Tracking the Indian Summer Monsoon Onset Back to the Preinstrument Period

PAULINA ORDOÑEZ

Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Mexico City, Mexico

DAVID GALLEGO, PEDRO RIBERA, AND CRISTINA PEÑA-ORTIZ

Departamento de Sistemas Físicos, Químicos y Naturales, Universidad Pablo de Olavide, Seville, Spain

RICARDO GARCÍA-HERRERA

Departamento de Física de la Tierra, Astronomía y Astrofísica II, Facultad de Ciencias Físicas,
Universidad Complutense de Madrid, and Instituto de Geociencias, Universidad Complutense
de Madrid and Consejo Superior de Investigaciones Científicas, Madrid, Spain

(Manuscript received 5 November 2015, in final form 13 July 2016)

ABSTRACT

The Indian summer monsoon onset is one of the most expected meteorological events of the world, affecting the lives of hundreds of millions of people. The India Meteorological Department has dated the monsoon onset since 1901, but its original methodology was considered subjective and it was updated in 2006. Unfortunately, the new method relies on OLR measurements, which impedes the construction of an objective onset series before the 1970s. An alternative approach is the use of the wind field, but the development of such an index is limited to the period covered by reanalysis products. In this paper historical wind records taken on board ships are used to develop a new onset series using only wind direction measurements, providing an objective record of the onset since the late nineteenth century. The new series captures the rapid precipitation increase associated with the onset, correlates well with previous approaches, and is robust against anomalous (bogus) onsets. A tendency for later-than-average onsets during the 1900–25 and 1970–90 periods and earlier-than-average onsets between 1940 and 1965 was found. A relatively stable relationship between ENSO and Indian monsoon onset dates was found; however, this link tends to be weaker during decades characterized by prevalent La Niña conditions. Furthermore, it was found that the link between the Pacific decadal oscillation (PDO) and the onset date is limited to the phases characterized by a shift from negative to positive PDO phases.

1. Introduction

The date of the Indian summer monsoon (ISM) onset is probably the most important meteorological event impacting the lives of the Indian people, particularly the agricultural planning that affects food production for over one billion people (Wang et al. 2009). Small changes in the onset date may have a great effect on highly vulnerable regions, even when the final average rainfall of the monsoon season as a whole is normal (Raju et al. 2007).

At the surface, the ISM onset is perceived as a fast and sustained increase in rainfall starting at the southernmost tip of the Indian subcontinent, typically in late May or early June (Fasullo and Webster 2003). The monsoonal rains then progress northward to cover the whole of India by late June. Traditionally, the official date of the ISM onset has been considered as the date of the monsoon onset over Kerala (MOK), which marks the beginning of a roughly 4-month period of rains in India, during which the country receives the bulk of its annual rainfall (Soman and Krishna Kumar 1993). The MOK is linked to profound changes in the atmospheric circulation. Accompanying such rainfall, the tropospheric westerly wind over Kerala is usually strong in the lowest 5 km of the troposphere and the relative humidity of the air is high from the surface to at least 500 hPa (Rao 1976). At the upper troposphere, the

---

Corresponding author address: Paulina Ordoñez Perez, Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Delegación Coyoacán, México City D. F. 04510, Mexico.
E-mail: orpep@atmosfera.unam.mx

Publisher's Note: This article was revised on 29 November 2016 to correct the affiliations of the fifth coauthor.

DOI: 10.1175/JCLI-D-15-0788.1

onset is generally linked to a northward shift of the subtropical jet stream to the north of the Tibetan Plateau (Yin 1949) and to the appearance of the tropical easterly jet over southern India (Koteswaram 1958). Notwithstanding this general behavior, the seasonal monsoon transitions can unfold at different times and in different spatial patterns, including the possibility of having a short but strong rain spell followed by hot and dry conditions previous to the monsoonal rains, which is known as ‘‘bogus’’ onset (Fieux and Stommel 1977; Flatau et al. 2001).

In view of this complex scheme, it is clear that designing an objective methodology to determine the MOK date covering all the possibilities constitutes a challenge. For this reason, the first methodology used to officially declare the MOK (up to 2005) was based on subjective criteria. The India Meteorological Department (IMD) determined the MOK date every year and for more than 100 years by qualitatively judging meteorological reports from stations of the Indian synoptic network. For the MOK, rainfall had to be widespread spatially over Kerala and persistent for a few days, but there was no specific threshold required to be surpassed (Joseph 2012). The lack of quantitative thresholds in this methodology and its susceptibility to fail during years of bogus onset (Rao 1976; Webster et al. 1998) led the IMD to adopt a new criterion for declaring the MOK in 2006. This new method is also based on the arrival of sustained rainfall to a set of stations (first step) and the finding of a wind field in accordance with the expected pattern associated with a well-established monsoon (second step). In addition, in order to filter bogus onsets, a third step makes use of OLR data to verify that on the selected date there is widespread convection around Kerala. The last requirement makes this method quite robust but dependent on the availability of OLR data, thus limiting the possibility of computing onset dates for the presatellite era. In an interesting exercise, Pai and Rajeevan (2009) reworked the MOK dates since 1971 by applying the new criteria. While both methods show reasonably consistent results (an average difference between 2 and 3 days), there are several years characterized by large differences in MOK dates that can reach up to 10 days. Therefore, it is not possible to obtain a long and homogeneous MOK date record by simply blending the old (presatellite era) and new (satellite era) series. As a result, all century-long MOK studies still rely on the IMD subjective series.

To get longer series of ISM onset dates based on a uniform objective methodology, several works have addressed the possibility of defining the onset through algorithms based on variables available for the presatellite era, such as precipitation, moisture transport, or wind. Fasullo and Webster (2003) defined the onset using vertically integrated moisture transport. Taniguchi and Koike (2006) found that the rapid enhancement of the wind speed was well related to the abrupt beginning of the rainy season. Xavier et al. (2007) defined the onset as the day when the tropospheric heat source shifts from south to north. Wang et al. (2009) were able to find a comparatively simpler but effective definition of the MOK date by using the wind field at 850 hPa alone. More recently, Carvalho et al. (2016) also defined the onset of the ISM based on a large-scale index of the Indian monsoon system. These authors obtained this large-scale monsoon index by performing a combined empirical orthogonal function (CEOF) analysis of precipitation, specific humidity and temperature at 2 m, and zonal and meridional winds at 10 m.

All these objective methods using large-scale circulation features have an advantage for verification of numerical models, but some existing indices do not have sufficiently high correlation with the IMD definition (MOK) (e.g., Fasullo and Webster 2003); thus, these indices have not been widely recognized as a faithful representation of the ISM onset (Wang et al. 2009). Traditionally, the IMD onset series has been used to study different aspects of the MOK, including its interannual (Raju et al. 2007) or intraseasonal variability (Bhatla et al. 2016) and its predictability (Puranik et al. 2013).

It could be argued that the use of wind measurements over an oceanic area is not a great advantage when the goal is the construction of a long-term MOK series, because of the lack of direct observational data apart from the reanalysis products. However, since early times, a lot of ship routes from Europe to India and the Far East traveled through the Arabian Sea and the northern Indian Ocean. Most of these ships took very precise wind measurements that were preserved in maritime reports and logbooks that are stored in historical archives (García-Herrera et al. 2005; Wheeler and García-Herrera 2008). The idea of using marine reports to track the monsoon onset was first proposed by Fieux and Stommel (1977), but despite the interest in their proposal, the small database of marine reports available limited their study to a few cases. Since that year, several international projects have recovered and digitized millions of these reports and meteorological observations contained in old logbooks (Woodruff et al. 2011 and references therein). Ribera et al. (2011) analyzed and reconstructed the occurrence of coastal storms with the help of these historical records of wind direction. More recently, Barriopedro et al. (2014) and Gallego et al. (2015) found that these data could be used to construct useful climate indices, including a consistent reconstruction of the West African monsoon back to 1839 (Gallego et al. 2015).

In this paper we aim to derive a new methodology, based on instrument-based wind direction data taken on board ships, capable of producing the longest possible series of the MOK. The generated onset index is not conceived as an alternative to current state-of-the-art MOK detection schemes, but it is thought to be used to study the stability of interannual variability and interdecadal and multidecadal variability. In addition, this methodology can help extend backward in time the MOK series when new logbook data over the affected area are digitized.

2. Data and methodology

Historical meteorological records taken on board ships usually provide, among other variables, wind strength and wind direction (García-Herrera et al. 2005). For historical records, the wind strength was usually taken by estimation and therefore not by instrument (Fieux and Stommel 1977; Gallego et al. 2007). On the other hand, wind direction was usually measured with a 16- or 32-point compass and it can be considered an instrument-based observation, even for the oldest records. The possibility of developing a homogeneous climatic index of instrument-based nature using wind data measured on board ships relies on the fact of using only the wind direction. For this work we made use of the International Comprehensive Ocean–Atmosphere Data Set (ICOADS; Woodruff et al. 2011), which currently stands as the most complete dataset of instrument-based and historical meteorological observations taken on board ships. At the time of writing, ICOADS release 2.5 contains over 261 million records, starting in 1662, and most of them include wind direction information.

The wind force field has been an essential part of any previous index intended to characterize the ISM. However, during late May or early June, the onset of the ISM is related to profound changes in the direction of the low-level jet over the Arabian Sea (Fieux and Stommel 1977; Joseph et al. 2006). Making use of this link, our aim has been to produce an index sensitive to the sudden change of the zonal flow in the northern Indian Ocean and, at the same time, dependent upon the persistence of such wind anomaly over time. The first condition requires searching for a strong shift in the prevalent wind direction in an adequately selected region. Once the shift has been detected, the second condition involves the establishment of a minimum threshold in the number of days in which the wind keeps its modified direction. To fulfill both requirements with the strongly space- and time-dependent coverage of ICOADS (raw observations are available only along a particular ship's route), we first selected an area just to the west of Kerala (7°–11°N, 60°–80°E). This area was defined as the "optimal area" because it maximizes the correlation between the percentage of west-southwest (WSW) winds (wind blowing between 225° and 270° from true north) and precipitation over India during May. The selected area lies across one of the main ship's routes (Fig. 1a) and consequently it has good data coverage. However, several sensitive tests were performed showing slight differences between this area and nearby rectangular areas that also were found to adequately represent the Indian monsoon precipitation during May. Figure 1b shows the temporal evolution of the total number of available records in ICOADS release 2.5 for May and June in the selected area. ICOADS coverage is almost insignificant up to 1870, growing progressively up to the turn of the twentieth century. From 1901 onward, the data coverage substantially improves, although some intervals still display less than 500 observations per year: 1917–1918 (the end of the World War I), 1939–1943 (during World War II), and the years around 1971 (Indo-Pakistani War).

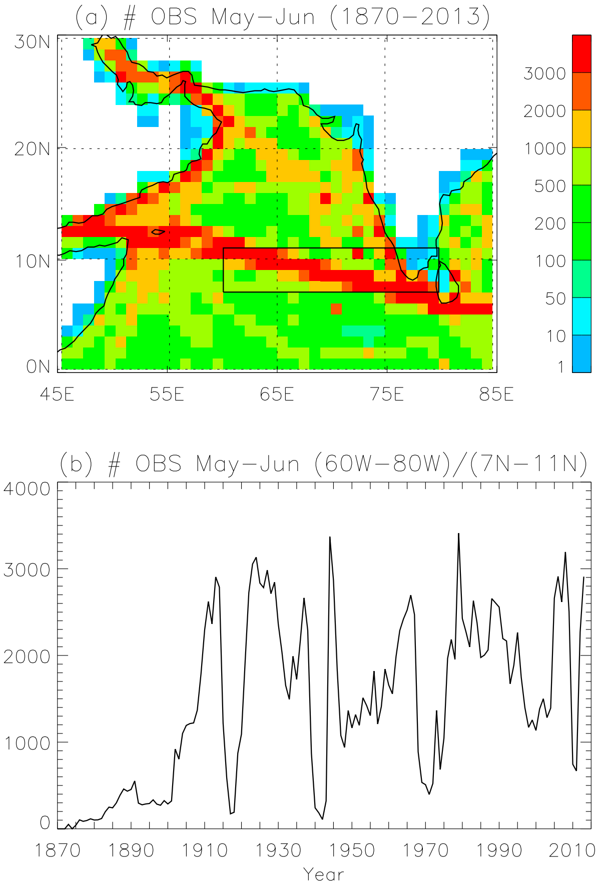

FIG. 1. (a) Area of analysis (black rectangle in the map) and total number of raw ICOADS release 2.5 data for May–June during 1870–2013 in a 1° × 1° grid (shading). (b) Total number of wind direction observations for May–June inside the rectangle in (a).

For the available data, we implemented a double-averaging procedure. In the first step, we computed for any given day N, the percentage of WSW winds (wind blowing between 225° and 270° from true north) over the total number of available wind observations in the area. We required at least five observations in a given day to

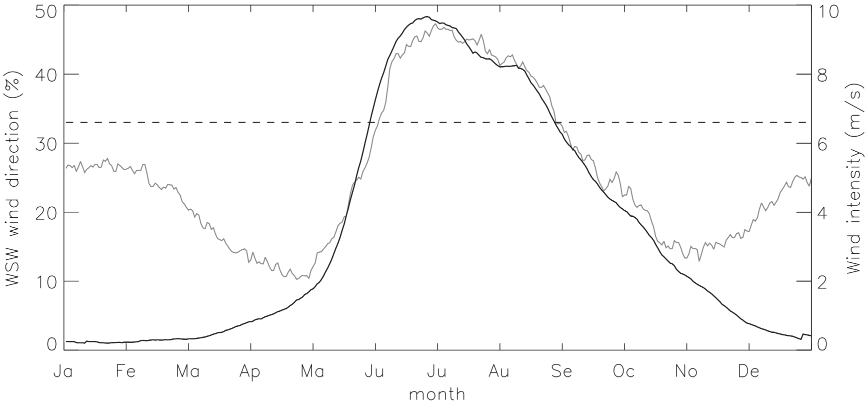

FIG. 2. Seasonal cycle of the WCI (black line) and average wind speed from the NCEP–NCAR reanalysis (gray line) for 1948–2013 inside the area 7°–11°N, 60°–80°E.

compute this percentage; otherwise, the day N was tagged as missing. In the second step, we defined the so-called western circulation index (WCI) as the 21-day running average (starting from day $N - 10$ to day $N + 10$) of the series created in the first step. WCI does not include direct information of wind speed, on which the moisture transport is very dependent. However, for the selected square, we found that the wind speed changes associated with the arrival of the monsoon are directly related to the changes in wind direction. Figure 2 compares the annual cycle of the WCI with the average wind speed based on the wind field at the sigma level 0.995 from the NCEP–NCAR reanalysis database (Kalnay et al. 1996) averaged over the same area used to compute the WCI. The wind speed shows two clear maxima in January and July. The first one is related to the northeasterly flow characteristic of the winter monsoon, and the second and stronger one is related to the southwesterly summer monsoon. Interestingly, from mid-February to late April the wind velocity slowly decreases to reach the annual minimum during the first days of May. After this minimum, the wind speed experiences a very rapid increase lasting until late June, the period when MOK occurs. Interestingly, when the seasonal evolution of the WCI is displayed (Fig. 2, black line), a single minimum is evidenced in January related to the absence of wind with a WSW component during the winter, followed by a slow increase lasting until late April. Remarkably, the rapid increase in the frequency of the WSW winds during May and June closely mimics that of the wind speed. This result suggests that the frequency of WSW winds could be used to define the MOK date from the WCI series. It is worth noting that the evolution of the WCI around its maximum in July is asymmetric, which is evident by its slower decrease from early September to early November. This is consistent with the monsoon retreat being slower than the onset, a characteristic that has been widely documented (Fasullo and Webster 2003) and is also evident in the evolution of the average wind speed (Fig. 2). Therefore, while the strong trend of the WCI during the monsoon establishment allows for a precise determination of the MOK, this was not the case for the withdrawal, because of the very gradual decrease of the WCI, resulting in uncertain withdrawal dates.

As the official IMD series has traditionally been the reference for most subsequent MOK definitions (e.g., Raju et al. 2007; Ramesh Kumar et al. 2009; Puranik et al. 2013, Bhatla et al. 2016), we used the official IMD series as a calibration to select optimal thresholds in the WCI aimed to identify the monsoon starting date. The maximum correlation between the MOK series produced by the IMD and those based on the WCI was obtained when the date of the WCI-based MOK was defined as the first of seven consecutive days for which the WCI over the selected domain is higher than 33%. This particular percentage of the WCI corresponds to climatological winds speed around $6.6\,\mathrm{m\,s^{-1}}$ (see Fig. 2), a value quite similar to those used in MOK detection schemes based on the wind speed (see, e.g., Wang et al. 2009), suggesting that the 33% value is a reasonable choice. A sensitivity test to changes in this value and the requirement of having at least five wind direction observations in the selected area to compute the WCI were carried out, confirming that the correlations of the WCI-based MOK and the IMD MOK remained significant and stable (not shown), confirming the robustness of the definition.

3. Results

Table 1 and Fig. 3 show the resulting MOK dates. Our approach indicates that 28 May is the long-term mean MOK date, with standard deviation of 7.3 days, which is very similar to other methodologies. The interannual

TABLE 1. Monsoon onset dates based on the WCI.

Monsoon onset dates				
30 May 1887	29 May 1922	21 May 1949	23 May 1970	4 Jun 1992
20 May 1888	4 Jun 1923	24 May 1950	13 Jun 1972	31 May 1993
26 May 1889	31 May 1924	27 May 1951	26 May 1973	26 May 1994
3 Jun 1890	15 May 1925	19 May 1952	22 May 1974	3 Jun 1995
28 May 1902	29 May 1926	7 Jun 1953	29 May 1975	3 Jun 1996
5 Jun 1903	21 May 1927	25 May 1954	25 May 1976	9 Jun 1997
31 May 1904	29 May 1928	18 May 1955	25 May 1977	30 May 1998
2 Jun 1905	24 May 1929	11 May 1956	19 May 1978	12 May 1999
1 Jun 1906	25 May 1930	30 May 1957	9 Jun 1979	14 May 2000
24 May 1907	30 May 1931	4 Jun 1958	28 May 1980	18 May 2001
29 May 1908	2 Jun 1932	27 May 1959	25 May 1981	23 May 2002
27 May 1909	17 May 1933	18 May 1960	3 Jun 1982	5 Jun 2003
1 Jun 1910	7 Jun 1934	17 May 1961	8 Jun 1983	13 May 2004
28 May 1911	3 Jun 1935	26 May 1962	28 May 1984	8 Jun 2005
4 Jun 1912	22 May 1936	29 May 1963	22 May 1985	21 May 2006
27 May 1913	8 Jun 1937	31 May 1964	6 Jun 1986	9 Jun 2007
22 May 1914	18 May 1938	27 May 1965	2 Jun 1987	25 May 2008
15 Jun 1915	29 May 1939	2 Jun 1966	29 May 1988	31 May 2010
26 May 1916	18 May 1946	4 Jun 1967	24 May 1989	26 May 2011
6 Jun 1919	28 May 1947	2 Jun 1968	15 May 1990	11 Jun 2012
29 May 1920	25 May 1948	25 May 1969	31 May 1991	22 May 2013
28 May 1921				

variability of the MOK date oscillates for a wide range of 46 days between the earliest day (11 May 1956) and the latest day (25 June 1915). The 15-yr moving average shows some interdecadal fluctuations. In general, the first and the last third of the twentieth century were characterized by later MOK dates, while the period between 1940 and 1965 was mostly dominated by early onsets.

Changes in large-scale circulation related to our definition of MOK have been analyzed by computing the differences in the horizontal wind of the NCEP–NCAR reanalysis for the period 7 days after the monsoon onset minus the period 7 days before it at both the sigma level of 0.995 and 200-hPa level from 1948 to 2013 (Fig. 4). The observed results are consistent with those expected for an average MOK. The wind pattern at the near-surface level (Fig. 4a) confirms that the cross-equatorial low-level jet is strongly reinforced around the MOK, with a clear increase of the monsoon westerlies over the southern Arabian Sea. At upper levels, Fig. 4b shows evidence that the tropical easterly jet stream is also strongly reinforced during the MOK in response to the troposphere heat source generated over South Asia by the low-level monsoon current.

To test whether our MOK adequately reflects the expected changes in precipitation over Kerala around the new onset, we used precipitation data from the Asian Precipitation–Highly-Resolved Observational Data Integration Toward Evaluation of Water Resources at 0.5° × 0.5° resolution from 1951 to 2007 (Yatagai et al. 2012). This database is considered to be a reliable representation of the rainfall along the Western Ghats in India because it uses a dense gauge network (Yatagai et al. 2012). Figure 5 shows the area-averaged daily rainfall for Kerala averaged over the period 1951–2007 for each day of a 21-day period from day −10 (10 days before the onset) to day 10 (10 days

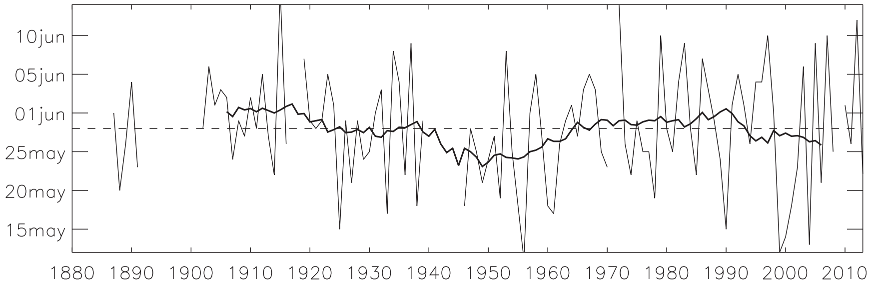

FIG. 3. MOK series (gray line) and 15-yr running average (black line).

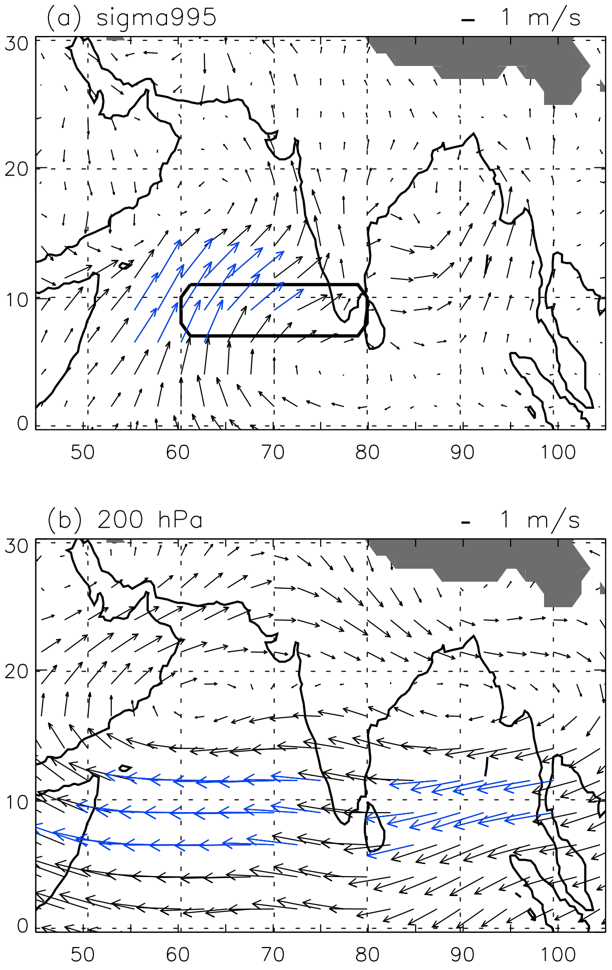

FIG. 4. Composite of the differences between the horizontal wind for one week after the onset minus the horizontal wind one week before the onset at (a) the sigma level 0.995 and (b) the 200-hPa level based on the NCEP–NCAR reanalysis from 1948 to 2013. Blue arrows indicate statistically significant correlation at $p < 0.05$. Gray shaded areas represent altitudes above 2000 m (the Himalayas).

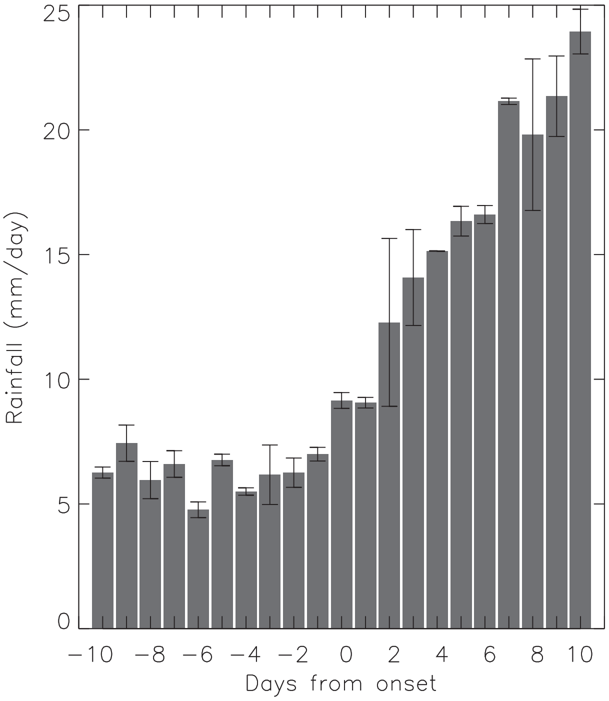

FIG. 5. Mean daily rainfall and standard deviation (error bars) for 1951–2007 over Kerala (mm day^{-1}) with respect to the MOK (day 0).

after the onset), where day 0 is the onset date for each year as estimated by our methodology. The intense and sustained transition in the rainfall regime signaled by the MOK is clearly seen. This transition is characterized by an increase in the precipitation over the Kerala coast from 5 mm day^{-1} to more than 20 mm day^{-1} in approximately 7 days from the onset date (day 0). This result suggests that on average, the MOK computed from the WCI series adequately represents the start of the monsoon in terms of rainfall amounts. It is worth mentioning that the increase in precipitation from 5 to 20 mm day^{-1} is a typical characteristic of the MOK as previously described by Fasullo and Webster (2003).

The composition of meteorological fields around the onset date as discussed in the previous paragraphs offers a first insight into the representativeness of the methodology. However, as stated in the introduction, one of the most common problems of objective methods developed to track the MOK is the large year-to-year variability in the specific type of onset, which can be abrupt, gradual, or double onset (Fieux and Stommel 1977). This fact implies that while most methods give quite similar results for most MOK dates, in some years discrepancies can be large. To test the performance or our new methodology we examined the joint evolution of WCI and precipitation for every year in the available series. Figure 6 shows a summary of the results. The threshold to declare the MOK is represented by a discontinuous horizontal line in each graph. Figures 6a and 6b illustrate the situation typical for most years (75% of cases). Onsets characterized by an abrupt increase in the precipitation over Kerala (Fig. 6a) are easily identified by the WCI, which is typically the 33% WCI threshold that is concurrent with the sudden precipitation increase. In cases of a slower progression of the onset, as shown in Fig. 6b, the threshold can be reached a few days prior to the widespread development of the precipitation in Kerala. Figure 6c shows the case in which the WCI evolution is most strongly disconnected from the rain development over Kerala. In this case, corresponding to the year 2000, the WCI reaches the critical value on 14 May, but there is

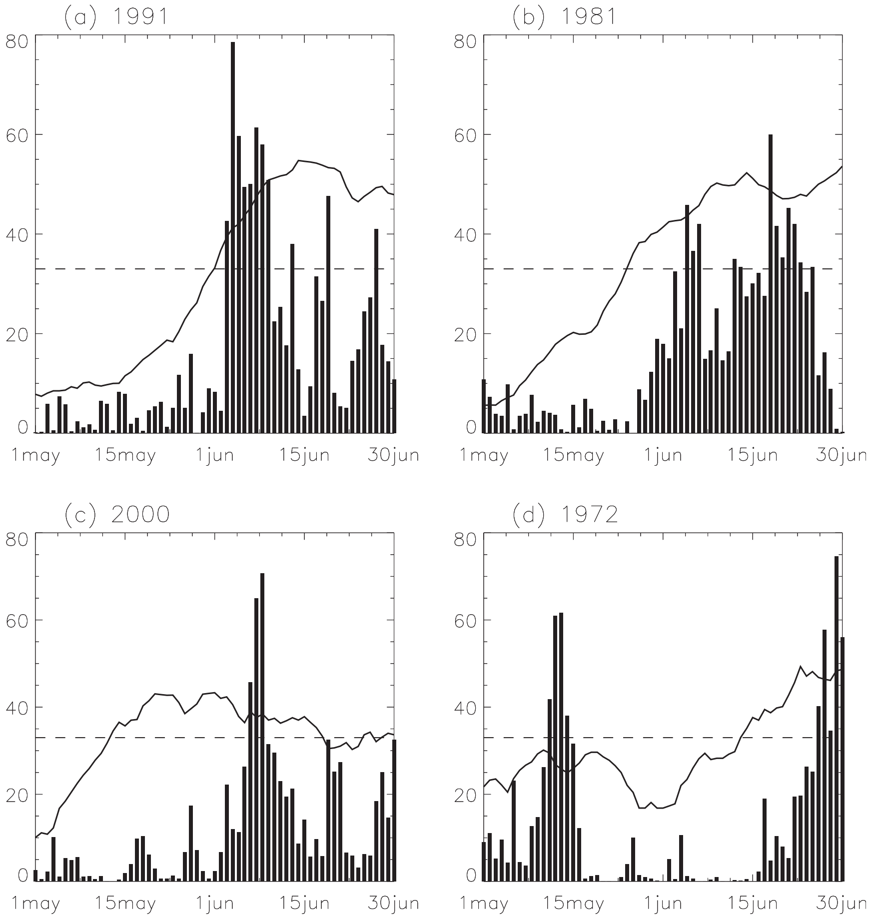

FIG. 6. Area-averaged daily rainfall (mm day^{-1}) over Kerala during May–June (bar plots) for (a) 1991, a year of abrupt monsoon onset; (b) 1981, a year of gradual monsoon onset; (c) 2000, the year of the largest discrepancies between large-scale circulation and rainfall behavior; and (d) 1972, a year with bogus onset. The solid line shows the value of WCI (percent of wind from the WSW), and the dashed horizontal line represents the 33% threshold of WCI to declare the MOK.

very little rain over Kerala until early June. The lack of coupling between the wind field and the rainfall for this year affects every onset index based on the wind field. The anomalous year 2000 was studied in detail by Wang et al. (2009), and they found that the strong low-level wind established around 15 May southwest of Kerala (which triggered the detection of an onset in our approach) was primarily driven by convection over the southeastern Bay of Bengal rather than by the convection over the southern tip of India. The precipitation did not develop inland until two weeks after the event. The last case shown in Fig. 6d (1972) corresponds to a bogus onset. A bogus onset is characterized by the existence of enhanced convection over land and the presence of westerly surface winds similar to those occurring during a typical onset but lasting a week or less.

The Madden–Julian oscillation (MJO) has been suggested to trigger the monsoon onset (Bhatla et al. 2016) and its timing to develop bogus onsets (Flatau et al. 2001). In this sense, Flatau et al. (2001) reported the existence of six cases for the years 1967, 1972, 1979, 1986, 1995, and 1997. Our methodology seems to be robust against five of these cases, as in the example shown in Fig. 6d. During a bogus onset, the WCI tends to increase but never exceeds the 33% threshold that is reached by 13 June. One exception is the year 1997, when the bogus onset and the real onset occurred within a short time difference and the WCI exceeds the threshold in the middle of the bogus onset.

a. Comparison with other onset series

There are several ISM onset series available with which to compare. As described in the introduction, the "official" but in some cases bogus MOK dates until 2005 were provided by the IMD and the corresponding previous

TABLE 2. Source of information and available dates of ISM onset derived by different methods.

ISM onset index	Data source	Dates available
IMDold	Joseph et al. (1994)	1901–2005
IMDup	Pai and Rajeevan (2009)	1971–present
WA09	Wang et al. (2009)	1948–2007
JO06	Joseph et al. (2006)	1971–2003
FW03	Fasullo and Webster (2003)	1948–2000
XA07	Xavier et al. (2007)	1950–2003

TABLE 3. Linear correlation between monsoon onset dates by different methods. The correlation coefficients are computed using data for the common period 1971–2000. The correlation coefficients are all significant at $p < 0.01$.

	This work	IMDold	IMDup	WA09	JO06	FW03	XA07
This work	1	0.81	0.85	0.89	0.75	0.56	0.78
IMDold		1	0.91	0.79	0.84	0.59	0.68
IMDup			1	0.79	0.90	0.56	0.59
WA09				1	0.81	0.69	0.84
JO06					1	0.61	0.60
FW03						1	0.59
XA07							1

IMD series (IMDold) is available from 1901 to 2005. From 2006 forward, the IMD modified the methodology to declare the MOK. The IMDold onsets have been reworked starting from 1971 using the new criteria (Pai and Rajeevan 2009), resulting in an updated MOK chronology (IMDup). We also compared our results with the MOK series of Joseph et al. (2006) and Wang et al. (2009) and with the ISM onset of Fasullo and Webster (2003) and Xavier et al. (2007). The available periods for each case and the acronyms used in the following discussion are given in Table 2.

We first compared the climatological mean onset dates as derived by different methods for their common period of availability (1971–2000). Our mean onset for this period is 29 May, a few days earlier than most of the other methods except for XA07 (which occurs on 29 May as well). The average onsets for the IMDold, IMDup, WA09, and JO06 series are detected between 1 and 3 June, while for the FW03 the mean onset is 6 June. The small advance in our average detection is related to the tendency of the WCI to rise faster than the precipitation in the cases of a slow and gradual increase in convection over southern India (the case shown in Fig. 6b).

The correlation coefficients for the period 1971–2000, common to all indices, are shown in Table 3. The IMDup onset has the highest correlations with IMDold ($r = +0.91$, $p < 0.01$) and JO06 ($r = +0.90$, $p < 0.01$) onsets. This is not surprising because the new objective method of the IMD shares common criteria both with the IMDold onset definition (rainfall over the Kerala state) and the JO06 definition (OLR and wind). For the same period, the correlation coefficient between the IMDup and our definition is quite high, reaching $r = +0.85$ ($p < 0.01$). With the exception of the FW03 ($r = +0.56$, $p < 0.01$) onset, the correlation coefficients of all onset series and our MOK dates are greater or equal than 0.75. In this sense, it must be pointed out that the FW03 index is not fully representative of the MOK because it was defined to represent all of India, which explains why this index shows relatively low correlations with all the other onset series. It is worth mentioning that the correlation between the IMDold and our series along their entire common period (1902–2005) is also high, reaching $r = +0.76$ ($p < 0.01$, value not shown in Table 3).

b. Long-term relation of MOK with ENSO and Pacific decadal oscillation

It is widely recognized that the precipitation during the Indian summer monsoon is modulated by the ENSO cycle (Fasullo and Webster 2003; Goswami and Xavier 2005; Xavier et al. 2007; Ju and Slingo 1995). However, there is an active debate concerning the origin and stability of this relationship (Li and Ting 2015). While some studies suggest a weakening of the ENSO–monsoon relationship and relate it to recent global warming (Kumar et al. 1999; Kinter et al. 2002), other authors suggest that the ENSO–monsoon relationship could be characterized by decadal fluctuations mostly due to internal climate variability (Krishnamurthy and Goswami 2000; Robinson et al. 2008; Li and Ting 2015). It is accepted that the ENSO modulation of the total monsoon rainfall essentially occurs through displacements of the ascending and descending branches of the Walker circulation linking the Indo-Pacific regions (Ropelewski and Halpert 1987; Kumar et al. 2006). However, ENSO also impacts the onset date with late (early) onsets typically associated with El Niño (La Niña) conditions. In this regard, Goswami and Xavier (2005) showed that a large part of the monsoonal variability associated with the ENSO cycle was as a result of changes in the length of the rainy season and thus the MOK date. Xavier et al. (2007) demonstrated that the physical link between ENSO and the monsoonal rains over India could be explained in terms of ENSO impacts on tropospheric temperatures over Eurasia, which would consequently modify the onset date. In a recent work, Adamson and Nash (2013) analyzed the ENSO–monsoon onset long-term variability between 1781 and 2011 by merging documentary evidence with official IMD data, finding a relatively stable relationship. However, these authors point out that the onset chronologies based on rainfall probably tended to

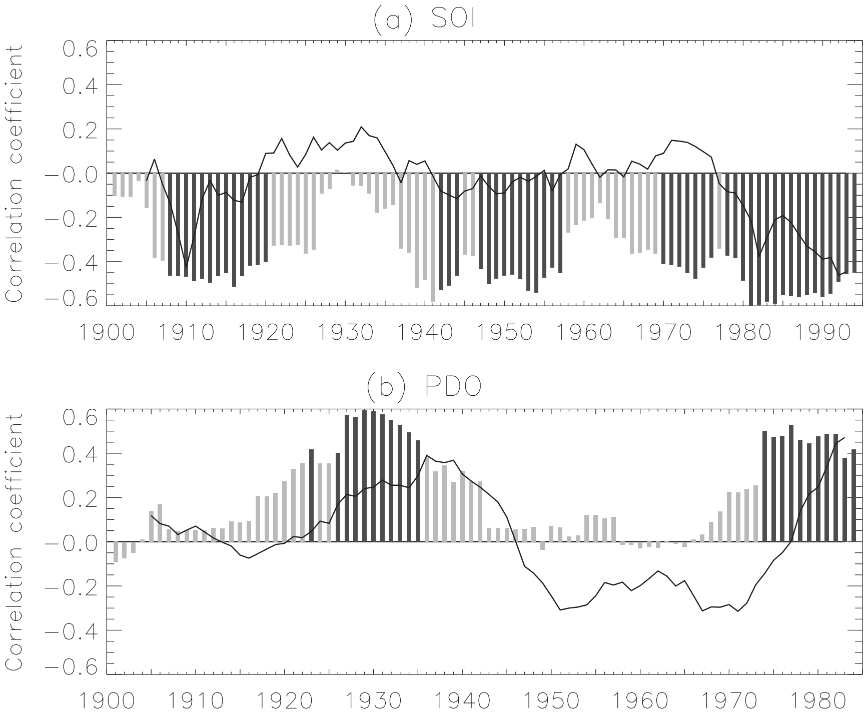

FIG. 7. Correlation coefficients between the monsoon onset and the 3-month (March–May) averaged (a) SOI in a 20-yr moving window and (b) PDO in a 30-yr moving window. Horizontal axis represents the starting year of each moving period. The PDO and SOI have been smoothed to highlight the interdecadal variability component; however, the analysis uses the unsmoothed time series. The smoothed PDO and SOI series have been multiplied by 0.4 and 0.6, respectively, to facilitate comparison. Dark bars indicate that the correlation is statistically significant at the 90% confidence level.

misrepresent the ENSO–monsoon onset relationship and recommended the generation of new objective methodologies based on dynamic variables to track the MOK before the 1950s. In this sense, the new MOK definition developed by this work offers an interesting opportunity to analyze the variability of the ENSO–Indian monsoon relationship using a longer and objective MOK series. To define the ENSO phase during the season prior to the MOK, we used the March–May SOI averages (Ropelewski and Jones 1987). Running correlations between the SOI and the MOK for windows of 20, 25, and 30 yr were computed. Figure 7a shows the case for the 20-yr window. The sign of the correlation is stable and negative during the entire twentieth century, indicating—as found by previous studies—that late MOK dates tends to occur under negative SOI (El Niño) conditions during the three months prior to the MOK. However, for this 20-yr window, the correlation reaches statistical significance ($p < 0.1$) only in the periods 1905–20, 1940–57, and 1970–2000. During the periods 1921–39 and 1958–69, the link between the SOI and MOK was weaker. Interestingly, when the 20-yr running average of the SOI is displayed (black line in Fig. 7a) it is evident that during periods of significant (nonsignificant) SOI–MOK correlations the SOI tended to be predominantly in a negative (positive) phase. This result suggests that the link between the monsoon onset in India and the ENSO cycle could be modulated by the decadal-scale fluctuations of the ENSO cycle and supports the idea that during the twentieth century natural variability dominated the ENSO–monsoon onset relationship (Li and Ting 2015). Equivalent analyses for the 25- and 30-yr windows (not shown) lead to the same conclusions.

The relationship between the Indian monsoon and the Pacific decadal oscillation (PDO; Mantua et al. 1997; Mantua and Hare 2002) has also been widely recognized. In this sense, Krishnamurthy and Krishnamurthy (2014) explored the relationship between the strength of the Indian monsoon and the PDO using both observations and model data. They found a consistent association between a warm (cold) PDO phase and a deficit (excess) of precipitation, but their study did not analyze possible changes in MOK. Recently, Watanabe and Yamazaki

(2014) explored the possibility of a consequent change in the timing of the Indian monsoon onset, finding significant changes in MOK during the transition from negative to positive PDO phases. Although they found a consistent propagation mechanism based on the differential wave train propagation, they recognized that their study relied on the analysis of only one complete PDO cycle and recommended a reevaluation of their findings with longer-term data. Following their methodology, we used a PDO index from the Joint Institute for the Study of the Atmosphere and Ocean website (http://jisao.washington.edu/pdo/) and computed the 30-yr running correlation coefficients between the PDO and our monsoon onset (Fig. 7b). We also display the 30-yr running average of the PDO index to highlight its interdecadal component. We found the same increase in the correlation as reported by Watanabe and Yamazaki (2014) during the early 1970s, with values reaching statistical significance by 1975. This increase is concurrent with the strong upward trend of the 30-yr filtered PDO. Our MOK series allows for the study of the previous period of upward PDO trend (1915–35), and we found the same increase in PDO–MOK correlation, reaching significant values by the mid-1920s and peaking in 1930 with $r = 0.59$ ($p < 0.1$). A decrease in the correlation values is found during phases of negative PDO trend (1936–50), and nonsignificant correlation during neutral (1900–1914) or negative (1951–70) PDO phases has been found.

Both analyses were also performed using the IMDold series and very similar results were obtained for both ENSO and PDO. These similarities are particularly interesting in the case of the PDO since they indicate that our newly developed series and the IMDold series share the interdecadal signal with this global climate mode.

4. Summary and remarks

While the IMD has determined the date of ISM onset every year for more than 100 years, its methodology is considered subjective, and during the last decade great effort has been devoted to the generation of objective monsoon onset series to better characterize the interannual and multidecadal variability of this essential part of the climatic system. Apart from the MOK series based on observed precipitation as that used by IMD, currently most of the objective ISM onset definitions are computed from reanalyzed databases. Fasullo and Webster (2003), Joseph et al. (2006), Xavier et al. (2007), and Wang et al. (2009) produced their onset series using daily data from the NCEP–NCAR reanalysis, and, consequently, they are limited to the second half of the twentieth century. To our knowledge, there have been no attempts to use the Twentieth Century Reanalysis data (Compo et al. 2011) to compute longer monsoon onset series. Although the current version of this database (version 2c) starts in 1851 and contains subdaily zonal and meridional winds at global scale, it is produced from sea level pressure observations alone. Thus, subtle details of the wind field as those necessary to define the MOK would presumably be rather uncertain in this database during the first half of the twentieth century. On the contrary, ICOADS release 2.5 currently contains thousands of direct observations of wind data since the last part of the nineteenth century. In this paper we show that the persistence of WSW winds over 7°–11°N, 60°–80°E measured using only the wind direction field can be used to define a meaningful MOK series. This new MOK series significantly correlates with previous definitions in their common periods. Unfortunately, ICOADS has some important gaps in this area during the twentieth century (mostly because of the absence of data during periods of war), and the coverage before 1880 is essentially negligible. Despite these caveats, the analysis of our MOK series can be summarized as follows:

1) The wind and precipitation fields before and after our MOK date well represent the changes in the lower and upper troposphere typically associated with the monsoon onset. A rapid increase in precipitation from an average of about 5 mm day^{-1} to over 20 mm day^{-1} in a 7-day period after the onset is evident. The correlation among our onset and those of previous works is high and significant. Onset dates exhibit a high interannual variability with a difference of 46 days between the earliest and the latest onset dates of our study period.
2) Abrupt and gradual onsets are properly determined. Our methodology seems rather robust against "bogus" onset detections. However, we found a few cases with discrepancies between the onset found by our method and the corresponding increase in precipitation. Notwithstanding, during the particular year with the greatest discrepancy (year 2000) the atmospheric behavior is considered anomalous (as compared with the rest of the onsets of the twentieth century; Wang et al. 2009). It would be interesting to study the possibility of more of such events occurring in the future to evaluate the possible impact on all the onset detection algorithms currently used.
3) In the complete twentieth century, the evolution of MOK dates does not exhibit significant trends. However, we found evidences of later-than-average onset during the 1900–25 and 1970–90 periods and earlier-than-average onset between 1940 and 1965. It is

interesting to compare these results with those of Krishnakumar et al. (2009), who computed the rainfall trends over Kerala separately for the premonsoon season (defined as March–May) and the southwest monsoon season (defined as June–September). They did not find evidence of periodicities in the June–September precipitation (a period typically used to compute monsoonal precipitation trends for India), but for the March–May precipitation they found lower-than-average values during 1900–25 and 1970–2000, roughly the periods that we have characterized as late onsets (after 1 June). On the contrary, the March–May precipitation in Kerala was above its long-term average value between 1925 and 1970, a period largely characterized by early onsets (before 1 June). This suggests that what it is traditionally interpreted as premonsoon season can include precipitation events corresponding to the monsoon itself (i.e., after the onset). In this sense Goswami and Xavier (2005) point out that for the study of the monsoon–ENSO relationship, the noninclusion of rainfall events outside the traditional definition of the South Asian monsoon season (June–September) may bias the interpretation of the results, and they recommended defining the rainfall corresponding to a monsoon season as the total precipitation between the onset and the withdrawal of the monsoon rather than that received between fixed dates. Our results strongly support this conclusion and extend it to the evaluation of the precipitation trends in India.

4) The correlation between the ENSO cycle as measured by the SOI and the MOK has been negative during the entire twentieth century; however, the strength of this relationship is not stable and two periods of nonsignificant correlation have been found for 1920–40 and 1960–73. These periods were characterized by the prevalent positive SOI phase (La Niña conditions).
5) The recently described relationship between the PDO and MOK limited to the phases characterized by the shift from negative to positive PDO phases has been confirmed for two complete PDO cycles. Watanabe and Yamazaki (2014) postulated that a stationary wave train plays an active role in connecting both the North Pacific Ocean and South Asian monsoon region. Because of the limitation of their study to the period 1948–2011, they suggested that this link should be further addressed by using longer-term data. Our results support their findings.

Finally, we would like to stress the fact that the aim of this research is not to provide a definition of the MOK better or more reliable than those based on the method currently operational at the IMD. Our main objective was to evaluate the feasibility of using historical wind measurements to build an instrument-based MOK series capable of being extended to the past, when no other instrument-based observations are available in the area affected by the ISM. This capability of being extended to the past is grounded in the absence of the wind force in our definition, a field usually problematic for historical observations. With our methodology we have been able to develop an MOK series starting in 1887 by using the data stored in the ICOADS release 2.5 database. In this sense, according to Wheeler and García-Herrera (2008), a large amount of wind observations for periods prior to the twentieth century remains unexplored in thousands of logbooks preserved in several British archives. It is also possible that data corresponding to periods of war are available in military archives not yet made public. Because of the monsoon's impact, it is difficult to exaggerate the interest to unveil these unexplored data to track the MOK by means of an instrument-based, homogenous, and objective methodology for more than 200 years. No doubt, it would largely justify the time and economic costs of its digitization.

Acknowledgments. This work is a contribution to the project INCITE (CGL2013-44530-P) and to the research group RNM-356 belonging to the Plan Andaluz de Investigación Desarrollo e Innovación.

REFERENCES

Adamson, G. C. D., and D. J. Nash, 2013: Long-term variability in the date of monsoon onset over western India. *Climate Dyn.*, **40**, 2589–2603, doi:10.1007/s00382-012-1494-x.

Barriopedro, D., D. Gallego, M. C. Álvarez-Castro, R. García-Herrera, D. Wheeler, C. Peña-Ortiz, and S. M. Barbosa, 2014: Witnessing North Atlantic westerlies variability from ships' logbooks (1685–2008). *Climate Dyn.*, **43**, 939–955, doi:10.1007/s00382-013-1957-8.

Bhatla, R., M. Singh, and D. R. Pattanaik, 2016: Impact of Madden-Julian oscillation on onset of summer monsoon over India. *Theor. Appl. Climatol.*, doi:10.1007/s00704-015-1715-4, in press.

Carvalho, L. M. V., C. Jones, F. Cannon, and J. Norris, 2016: Intraseasonal-to-interannual variability of the Indian monsoon identified with the large-scale index for the Indian monsoon system (LIMS). *J. Climate*, **29**, 2941–2962, doi:10.1175/JCLI-D-15-0423.1.

Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, doi:10.1002/qj.776.

Fasullo, J., and P. J. Webster, 2003: A hydrological definition of Indian monsoon onset and withdrawal. *J. Climate*, **16**, 3200–3211, doi:10.1175/1520-0442(2003)016<3200a:AHDOIM>2.0.CO;2.

Fieux, M., and H. Stommel, 1977: Onset of the southwest monsoon over the Arabian Sea. *Mon. Wea. Rev.*, **105**, 231–236, doi:10.1175/1520-0493(1977)105<0231:OOTSMO>2.0.CO;2.

Flatau, M. K., P. J. Flatau, and D. Rudnick, 2001: The dynamics of double monsoon onsets. *J. Climate*, **14**, 4130–4146, doi:10.1175/1520-0442(2001)014<4130:TDODMO>2.0.CO;2.

Gallego, D., R. García-Herrera, N. Calvo, and P. Ribera, 2007: A new meteorological record for Cádiz (Spain) 1806–1852: Implications for climatic reconstructions. *J. Geophys. Res.*, **112**, D12108, doi:10.1029/2007JD008517.

——, P. Ordoñez, P. Ribera, C. Peña-Ortiz, and R. García-Herrera, 2015: An instrumental index of the West African Monsoon back to the nineteenth century. *Quart. J. Roy. Meteor. Soc.*, **141**, 3166–3176, doi:10.1002/qj.2601.

García-Herrera, R., G. P. Können, D. Wheeler, M. R. Prieto, P. D. Jones, and F. B. Koek, 2005: CLIWOC: A climatological database for the world's oceans 1750–1854. *Climatic Change*, **73**, 1–12, doi:10.1007/s10584-005-6952-6.

Goswami, B. N., and P. R. Xavier, 2005: ENSO control on the south Asian monsoon through the length of the rainy season. *Geophys. Res. Lett.*, **32**, L18717, doi:10.1029/2005GL023216.

Joseph, P. V., 2012: Onset, advance and withdrawal of monsoon. *Monsoon Monograph*, Vol. 1, A. Tyagi et al., Eds., India Meteorological Department, Ministry of Earth Sciences, Government of India, 284–329.

——, J. Eisheid, and R. J. Pyle, 1994: Interannual variability of the onset of the Indian summer monsoon and its association with atmospheric features, El Niño, and sea surface temperature anomalies. *J. Climate*, **7**, 81–105, doi:10.1175/1520-0442(1994)007<0081:IVOTOO>2.0.CO;2.

——, K. P. Sooraj, and C. K. Rajan, 2006: The summer monsoon onset process over South Asia and an objective method for the date of monsoon onset over Kerala. *Int. J. Climatol.*, **26**, 1871–1893, doi:10.1002/joc.1340.

Ju, J., and J. Slingo, 1995: The Asian summer monsoon and ENSO. *Quart. J. Roy. Meteor. Soc.*, **121**, 1133–1168, doi:10.1002/qj.49712152509.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kinter, J. L., K. Miyakoda, and S. Yang, 2002: Recent changes in the connection from the Asian monsoon to ENSO. *J. Climate*, **15**, 1203–1214, doi:10.1175/1520-0442(2002)015<1203:RCITCF>2.0.CO;2.

Koteswaram, P., 1958: The easterly jet stream in the tropics. *Tellus*, **10A**, 43–57, doi:10.1111/j.2153-3490.1958.tb01984.x.

Krishnakumar, K. N., G. S. L. H. V. Prasada Rao, and C. S. Gopakumar, 2009: Rainfall trends in twentieth century over Kerala, India. *Atmos. Environ.*, **43**, 1940–1944, doi:10.1016/j.atmosenv.2008.12.053.

Krishnamurthy, L., and V. Krishnamurthy, 2014: Influence of PDO on South Asian summer monsoon and monsoon–ENSO relation. *Climate Dyn.*, **42**, 2397–2410, doi:10.1007/s00382-013-1856-z.

Krishnamurthy, V., and B. N. Goswami, 2000: Indian monsoon–ENSO relationship on interdecadal timescale. *J. Climate*, **13**, 579–595, doi:10.1175/1520-0442(2000)013<0579:IMEROI>2.0.CO;2.

Kumar, K. K., B. Rajagopalan, and M. A. Cane, 1999: On the weakening relationship between the Indian monsoon and ENSO. *Science*, **284**, 2156–2159, doi:10.1126/science.284.5423.2156.

——, ——, M. Hoerling, G. Bates, and M. Cane, 2006: Unraveling the mystery of Indian monsoon failure during El Niño. *Science*, **314**, 115–119, doi:10.1126/science.1131152.

Li, X., and M. Ting, 2015: Recent and future changes in the Asian monsoon-ENSO relationship: Natural or forced? *Geophys. Res. Lett.*, **42**, 3502–3512, doi:10.1002/2015GL063557.

Mantua, N. J., and S. R. Hare, 2002: The Pacific decadal oscillation. *J. Oceanogr.*, **58**, 35–44, doi:10.1023/A:1015820616384.

——, ——, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1079, doi:10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2.

Pai, D. S., and M. N. Rajeevan, 2009: Summer monsoon onset over Kerala: New definition and prediction. *J. Earth Syst. Sci.*, **118**, 123–135, doi:10.1007/s12040-009-0020-y.

Puranik, S. S., K. C. Sinha Ray, P. N. Sen, and P. Pradeep Kumar, 2013: An index for predicting the onset of monsoon over Kerala. *Curr. Sci.*, **105**, 954–961.

Raju, P. V. S., U. C. Mohanty, and R. Bhatla, 2007: Interannual variability of onset of the summer monsoon over India and its prediction. *Nat. Hazards*, **42**, 287–300, doi:10.1007/s11069-006-9089-7.

Ramesh Kumar, M. R., S. Sankar, and C. Reason, 2009: An investigation into the conditions leading to monsoon onset over Kerala. *Theor. Appl. Climatol.*, **95**, 69–82, doi:10.1007/s00704-008-0376-y.

Rao, Y. P., 1976: *Southwest Monsoon. Synoptic Meteor. Monogr.*, Vol. 1/1976, India Meteorological Department, Government of India, 367 pp.

Ribera, P., D. Gallego, C. Pena-Ortiz, L. Del Rio, T. A. Plomaritis, and J. Benavente, 2011: Reconstruction of Atlantic historical winter coastal storms in the Spanish coasts of the Gulf of Cadiz, 1929–2005. *Nat. Hazards Earth Syst. Sci.*, **11**, 1715–1722, doi:10.5194/nhess-11-1715-2011.

Robinson, L. F., V. H. De la Peña, and Y. Kushnir, 2008: Detecting shifts in correlation and variability with application to ENSO-monsoon rainfall relationships. *Theor. Appl. Climatol.*, **94**, 215–224, doi:10.1007/s00704-007-0351-z.

Ropelewski, C. F., and M. S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606–1626, doi:10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2.

——, and P. D. Jones, 1987: An extension of the Tahiti–Darwin Southern Oscillation index. *Mon. Wea. Rev.*, **115**, 2161–2165, doi:10.1175/1520-0493(1987)115<2161:AEOTTS>2.0.CO;2.

Soman, M. K., and K. Krishna Kumar, 1993: Space-time evolution of meteorological features associated with the onset of Indian summer monsoon. *Mon. Wea. Rev.*, **121**, 1177–1194, doi:10.1175/1520-0493(1993)121<1177:STEOMF>2.0.CO;2.

Taniguchi, K., and T. Koike, 2006: Comparison of definitions of Indian summer monsoon onset: Better representation of rapid transitions of atmospheric conditions. *Geophys. Res. Lett.*, **33**, L02709, doi:10.1029/2005GL024526.

Wang, B., Q. Ding, and P. V. Joseph, 2009: Objective definition of the Indian summer monsoon onset. *J. Climate*, **22**, 3303–3316, doi:10.1175/2008JCLI2675.1.

Watanabe, T., and K. Yamazaki, 2014: Decadal-scale variation of South Asian summer monsoon onset and its relationship with the Pacific decadal oscillation. *J. Climate*, **27**, 5163–5173, doi:10.1175/JCLI-D-13-00541.1.

Webster, P. J., V. O. Magana, T. N. Palmer, J. Shukla, R. A. Tomas, M. Yanai, and T. Yasunari, 1998: Monsoons: Processes, predictability and the prospects for prediction. *J. Geophys. Res.*, **103**, 14 451–14 510, doi:10.1029/97JC02719.

Wheeler, D., and R. García-Herrera, 2008: Ships' logbooks in climatological research. *Ann. N. Y. Acad. Sci.*, **1146**, 1–15, doi:10.1196/annals.1446.006.

Woodruff, S. D., and Coauthors, 2011: ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951–967, doi:10.1002/joc.2103.

Xavier, P. K., C. Marzin, and B. N. Goswami, 2007: An objective definition of the Indian summer monsoon season and a new perspective on the ENSO–monsoon relationship. *Quart. J. Roy. Meteor. Soc.*, **133**, 749–764, doi:10.1002/qj.45.

Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, 2012: APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull. Amer. Meteor. Soc.*, **93**, 1401–1415, doi:10.1175/BAMS-D-11-00122.1.

Yin, M. T., 1949: A synoptic-aerologic study of the onset of the summer monsoon over India and Burma. *J. Meteor.*, **6**, 393–400, doi:10.1175/1520-0469(1949)006<0393:SASOTO>2.0.CO;2.