exceed current operational level (Tab. 2). Similar to the 10-30d rainfall prediction, stationary high-skill period is hard to acquire, and temporal distribution of high-skill periods is relatively dispersive (January 1st to January 25th, March 29th to April 11th, May 24th to July 8th, July 25th to August 18th, September 9th to September 28th, and October 5th to November 6th) (Fig. 8a-c). Temporal evolution of SCC can be divided into two stages: from February to July, SCC mainly displays positive skill and generally keeps stable, however, from August to January, SCC turns to changeable and occurrence of ensembles with negative skill becomes frequent (Fig. 8d). Also, daily evolution of RMSE does resemble that in the 10-30d case, characterizing by large and small bias occurring during May-September and October-April respectively with a small peak in November (Fig. 8e), which may suggest it is hard for climate models to predict extreme precipitation[46].

4 CONCLUSION AND DISCUSSION

In this paper, a downscaling strategy correcting systematic bias of climate model and its real-time forecast performance in an operational platform are introduced. Various assessment schemes of the Ps score, Pc score, ACC, SCC, RMSE, AB, RB, and SC are applied onto long-term temperature and precipitation verification for two forecast prescriptions of 1-30d and

10-30d. Given the behavior of 335 independent forecast ensembles from January 1st to November 30th in 2019, prediction ability of the downscaling-model is basically satisfying. Result of the 1-30d forecast is moderately superior to that of the 10-30d case. While the capability of 1-30d temperature prediction is close to current operational level, the result of monthly-scale precipitation prediction is statistically better than subjective forecast. Therefore, based on the CFSv2 forecast products, an operational prediction platform is built with the downscaling method, whose actual performance demonstrates this system is practically useful and valuable.

Although the climatology-recovery strategy is effective in reducing modeling-bias, systematic errors have not been totally removed (e.g. Fig. 2b, d). In order to further improve modeling forecast-level, other statistical and physical means such as cumulative distribution function transform[47], predictable components extraction[48], and multi-model ensembles integration[49] can be brought in and more forecast experiments need to carry out, which will be our future work.

Acknowledgments. This work is jointly supported by Zhejiang Province Basic Public Welfare Program (grant number LGF19D050001), National Key Research and Development Program of China (grant number 2018YFC1505601), Key Research and Development Program of Zhejiang Province (grant number 2021C02036), and Key Program of Zhejiang Meteorological Bureau (grant number 2020ZD14). We are indebted to Drs. Yiping Yao, Zhengquan Li, and Kuo Wang for helpful discussions. We also want to thank Mr. Zhiqiang Zhao for calculating some statistical results.

REFERENCES

[1] Zhao, Z., and H. Liu. (2003) Operational technology development of short-term climate prediction in China. Zhejiang Meteorology, 24: 1-6 (in Chinese).

[2] Chang, C., Z. Zhao, and X. Xu. (1978) On the method of factor combination in the statistacal weather prediction. Chinese Journal of Atmospheric Sciences, 2: 48-54 (in Chinese).

[3] Wei, F., Y. Xie, and M. E. Mann. (2008) Probabilistic trend of anomalous summer rainfall in Beijing: Role of interdecadal variability. Journal of Geophysical Research: Atmospheres, 113, doi: 10.1029/2008JD010111.

[4] Wei, F. (2011) Physical basis of short-term climate prediction in China and short-term climate objective prediction methods. Quarterly Journal of Applied Meteorology, 22: 1-11 (in Chinese).

[5] Wu, B., and T. Zhou. (2012) Prediction of decadal variability of sea surface temperature by a coupled global climate model FGOALS_gl developed in LASG/IAP. Chinese Science Bulletin, 57: 2453-2459.

[6] Ren, H., J. Wu, C. Zhao, et al. (2016) MJO ensemble prediction in BCC-CSM1.1(m) using different initialization schemes. Atmospheric and Oceanic Science Letters, 9: 60-65.

[7] Jie, W., F. Vitart, T. Wu, et al. (2017) Simulations of the Asian summer monsoon in the sub-seasonal to seasonal prediction project (S2S) database. Quarterly Journal of the Royal Meteorological Society, 143, doi: 10.1002/qj.3085.

[8] Xin, X., F. Gao, M. Wei, et al. (2018) Decadal prediction skill of BCC-CSM1.1 climate model in East Asia. International Journal of Climatology, 38: 584-592.

[9] Xin, X., M. Wei, Q. Li, et al. (2019) Decadal prediction skill of BCC-CSM1.1 with different initialization strategies. Journal of the Meteorological Society of Japan, 97, doi:

10.2151/jmsj.2019-043.
[10] Feng, G., H. Cao, X. Gao, et al. (2001) Prediction of precipitation during summer monsoon with self-memorial model. Advances in Atmospheric Sciences, 18: 701-709.
[11] Gu, X., and J. Jiang. (2005) A complex autoregressive model and application to monthly temperature forecasts. Annales Geophysicae, 23: 3229-3235.
[12] Kang, H., K. An, C. Park, et al. (2007) Multimodel output statistical downscaling prediction of precipitation in the Philippines and Thailand. Geophysical Research Letters, 34: 87-101.
[13] Ke, Z., P. Zhang, L. Chen, et al. (2011) An experiment of a statistical downscaling forecast model for summer precipitation over China. Atmospheric and Oceanic Science Letters, 4: 270-275.
[14] Liu, Y., and K. Fan. (2012) Prediction of spring precipitation in China using a downscaling approach. Meteorology and Atmospheric Physics, 118: 63-79.
[15] Lian, C., Z. Zeng, W. Yao, et al. (2014) Extreme learning machine for the displacement prediction of landslide under rainfall and reservoir level. Stochastic Environmental Research and Risk Assessment, 28: 1957-1972.
[16] Wu, Z., F. Wu, J. Chai, et al. (2019) Prediction of daily precipitation based on deep learning and broad learning techniques. 2019 IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), IEEE.
[17] Peng, T., X. Zhi, Y. Ji, et al. (2020) Prediction skill of extended range 2-m maximum air temperature probabilistic forecasts using machine learning post-processing methods. Atmosphere, 11, 823, doi: 10.3390/atmos11080823.
[18] Vitart, F., C. Ardilouze, A. Bonet, et al. (2017) The subseasonal to seasonal (S2S) prediction project database. Bulletin of the American Meteorological Society, 98: 163-173.
[19] Trenberth, K. E., B. Moore, T. R. Karl, et al. (2006) Monitoring and prediction of the earth's climate: A future perspective. Journal of Climate, 19: 5001-5008.
[20] Saha, S., S. Nadiga, C. Thiaw, et al. (2006) The NCEP Climate Forecast System. Journal of Climate, 19: 3483-3517.
[21] Saha, S., S. Moorthi, X. Wu, et al. (2014) The NCEP Climate Forecast System version 2. Journal of Climate, 27: 2185-2208.
[22] Saha, S., S. Moorthi, H.-L. Pan, et al. (2010) The NCEP Climate Forecast System Reanalysis. Bulletin of the American Meteorological Society, 91: 1015-1057.
[23] Liu, Y., K. Fan, and Y. Zhang. (2013) A statistical downscaling model for summer rainfall over China stations based on the Climate Forecast System. Chinese Journal of Atmospheric Sciences, 37: 1287-1296 (in Chinese).
[24] Luo, L., W. Tang, Z. Lin, et al. (2013) Evaluation of summer temperature and precipitation predictions from NCEP CFSv2 retrospective forecast over China. Climate Dynamics, 41: 2213-2230.
[25] Zuo, Z., S. Yang, Z.-Z. Hu, et al. (2013) Predictable patterns and predictive skills of monsoon precipitation in Northern Hemisphere summer in NCEP CFSv2 reforecasts. Climate Dynamics, 40: 3071-3088.
[26] Lang, Y., A. Ye, W. Gong, et al. (2014) Evaluating skill of seasonal precipitation and temperature predictions of NCEP CFSv2 forecasts over 17 hydroclimatic regions in China. Journal of Hydrometeorology, 15: 1546-1559.
[27] Tippett, M. K., M. Almazroui, and I.-S. Kang. (2015) Extended-range forecasts of areal-averaged rainfall over Saudi Arabia. Weather and Forecasting, 30: 1090-1105.
[28] Guo, Y., J. Li, and Y. Li. (2014) Seasonal forecasting of North China summer rainfall using a statistical downscaling model. Journal of Applied Meteorology and Climatology, 53: 1739-1749.
[29] Jin, L., J. Zhu, Y. Huang, et al. (2015) A nonlinear statistical ensemble model for short-range

rainfall prediction. Theoretical and Applied Climatology, 119: 791-807.
[30] Liu, N., and S. Li. (2015) Short-term climate prediction for summer rainfall based on time-scale decomposition. Journal of Applied Meteorological Science, 26: 328-337 (in Chinese).
[31] Liu, Y., and H.-L. Ren. (2015) A hybrid statistical downscaling model for prediction of winter precipitation in China. International Journal of Climatology, 35: 1309-1321.
[32] Liu, Y., and H.-L. Ren. (2017) Improving ENSO prediction in CFSv2 with an analogue-based correction method. International Journal of Climatology, 37, doi: 10.1002/joc.5142.
[33] Tian, B., and K. Fan. (2019) Seasonal climate prediction models for the number of landfalling tropical cyclones in China. Journal of Meteorological Research, 33: 837-850.
[34] Dai, H., and K. Fan. (2020) Skillful two-month-leading hybrid climate prediction for winter temperature over China. International Journal of Climatology, 40, doi: 10.1002/joc.6497.
[35] Lu, Z., Y. Guo, J. Zhu, et al. (2020) Seasonal forecast of early summer rainfall at stations in South China using a statistical downscaling model. Weather and Forecasting, 35: 1633-1643.
[36] Liu, Y., K. Fan, L. Chen, et al. (2021) An operational statistical downscaling prediction model of the winter monthly temperature over China based on a multi-model ensemble. Atmospheric Research, 249, doi: 10.1016/j.atmosres.2020.105262.
[37] Krishnamurthy, V. (2018) Predictability of CFSv2 in the tropical Indo-Pacific region, at daily and subseasonal time scales. Climate Dynamics, 50: 3931-3948.
[38] Yu, Z., L. Wu, D. Gao, et al. (2016) Quality control of automatic meteorological observation data. Journal of the Meteorological Sciences, 36: 703-708 (in Chinese).
[39] Sun, Z., and H. Chen. (2020) Short-term climate prediction: A review for the related work in NUIST in the past 60 years and an outlook for the future. Transactions of Atmospheric Sciences, 43: 745-767 (in Chinese).
[40] Zhu, Y., X. Yu, L. Zhao, et al. (2013) Extended-range weather forecasts of 10-30 days and strategy thinking. Desert and Oasis Meteorology, 7: 38-44 (in Chinese).
[41] Jiang, X., X. Liu, F. Huang, et al. (2010) Comparison of spatial interpolation methods for daily meteorological elements. Chinese Journal of Applied Ecology, 21: 624-630.
[42] He, H., Q. Li, T. Wu, et al. (2014) Temperature and precipitation evaluation of monthly Dynamic Extended Range Forecast operational system DERF2.0 in China. Chinese Journal of Atmospheric Sciences, 38: 950-964 (in Chinese).
[43] Palmer, T. N. (1988) On the prediction of forecast skill. Monthly Weather Review, 16: 2453-2480.
[44] Chen, S., B. Liu, X. Tan, et al. (2020) Inter-comparison of spatiotemporal features of precipitation extremes within six daily precipitation products. Climate Dynamics, 54: 1057-1076.
[45] Ma, H., C. Xu, and J. Zhang. (2015) A systematic evaluation and some thoughts for short-term climate prediction quality for Zhejiang province during 2007-2014. Zhejiang Meteorology, 36: 8-16 (in Chinese).
[46] Zhang, L., and Y. Ding. (2008) Evaluation of extreme heavy precipitation in coupled ocean-atmosphere general circulation models. Journal of Applied Meteorological Sciences, 19: 760-769 (in Chinese).
[47] Vigaud, N., M. Vrac, and Y. Caballero. (2013) Probabilistic downscaling of GCM scenarios over southern India. International Journal of Climatology, 33: 1248-1263.
[48] Wang, Q., J. Chou, and G. Feng. (2014) Extracting predictable components and forecasting techniques in extended-range numerical weather prediction. Science China Earth Science, 57: 1525-1537.
[49] Kharin, V. V., and F. W. Zwiers. (2002) Climate predictions with multimodel ensembles. Journal

of Climate, 15: 793-799.