Journal of the Meteorological Society of Japan

EARLY ONLINE RELEASE

The DOI for this manuscript is
DOI:10.2151/jmsj.2022-042
J-STAGE Advance published date: July 7th, 2022

A Comparison of Two 20th Century Reanalysis Datasets from the Perspective of Cross-equatorial Flows

Linhui Li

Department of Atmospheric Sciences, School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

and

Shuanglin Li

Department of Atmospheric Sciences/Centre for Severe Weather and Climate and Hydro-geological Hazards, School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

Climate Change Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

Corresponding author: Shuanglin Li, Department of Atmospheric Sciences, China University of Geosciences, 68 Jinchengjie St., East Lake High-Tech Development Zone, Wuhan 430074, China

E-mail: shuanglin.li@mail.iap.ac.cn

Abstract

Twentieth-century atmospheric reanalysis datasets are substantially important for understanding climate in the early era of the century. This paper first compares two sets of the twentieth-century atmospheric reanalyses, the NOAA-CIRES-DOE 20th Century Reanalysis Version 3 (20CRv3) and the ECMWF 20th century reanalysis (ERA20C), as far as the summer low-level cross-equatorial flows (CEFs) over the Asian-Australian monsoon region are concerned. The results show evident regional differences in intensity of individual branches of CEFs between the two reanalyses, in spite of an overall agreement in climatological seasonal mean and variability. At interannual timescale, significant differences are seen prior to 1925 and in the 1940s. During the two periods there are often opposite variations in Somali CEF in the two datasets, along with obvious different amplitudes (variances) in the Bay of Bengal (BOB) and Australian CEFs. At interdecadal timescale, the two datasets have different periodicities in Somali CEF, and have a greater fluctuation of BOB CEF after 1925 in ERA20C than 20CRv3, as well as an opposite decadal variation in the Australian CEF prior to 1940 and in the 1960s. As for the long-term trend, both the Somali and BOB CEFs exhibit intensification in both the datasets, but the intensification amplitude is bigger in 20CRv3 than ERA20C for Somali CEF; the Australian CEF exhibits a weakening trend in both the datasets, but is less evident in 20CRv3. To figure

out which of the two datasets is relatively more reliable, the observed cross-equatorial meridional gradient of sea-level pressure index and the Indian summer monsoon rainfall index, which both have longer instrumental records, are used as benchmarks to validate the CEFs in view of their close connections. The results suggest that ERA20C is more reliable, and thus more suitable for investigating decadal climate variability of the 20th century across the hemispheres.

Keywords 20CRv3; ERA20C; cross-equatorial flows; reanalysis comparison; proxy index

1. Introduction

Atmospheric reanalysis produces datasets for synoptic and climate research through a fixed data assimilation system and atmospheric general circulation model ingesting all available observed variables every several hours. It provides an estimate of the climate state at each time step, and the variable outputs from the system are dynamically consistent with one another. The reanalysis datasets are thus an important resource for the research community (Deng et al. 2010; Zhao et al. 2010).

The first generation of atmospheric reanalysis dataset was developed in the mid-1990s by the joint efforts of the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) (NCEP/NCAR) (Kalnay et al. 1996), which extended back to 1948 and

continues to the present. During the recent decades, various global atmospheric reanalysis products have been developed including the fifth-generation atmospheric reanalysis (ERA5) (Hersbach et al. 2020) and the interim reanalysis (ERA-Interim) (Dee et al. 2011) from the European Centre for Medium-Range Weather Forecasts (ECMWF), the second version of the Modern-Era Retrospective Analysis for Research and Applications (MERRA-2) (Gelaro et al. 2017) from the National Aeronautics and Space Administration (NASA), Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010) from NCEP, Japanese 55-year Reanalysis (JRA-55) (Kobayashi et al. 2015) generated by the Japan Meteorological Agency (JMA) , and the Global Atmospheric Reanalysis (CRA-40) (Wang et al. 2018) from the China Meteorological Administration (CMA). The length period of most of these datasets covers just after 1950, especially after 1979 when the satellite era began. However, as the decadal variability issues cause more and more attention due to their involvement into global warming, a demand for longer observational series beyond decades has become more and more urgent. Thus, researchers began to develop century-long reanalyses.

There are three sets of century-long reanalyses: the first is the NOAA-CIRES-DOE 20th Century Reanalysis (20CR) (Compo et al. 2011), which assimilates the surface pressure and extends the period to 1836, and updates the latest version 3, 20CRv3 (Slivinski et al. 2019). The two else are from ECMWF including the ECMWF 20th century reanalysis (ERA20C) (Poli et

al. 2016) and the ECMWF Coupled 20th Century Reanalysis (CERA20C) (Laloyaux et al. 2018). All the two ECMWF datasets have a period spanning 1900/1901 to 2010. Among them, ERA20C assimilates surface and mean sea level pressures as well as marine surface winds; CERA20C uses the coupled model, which simultaneously ingests atmospheric and ocean observations. These century-long reanalyses are employed for various climate researches (Bett et al. 2017; Polonskii et al. 2017; Stankunavicius et al. 2017; Welker and Martius 2014). However, the results when using different reanalyses are often inconsistent with each other. For example, the wind speeds in the North Atlantic and North Pacific in ERA20C and CERA20C show a significant increase, but no obvious trend is seen in 20CR (Wohland et al. 2019); Correspondently, there is an obvious upward trend in the Arctic Oscillation mode and wind speed along the northern storm path due to a biased weakening trend in the sea level pressure over the Arctic in ERA20C (Bloomfield et al. 2018; Rohrer et al. 2019); Also, ERA20C and 20CR show a difference in the trend of extratropical cyclones and storms even in their signs and magnitudes (Befort et al. 2016); There is an inconsistency in regional extreme precipitation trend among them (Donat et al. 2016). All these differences pose a challenge to decadal studies: which reanalysis is more reliable and which dataset can be selected for use. Therefore, assessing the quality of the available century-long reanalyses is necessary.

On the other hand, as the exchange channels of momentum, water vapor

and energy across the hemispheres, cross-equatorial flows (CEFs) over the Asian-Australian monsoon region play an important role in global teleconnection (Wang and Xue 2003; Zurita-Gotor 2020) and Asian monsoon climate (Fan et al. 2018; Jain et al. 2021; Lei and Yang 2008). They are also substantially important for global climate (Hoskins et al. 2020; Zeng and Li 2002). There are primarily three branches of CEFs in boreal summer in the lower atmosphere of the Asian-Australian monsoon region (Fig. 1), namely the Somali CEF, the Bay of Bengal (BOB) CEF, and the Australian CEF from west to east. The Australian CEF is composed of three sub-branches, i.e. the South China Sea, the Celebes Sea, and the New Guinea sub-branches, respectively. Among them, Somali CEF and BOB CEF supply a majority of moisture for the South Asian summer monsoon (Halpern and Woiceshyn 2001), which can induce a drought or flood. They also influence the East Asia summer monsoon by providing partly water vapor and contributing to the eastward withdrawal of the Northwest Pacific subtropical high (Gao and Xue 2006; Li and Wu 2002). The variation of Australian CEF is directly related to the outbreak of the South China Sea summer monsoon, adjusting the advance and retreat of the Northwest Pacific subtropical high, and then affecting the distribution of the rain belt in East Asia (Li and Li 2014, 2016). The above three CEFs have a similar distribution of climatological mean and variability in ERA20C and 20CRv3 compared to ERA5 and NCEP/NCAR as well as JRA55 in the period of 1958-2010 (Fig. 1).

To understand the decadal variability of CEFs, century-long datasets are needed. Researchers often used two sets of the 20th-century reanalysis, ERA20C and 20CRv3, to serve this purpose (Attada et al. 2019; Huang et al. 2018; Huang et al. 2019; Prasanna 2016; Wu and Mao 2019). Although there is another reanalysis, CERA20C, it does not exhibit a significant improvement in Asian monsoon relative to ERA20C and is thus seldom used (Laloyaux et al. 2018). However, there is a large uncertainty in the quality of both ERA20C and 20CRv3 in the early stage due to the scarcity of instrumental records particularly in the Indian Ocean (McPhaden et al. 2009), also due to the bias of both model and assimilation system used to produce the reanalysis products (Ajayamohan 2007; Martin et al. 2000; Prasanna et al. 2020). Therefore, an inter-comparison and assessment of these two reanalyses, ERA20C and 20CRv3, are needed, and constitute the motivation of the present study.

2. Datasets and methods

Monthly meridional wind components and zonal wind components are from 20CRv3 and ERA20C. Among them, 20CRv3 is an ensemble of 80 members, which is produced by the NCEP Global Forecast System (GFS) coupled atmosphere, Noah land surface and thermodynamic ice model with the assimilation of surface pressure observations from the International Surface Pressure Databank (ISPD) (Cram et al. 2015) version 4.7. It consists of two sub-versions: 20CRv3si (1836-1980) and 20CRv3mo (1981-2015),

155 which use different datasets for SST boundary conditions, respectively are
156 Simple Ocean Data Assimilation with Sparse Input version 3 (SODAsi.3)
157 (Giese et al. 2016) and HadISST2.2. ERA20C is a single-member reanalysis,
158 which is produced by ECMWF's Integrated Forecast System (IFS) with a
159 coupled model of Atmosphere, Land-surface and Ocean-waves, assimilating
160 the observations of surface and mean sea level pressures from ISPD version
161 3.2.6 and the International Comprehensive Ocean-Atmosphere Data Set
162 (ICOADS) (Woodruff et al. 2011) version 2.5.1, as well as surface marine
163 winds from ICOADS. The HadISST version 2.1.0.0 provides its SST and sea
164 ice boundary conditions. Both of 20CRv3 and ERA20C have a horizontal
165 resolution of 1°×1°. Summer is referred to as June to August in this study.

166 Because no benchmark can be referred to as a realistic observation, here
167 we use two proxies, the Cross-Equatorial large-scale meridional sea-level
168 Pressure (SLP) Gradient (CEPG) over the tropical Indian Ocean and the
169 Indian summer monsoon rainfall index (ISMR), to validate the CEFs, in view of
170 their substantial connections with CEFs and relatively higher quality in addition
171 to longer availability. The SLP used is from the Hadley Centre (HadSLP2)
172 spanning 1850 to 2012 with a horizontal resolution of 5°×5° (Allan and Ansell
173 2006). CEPG is defined as the difference of June-August averaged sea-level
174 pressure in a southern box (5-10°S 38-60°E) minus that in a northern box
175 (5-10°N, 38-60°E). Here the longitudinal range of the two boxes corresponds
176 to the region the Somali CEF dominates. The rainfall dataset used to calculate

ISMR is from the Indian daily rainfall data (IMD4) covering the 20th century with a horizontal resolution of 0.25°×0.25° (Pai et al. 2014). ISMR is defined as the June-August averaged rainfall in the Indian monsoon core region (18°N-24°N, 76°E-87°E).

A common period of 1901-2010 among these two datasets is focused on. The CEF comparison is carried out from the spatial distribution of climatological meridional wind components and their standard deviation, as well as the evolution of different timescale components, including the long-term trend, interdecadal and interannual variability. The long-term trend is obtained by linear regression, the decadal component is calculated from the 11-year running mean, and the interannual component is from a 10-year high-pass filter, respectively. The consistency between the two reanalyses is evaluated through their temporal correlation. The deviation between them is measured by the root-mean-square error (RMSE) as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}}.$$

Where x_i and y_i denote the members of samples *x* and *y*, respectively, and *n* is the length of the sampling. Since the degree of freedom has changed after the filters, the Monte Carlo Method is applied to calculate the critical value of the correlation coefficient with significance shown in Table 1 (Shi et al. 1997; Zhao and Han 2005; Thompson 1979). The significance test for the difference between the reanalyses in terms of climatological mean and standard deviations are the Student's *t* test and the *F* test, respectively.

198 **3. Results**

199 *3.1 Differences between the two reanalyses, ERA20C and 20CRv3*

200 The vertical distribution of climatological CEFs expressed as the
201 meridional wind along the equator (average of 5°S-5°N) and their standard
202 deviations are displayed in Fig. 2. A consistence in both reanalyses is that the
203 Somali CEF is the strongest, followed by the Australian CEF, and the BOB
204 CEF is the weakest. Another consistence exhibits in their variability distribution,
205 for the Australian CEF has a violent fluctuation while the Somali and BOB
206 CEFs are relatively stable. The Somali CEF maximum is located at 850-925
207 hPa, and the other two are at 925-1000 hPa. To quantify the CEF, we refer to
208 the longitudinal range of 38°-60°E for Somali CEF, 79°-94°E for BOB CEF, the
209 three separated longitudinal ranges of 102°-112°E, 121°-132°E and
210 143°-150°E for the three sub-branches of Australian CEF. The latitude range
211 used to calculate CEFs is limited to 5°S-5°N.

212 From Fig. 2c an inconsistence in the climatological mean CEFs is seen, in
213 that the maximum intensity of Somali CEF is weaker in 20CRv3 than ERA20C,
214 along with a greater intensity in its eastern side; BOB CEF is also stronger in
215 20CRv3. A similarly stronger amplitude is seen in the two western
216 sub-branches of Australian CEF in 20CRv3, along with a weaker amplitude of
217 their eastern counterpart. The difference in the CEF variability is also evident
218 (Fig. 2f). There is a spatial contrast in the Somali CEF as ERA20C has greater
219 variability on the eastern side of Somali CEF, and 20CRv3 displays stronger

fluctuations on the west side. It is opposite for the east sub-branch of the Australian CEF. All the rest CEF branches show greater variability in 20CRv3.

To scrutiny these regional differences between the two reanalyses, we define the CEF index (Fig. 3) as the average meridional wind at 925 hPa within their corresponding range, and the Australian CEF index is the average of its three sub-branches. The Somali (Australian) CEF has a similar intensity of 8.85 m s^{-1} and 8.87 m s^{-1} (2.90 m s^{-1} and 2.95 m s^{-1}) in the two reanalyses, respectively. In contrast, the BOB CEF is stronger in 20CRv3 with a value of 3.18 m s^{-1}, while this value is 2.70 m s^{-1} in ERA20C. Also, the standard deviation of the BOB CEF (0.47 m s^{-1} and 0.46 m s^{-1}) and Australian CEF (0.76 m s^{-1} and 0.79 m s^{-1}) is very close to each other in the two reanalyses. However, the Somali CEF has greater variability in 20CRv3 as its standard deviation is 0.43 m s^{-1}, greater than 0.37 m s^{-1} in ERA20C.

From the year-to-year evolution of three CEFs, one can see an evident inconsistence between the two reanalyses (Fig. 3), especially for the Somali CEF which in ERA20C is stronger from 1901 to 1925, but turns weaker during the period of 1960-1980. The BOB CEF and Australian CEF show an overall consistency in the two reanalyses, with the temporal correlation of 0.59 and 0.76 between the two datasets for the two CEFs, respectively. But the BOB CEF in ERA20C is constantly lower than that in 20CRv3 throughout the century, and the amplitude gap prior to 1950 in Australian CEF is still un-negligible. Moreover, all these three CEFs show multi-timescale variation,

242 which contributes to the above differences. Therefore, we investigate their
243 difference with respect to timescales.

244 For the interannual timescale (Fig. 4a, c, e), there is a gap between the
245 Somali CEFs in the two datasets, but a consistency in the other two CEFs.
246 This is also seen from their distinct correlations with the coefficients of 0.25,
247 0.64 and 0.80 for Somali CEF, BOB CEF and Australian CEF, respectively. In
248 order to isolate the major periods with a large gap between the datasets, an
249 11-year running correlation and 11-year running root-mean-square error are
250 calculated between the reanalyses (Fig. 4b, d, f). For the Somali CEF, there
251 are three periods with a large gap, prior to 1925, 1945-1955 and around 2000,
252 during which the CEF's correlations between the reanalyses are weak or even
253 negative and the CEF anomalies are often opposite to each other in the two
254 datasets. The large gap periods for BOB CEF are pre-1920, the 1940s and the
255 1970s, respectively. For Australian CEF, they are pre-1950 and around 1990.
256 Nevertheless, these gaps are mainly caused by their amplitude. An overall
257 agreement in the interannual variability of these two CEFs is basically kept,
258 which is seen from a steady correlation above the 90% significant level
259 throughout the whole period.

260 The gaps in the CEFs interannual variability are mainly in the period prior
261 to 1950. This is understandable because the observations available for
262 assimilation are relatively few (Cram et al. 2015; Hersbach et al. 2015;
263 Woodruff et al. 2011), due to the limitations of observation technology and the

impacts of the World War I (1914–1918) as well as the World War II (1939–1945). Therefore, the biases caused by models cannot be well removed to yield reality, leading to a large divergence between the reanalyses.

For the interdecadal timescale, all the three CEFs vary inconsistently in the two reanalyses (Fig. 5), albeit the consistence in Australian CEF is relatively high with a correlation of 0.52. To understand the gap in depth, the power spectrum analysis is taken into account first. As shown in Fig. 6, the Somali CEF in the two reanalyses displays a different variation cycle. The dominant periodicity in ERA20C is around 50-year, in accompany with the positive anomalies in the periods before 1920, 1933-1952 and after 1988 and the negative anomalies in 1921-1932 and 1953-1987 (Fig. 5). In comparison, the Somali CEF in 20CRv3 has two cycles of 30 and 60-80 years, respectively, being in positive phase in 1906-1910 and 1925-1975 and in negative phase in 1911-1924 and 1976-2005. For the BOB CEF, even though it shares the same cycles of 20 and 40 years in the two datasets, there are still differences in intensity and phase period. The fluctuation in 20CRv3 becomes weak after 1925, while in ERA20C, the 40-year cycle lasts the whole century and the 20-year cycle is also evident before 1970. As a consequence, the decadal component of the BOB CEF in 20CRv3 basically stays positive from 1925 to 1990, but in ERA20C there are two positive periods in 1936-1950 and 1975-1995 and three negative periods in 1906-1935, 1951-1975 and after 1995, respectively. The major noticeable gaps in the Australian CEF between

the two reanalyses appear before 1940 and in the 1960s. In the first period, the Australian CEF in 20CRv3 displays a periodic fluctuation of 15-20 years and maintains positive. However, in ERA20C it basically stays in the negative phase and shows more frequent fluctuations with three cycles of 8-15 years, around 25 years, and 80-100 years, respectively, albeit the last one is less significant. In the second period, the anomaly of the Australian CEF is positive in ERA20C but negative in 20CRv3.

For the long-term trend, both ERA20C and 20CRv3 show an enhancement in the Somali and BOB CEFs, but a weakening in the Australian CEF (Fig. 7). As for the amplitude, the BOB CEF in both the reanalyses has the same trend of 0.5 m s^{-1} from 1901 to 2010, but the Somali CEF has evidently amplified intensification with the increment of 1.0 m s^{-1} in 20CRv3, much greater than that in ERA20C which has the value of 0.2 m s^{-1}. In contrast, 20CRv3 yields a smaller weakening trend in the Australian CEF with the southerly wind speed declining by 0.08 m s^{-1} during the 20th century, while the value is 0.4 m s^{-1} in ERA20C. In general, both the two reanalyses yield a consistent trend in all the three CEFs during the past century, but with a different amplitude.

3.2 Validation of the reliability of CEFs in the two reanalyses

a. The benchmark indices for validation

The above comparison suggests that both ERA20C and 20CRv3 bear an overall consistence in the interannual variability and long-term trend of the

three CEFs, but a substantial gap in the interdecadal variability. It does not imply which one of the two reanalysis datasets is more realistic or reliable. Due to the lack of instrumental records in the early 20th century, it is challenging to validate the CEFs in these reanalyses. Therefore, we investigate the accuracy of CEFs in the two reanalyses based on climate dynamic teleconnection.

We seek proxy indicators to validate CEFs. The indicators should have a clear physical meaning and stable dynamical connection with the CEF, in addition to having a long record. It is well known that the cross-equatorial meridional gradient of sea-level pressure (SLP) is the primary driver of the CEF, meanwhile SLP (say HadSLP2) has a longer instrumental record. Thus, CEPG can act as one proxy indicator for the CEFs. Besides, the Indian summer monsoon rainfall is greatly affected by the Somali CEF, and they bear a steady, solid correlation. Thus, ISMR can be used as another indicator for Somali CEF. Therefore, we consider these two variables and explore their feasibility as the benchmark index.

The validation is conducted by using the period of 1981-2010 during which the quality of the reanalyses is greatly guaranteed. One low-pass filter is used to obtain the CEF interdecadal components. Five reanalyses, ERA5, ERA-interim, JRA55, MERRA-2 and CFSR, are used to derive the CEF indices. As shown in Fig. 8, one substantial consistence exists in each CEF among the five reanalyses, even if one tiny divergence at the late 20th century. Therefore, we take their ensemble mean as the reality of the CEFs.

Comparing the individual CEPG with the corresponding CEF index, one substantial agreement is seen in the BOB and Australian CEFs, albeit a few discrepancies in the Somali CEF in the last several years. Thus, the CEPG index defined here is a reasonable indicator for the CEF interdecadal variability.

The Indian summer rainfall is closely connected to Somali CEF since its water vapor is mainly from the transport of Somali CEF (Kumar et al. 1999; Ordonez et al. 2013; Zhu 2012). Although some studies illustrated that factors such as mid-latitude disturbance and topographic distribution exert a significant impact on the Indian summer rainfall and lead to a considerable spatial variability (Kumar et al. 1995; Parthasarathy et al. 1993), here we consider the core area of Indian summer monsoon rainfall. Figure 9 shows the spatial distribution of the interdecadal correlation between the Indian summer rainfall and the Somali CEF in the five reanalyses as well as their ensemble mean. It is easy to see that the rainfall in the Indian monsoon core region (purple box) has a strongly positive correlation with the Somali CEF. This is well reflected in previous researches, since the variation of the rainfall in this area corresponds to the active and break phase of the monsoon, and is always accompanied by the intensity adjustment of the Somali CEF (Joseph and Sijikumar 2004; Rajeevan et al. 2010; Raman et al. 2011). Therefore, the ISMR index can be taken as another proxy indicator for the Somali CEF.

It is worthwhile noting that the early air pressure observations over the

Indian Ocean were relatively deficient, which affects the reliability of the Somali CEPG index. We compared it with the ISMR index and found that they verify each other (Fig. 10), with a correlation of 0.58. In comparison, the early observation over the subtropical northwest Pacific is relatively intense, which implies a relatively greater credibility of the Australian CEPG index. Therefore, we have sufficient confidence in the reference value of these indicators, but it is slightly weaker in the BOB CEPG index.

b. Validation result

Table 2 compares the correlations of the individual CEF indices derived from the two reanalyses with the correspondent proxy index. Visually, ERA20C outperforms 20CRv3 in all the three CEFs. This is also from one comparison of the interdecadal variation of the CEFs in the two reanalyses with the proxy index (Figs. 10 and 11). Both the variations of the Somali CEPG and the ISMR index are similar to the Somali CEF index derived from ERA20C, with a cycle of 50 years, but different from that in 20CRv3. The latter has two cycles of 30 years and 60-80 years. Besides, the Somali CEF index derived from ERA20C reproduces the increase in the proxy indexes in the 1920s and 1980s and the decrease in the 1960s, albeit some discrepancies prior to the 1920 and after 1990. In comparison, the evolution of the Somali CEF index in 20CRv3 missed all these shift points after 1920.

For the BOB CEF, its decadal variability in both reanalyses is less connected with the BOB CEPG index with the insignificant correlation of 0.33

and 0.17. This is verified by their different variation periodicities. The BOB CEPG index has a periodicity of 40-50 years before 1970 and then 15 and 30 years. This is different from the BOB CEF in the two reanalyses which both have two periodicities of 20 and 40 years, even evident before 1930 in 20CRv3. Nonetheless, the BOB CEF derived from ERA20C is obviously different from the proxy index in the two periods. The first is 1920-1950 when the BOB CEPG index stays positive, but the BOB CEF in ERA20C exhibits a strong 20-year cycle fluctuation. Another period is after 1980, during which the BOB CEF in ERA20C is anomalously positive until 1995, while the BOB CEPG index turns negative in the mid-1980s but positive in the early 21st century. The only consistent period between the BOB CEF in 20CRv3 and the BOB CEPG index is after 1975.

For the Australian CEF, the CEF index derived from ERA20C and 20CRv3 resembles the Australian CEPG index unanimously with the correlation coefficient of 0.67 and 0.59, respectively. But during the two gap periods of the two reanalyses mentioned above, ERA20C shows a better consistency with the proxy CEPG index. For example, the Australian CEF index and the CEPG index have a similar 20-year cycle before 1940. But the Australian CEF index derived from 20CRv3 is opposite to the CEPG index, except for a short period 1920-1930 with the same positive anomaly. Also, in the 1960s both the CEPG index and the CEF index derived from ERA20C are positive, but the CEF index in 20CRv3 is negative, in spite of a consistent enhancement in the CEF index

in both the reanalyses and the Australian CEPG index in the preceding period. In brief, all these validations of the interdecadal variability suggest that ERA20C is more closely similar to the proxy, Thus is more credible than 20CRv3.

4. Summary and Discussions

This paper first compared the two 20th century atmospheric reanalysis datasets, ERA20C and 20CRv3, as far as the boreal summer low-level cross-equatorial flows over the Asian-Australian monsoon region are concerned. The results suggest an overall consistence with respect to the spatial structure of the CEF climatology and their variability expressed as standard deviation. Also, substantial gaps between the two datasets are seen as follows.

1) Climatologically, the intensity of BOB CEF and the variability of Somali CEF are stronger in 20CRv3 than ERA20C. For the year-to-year evolution in ERA20C compared to that in 20CRv3, the Somali CEF is stronger from 1901 to 1925, but is weaker during the period of 1960-1980; the BOB CEF is constantly lower throughout the century and the Australian CEF has a weaker amplitude prior to 1950.

2) At interannual timescale, significant differences between the two datasets are seen in 1901-1950, with frequently opposite anomalies in Somali CEF and different amplitude in the BOB and Australian CEFs.

3) At interdecadal timescale, the two reanalyses show weak correlations

418 over three CEFs, with different decadal variation cycles in Somali CEF, along
419 with the weaker periodic fluctuation of the BOB CEF in 20CRv3, and the
420 opposite anomaly in the Australian CEF before 1940 and in the 1960s.

421 4) As for the long-term trend, the two reanalyses show a consistence in
422 the signal, with the upward trend in Somali CEF and BOB CEF and the
423 downward trend in Australian CEF. But there are gaps in the trend amplitude,
424 with a stronger strengthening in the Somali CEF but a less evident weakening
425 in the Australian CEF in 20CRv3.

426 The gaps in interdecadal variability of all the CEFs between the
427 reanalyses raise one issue: which one is more reliable and which dataset can
428 be selected for use? It is difficult to address this issue directly, since there are
429 no sufficient instrumental wind records in the early 20th century. Here this is
430 conducted by comparing CEFs in the two datasets with two proxy indicators.
431 The first is the cross-equatorial meridional gradient of sea-level pressure, and
432 the second is the Indian monsoon core region rainfall. They both have
433 substantially dynamic connections with the CEFs and longer instrumental
434 records. In addition, they verify each other during the 20th century. The
435 validation results suggest that all the CEFs in ERA20C are more reasonable
436 relative to 20CRv3, since the Somali CEF has the same periodicity variation as
437 the two proxy indicators, and the Australian CEF has a consistent fluctuation
438 with the proxy indicator. Thus, ERA20C is more suitable for investigating
439 interdecadal climate variability associated with CEFs.

The present finding that ERA20C is a more reliable dataset is of important meaning for understanding the interdecadal variability of Asian-Australian summer monsoon, as well as for realizing global climate change. No doubt that the exchange of mass, momentum, water moisture and energy across the hemispheres are exclusively important for global climate distribution. Besides, the significant gap in interdecadal variability of the CEFs between the reanalyses provides some clues for developing next-generation atmospheric reanalysis products. The evident lag in the decadal variation of Somali CEF before 1920 and after 1980 in ERA20C relative to 20CRv3 and the two proxy indicators implies an urgent need to assimilate the oceanic observations into reanalysis dynamic system.

There are some deficiencies in this study. First, the benchmark indexes used for the CEFs validation are only approximate and cannot fully explain the variability of the CEFs. Second, the reason for the gap in inter-decadal variation of three CEFs in the two reanalyses is unclear. According to their similar variation periodicity to AMO and PDO, one may guess that the model used for the two reanalyses may have deficiencies in capturing decadal variability associated with PDO and AMO. The other possible reason is that the SST datasets applied for boundary conditions in the two reanalyses are different, leading to the bias of cross-equatorial SST gradient and contributing to the difference of CEFs between the reanalyses. Nonetheless, this issue deserves further investigation.

Acknowledgments

This work is jointly supported by the Strategic Project of the Chinese Academy of Science (Grant XDA19070402) and NSFC (Grant 41731177). 20CRv3 dataset is provided by the Office of Science Biological and Environmental Research (BER), US Department of Energy, and the Climate Program Office, National Oceanic and Atmospheric Administration, which is downloaded from the NOAA Physical Sciences Laboratory. ERA20C dataset is provided by ECMWF.

References

Ajayamohan, R. S., 2007: Simulation of south-Asian summer monsoon in a GCM. *Pure Appl. Geophys.*, **164**, 2117-2140.

Allan, R. and T. Ansell, 2006: A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *J. Climate*, **19**, 5816-5842.

Attada, R., H. P. Dasari, A. Parekh, J. S. Chowdary, S. Langodan, O. Knio and I. Hoteit, 2019: The role of the Indian summer monsoon variability on Arabian Peninsula summer climate. *Climate Dyn.*, **52**, 3389-3404.

Befort, D. J., S. Wild, T. Kruschke, U. Ulbrich and G. C. Leckebusch, 2016: Different long-term trends of extra-tropical cyclones and windstorms in ERA-20C and NOAA-20CR reanalyses. *Atmos. Sci. Lett.*, **17**, 586-595.

484 Bett, P. E., H. E. Thornton and R. T. Clark, 2017: Using the twentieth century
485 reanalysis to assess climate variability for the European wind industry.
486 *Theor. Appl. Climatol.*, **127**, 61-80.
487 Bloomfield, H., L. Shaffrey, K. Hodges and P. Vidale, 2018: A critical
488 assessment of the long-term changes in the wintertime surface Arctic
489 Oscillation and northern hemisphere storminess in the ERA20C
490 reanalysis. *Environ. Res. Lett.*, **13**, 094004,
491 doi:10.1088/1748-9326/aad5c5.
492 Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin,
493 B. E. Gleason, R. S. Vose, G. Rutledge and P. Bessemoulin, 2011: The
494 twentieth century reanalysis project. *Quart. J. Roy. Meteor. Soc.*, **137**,
495 1-28.
496 Cram, T. A., G. P. Compo, X. G. Yin, R. J. Allan, C. McColl, R. S. Vose, J. S.
497 Whitaker, N. Matsui, L. Ashcroft, R. Auchmann, P. Bessemoulin, T.
498 Brandsma, P. Brohan, M. Brunet, J. Comeaux, R. Crouthamel, B. E.
499 Gleason, P. Y. Groisman, H. Hersbach, P. D. Jones, T. Jonsson, S.
500 Jourdain, G. Kelly, K. R. Knapp, A. Kruger, H. Kubota, G. Lentini, A. Lorrey,
501 N. Lott, S. J. Lubker, J. Luterbacher, G. J. Marshall, M. Maugeri, C. J.
502 Mock, H. Y. Mok, O. Nordli, M. J. Rodwell, T. F. Ross, D. Schuster, L.
503 Srnec, M. A. Valente, Z. Vizi, X. L. Wang, N. Westcott, J. S. Woollen and S.
504 J. Worley, 2015: The international surface pressure databank version 2.
505 *Geosci. Data J.*, **2**, 31-46.

506 Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U.
507 Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M.
508 Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M.
509 Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Holm,
510 L. Isaksen, P. Kallberg, M. Kohler, M. Matricardi, A. P. McNally, B. M.
511 Monge-Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C.
512 Tavolato, J. N. Thepaut and F. Vitart, 2011: The ERA-Interim reanalysis:
513 Configuration and performance of the data assimilation system. *Quart. J.*
514 *Roy. Meteor. Soc.*, **137**, 553-597.
515 Deng, X. H., P. M. Zhai and C. H. Yuan, 2010: Comparative analysis of
516 NCEP/NCAR, ECMWF and JMA reanalysis. *Meteor. Sci. Technol.*, **38**, 1-8.
517 (in Chinese)
518 Donat, M. G., L. V. Alexander, N. Herold and A. J. Dittus, 2016: Temperature
519 and precipitation extremes in century-long gridded observations,
520 reanalyses, and atmospheric model simulations. *J. Geophys. Res. Atmos.*,
521 **121**, 11174-11189.
522 Fan, Q. Y., Z. X. Li and J. Z. Wang, 2018: Contact between precipitation
523 anomaly in Asian-Australian monsoon region and low-level
524 cross-equatorial flow variation over South China Sea in summer. *Trans.*
525 *Atmos. Sci.*, **41**, 684-691. (in Chinese)
526 Gao, H. and F. Xue, 2006: Seasonal variation of the cross-equatorial flows and
527 their influences on the onset of South China Sea summer monsoon.

Climatic. Environ. Res., **1**, 57-68. (in Chinese)

Gelaro, R., W. McCarty, M. J. Suarez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A. M. da Silva, W. Gu, G. K. Kim, R. Koster, R. Lucchesi, D. Merkova, J. E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S. D. Schubert, M. Sienkiewicz and B. Zhao, 2017: The Modern-Era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419-5454.

Giese, B. S., H. F. Seidel, G. P. Compo, and P. D. Sardeshmukh, 2016: An ensemble of ocean reanalyses for 1815-2013 with sparse observational input. *J. Geophys. Res.: Oceans*, **121**, 6891-6910.

Halpern, D. and P. M. Woiceshyn, 2001: Somali jet in the Arabian Sea, El Niño, and India rainfall. *J. Climate*, **14**, 434-441.

Hersbach, H., P. Poli and D. Dee, 2015: *The observation feedback archive for the ICOADS and ISPD data sets.* ERA Report Series, ECMWF, Reading, UK, 18 pp. [Available at https://www.ecmwf.int/en/publications]

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horanyi, J. Munoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R.

J. Hogan, E. Holm, M. Janiskova, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume and J. N. Thepaut, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999-2049.

Hoskins, B. J., G. Y. Yang and R. M. Fonseca, 2020: The detailed dynamics of the June-August Hadley Cell. *Quart. J. Roy. Meteor. Soc.*, **146**, 557-575.

Huang, B., U. Cubasch and Y. Li, 2018: East Asian summer monsoon representation in re-analysis datasets. *Atmos.*, **9**, 235, doi:10.3390/atmos9060235.

Huang, X., T. J. Zhou, W. X. Zhang, J. Jiang, P. X. Li and Y. Zhao, 2019: Northern hemisphere land monsoon precipitation changes in the twentieth century revealed by multiple reanalysis datasets. *Climate Dyn.*, **53**, 7131-7149.

Jain, S., S. K. Mishra, A. Anand, P. Salunke and J. T. Fasullo, 2021: Historical and projected low-frequency variability in the Somali jet and Indian summer monsoon. *Climate Dyn.*, **56**, 749-765.

Joseph, P. V. and S. Sijikumar, 2004: Intraseasonal variability of the low-level jet stream of the Asian summer monsoon. *J. Climate*, **17**, 1449-1458.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White and J. Woollen, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-472.

Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H.

572 Kamahori, C. Kobayashi, H. Endo, K. Miyaoka and K. Takahashi, 2015:
573 The JRA-55 reanalysis: General specifications and basic characteristics. *J.*
574 *Meteor. Soc. Japan.*, **93**, 5-48.
575 Kumar, K. K., M. K. Soman and K. R. Kumar, 1995: Seasonal forecasting of
576 Indian summer monsoon rainfall: a review. *Wea.*, **50**, 449-467.
577 Kumar, M. R. R., S. S. C. Shenoi and P. Schluessel, 1999: On the role of the
578 cross equatorial flow on summer monsoon rainfall over India using
579 NCEP/NCAR reanalysis data. *Meteor. Atmos. Phys.*, **70**, 201-213.
580 Laloyaux, P., E. de Boisseson, M. Balmaseda, J. R. Bidlot, S. Broennimann, R.
581 Buizza, P. Dalhgren, D. Dee, L. Haimberger, H. Hersbach, Y. Kosaka, M.
582 Martin, P. Poli, N. Rayner, E. Rustemeier and D. Schepers, 2018:
583 CERA-20C: A coupled reanalysis of the twentieth century. *J. Adv. Model.*
584 *Earth Syst.*, **10**, 1172-1195.
585 Lei, X. C. and X. Q. Yang, 2008: Interannual variation characteristics of eastern
586 hemispheric cross-equatorial flow and its concurrent relationships with
587 temperature and rainfall in China. *J. Trop. Meteor.*, **14**, 97-100.
588 Li, C. and S. L. Li, 2014: Interannual seesaw between the Somali and the
589 Australian cross-equatorial flows and its connection to the East Asian
590 summer monsoon. *J. Climate*, **27**, 3966-3981.
591 Li, C. and S. L. Li, 2016: Connection of the interannual seesaw of the
592 Somali-Australian cross-equatorial flows with China summer rainfall. *Chin.*
593 *Sci. Bull.*, **61**, 1453-1461. (in Chinese)

Li, C. Y. and J. B. Wu, 2002: Important role of the Somali cross-equator flow in the onset of the South China Sea summer monsoon. *Chin. J. Atmos. Sci.*, **2**, 185-192. (in Chinese)

Martin, G. M., K. Arpe, F. Chauvin, L. Ferranti, K. Maynard, J. Polcher, D. B. Stephenson and P. Tschuck, 2000: Simulation of the Asian summer monsoon in five European general circulation models. *Atmos. Sci. Lett.*, **1**, 37-55.

McPhaden, M. J., G. Meyers, K. Ando, Y. Masumoto, V. S. N. Murty, M. Ravichandran, F. Syamsudin, J. Vialard, L. Yu and W. Yu, 2009: The research moored array for African-Asian-Australian monsoon analysis and prediction. *Bull. Amer. Meteor. Soc.*, **90**, 459-476.

Ordonez, P., P. Ribera, D. Gallego and C. Pena-Ortiz, 2013: Influence of Madden-Julian Oscillation on water budget transported by the Somali low-level jet and the associated Indian summer monsoon rainfall. *Water Resour. Res.*, **49**, 6474-6485.

Pai, D. S., L. Sridhar, M. Rajeevan, O. P. Sreejith, N. S. Satbhai and B. Mukhopadhyay, 2014: Development of a new high spatial resolution (0.25× 0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, **65**, 1-18.

Parthasarathy, B., K. R. Kumar and A. A. Munot, 1993: Homogeneous Indian monsoon rainfall: variability and prediction. *Proc. Indian Acad. Sci.*, **102**,

121-155.

Poli, P., H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. Tan, C. Peubey and J.-N. Thépaut, 2016: ERA-20C: An atmospheric reanalysis of the twentieth century. *J. Climate*, **29**, 4083-4097.

Polonskii, A. B., S. B. Krasheninnikova and D. V. Basharin, 2017: Interdecadal variability of the meridional Ekman heat and mass transport in the North Atlantic and its relation to the Atlantic Multidecadal Oscillation. *Russ. Meteor. Hydrol.*, **42**, 653-660.

Prasanna, V., 2016: South Asian summer monsoon rainfall variability and trend: Its links to Indo-Pacific SST anomalies and moist processes. *Pure Appl. Geophys.*, **173**, 2167-2193.

Prasanna, V., B. Preethi, J. Oh, I. Kim and S. Woo, 2020: Performance of CMIP5 atmospheric general circulation model simulations over the Asian summer monsoon region. *Glob. Planet. Change*, **194**, 103298, doi:10.1016/j.gloplacha.2020.103298.

Rajeevan, M., S. Gadgil and J. Bhate, 2010: Active and break spells of the Indian summer monsoon. *J. Earth Syst. Sci.*, **119**, 229-247.

Raman, M. R., M. V. Ratnam, M. Rajeevan, V. Rao and S. V. B. Rao, 2011: Intriguing aspects of the monsoon low-level jet over Peninsular India revealed by high-resolution GPS radiosonde observations. *J. Atmos. Sci.*, **68**, 1413-1423.

638 Rohrer, M., S. Brönnimann, O. Martius, C. C. Raible and M. Wild, 2019:
639 Decadal variations of blocking and storm tracks in centennial reanalyses.
640 *Tellus A: Dyn. Meteor. Oceanogr.*, **71**, 1586236,
641 doi:10.1080/16000870.2019.1586236
642 Saha, S., S. Moorthi, H. L. Pan, X. R. Wu, J. D. Wang, S. Nadiga, P. Tripp, R.
643 Kistler, J. Woollen, D. Behringer, H. X. Liu, D. Stokes, R. Grumbine, G.
644 Gayno, J. Wang, Y. T. Hou, H. Y. Chuang, H. M. H. Juang, J. Sela, M.
645 Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J.
646 Meng, H. L. Wei, R. Q. Yang, S. Lord, H. Van den Dool, A. Kumar, W. Q.
647 Wang, C. Long, M. Chelliah, Y. Xue, B. Y. Huang, J. K. Schemm, W.
648 Ebisuzaki, R. Lin, P. P. Xie, M. Y. Chen, S. T. Zhou, W. Higgins, C. Z. Zou,
649 Q. H. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge and
650 M. Goldberg, 2010: The NCEP climate forecast system reanalysis. *Bull.*
651 *Amer. Meteor. Soc.*, **91**, 1015-1057.
652 Shi, N., F. Y. Wei, G. L. Feng and T. L. Shen, 1997: Monte Carlo test used in
653 correlation and composite analysis of meteorological fields. *J. Nanjing Inst.*
654 *Meteor.*, **3**, 355-359. (in Chinese)
655 Slivinski, L. C., G. P. Compo, J. S. Whitaker, P. D. Sardeshmukh, B. S. Giese,
656 C. McColl, R. Allan, X. Yin, R. Vose and H. Titchner, 2019: Towards a more
657 reliable historical reanalysis: Improvements for version 3 of the twentieth
658 century reanalysis system. *Quart. J. Roy. Meteor. Soc.*, **145**, 2876-2908.
659 Stankunavicius, G., D. Basharin, R. Skorupskas and G. Vivaldo, 2017:

660 Euro-Atlantic blocking events and their impact on surface air temperature
661 and precipitation over the European region in the 20th century. *Climate*
662 *Res.*, **71**, 203-218.
663 Thompson, R., 1979: Coherence significance levels. *J. Atmos. Sci.*, **36**,
664 2020-2021.
665 Wang, H. J. and F. Xue, 2003: The interannual variability of Somali jet and its
666 influences on the inter-hemispheric water vapor transport and the East
667 Asian summer rainfall. *Chin. J. Geophys.*, **46**, 11-20.
668 Wang, M., S. Yao, L. Jiang, Z. Liu, C. Shi, K. Hu, T. Zhang, Z. Zhang and J. Liu,
669 2018: Collection and pre-processing of satellite remote-sensing data in
670 CRA-40 (CMA's Global Atmospheric Reanalysis). *Adv. Meteor. Sci.*
671 *Technol.*, **8**, 158-163. (in Chinese)
672 Welker, C. and O. Martius, 2014: Decadal-scale variability in hazardous winds
673 in northern Switzerland since end of the 19th century. *Atmos. Sci. Lett.*, **15**,
674 86-91.
675 Wohland, J., N. E. Omrani, D. Witthaut and N. S. Keenlyside, 2019:
676 Inconsistent wind speed trends in current twentieth century reanalyses. *J.*
677 *Geophys. Res. Atmos.*, **124**, 1931-1940.
678 Woodruff, S. D., S. J. Worley, S. J. Lubker, Z. H. Ji, J. E. Freeman, D. I. Berry,
679 P. Brohan, E. C. Kent, R. W. Reynolds, S. R. Smith and C. Wilkinson, 2011:
680 ICOADS release 2.5: Extensions and enhancements to the surface
681 marine meteorological archive. *Int. J. Climatol.*, **31**, 951-967.

682 Wu, X. F. and J. Y. Mao, 2019: Decadal changes in interannual dependence of
683 the Bay of Bengal summer monsoon onset on ENSO modulated by the
684 Pacific Decadal Oscillation. *Adv. Atmos. Sci.*, **36**, 1404-1416.
685 Zeng, Q. C. and J. P. Li, 2002: Interactions between the northern and southern
686 hemispheric atmospheres and the essence of monsoon. *Chin. J. Atmos.*
687 *Sci.*, **26**, 433-448. (in Chinese)
688 Zhao, J. and Y. B. Han, 2005: Estimation of correlation significance levels after
689 moving average. *J. Beijing Natl. Univ.: Nat. Sci.*, **2**, 139-141. (in Chinese)
690 Zhao, T. B., C. B. Fu, Z. J. Ke and W. D. Guo, 2010: Global atmosphere
691 reanalysis datasets: Current status and recent advances. *Adv. Earth Sci.*,
692 **25**, 242-254. (in Chinese)
693 Zhu, Y. L., 2012: Variations of the summer Somali and Australia
694 cross-equatorial flows and the implications for the Asian summer
695 monsoon. *Adv. Atmos. Sci.*, **29**, 509-518.
696 Zurita-Gotor, P., 2020: The impact of divergence tilt and meridional flow for
697 cross-equatorial eddy momentum transport in Gill-Like settings. *J. Atmos.*
698 *Sci.*, **77**, 1933-1953.
699
700
701
702

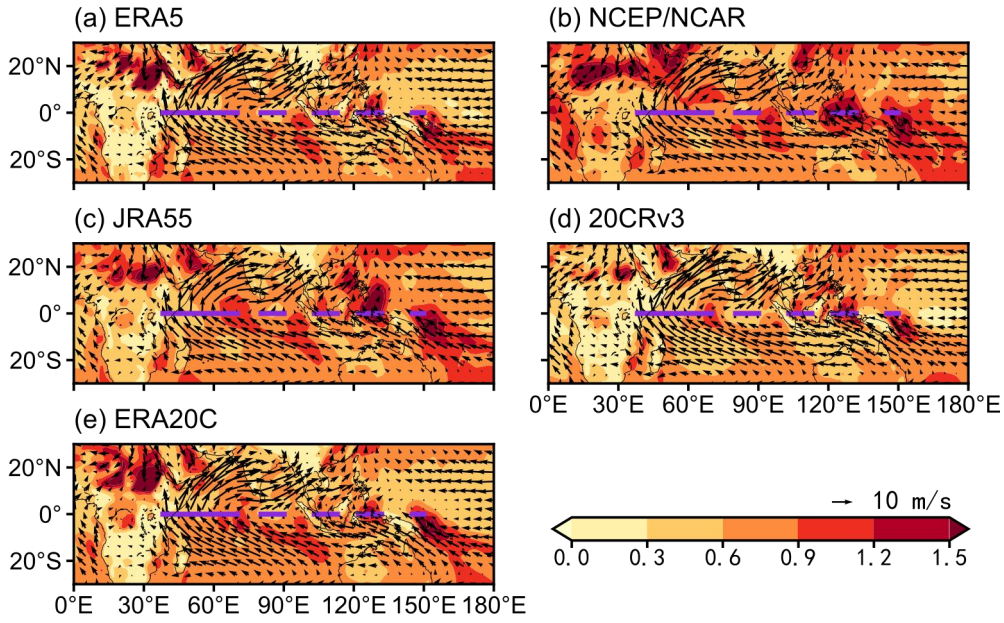

Fig. 1. The horizontal distribution of summer (June to August) climatological horizontal wind (arrows; m s^{-1}) and standard deviation of meridional wind in 925 hPa over the eastern hemisphere (color shading) in (a) ERA5, (b) NCEP/NCAR, (c) JRA55, (d) 20CRv3 and (e) ERA20C. The purple lines from west to east indicate the position of the Somali CEF, the BOB CEF and three sub-branches of the Australian CEF. The climatology is calculated as the mean through 1958-2010.

703

704

705

706

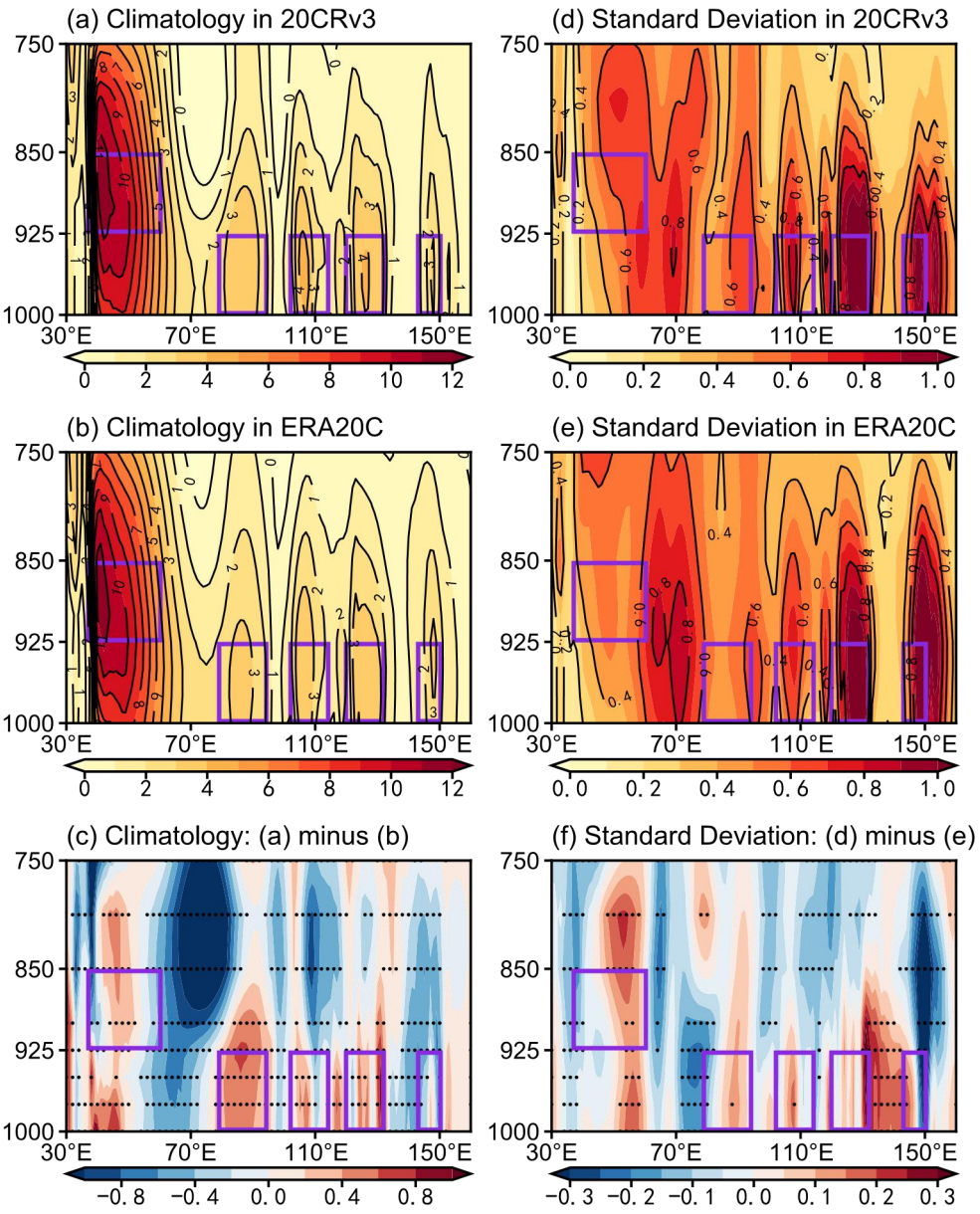

Fig. 2. (a, b) Vertical distribution of the summer climatologic cross-equatorial meridional wind (m s^{-1}) in 20CRv3 and ERA20C during 1901-2010. (d, e) as (a, b) but for the standard deviation. (c, f) the difference between 20CRv3 and ERA20C. The purple box marks the maximum of the CEFs, which is in the 850-925 hPa for Somali CEF, but in 925-1000 hPa for others. The black dots represent the significance over the 95% level. The climatology is calculated as the mean through 1958-2010.

707

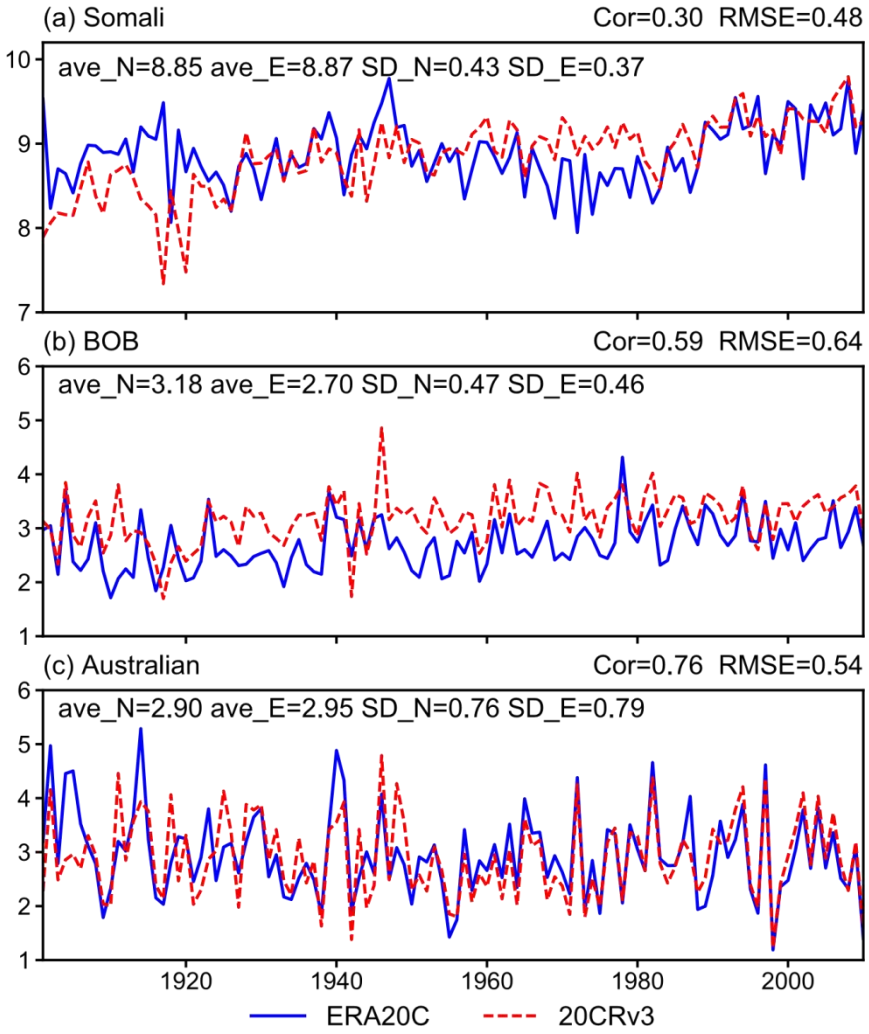

Fig. 3. The series of (a) Somali CEF index, (b) BOB CEF index, and (c) Australian CEF index (m s^{-1}) in ERA20C and 20CRv3. ave_N (ave_E) and SD_N (SD_E) indicate the mean value and standard deviation of the CEF indexes in 20CRv3 (ERA20C), respectively, and the gaps between the two reanalyses in each CEF is measured by the Cor (correlation coefficient) and RMSE (root-mean-square error).

708

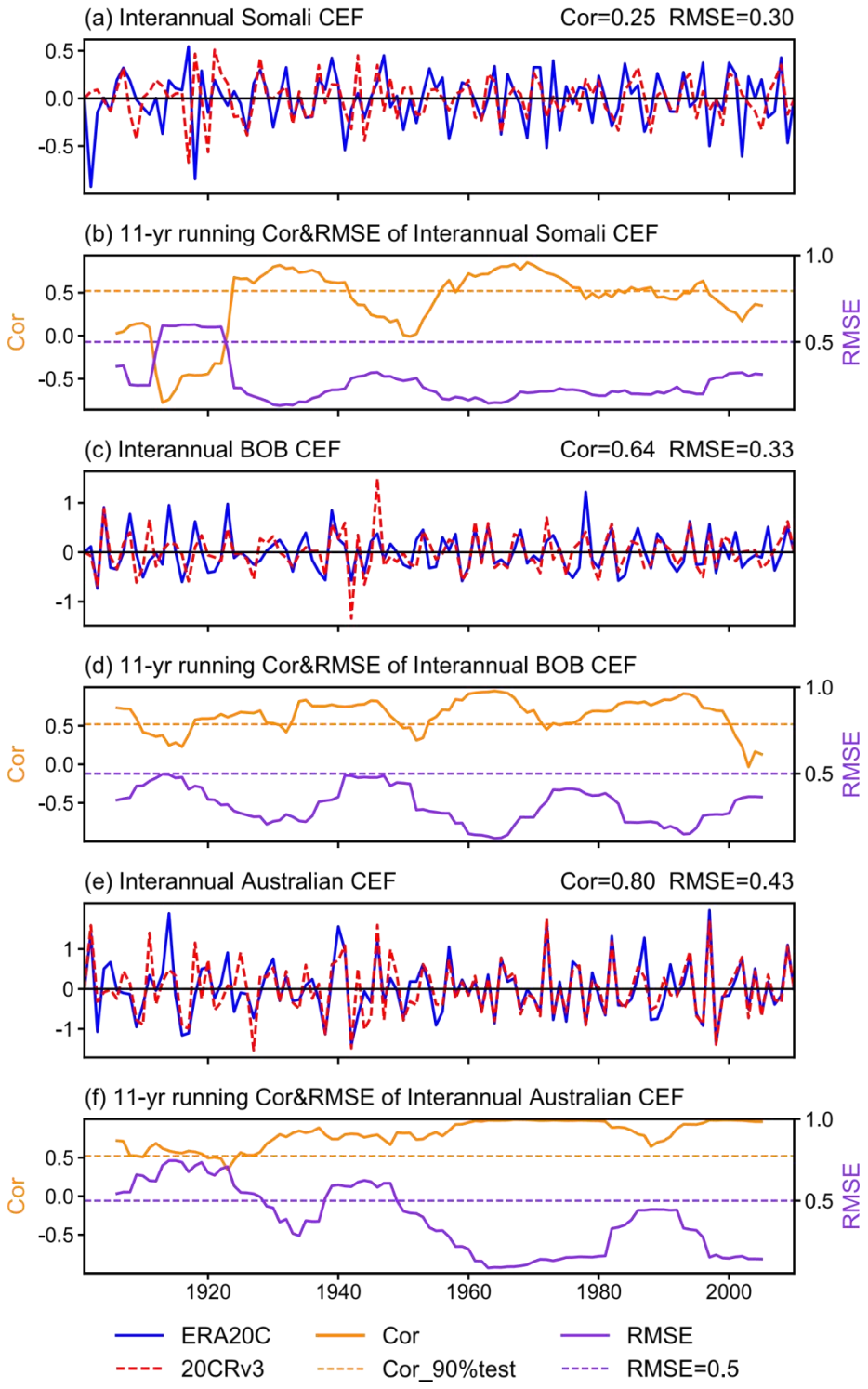

Fig. 4. The interannual variability of (a) Somali CEF index, (c) BOB CEF index, and (e) Australian CEF index (m s^{-1}) in ERA20C and 20CRv3. (b), (d), (f) present the correlation (Cor) and root-mean-square error (RMSE) of the two reanalyses in the 11-yr running window over the three CEFs.

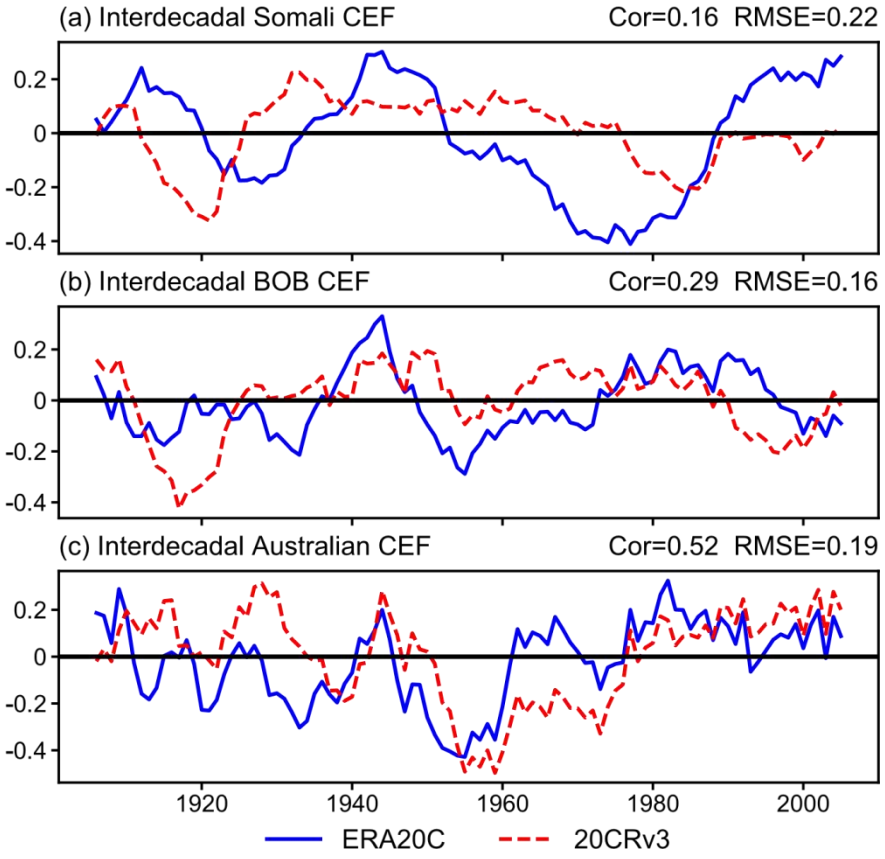

Fig. 5. The interdecadal variability of (a) Somali CEF index, (b) BOB CEF index, and (c) Australian CEF index (m s^{-1}) in ERA20C and 20CRv3.

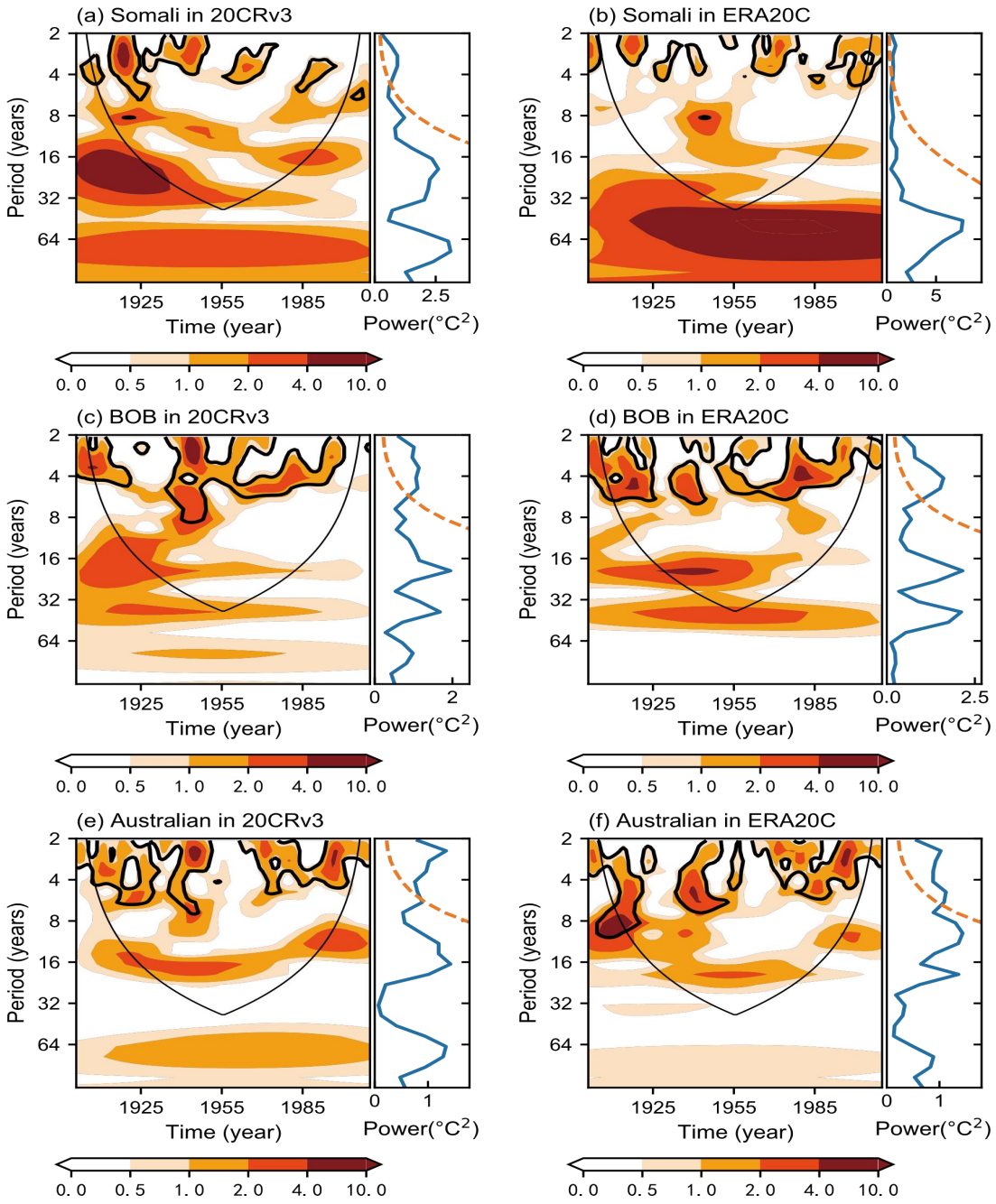

Fig. 6. Wavelet power spectrum (the left part in the shade) and global power spectrum (the right part with blue line) of (a) Somali CEF, (c) BOB CEF and (e) Australian CEF in 20CRv3. (b), (d), (f) are the same as (a), (c), (e), but for the CEFs in ERA20C; The thick black lines and the orange dashed lines indicate the significance at 95% level and the thin black lines are the cones of edge influence.

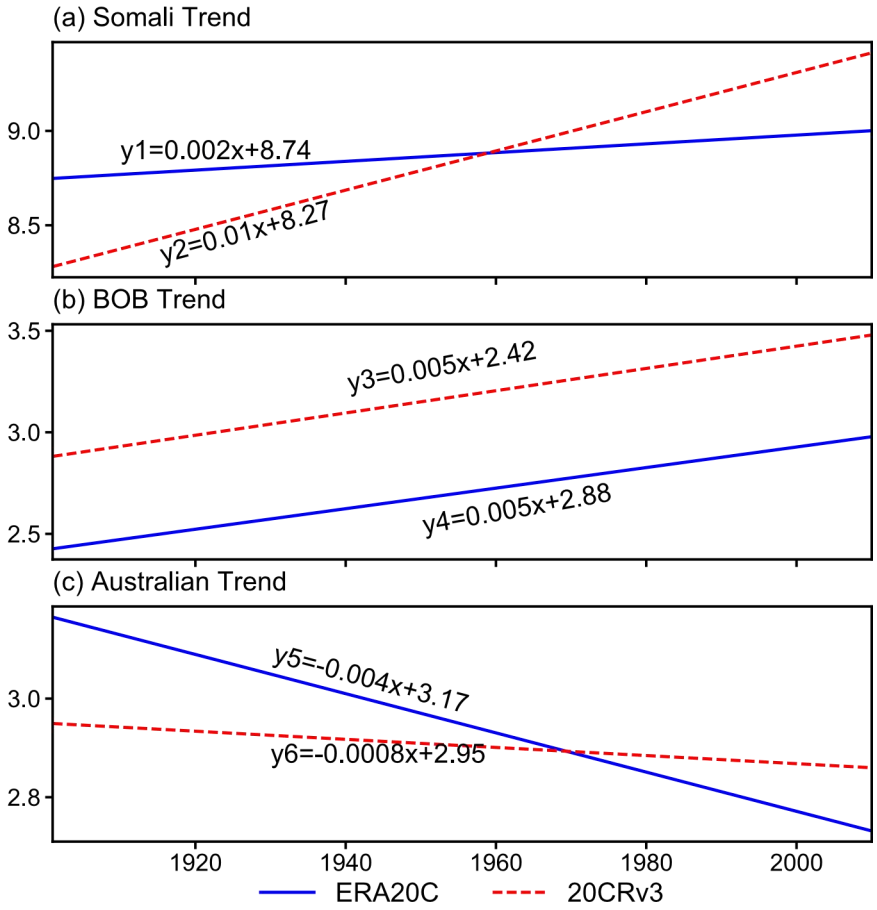

Fig. 7. The long-term trend of (a) Somali CEF, (b) BOB CEF and (c) Australian CEF in 20CRv3 and ERA20C with the corresponding linear regression equation.

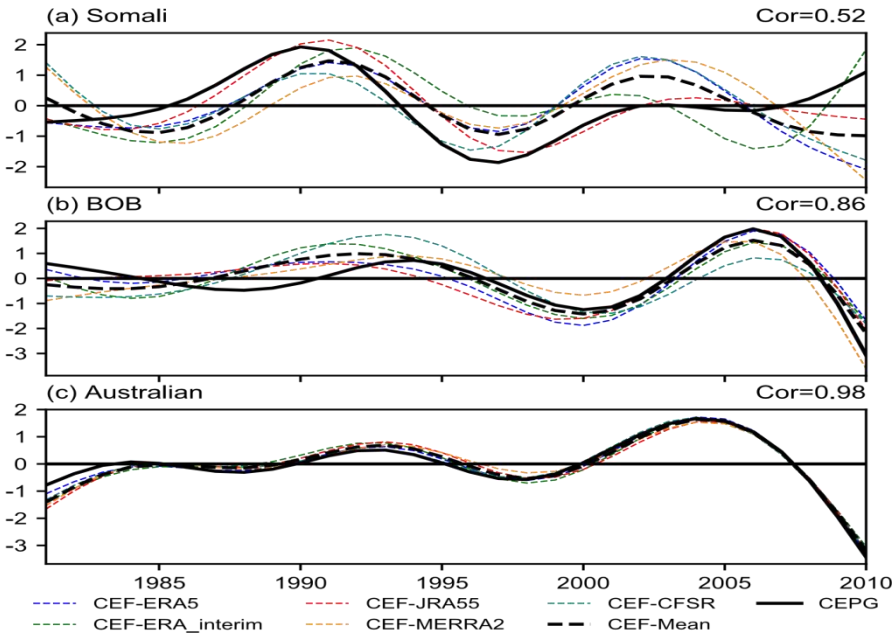

Fig. 8. The interdecadal variability of (a) Somali CEF, (b) BOB CEF and (c) Australian CEF during 1981-2010 in five reanalyses and their ensemble mean, as well as the corresponding CEPG index, obtained by 10-year low-pass filtering. The correlation (Cor) of the ensemble mean CEF index and the corresponding CEPG index is present in their upper-right corner.

724

725

726

727

728

729

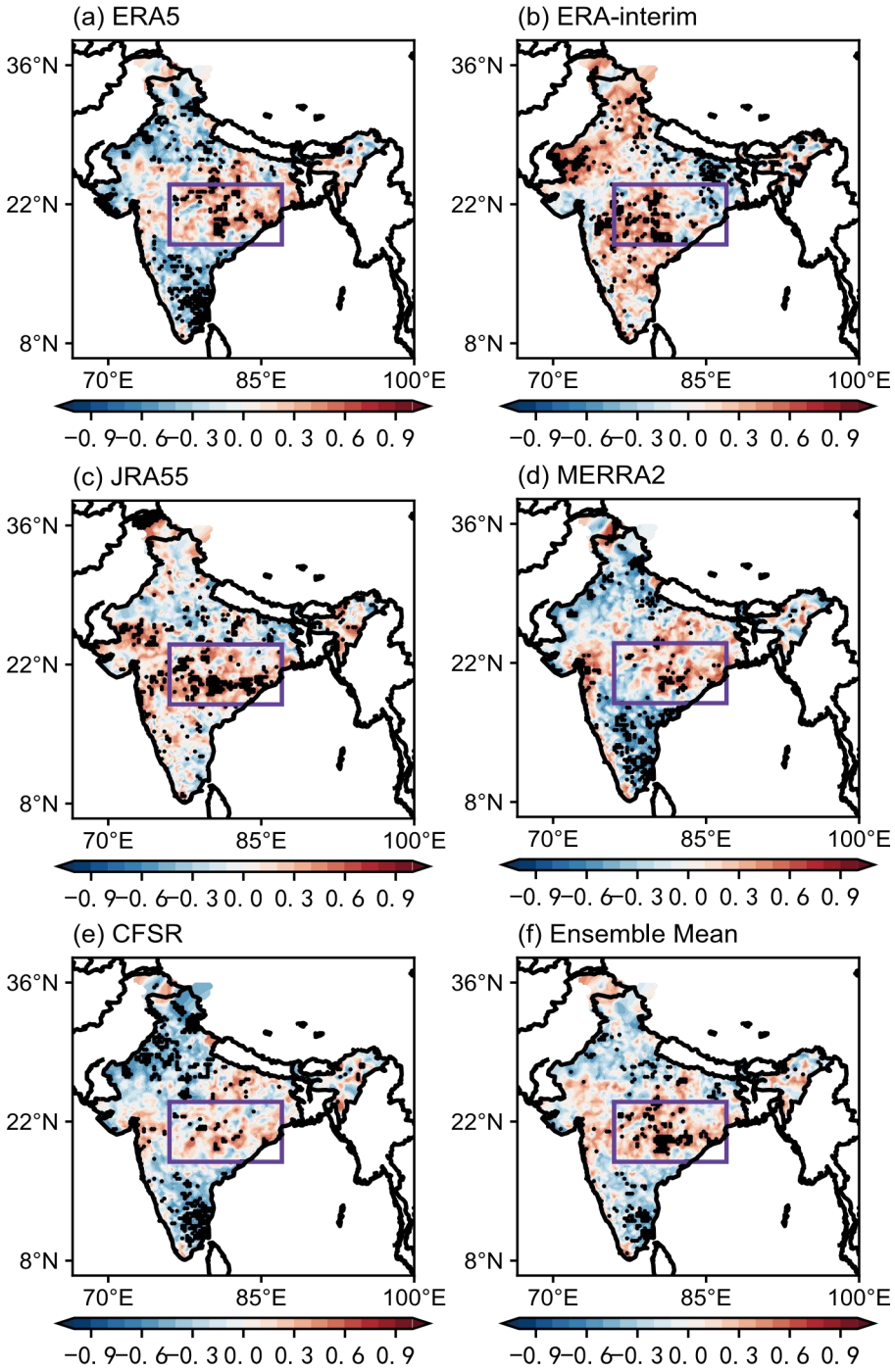

Fig. 9. The spatial distribution of interdecadal correlation between Indian summer rainfall and Somali CEF during 1981-2010 in five reanalyses and their ensemble mean. The dotted areas passed the significance test in the 95% level, and the areas within the purple box are the Indian monsoon core region.

730

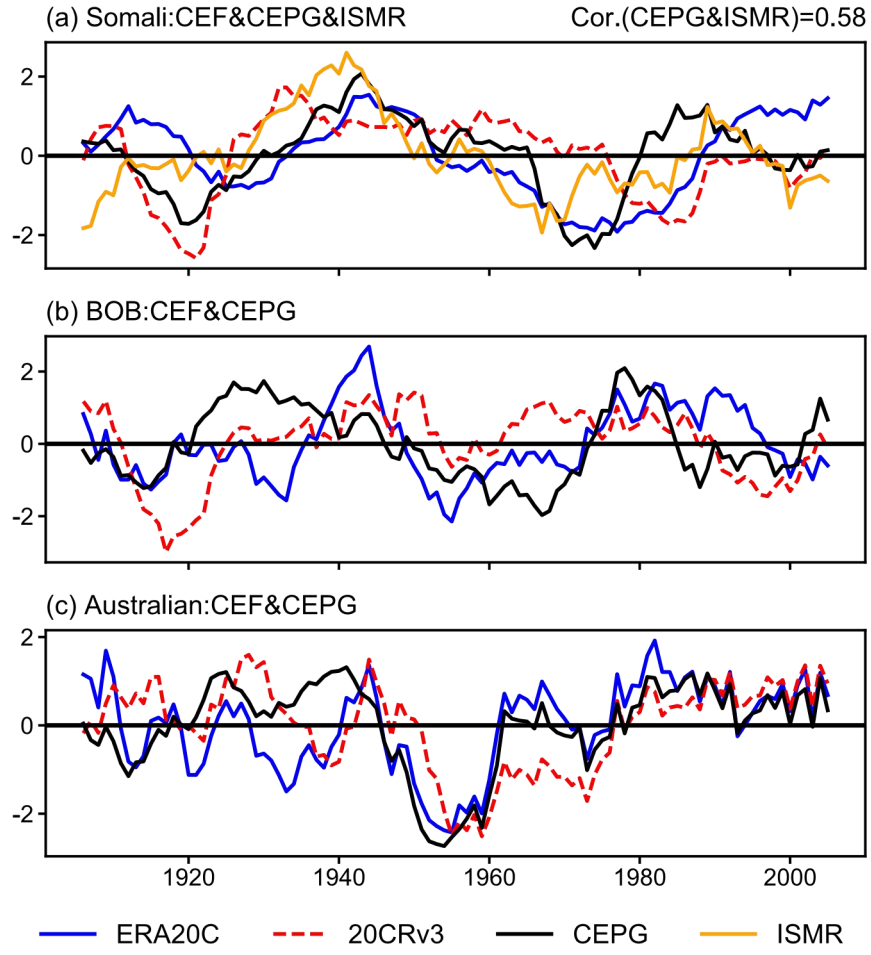

Fig. 10. (a) The normalized interdecadal Somali CEF index in the two reanalyses and its benchmark indexes (Somali CEPG and ISMR). (b) and (c) are the same as (a) but for the BOB CEF and Australian CEF, as well as their corresponding benchmark indexes (BOB CEPG and Australian CEPG). The correlation coefficient (Cor) between the Somali CEPG index and the ISMR index is 0.58.

731

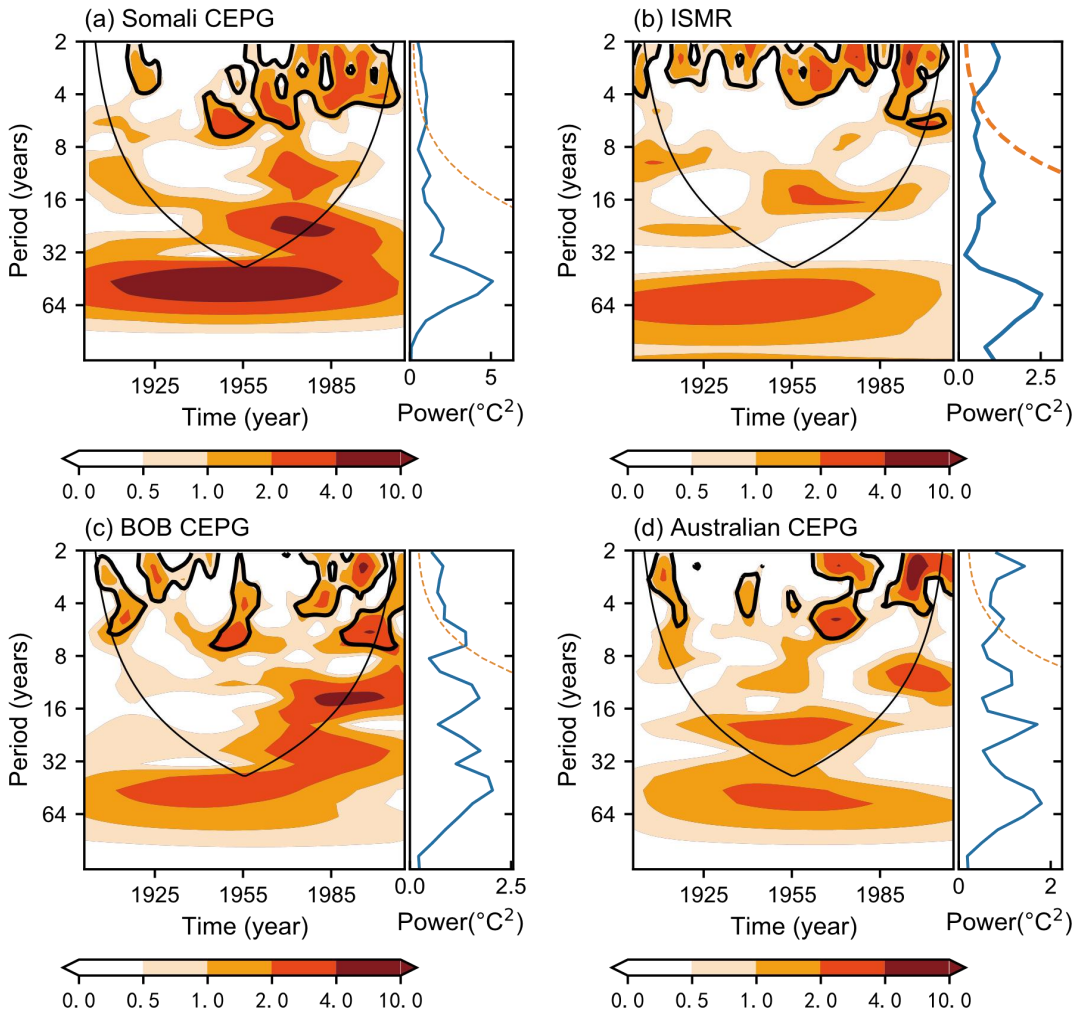

Fig. 11. Same as Fig. 6, but for the benchmark indexes, including (a) Somali CEPG index, (b) ISMR index, (c) BOB CEPG index and (d) Australian CEPG index.

Table 1 Critical value of correlation coefficient at the different significance levels (90%,95%,99%) for the CEF series (n=110) in different time scales, calculated by the Monte Carlo method.

Time Scale	α=0.1	α=0.05	α=0.01
Original	0.16	0.19	0.24
Interannual	0.17	0.21	0.27
Interdecadal	0.48	0.55	0.66

Table 2 The correlation coefficients of the Somali CEF, BOB CEF as well as the Australian CEF with their corresponding benchmark indexes in 20CRv3 and ERA20C. The benchmark index for Somali CEF includes Somali CEPG index and ISMR index, but for BOB CEF and Australian CEF is only the corresponding CEPG index. The coefficient marked with a (two) star passed the 90% (95%) significance test.

	Somali		BOB	Australian
	CEPG	ISMR	CEPG	CEPG
20CRv3	0.38	0.34	0.17	0.59**
ERA20C	0.53*	0.43	0.33	0.67**