Evaluating precipitation prediction skill for the pre- and postrainy seasons in South China in ECMWF subseasonal forecasts

Yanan Liu1, Qiong Wu2, Yizhi Zhang1* and Lujun Jiang1

Abstract
The rainy season in South China is divided into two phases, the pre- and postrainy seasons, according to the seasonal progression of the East Asian summer monsoon. The precipitation prediction skills for the two rainy seasons are investigated using subseasonal-to-seasonal (S2S) hindcast data from the European Centre for Medium-Range Weather Forecasts (ECMWF) for 2001–2019. The precipitation prediction skills and biases differ between the two rainy seasons, although some similar characteristics exist regarding circulation patterns and their influence on precipitation. During the two rainy seasons, the prediction ability of circulation at 850 hPa in key areas is relatively high, and the influence of circulation on precipitation is well captured; additionally, the relationship between circulation in key areas at 500 hPa and precipitation is less accurately constrained. Moreover, the precipitation prediction skill in the prerainy season is higher than that in the postrainy season. The main bias is that the 200 hPa westerly winds provide favorable divergence conditions for prerainy season precipitation (preprecipitation), while the postrainy season precipitation (postprecipitation) displays almost no correlation with the circulation in the reanalysis product; however, the simulated circulation at 200 hPa is closely connected to the precipitation in both rainy seasons; therefore, the lower prediction skill in the postrainy season is likely associated with overestimation of the complex physical mechanism of the upper-level circulation in the model.

Keywords  Climate prediction, Prediction skill, Pre- and postrainy seasons in South China

Introduction
Huanan refers to the area south of 28°N in China and east of the Yunnan-Guizhou Plateau. The rainy season in South China (Huanan) occurs from April to September (Ramage 1952), and both hourly and longer-scale heavy rainfall occur frequently (Luo et al. 2016; Zheng et al. 2016), resulting in high flash flood and urban waterlogging risks in South China (Hallegatte et al. 2013; Luo 2017). With the seasonal progression of the East Asian summer monsoon circulation and precipitation, the rainy season in the Huanan area is often divided into two phases: the pre- and postrainy seasons (Li et al. 2011; Yuan et al. 2010). The prerainy season in South China (April–June) is affected by large-scale systems, which account for approximately half of the total annual precipitation in China (Peng et al. 2006), and the overall trend is decreasing. In contrast, the postrainy season in South China (July–September) is mostly affected by complex tropical weather systems, such as tropical cyclones and the intertropical convergence zone (ITCZ), which are comparatively unpredictable. Thus, the predictability of the two rainy seasons may differ. Extreme precipitation events, such as floods and droughts during the pre- and postrainy seasons in South China have severely affected economic development and people’s safety in
models is mainly associated with the fact that they can -
2013, 2019), and the low prediction skill of these climate
10–30 day predictive skill between the pre- and postrainy
causes is important for improving predictions.

The differences in prediction skills and the corresponding
deviations (Liu et al. 2021). Thus, understanding
impacts of some key circulations on precipitation exhibit
mechanisms in the model exhibit discrepancies, and the
models (Zeng et al. 2012). In addition, some physical
lation cannot be well simulated or predicted in climate
eastern China is low, partly because the regional circu -
seasonal prediction skill of summer precipitation over

Due to differences in geographical location and mon -
and mid-to-high-level circulations (Chen et al. 2019).
Due to differences in geographical location and mon-
soon stage, there are also differences in the contributions
of low-latitude and mid-to-high-latitude circulations.
Generally, low-latitude circulations are more predictable
than mid-to-high-latitude circulations (Ren et al. 2017;
Wu et al. 2017). Consequently, it is necessary to further
understand the predictability of precipitation in the two
different rainy seasons.

The development of dynamic models has contributed
to important advances in subseasonal to seasonal pre-
dictions (Saha et al. 2006, 2014; Vitart et al. 2017). Sev-
eral models have shown reasonable skill in predicting
the interannual variability in Asian monsoon intensity
(Jiang et al. 2013; Liu et al. 2014, 2015; Zhu and Shukla
2013). Nevertheless, due to model errors, predicting
regional precipitation across China remains a great chal-
lenge for dynamical models (Liang et al. 2019; Liu et al.
2013, 2019), and the low prediction skill of these climate
models is mainly associated with the fact that they can-
not effectively simulate physical mechanisms. The sub-
seasonal prediction skill of summer precipitation over
eastern China is low, partly because the regional circu-
lotation cannot be well simulated or predicted in climate
models (Zeng et al. 2012). In addition, some physical
mechanisms in the model exhibit discrepancies, and the
impacts of some key circulations on precipitation exhibit
notable deviations (Liu et al. 2021). Thus, understanding
the differences in prediction skills and the corresponding
causes is important for improving predictions.

In this study, the difference in extended-range of the
10–30 day predictive skill between the pre- and postrainy
seasons in South China is analyzed. The ability of the
model (the S2S dataset produced by the ECMWF) to pre-
dict the key circulation systems that affect precipitation
in the two seasons and the corresponding relationships
with precipitation are assessed. The possible reasons for
the differences in the prediction skills between the two
seasons are discussed. Specifically, the data and meth-
ods used in this study are described in “Data and meth-
ods Sect.” “Skill in precipitation predictions between the
pre- and postrainy seasons in South China Sect.” comp-
ares the precipitation predictions between the pre- and
postrainy seasons in South China. In Prediction of gen-
eral circulation associated with precipitation during the
pre- and postrainy seasons in South China” Sect., the
ability of the model to simulate the general circulations
associated with precipitation is evaluated. The sources
of prediction bias in the two rainy seasons are discussed in
“Sources of prediction bias between the pre-and po-
trainy seasons in South China” Sect. A summary and dis-
cussion are given in “Conclusions and discussion” Sect.

Data and methods

Datasets

The subseasonal to seasonal prediction (S2S) database
from 11 operational centers is useful for assessing predic-
tions in the S2S time range. Considering the better fore-
casting ability of the ECMWF product (Lin 2019), in the
present analysis, we use ensemble forecasting data from
the ECMWF. These data are obtained with the Integrated
Forecasting System (IFS) version CY46R1. Operationally,
the S2S dataset produced by the ECMWF is generated on
the fly twice a week (Thursday and Monday). The system
provides daily ensemble forecasts for a wide variety of
atmospheric variables (total precipitation, geopotential
height, wind field, etc.) with a 46 day lead time. The rea-
alysis forecasts are used to calibrate the real-time fore-
casts by running an 11-member ensemble for the same
start date as the 33-member real-time forecast but over
the last 19 years from 2001 to 2019, and the ERA-Interim
product is used to establish the initial conditions. Addi-
tional details are available at https://confluence.ecmwf.
int/display/S2S/ECMWF+Model. The reanalysis data
used for model evaluation are the ERA-Interim daily
precipitation and atmospheric variable data. The cumulative
precipitation reanalysis product is produced in meters
(m) at a 12 h temporal step and is converted to millimeters
(mm). More details are available at https://apps.
ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/.
Precipitation with a horizontal resolution of 1.0° × 1.0°
was used in both the model and reanalysis datasets, and
the 500 hPa geopotential height with a horizontal resolu-
tion of 2.5° × 2.5° was used.
Durations and areas of the two types of seasons
With the seasonal progression of the East Asian summer monsoon circulation and precipitation, the rainy season in South China is often divided into two phases: the pre- and postrainy seasons; thus, the first month (5/21–6/20) after the South China Sea monsoon breaks is chosen as the study period for the pre-rainy season in South China (Jiang et al. 2003; Li et al. 2021; Wang et al. 2005; Zheng et al. 2009a, b). Following the relevant monitoring regulations of the China Meteorological Administration (QX/T 395–2017) and taking a comparative analysis of pre- and postrainy seasons into account, the duration (7/13–8/12) is selected for the postrainy period. Moreover, these two sectors (5/21–6/20 and 7/13–8/12) can guarantee forecasts with different lead times (0, 3, and 7 days ahead, Fig. 1).

That is, forecasts beginning at each start time are run for precipitation or circulation variables during the same pre- (May 21 to June 20) or postrainy (July 13 to August 12) periods in South China. For example, forecasts for the prerainy period in South China with a 0-day lead date represent the means of precipitation or circulation variables from May 21 to June 20 initialized on May 21. Similarly, 3 days lead forecasts represent the same prerainy period but initialized on May 18, and 7 days lead forecasts are similar but initialized on May 14; forecasts for the postrainy period in South China with 0-day lead represent the means of precipitation or circulation variables from July 13 to August 12 initialized on July 13. Moreover, the studied area is located between 105°–117°E and 18°–27°N, and the forecast targets are both one-month average values.

Methods
The percentage anomaly between the reanalysis and predicted variables is defined as the departure from the multiyear mean climatologies for the two rainy seasons (2001 to 2019) and is calculated as follows:

\[
DIFF(\%) = \frac{F - X}{X} \times 100
\]

where \( F \) is the ensemble forecast value and \( X \) is the averaged reanalysis value.

The forecasting skill of the model is assessed based on the anomaly correlation coefficient (ACC), temporal correlation coefficient (TCC) and root-mean-square error (RMSE) (Wu et al. 2023; Yue et al. 2022; Zhi et al. 2018).

\[
\text{ACC}_i = \frac{\sum_{i=1}^{M} (\Delta x_{ij} - \overline{x}_i) \times (\Delta y_{ij} - \overline{y}_j)}{\sqrt{\sum_{i=1}^{M} (\Delta x_{ij} - \overline{x}_i)^2 \times \sum_{j=1}^{N} (\Delta y_{ij} - \overline{y}_j)^2}}
\]

(2)

\[
\text{TCC}_i = \frac{\sum_{j=1}^{N} (x_{ij} - \overline{x}_j) \times (y_{ij} - \overline{y}_j)}{\sqrt{\sum_{j=1}^{N} (x_{ij} - \overline{x}_j)^2 \times \sum_{j=1}^{N} (y_{ij} - \overline{y}_j)^2}}
\]

(3)

\[
\text{RMSE}_i = \sqrt{\frac{\sum_{j=1}^{N} (x_{ij} - y_{ij})^2}{N}}
\]

(4)

where \( x_{ij} \) and \( y_{ij} \) represent the reanalysis field and prediction field, respectively; \( i = 1, 2, 3, \text{etc.} \); \( M \) represents the grid number in the study region; and \( j = 1, 2, 3, \text{etc.} \). \( N \) represents the time series. Additionally, \( \overline{x}_i = \frac{1}{N} \sum_{j=1}^{N} x_{ij} \), \( \overline{y}_j = \frac{1}{N} \sum_{i=1}^{N} y_{ij} \), \( \Delta x_{ij} = x_{ij} - \overline{x}_i \), and \( \Delta y_{ij} = y_{ij} - \overline{y}_j \).

The definition of a blocking high is determined according to the regulations of the NCC for the real-time monitoring of blocking highs in the Northern Hemisphere; this definition is based on the height field gradient difference approach and was improved upon by Tibaldi and Molteni (Zhao 2000). A certain longitude index can be calculated as follows:

\[
\text{GHG} = \frac{Z(\varphi_1) - Z(\varphi_2)}{\varphi_1 - \varphi_2}
\]

(5)

where \( Z \) is the geopotential height; \( \varphi_1 \) and \( \varphi_2 \) are the latitudinal angle; \( \varphi_1 = 60 + \delta \); and \( \varphi_2 = 40 + \delta \). In this case, \( \delta \) can be \(-5, 0, \text{or } 5\). The blocking pressure in a certain key area is the average blocking pressure within the longitudinal interval of the key area.

The Eurasian teleconnection (EU) index can be defined as follows:

\[
I = 0.25Z^{(60°N,20°E)} - 0.5Z^{(62.5°N,60°E)} + 0.25Z^{(60°N,110°E)}
\]

(6)

where \( Z' = Z - \overline{Z}, Z^* = Z \sin 45°/\sin \psi, Z \) is the 500 hPa geopotential height, \( \overline{Z} \) is the climatological mean of the 500 hPa geopotential height, and \( \psi \) is the latitude (Wallace and Gutzler 1981, inverted to represent the increased postprecipitation).

Regression analysis is used to assess the relationship between an outcome variable and one or more factors or confounding variables, and correlation analysis is used to test the significance of the relationship.
Results
Skill in precipitation predictions between the pre- and postrainy seasons in South China
The skill of precipitation predictions between the pre- and postrainy seasons in South China was compared, and Fig. 2 shows the ACC distributions for the ERA data and ECMWF predictions for different lead times and ensemble-averaged forecasts. By comparing forecasts with different lead times (0, 3, 7 and 10 days ahead), it is found that the precipitation forecasting ability is better for the prerainy season than for the postrainy season for lead times of 0, 3, and 7 days (Fig. 2a), both passing the 90% significance level. The ACC is the best for the 0 day lead time, and the multiyear average ACC for the prerainy season...
season is 0.26, which is significant at the 99% significance level, and 0.08 for the postrainy season is not significant. A decrease in the ACC is associated with an increase in lead time, although the rate of decrease varies for each season. Based on the ensemble forecasts of 33 members with lead times of 0, 3, and 7 days (for similar skill in the prerainy season of 10 day lead time), obvious interannual differences are observed in the precipitation forecasting skill in the two seasons (Fig. 2b). The precipitation forecasting ability during the prerainy season from 2001 to 2019 is generally better than that during the postrainy season. According to the multiyear averages, the ACC of the ensemble forecast is 0.30 during the prerainy season and 0.09 during the postrainy season. That is, there are obvious differences between the two rainy seasons in terms of the skill of ensemble precipitation forecasts.

For leads of 0, 3, and 7 days, the predictions skill in prerainy season is obviously better than in postrainy season, in order to find the possible cause for the difference that the rainy season predictions with leads of 0, 3, and 7 days are used in the following analysis. Considering the samples more comprehensive, the rainy season predictions for 33 members with leads of 0, 3, and 7 days (the same 11 members at each lead time) initialized with slightly different initial atmospheric conditions are used in the following analysis.

**Fig. 4**  The mean circulation a–c and synoptic model d large-scale circulations during the prerainy season in South China. a The 500 hPa geopotential height field (contours, unit: dagpm) and wind field (vector, unit: m/s), b the 850 hPa wind field (wind bar, 4 m/s) and pseudo-equivalent potential temperature (contours unit: K), c Sea-level pressure field (contours, unit: hPa). The shading indicates the terrain height (unit: m).
smaller in most areas in the postrainy season (generally 100–120 mm). This difference is likely related to the greater precipitation in the prerainy season than in the postrainy season. From the TCC between the reanalysis and model-predicted precipitation values in the two rainy seasons (Fig. 3c and f), there are obvious differences in different areas in the prerainy season. Specifically, the forecasting ability is best in southermost China and adjacent areas, and in most areas with TCCs between 0.4 and 0.6, the forecasts are significantly accurate at the 95% confidence level. In general, the TCC is low in the postrainy season, with weak TCCs of 0.1–0.3 in most areas, and the forecasts in almost all areas are not significant at the 95% significance level.

**Prediction of general circulation associated with precipitation during the pre- and postrainy seasons in South China**

**The prerainy season in South China**

Figure 4 shows the mean circulation (a-c) and conceptual map (d) of large-scale climatological circulations during the prerainy season in South China, and the circulation concept map is summarized based on the analysis of the high, low-level and surface circulation characteristics of the prerainy process (Fig. 4d). At 500 hPa (Fig. 4a), the Lake Baikal high is blocked, and the stable East Asian low trough extends from the south of the blocking high to the middle latitudes; these systems are advantageous for bringing cold air north of the Huanan area and for promoting interactions with the East Asian summer monsoon (Ding et al. 2020), while the WPSH is distributed zonally above the ocean to the east of Taiwan and south of Japan. Huanan is located at the bottom of the East Asian trough and northwest of the WPSH, where it forms a favorable situation of “high in the east and low in the west” at 850 hPa, and southwesterly winds prevail over South China (Fig. 4b). Moreover, a low-pressure zone from South China to Northeast China is found in the sea level pressure level field (Fig. 4c); the cold air gradually moves south from the back of the low, and the front system moves south after the rainy season begins. Moreover, the 200 hPa SAH moves to the Indo-China Peninsula with the upper-level westerlies covering Huanan, thus providing upper-level divergent conditions for heavy rainfall during the prerainy season.

The ECMWF prediction ability at the 500 hPa geopotential height is quite high, as shown in Fig. 5a, and the correlation coefficients between the ERA-Interim...
and ECMWF data are almost significant at the 95% significance level. Then, according to the synoptic model shown above, the responses of the circulations at different levels to precipitation in the prerainy season were further analyzed. To key out the key areas of circulation at 500 hPa that affect precipitation, the reanalysis (model-predicted) precipitation index is used to regress the reanalysis (model-predicted) 500 hPa geopotential height. The circulation variables in Fig. 5(e), (f), and (g) are the ensemble means of the ECMWF forecasts with lead times of 0, 3, and 7 days, respectively, during the same prerainy period in South China (May 21 to June 20). The regressed reanalysis field (Fig. 5b) indicates that the positive areas are located in high-latitude regions, such as those near Lake Baikal, and the negative areas are located in middle-latitude areas from Mongolia to Inner Mongolia, which highlights the positive influence of the active trough on precipitation. The regressed geopotential heights at high and middle latitudes are basically consistent with those based on the reanalysis product (Fig. 5e), and the regressions are positive, negative and positive from high to low latitudes.

Furthermore, the reanalysis (modeled) precipitation index is used to regress the reanalysis (modeled) 850 hPa wind field. The regressed reanalysis field (Fig. 5c) indicates that the areas that passed the significance level were mainly located in southernmost China (southwesterly) and northernmost China (southeasterly). However, the significant regression areas are distributed from high to low latitudes in the regressed model field (Fig. 5f): the
Based on the above analysis, the low-, middle- and high-level circulation indices are obtained. First, the effects of two key circulation configurations (the high-latitude blocking high pressure (BHP) and geopotential height anomaly over middle-latitude ocean areas at 500 hPa) on precipitation during the prerainy season are discussed. According to the reanalysis products and model performance, δ is set to -5 in formula 5 for the BHP index. The geopotential height difference between the two different latitudes in the longitude range from east of Lake Baikal to the Okhotsk Sea (110°–140° E; the area between the lines in Fig. 6a and c) was selected to represent the BHP in this key area. Figure 6a shows the 500 hPa height anomaly field regressed from the BHP in the reanalysis product. In the central area, a single BHP can be observed. The model regression results also highlights key areas, and the “+ −” distribution in the high...
and middle latitudes is significant (Fig. 6c). In the reanalysis precipitation field regressed based on the BHP, a significant positive center is found in the southern region, and a negative correlation is observed in a small part of the northern region (Fig. 6b). In the model precipitation field regressed based on the BHP, northern Huanan is the center of the positive area, and the magnitude gradually decreases from the center, although not significantly at the 95% significance level (Fig. 6d).

To discuss the effect of the low-level 850 hPa wind anomaly on precipitation during the prerainy season, the significant positive area is selected from the most obvious southwesterly area (15°–20°N, 100°–110°E) in Fig. 5c, f. The 850 hPa wind index (U850) is the average zonal wind anomaly within the study area (black box in Fig. 7a and c). Figure 7a and c show the 850 hPa wind distributions regressed based on U850 in the reanalysis and model sets, respectively. Both indices effectively highlight the key areas with anomalous southwesterlies in South China. In the reanalysis precipitation field regressed with U850 (Fig. 7b), significant positive areas can be seen across all of Huanan and the adjacent sea. In the regression field of modeled precipitation based on U850 (Fig. 7d), the overall distribution of precipitation is similar to that in the reanalysis regression field. However, the results are significant in comparatively fewer regions.

Ding et al. (2011) and Du and Chen (2019) suggested that the interaction between the upper-level westerly jet and the low-level southwesterly jet during the prerainy season in South China is significant and can be
conducive to coupling upper- and low-level jets to facilitate the occurrence of heavy rain. Combined with the significant response areas in Fig. 5d and g, the 200 hPa wind index (U200) is constructed as the average 200 hPa zonal wind anomaly in the region (22.5°–30°N, 110°–120°E), black box in Fig. 8a and c) and is standardized to regress the 200 hPa wind anomaly field (Fig. 8a), which displays an eastern-western-eastern zonal wind distribution at high (40°N), middle (30°N), and low (10°N) latitudes. The regression field of the model displays a similar distribution along the high (45°N), middle (35°N) and low (15°N) latitudinal belts, with values passing the 95% confidence test in both cases (Fig. 8c). In the reanalysis precipitation field regressed based on U200, there is a significant positive center in the western part of the region (Fig. 8b). In the regressed field of modeled precipitation, the positive significant center is along the northwestern boundary of Huanan (Fig. 8d), and the values in relatively few regions passed the significance test.

The postrainy season in South China

Previous studies have shown that the zonal “− + −” wave train structure in the Ural Mountains of Eurasia, the area between Lake Balkhash and Lake Baikal, and the Sea of Okhotsk is conducive to decreased precipitation in South China and vice versa (Wang and He 2015; Wei et al. 2003; Zhang et al. 2003; Zou et al. 2013). To analyze the key areas of circulation affecting precipitation during the postrainy season in South China, the reanalysis (model-predicted) precipitation index is used to regress the reanalysis (model-predicted) 500 hPa height field. The regressed field (Fig. 9a) indicates that the regressions over the Nordic continent and in the Ural Mountains, the Baikal region, and the Sea of Okhotsk display “+ − + −” wave trains, similar to the EU pattern. In the regressed model field (Fig. 9d), similar “− + −” wave train distributions can be observed in the Ural Mountains, Northeast China and the Sea of Okhotsk, but the positive center above the Nordic continent nearly disappears.

Combined with the significant response areas in Fig. 9a, the EU index is constructed (Formula 6) and standardized to regress the reanalysis (model-predicted) 500 hPa height anomaly field (Fig. 9b and c). The comparable high-latitude fields show similar anomalous negative distributions above the Ural Mountains and the Sea of Okhotsk and positive distributions from Lake Baikal to the Siberian Plain; these distributions resemble the distributions of key areas in the 500 hPa height field regressed based on the precipitation index, but both pass the 95% confidence test. At middle latitudes, the negative areas are both nonsignificant. In the reanalysis precipitation field regressed based on the EU product, there is a positive anomaly in most areas that is significant (Fig. 9c), and significant positive areas are...
uncommon in the regressed field of model precipitation (Fig. 9f). There is a difference between the reanalysis and model results in terms of the significant relationship between EU and postprecipitation.

Figure 10a shows the regression between the 850 hPa wind index (U850) and the 850 hPa wind field anomaly in the reanalysis, which is the same as that in Fig. 10c but in the model; key areas with significant anomalous southwesterlies at 850 hPa can be observed over the South China Sea in both figures; thus, one was induced by the cyclone in the northwest Pacific Ocean, and the other was induced by the interaction of the cyclone in the northwest Pacific Ocean and anticyclones in southeast China. That is, the influence of 850 winds on precipitation is well captured; however, the simulation of the circulation system position needs to be further improved. Then, in the reanalysis precipitation field regressed based on U850 (Fig. 10b), a significant positive anomaly can be seen over Huinan, except for the northeastern region. In the modeled precipitation field regressed based on U850 (Fig. 10d), the overall distribution of precipitation is similar to that obtained via reanalysis regression. However, the values were significant in more regions.

The distribution of the regression between U200 and the corresponding wind field anomalies in the postrainy season in South China is shown in Fig. 11a and c. The regression field of the model (Fig. 11c) illustrates the
distribution of the latitudinal wind belts at high (40°N)-
middle (30°N)-low (10°N) latitudes, all passing the 95% confidence test, while no areas display significant values in the reanalysis regression field (Fig. 11a). Moreover, Huanan was influenced by the upper-level westerly divergence condition in the model, but there was no signal in the reanalysis product. In the modeled precipitation field regressed with U200 (Fig. 11d), the positive centers in the northwestern and southeastern coastal regions generally passed the significance test.

Based on the above analysis, there is a large difference between the reanalysis and model results regarding the relationship between the U200 and precipitation in the postrainy season in South China, revealing that the physical mechanisms of the upper-level wind and precipitation in the model are likely overestimated. To further study the relationship between upper-level circulation and precipitation during the postrainy season in South China, the precipitation index is used to regress the vertical cross-section of the meridional wind field along 117°E (Fig. 12). The anomalous southerlies in central Huanan extend from the near-surface levels to the upper levels at 200 hPa and are significant at all levels in the modeled vertical cross section of the wind field (Fig. 12b); additionally, significant southerly anomalies exist only at low levels in the reanalysis regression field (Fig. 12a), which further indicates that the physical mechanisms of

---

**Fig. 11** Regression of the 200 hPa wind field anomaly field a and the precipitation field b with the 200 hPa wind index during the Huanan postrainy based on the ERA-Interim product. Panels c and d are the same as panels a and b but for the ECMWF predictions. The red shading in a and c and the dots in b and d represent areas with values that are significant at the 95% confidence level.
upper-level wind and precipitation in the model are different and overestimated.

Sources of prediction bias between the pre- and postrainy seasons in South China

The difference in circulation in the model will inevitably lead to different simulation effects when estimating the circulation indices in each key area. As shown in Table 1, the simulation effect of the BHP index at 500 hPa is the greatest, and the corresponding correlation coefficient with the reanalysis product is 0.63; additionally, the correlation coefficient with U850 is 0.59, and the simulation effect for U200 is comparatively worse (correlation coefficient of 0.41). Thus, the three indices at low, middle and high atmospheric levels, with correlation coefficients ranging from 0.4 to 0.6, are reasonable during the prerainy season. The simulation effect based on the U850 index in the postrainy season is the best, with a correlation coefficient of 0.76, and the EU simulation effect is good, with a coefficient of 0.59. Similar to those in the prerainy season, the key areas of the 500 hPa and 850 hPa circulation indices in the postrainy season are effectively simulated, with correlation coefficients of 0.6–0.8, and the coefficient for U200 is 0.2. Notably, the U200 index is constructed as the average 200 hPa zonal wind anomaly in the same area (black box in Fig. 8a and c) and is used to constrain the upper-level circulation conditions, even though there are no areas with significant values in the reanalysis dataset.

The relationships between the circulation indices in the two rainy seasons and precipitation in the key areas in the reanalysis product and model results are further analyzed.
(Table 2). The correlation coefficient values between the BHP index and precipitation at 500 hPa during the pre-rainy season reach 0.61 and 0.35 for the ERA-Interim and ECMWF predictions, respectively. The correlation coefficients between the U850 index and precipitation are the highest (0.85 and 0.78), and values of 0.54 and 0.45 are obtained between the U200 index and precipitation. The relationships between the indices and precipitation in the reanalysis and model sets are consistent. In the postrainy season, the relationships between precipitation and the EU index at 500 hPa are 0.51 and 0.30, and the coefficients between the U850 index and precipitation are 0.63 and 0.56, respectively. For the reanalysis product, the correlation coefficient between the U200 index and precipitation is 0.20 and not significant. However, the correlation coefficient reaches 0.58 for the model results. The physical mechanism at 200 hPa in the postrainy season is overestimated by the ECMWF model in key areas, but the role of the EU is also underestimated, resulting in a lower precipitation prediction skill for the postrainy season.

Conclusions and discussion
In this study, the interannual skill of precipitation prediction in the pre- and postrainy seasons in South China is evaluated using reanalysis data (ERA-Interim) and ECMWF-S2S hindcast data for 2001–2019. In most years, the precipitation prediction ability is greater during the prerainy season than during the postrainy season in South China. During the Huanan rainy season, the TCC in most of the region reaches between 0.4 and 0.7. However, the TCC in most of the central region is less than 0.2 during the postrainy season in South China.

The precipitation prediction ability and corresponding bias differ between the two rainy seasons, and some similar characteristics exist in terms of the simulated circulations and their influence on precipitation. During the two rainy seasons, the circulation prediction ability at low levels (850 hPa) is high, and the relationship between circulation and precipitation is well captured. However, the prediction ability for circulations at middle levels (500 hPa) in key areas is relatively good, although the relationships between circulations and precipitation in key areas are less skillful in the model during the two rainy seasons.

Model biases are observed in the circulation simulation and influence precipitation at high levels (200 hPa). In the prerainy season in South China, the response to the upper-level westerlies in certain areas is strong, the 200 hPa circulation is well simulated, and the response of precipitation to U200 is effectively reflected in the model. During the postrainy season in South China, the prediction of circulation at 200 hPa is not captured, but the response of postprecipitation to U200 is fairly strong in the model. The main bias is that the 200 hPa westerly provides favorable divergence conditions for the prerainy period, but circulations are not present for the postrainy period in the reanalysis product. During the two rainy seasons, the simulated 200 hPa westerlies are tightly connected to the corresponding precipitation, and the lower prediction skill in the postrainy season is more likely due to the overestimation of the physical mechanism of the upper-level circulation in the model, while the underestimation of EU also result in lower prediction skills for postprecipitation.

Numerical models inevitably contain errors. It is becoming increasingly difficult to improve models to further increase their forecasting ability (Feng et al. 2013; Ren 2006; Saha 1992; Zheng et al. 2009a, b). Many studies have used statistical methods to improve the forecasting ability of dynamic models (Feddersen and Andersen 2005; Guo and Li 2012; Wei and Huang 2010), but the objective is to understand the sources of bias in these models. In this paper, the difference in precipitation predictability between the pre- and postrainy seasons in South China is discussed. The sources of bias in these two periods were analyzed, and the findings have important implications for the localized interpretation of models in the future. During the two rainy seasons, more attention should be given to the physical mechanisms associated with circulation and precipitation in key areas. Especially in the postrainy season, it is necessary to conduct statistical analyses of the upper-level circulation fields and surface precipitation in reanalysis products, and inappropriate physical effects should be removed. Therefore, it is necessary to explore the specific physical mechanisms of models for predicting precipitation in different rainy seasons in future work.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2S</td>
<td>Subseasonal-to-seasonal forecasts</td>
</tr>
<tr>
<td>ECMWF</td>
<td>The European Centre for Medium-Range Weather Forecasts</td>
</tr>
<tr>
<td>Preprecipitation</td>
<td>Prerainy season precipitation</td>
</tr>
<tr>
<td>Postprecipitation</td>
<td>Postrainy season precipitation</td>
</tr>
<tr>
<td>ACC</td>
<td>The anomaly correlation coefficient</td>
</tr>
<tr>
<td>TCC</td>
<td>The temporal correlation coefficient</td>
</tr>
<tr>
<td>RMSE</td>
<td>The root-mean-square error</td>
</tr>
<tr>
<td>ITCZ</td>
<td>The intertropical convergence zone</td>
</tr>
<tr>
<td>WPSH</td>
<td>The western Pacific subtropical high</td>
</tr>
<tr>
<td>SAH</td>
<td>The South Asian High</td>
</tr>
<tr>
<td>BHP</td>
<td>The blocking high pressure</td>
</tr>
<tr>
<td>EU</td>
<td>The Eurasian teleconnection</td>
</tr>
</tbody>
</table>

Acknowledgements

The authors thank the ECMWF for providing the S2S data used in this study and are extremely grateful to the anonymous reviewers for their valuable comments on this paper.
Author contributions
YL. Meteorological Sciences Institute, Jiangxi Meteorological Bureau, Nanchang, 330046, China. Performed the research, analyzed the data, and wrote the paper. QW. The Climate Center, Jiangxi Meteorological Bureau, Nanchang, 330046, China. Contributed to the central idea and collected the data. VZ. The Meteorological Sciences Institute, Jiangxi Meteorological Bureau, Nanchang, 330046, China. Contributed to refining the ideas and finalizing this paper. LJ. The Meteorological Sciences Institute, Jiangxi Meteorological Bureau, Nanchang, 330046, China, discussed the results.

Funding
This work is supported by the National Natural Science Foundation of China (41765001).

Availability of data and materials
The authors confirm that the data supporting the findings of this study are available within the article manuscript.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 12 June 2023 Accepted: 17 February 2024
Published online: 05 March 2024

References
Li ZH, Luo YL (2021) Statistical characteristics of pre-summer rainfall over South China before and after South China Sea monsoon onset from Torrent Rain and Disa 40(2):101–110

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.