Subseasonal Features of the Indian Monsoon

LUN DAI,a,b TAT FAN CHENG,a BIN WANG,b AND MENGQIAN LUa

a Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, China
b Department of Atmospheric Sciences and International Pacific Research Center, University of Hawaii at Manoa, Honolulu, Hawaii

ABSTRACT: The Indian monsoon is of utmost concern to agriculture, the economy, and the livelihoods of billions in South Asia. However, little attention has been paid to the possibility of distinct subseasonal episodes phase-locked in the Indian monsoon annual cycle. This study addresses this gap by utilizing the self-organizing map (SOM) method to objectively classify six distinct subseasonal stages based on the 850-hPa wind fields. Each subseasonal stage ranges from 23 to 90 days. The Indian summer monsoon (ISM) consists of three substages, the ISM-onset, ISM-peak, and ISM-withdrawal, altogether contributing to 82% of the annual precipitation. The three substages signify the rapid northward advance, dominance, and gradual southward retreat of southwesternlies from mid-May to early October. The winter monsoon also comprises three substages (fall, winter, and spring), distinguishable by the latitude of the Arabian Sea high pressure ridge and hydrological conditions. This study proposes two compact indices based on zonal winds in the northern and southern Arabian Sea to measure the winter and summer monsoons, respectively. These indices capture the development and turnabouts of the six SOM-derived stages and can be used for subseasonal monsoon monitoring and forecasts. The spring and the ISM-onset episodes are highly susceptible to compound hazards of droughts and heatwaves, while the greatest flood risk occurs during the ISM-peak stage. The fall stage heralds the peak season for tropical storms over the Arabian Sea and the Bay of Bengal. The annual start and end dates of the ISM-peak are highly correlated (0.6–0.8) with the criteria-based dates proposed previously, supporting the delineation of the Indian monsoon subseasonal features.

SIGNIFICANCE STATEMENT: This research explores the existence of subseasonal features in the Indian monsoon annual cycle. Through the use of machine learning, we discover that the Indian summer monsoon and winter monsoon each consist of three substages. These substages’ evolution can be measured by two compact indices proposed herein, which can aid in subseasonal monsoon monitoring and forecasts in South Asia. Pertaining to hazard adaptations, this work pinpoints the subseasonal episodes most susceptible to droughts, heatwaves, floods, and tropical storms. High correlations are obtained when validating the substages’ yearly start and end dates against those documented in the existing literature, offering credibility to the subseasonal features of the Indian monsoon.

KEYWORDS: Atmospheric circulation; Monsoons; Risk assessment; Clustering

1. Introduction

The marked land–sea thermal contrast between the Indian Ocean and the South Asian continent, plus the thermal and mechanical effects of the Tibetan Plateau, all render the Indian monsoon (IM) one of the most energetic monsoon systems on Earth (Wang and LinHo 2002; Webster 1987). As a tropical monsoon system, the IM is best known for its two distinct phases throughout the year, the wet summer and the dry winter, vital to the ecosystem services, agriculture, and the welfare of billions in South Asia (Wang and Ding 2008; Wang and LinHo 2002; Wang et al. 2017).

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-23-0099.s1.

Corresponding author: Mengqian Lu, mengqian.lu@ust.hk

DOI: 10.1175/JCLI-D-23-0099.1

© 2023 American Meteorological Society. This published article is licensed under the terms of the default AMS reuse license. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
et al. (2009) showed that a simple and objective definition based on the 850-hPa zonal wind fields could well represent the historical ISM onset dates derived from the IMD and the local changes over Kerala. Instead of determining the ISM onset over the southwest tip of the Indian subcontinent only, Fasullo and Webster (2003) proposed the hydrologic onset and withdrawal index (HOWI) to investigate the ISM onset over India based on the large-scale vertically integrated moisture transport. For the northeast monsoon, early studies suggested that southeastern India and Sri Lanka received up to 50% of annual precipitation due to moist onshore flows over the Bay of Bengal (BoB) from October through November (Kumar et al. 2004, 2007; Matsumoto 1990; Sengupta and Nigam 2019; Wang and LinHo 2002). This period also corresponds to the postmonsoon season, together with the premonsoon season (i.e., April and May), which is the tropical storm season over the Arabian Sea and the BoB (Balaguru et al. 2014; Mahala et al. 2015). From December to February, the winter monsoon could bring over a third of the annual total precipitation in northern India (Dimri et al. 2016). These clues hint at the often neglected subseasonal features of the winter monsoon season.

In this study, we aim to comprehensively understand the subseasonal features of the IM system by using an objective and systematic classification framework. We adopt the self-organizing map (SOM)-based approach proposed by Dai et al. (2021) to identify distinct stages in the IM annual cycle, derive the annual start and end dates for each stage, and evaluate the robustness of the stage delineation. SOM is a neural network–based dimension-reduction method (Kohonen 2001) that has been proven useful in a wide range of climatological studies (Chu et al. 2012; Johnson 2013; Chiang et al. 2017; Islam et al. 2019; Dai et al. 2020; Pan and Lu 2020). With the SOM-derived subseasonal stages, their characteristics, including the stage-mean hydrometeorological patterns, monsoon evolutions, and rapidity, are explored. To facilitate operational monsoon monitoring, forecasts, and hazard warnings, we propose two compact monsoon indices alongside a comprehensive picture of stage-wise hazards. Finally, we evaluate the fidelity of the annual start and end dates of each subseasonal IM stage by comparing them with the traditionally defined IM onset and withdrawal dates in the literature. These findings will be useful for optimizing agricultural practices, water resources management, and operational monsoon forecasts.

The structure of the paper is as follows. Section 2 provides details on the data source, the SOM-based approach used in the study, and the hazard risk assessment methodology. In section 3a, we present the classified subseasonal IM stages and examine the rapidity switch between the two stages, as well as the monsoon evolution. Representative indices are demonstrated in section 3b. Section 3c provides a stage-wise hazard risk assessment. The derivation and validation of the annual start and end dates of the subseasonal IM stages are provided in section 3d. A summary and discussions of the findings are provided in section 4.

2. Data and methods

a. Data

Surface- and pressure-level meteorological fields at 1° grid resolution are retrieved from the fifth generation of the European Centre for Medium-Range Weather Forecast (ECMWF) atmospheric reanalysis data (ERA5) for a 40-yr period (1979–2018) (Hersbach et al. 2020). The 850-hPa horizontal winds at a slightly coarser grid resolution (i.e., 1.5°) are adopted in the SOM-based approach to maintaining a relatively low dimensionality in SOM clustering. The equivalent potential temperature \( \theta_e \) and its vertical mean in the lower troposphere from 925 to 700 hPa (denoted as \( \langle \theta_e \rangle_{925-700} \)) are computed using empirical equations (Bolton 1980) (see Text S1 in the online supplemental material for details). The best track dataset for tropical storms (i.e., including all observations that are marked as “tropical storm” in “nature” flag) is retrieved from International Best Track Archive for Climate Stewardship (IBTrACS), and only those that occurred from 1979 to 2018 over the BoB and the Arabian Sea are considered (Kruk et al. 2010). The CPC Unified Gauge-Based Analysis of Daily Precipitation is adopted for hazard analysis (Chen et al. 2008; Xie et al. 2007).

b. A SOM-based approach to classify monsoon stages

Following the approach proposed by Dai et al. (2021), we determined the IM annual cycle by applying SOM on the 5-day-moving-mean daily climatology of the 850-hPa horizontal winds over the IM region (9°–30°N, 60°–90°E) (Fig. S2a). The 850-hPa winds fields are helpful in not only determining the onset of the ISM rainfall (Wang et al. 2009), but also being the most representative variable for depicting the seasonal patterns of the prevailing monsoonal weather and climate (Krishnamurti et al. 2013). The purpose of our SOM-based approach is to group the days that share similar wind patterns so as to identify any subseasonal monsoon stages (Figs. S2b,c).

Before the SOM training, grid cells at altitudes higher than 1500 m were removed from the input matrix, and the wind fields were weighted by the square root of the cosine of latitude to account for the difference in grid areas across latitudes (Dai et al. 2021; Johnson et al. 2008). The Euclidean distance was adopted to measure the similarity between the input vector (i.e., a daily wind field) and the weighting vector of the node (i.e., a SOM pattern). When fixing a latitudinal band between 9° and 30°N, we show that the stage delineation is insensitive to the longitudinal boundary (e.g., 40°–100°, 60°–100°, and 60°–90°E) (Fig. S2). For instance, there is a 9-day difference around late August, but all other deviations are less than 4 days (Fig. S2b). When extending the southern boundary of the IM to 5°N, only the cold-season stages are affected slightly (Fig. S2c). Thus, we hereafter choose the core area of the IM region (i.e., 9°–30°N, 60°–90°E) as our study region, for it has the lowest spatial dimensionality among the six regions (Fig. S2a) while being capable of representing the subseasonal features of the large-scale IM circulation system.

SOM also requires a prescribed number and topological ordering of its nodes. We adopted the linear topologies
configurations (i.e., $1 \times 2$ to $1 \times 8$ SOM node settings) to mimic the annual cycle in which two consecutive stages are generally more alike than the nonconsecutive pair. Their distinguishability of clustering results was evaluated by the false discovery rate test, which uses the Student's t test as a local test (Dai et al. 2021; Wilks 2006, 2011) (Fig. S3a). It turns out that the $1 \times 5$ SOM configuration captures the key patterns of the IM annual cycle using the least number of nodes (Fig. S3b) while maintaining a relatively low indistinguishable ratio averaged over all pairs of SOM nodes in the topology (Fig. S3a). Note that a monsoon pattern from the $1 \times 5$ SOM clustering appears twice in a year (Fig. S3b), resulting in a total of six IM stages. Detailed explanations of each stage follow in section 3a.

c. Hydrometeorological hazard risk indices

To assess the risk of hydrometeorological hazards on a stage-by-stage basis, we devise four indices that quantify extreme temperature and rainfall days. As a proxy for heatwaves, hot days at a given grid cell are defined as those with daily maximum temperature (based on hourly ERA5 data) exceeding the 90th percentile of the historical distribution (1979–2018). Conversely, cold days are defined as those with a daily minimum temperature lower than the 10th percentile of the historical distribution. To disregard the regions where the local extreme temperature is not substantial enough to cause hazards, the regions where the mean temperature of the hot day is lower than 35°C and mean temperature of the cold day is higher than 10°C are excluded from hazard analysis. Heavy precipitation days are defined as those with daily precipitation greater than the 90th percentile of the historical distribution of the wet-day precipitation (i.e., $>1$ mm day$^{-1}$) at the given grid cell. Dry days refer to those with daily precipitation less than 1 mm day$^{-1}$. We compute the stage-wise probability of extreme temperature and rainfall days by dividing the total number of relevant days by the stage duration, which enables us to estimate the likelihood of hazards at a given magnitude. This approach obviates the need for ad hoc hazard identification criteria (Pai et al. 2013).

3. Results

a. Six subseasonal Indian monsoon stages and rapidity around the stage start dates

Our study identifies six distinct subseasonal stages within the IM annual cycle using a SOM-based approach and sensitivity tests (as described in section 2b). To simplify our terminology, here we refer to these stages as the IM stages or monsoon stages. We begin by analyzing the stage-wise environmental conditions to reveal the spatiotemporal characteristics of the mean state for each stage.

1) MEAN CIRCULATION PATTERN

(i) ISM season: Onset, peak, and withdrawal

The most pluvial monsoon stage occurs in the peak ISM season (hereafter called ISM-peak), which lasts for 3 months from 8 June through 5 September and receives about 64.4% annual precipitation in the land region (Fig. 1b). Dynamically, the ISM-peak is characterized by the strongest southwesterly monsoon flows over the Arabian Sea (Fig. 1b). The strong southwesterly is established by the pronounced pressure gradient between the northern IM low and Mascarene high (Fig. 2b), which generates the strongest low-level northward cross-equatorial flow on Earth that turns northeastward under Coriolis force. While a uniformly deep pressure depression contains high ($\theta_v$)$_{925-700}$ (generally $>340$ K) air masses over the Indian subcontinent (Fig. 1b), the torrential precipitation (i.e., $>14$ mm day$^{-1}$) is confined over the western Ghats—a region located in the windward side of the southern Deccan Plateau due to orographic lifting—where a dry condition is found on the lee side (i.e., Tamilnadu), and northeast India sees moderate monsoonal rains (i.e., 6–14 mm day$^{-1}$) (Fig. 1b). Strong precipitation fields (i.e., $>14$ mm day$^{-1}$) are found in the northern BoB and the Himalayas, signifying a concurrent rainy episode there (Fig. 1b). The latent heat released in the BoB convection exciting ascending Rossby waves, largely contributes to maintaining the giant upper-tropospheric South Asian high (Fig. 2b). The strongest easterly vertical shear prevails in the BoB and the Arabian Sea, prohibiting the development of tropical storms.

Two transition stages exist before and after the pluvial ISM-peak, hereafter termed the ISM-onset and the ISM-withdrawal stages, respectively. The ISM-onset stage starts on 16 May and ends on 7 June (Fig. 1a); 16 May reflects the earliest ISM season onset, and the ISM-onset stage accommodates about 90% of the yearly ISM season onset occurrence (see Fig. 4 of Wang et al. 2009). The ISM-withdrawal stage extends from 6 September to 9 October and marks the gradual withdrawal of the ISM from northwest to southeast India. The mean circulation in both stages is characterized by moderate westerlies meandering around the 15°N south of the subcontinent. Yet, three differences are worth noting. First, the ISM-onset stage sees a significant monsoon low pressure system over northern Indian region (Fig. 2a). In contrast, the monsoon low is invisible during the ISM-withdrawal stage (Fig. 2c). Second, the upper-level South Asian high is centered at 18°N, 100°E during the ISM-onset (Fig. 2a). In comparison, it is located at 22°N, 90°E during the ISM-withdrawal stage (Fig. 2c), suggesting a delay in the southeastward retreat of South Asian high. Third, the ISM-withdrawal stage sees threefold as much land precipitation as that in the ISM-onset stage (Figs. 1a,c), due to the low-level depression in northeast India.

(ii) Fall, winter, and spring stages

The dry season of the IM annual cycle is characterized by three stages, namely the fall, winter, and spring stages. In general, the subcontinent is mostly in a dry climate during these three stages, with only 18% annual precipitation received (Figs. 1d–f). Precipitation is concentrated in the southwest BoB, Sri Lanka, and the southeast tip of India.

The fall stage covers 10 October–25 December, and around 7% of annual land rainfall is received. As the low-level continental high rapidly develops aloft the arid Arabian Peninsula...
due to a strong radiative cooling, the high pressure ridge extends to northern India around 22°N (Fig. 2d). To the south of the ridge, northeasterlies prevail between 10° and 20°N by geostrophic balance (Fig. 1d). Meanwhile, the 330-K \(\theta_{925-700}\) contour line—often viewed as the line of demarcation between the tropical and midlatitude air masses (Tomita et al. 2011)—resides at around 18°N over the Indian subcontinent and the BoB (Fig. 1d). In association with this is a rapid suppression of rainfall in most parts of the subcontinent, heralding the arrival of the first dry episode in the region (Fig. 1d). On the other hand, the cyclonic circulation controls the southern BoB with an elongated low pressure trough extending westward, conveying moisture from the BoB to the southeastern tip of the subcontinent and Sri Lanka, therefore kickstarting the wettest period there (Fig. 1d). This stage broadly represents the Indian northeast monsoon season (October to December) pictured in the literature (Kumar et al. 2004, 2007).

The winter stage begins on 26 December when the entire IM region is filled with relatively dry and cool air masses with \(\theta_{925-700}\) lower than 330 K (Fig. 1e). Albeit with a weakened intensity, the low-level continental high from the Middle East extends farther southeastward to influence the entire Indian subcontinent (Fig. 1e). Associated with the southward movement of this high pressure system is the stronger northwest monsoons in the north (Fig. 1e), mainly covering the Indian winter monsoon season (December–February) mentioned in the literature (Dimri et al. 2016). However, the clustering results suggest that the winter stage may last until mid-March.

The ensuing spring stage is rarely documented in the literature. It spans from mid-March to mid-May (Fig. 1f). During the spring stage, the Arabian high retreats westward with a southward shift of ridgeline. The prevailing northwesterlies fortify pronouncedly in northern India, while the northeasterlies weaken in the southern Arabian Sea (Fig. 1f). Meanwhile,
the 330-K contour line of \( \langle \theta_e \rangle_{925-700} \) resides north of 20°N over the Indian subcontinent, while the enhanced precipitation is found only in northeast India (Fig. 1f). Note that we adopted the CPC Unified Gauge-Based Analysis of Daily Precipitation to validate the land precipitation with ERA5, and it shows that both the pattern and the fraction of land precipitation are comparable (Fig. S4).

2) LATITUDE–TEMPORAL CHARACTERISTICS OF THE INDIAN MONSOON STAGES

Figure 3a presents the precipitation rate and the 850-hPa wind averaged over the central longitude of the IM (60°–90°E). It provides a complementary portrait of the subseasonal stages as a function of latitude and time. The ISM-onset stage is characterized by a northward progression of the southwest monsoon onset from 5°–10°N around 16 May to 20°–25°N around 7 June. The onset features a surge of monsoon rain and a reversal of northerly to southerly. In contrast, the ISM-withdrawal features a southward progression of the monsoon retreat from 20°–25°N around 6 September to 10°N around 9 October. As a result, the ISM-onset stage depicts the ISM onset progression from the south (10°N) to the north (25°N), while the ISM-withdrawal stage portrays its retreat process from 20°–25° to 10°N. The duration of the ISM-withdrawal stages is one-third more extended than the ISM-onset, consistent with the well-known feature of the ISM. The ISM-peak stage is marked by the maximum precipitation zone at 20°N, and the wind speed of the westerly exceeds 10 m s\(^{-1}\) south of 20°N throughout the entire stage.

The Indian winter monsoon is dominated by an anticyclonic circulation pattern in the latitude–time diagram with northwesterlies to the north and northeasterlies to the south of 20°N approximately. The ridge line migrates southward from 23°N in the fall stage to about 20°N in the winter stage and further south to about 15°N in the spring stage (i.e., orange line in Fig. 3a). Correspondingly, strong northeasterlies (4–7 m s\(^{-1}\)) are seen south of 15°N in the fall and winter stages while relatively strong northwesterlies (4–5 m s\(^{-1}\)) occur north of 20°N during the winter and spring stages. The fall stage is distinguished from the winter stage by its significant rainfall south of 10°N. The spring stage is distinguished from the winter stage by its phenomenal warming: the 330-K contour line of \( \langle \theta_e \rangle_{925-700} \) increases from 326 to 340 K almost simultaneously from 10° to 25°N (Fig. 3b). Interestingly enough, there exists a spring–fall asymmetry in which the spring stage tends to be relatively shorter (55 days) and drier than the fall stage (77 days) (Fig. 3a).

In summary, the latitudinal–temporal structures of the IM support the division of the six subseasonal stages with various lengths. Each stage exhibits a unique feature that distinguishes it from other stages: the spring stage’s sudden warming across a broad latitudinal band; the ISM-onset stage’s swift northward burst of the southwest monsoon and rainfall; the ISM-peak stage’s maximum rainfall rate (i.e., >8 mm day\(^{-1}\)) and the associated strongest southwesterly (i.e., >10 m s\(^{-1}\)); the ISM-withdrawal stage’s gradual southward retreat of the westerly and rainy season; the fall stage’s northeasterly-induced rainy period in southern India and Sri Lanka; and the winter stage’s strong northwesterly (northeasterly) north (south) of 20°N.
3) RAPIDITY AROUND THE STAGE START DATES

The rapid change across the start date of each stage provides additional evidence supporting the delineation of the six subseasonal stages. In this respect, for each stage, we compute the 10-day mean-state difference before and after each start date in precipitation, the 850-hPa wind fields, and the 925–700-hPa vertical mean equivalent potential temperature (θe)_{925-700} (contours; unit: K) over the land region. The zonal mean is taken in the latitudinal band of the IM region (60°–90°E). The color scale of the vectors is shown below the diagram. The dashed black lines indicate the first day of each of the six stages.

After the fall stage starts, most of the study region shows 2–4 mm day^{-1} rainfall reduction, but Sri Lanka and southeast India exhibit an increase of about 3 mm day^{-1} (Fig. 4d). In the winter stage that ensues, a negative (θe)_{925-700} field is found over the entire region after the stage starting. In contrast, the change of wind regimes is diminished compared with others (Fig. 4e). The arrival of the spring stage is signaled by the rapid increases in the 2-m temperature and the (θe)_{925-700} in the land region in response to the northward migration of the maximum solar radiation (Figs. 3 and 4f). The anticyclonic circulation over the Arabian Sea and cyclonic circulation over the Indian subcontinent (Fig. 4f) signifies the land–sea contrast in response to increased solar irradiance. Overall, the rapidity around the start of the three stages in the winter monsoon season is less intensive than in the ISM season. However, the stepwise southward progression of the Arabian Sea high pressure ridge line (the orange line in Fig. 3a) is evidence for the subseasonal stages.

b. Two circulation indices depicting six stages

With a clearer understanding of the six subseasonal stages (section 3a), it now becomes possible to devise simple indices to adequately represent the annual evolution of the IM, which would be particularly useful for operational monsoon monitoring. Here, we introduce two indices that describe the low-level wind fields over the Arabian Sea, an upstream region of the IM. The first index is the northern Arabian Sea...
flow (NASF) index, which is computed by averaging the 850-hPa zonal wind speed over the region of 15°–25°N, 60°–75°E. It helps describe the behavior of the Arabian Sea high pressure ridge during the three northeast monsoon stages (Figs. 1d–f). The second index, named the southern Arabian Sea flow (SASF) index, captures the westerly features over the southern Arabian Sea (5°–15°N, 50°–80°E) and characterizes the three southwest monsoon stages (Fig. 5).

It turns out that the NASF index successfully captures the reversal of winds from easterlies to westerlies associated with the southward migration of the ridge line from the fall to the spring stage (Figs. 3 and 4d–f). Specifically, the NASF index shows significant easterlies (−1 to −3 m s⁻¹) during the fall stage, wanders around zero (−1 to 1 m s⁻¹) later in the winter stage, and finally turns to moderate westerlies (1 to 4 m s⁻¹) during the spring stage (Fig. 5). Interestingly, the NASF index also captures the main circulation features of the three summer stages. The westerlies remain almost fixed at 4 m s⁻¹ during the ISM-onset stage until the start of the ISM-peak, right after which it rockets to over 6 m s⁻¹ and maintains that level throughout the stage over the northern Arabian Sea (Fig. 5). Unlike the ISM-onset case, the index drops drastically in the ISM-withdrawal stage from 6 to −1 m s⁻¹ within only a month, signifying the abrupt retreat of the southwest monsoon over the northern Arabian Sea.

In terms of the SASF index, it features the prevailing easterlies at 4 m s⁻¹ during the winter stage, then starts to weaken at the beginning of the spring stage over the southern Arabian Sea. Then SASF index crosses the zero-line 2 weeks before the ISM-onset stage starts. It is worth noting that both the NSA and SASF indices reach 4 m s⁻¹ at the start of the ISM-onset stage, which suggests that moderate westerlies dominate the Arabian Sea. Apart from the NASF index standing at 4 m s⁻¹, the SASF index is extremely enhanced to 10 m s⁻¹, suggesting a rapid acceleration process of westerlies in the southern Arabian Sea. Moreover, the enhancement process in the northern Arabian Sea starts near the end of the ISM-onset stage, at the same time, the acceleration process over the southern region suddenly slows down. During the entire ISM-peak stage, the westerly over the southern Arabian Sea
is greater than 8 m s$^{-1}$, then decreases moderately and crosses the zero line from positive to negative slightly later than the start of the fall stage. It is worth noting that, unlike the ISM-onset stage, the decreasing tendency is similar between the SASF and the NASF indices during the ISM-withdrawal stage. Finally, the decreasing tendency of easterlies is lower than before and finally close to the 4 m s$^{-1}$ (Fig. 5). Taken together, the results suggest that the NASF and SASF indices are capable of describing the annual cycle of the IM system.

Interestingly, the southern Arabian Sea features a westerly burst during the ISM-onset stage and a relatively slow decay of westerly during the ISM-withdrawal stage. In contrast, the northern Arabian Sea shows little change of the westerly during the onset stage while a swift decrease of westerly during the withdrawal stage. The skewed annual cycles in monsoon westerlies are opposite between the southern and northern Arabian Sea (Fig. 5).

c. Stage-wise hazards risk assessment

Given the exploration of the subseasonal characteristics of the six IM stages, it is important to examine the likelihood of extreme events, such as heavy precipitation days, dry days, hot days, and cold days, as well as the frequency of tropical storms during these stages.

Consistent with what we mentioned in section 3a(1), considering the local rainfall climatology, the ISM-peak stage features relatively high flooding potentials that stand out over both the windward side of the southern Deccan Plateau and northeast India, as they would experience heavy precipitation (i.e., exceeding the 90th percentile of the local wet-day precipitation distribution; section 2c) for over 25% of the time in the 3-month-long ISM stage [Fig. 6b(1)]. Before and after the ISM-peak stage, the circulation differences in the two transition stages (i.e., ISM-onset and ISM-withdrawal) cause distinctive hydrometeorological hazard risks. The ISM-onset is characterized by an astoundingly high probability of hot days and drier-than-normal conditions in the interior of the Indian subcontinent [Figs. 6a(2),(3)]. Given the antecedent dry soils and warm surface during the preceding spring stage, the ISM-onset episode is prone to heatwaves and droughts [Figs. 6a,f(2),(3)]. Further, considering the rather brief span of the ISM-onset stage (23 days) and the rapid intensification of the southwest monsoon and precipitation as the ISM-peak stage begins (Fig. 2a), extreme weather whiplash would be likely to occur and pose great challenges to agriculture and water resources management in the subcontinent. By comparison, the ISM-withdrawal stage appears to be the most pleasant period when the risks of heatwaves, droughts, and floods all remain relatively low [Figs. 6c(1)–(3)].

In the fall stage, we observe a contrasting risk of hazards across the subcontinent, where a greater flood risk is expected in the southern tip while wide-ranging droughts are in the rest of the region [Figs. 6d(1),(2)]. Meanwhile, this stage is also the peak season for tropical storms, as it sees over half of the annual tropical storms in the Arabian Sea and the BoB (Fig. 7). Proceeding into the coldest stage of the year (i.e., the winter stage), the northern subcontinent of India and bordered Pakistan face frequent strikes of cold waves [Fig. 6e(4)]. As the spring stage begins and the near-surface temperature increases, the occurrence of hot days soars in the southern Deccan Plateau [Fig. 6f(3)]. More importantly, after a half-year-long arid condition since the fall stage [Figs. 6d–f(2)], the spring stage is the episode when the entire subcontinent is highly prone to compound hazards of droughts and heatwaves, giving rise to higher likelihoods of water scarcity, food crisis, and heat stress issues.

d. Annual start and end dates of the IM stages

As highlighted in the introduction, the onset and withdrawal of the ISM have been the primary focus of research,
with other periods often described only by month and without accounting for interannual variations. The loosely defined premonsoon (April and May) and postmonsoon (October and November) periods are often overlooked, despite their significance. To address this gap, we propose an alternative approach to identify critical timings throughout the year based on the climatological annual cycle of the IM, thereby highlighting the importance of each subseasonal stage.

![Grid-based probability of occurrence](image-url)
In this study, we developed a revised method for defining the onset of the IM stages, building on the approach introduced by Dai et al. (2021). Specifically, based on the IM monsoon annual cycle (Fig. 1), we identified the three nearest IM stages for each day and calculated the Euclidean distances between the daily and mean-state 850-hPa wind fields of these three stages. The day was then assigned to the stage with the minimum Euclidean distance. To make the daily assignment more robust, if the difference between the stage with the minimum Euclidean distance and the climatological stage is less than 1% of the total of three distances, the day is assigned to the climatological stage. The second step is to determine the start and end dates. The starting day of a stage is identified when the stage first persists for at least 10 days, except that we use 5 days for the ISM-peak stage. A similar approach is adopted for identifying the end day of a stage (i.e., the last day when a stage has persisted for at least 10 days). A similar 5-day threshold is applied to the ISM-peak stage. We have conducted a sensitivity analysis regarding the choice of thresholds on the start and end dates of each stage (Table S1). Note that because the start dates of the ISM-onset and ISM-withdrawal stages are sensitive to the number of consecutive days selected due to their relatively short durations (results not shown), we therefore used the end day of the spring and the ISM-peak stages to represent their annual commencement. A table providing the annual start and end dates is provided in Table S2.

It should be noted that the mean ISM-peak stage start date (i.e., 7 June) derived herein corresponds to the moment when the Indian subcontinent south of 20°N declares the southwest monsoon onset (i.e., 10 June; Fig. S1a), instead of onset in Kerala. However, the Pearson correlations between the ISM-peak annual start dates derived herein and the ISM onsets in the previous studies (Ghanekar et al. 2019; Prasad and Hayashi 2005; Wang et al. 2009) reach as high as 0.6–0.8 (Table 1), suggesting strong relations of the interannual variability between the monsoon onset in Kerala and central India. On the other hand, the annual ISM-peak retreat presented herein (5 September) is similar to that measured by the HOWI (Fasullo and Webster 2003).

The fall stage starts on 9 October on a 42-yr average (Table 1), which marks the termination of westerlies over the north of 10°N (Fig. 3a). Notice that around the same time, the IMD declares the normal ISM withdrawal line (i.e., the westerly monsoon) would generally reside between 15° and 20°N (Fig. S1b). This is because the IMD withdrawal criteria require the cessation of rainfall activity over the area for 5 days in a row. Yet, the southeast Indian subcontinent is still rainy after 10 October. Note that this southeast Indian rainfall is not typical southwest monsoon rainfall from the circulation perspective. In addition, it is worth noting that the standard deviations of annual start and end dates of all stages are around 8 days only, except the winter and spring stages whose start dates exhibit an interannual variability of 12–16 days (Table 1). This suggests the inherent uncertainties in capturing the commencement of the winter and spring stages, as also noted in our monsoon rapidity analysis [section 3a(3)].

4. Discussion and conclusions

This study was motivated by the question of whether the IM annual cycle can be characterized by wet and dry seasons alone. To address this question, we conducted a SOM analysis of the 850-hPa circulation, which allowed us to objectively identify six subseasonal IM stages. These stages include three wet stages that cover the ISM season, namely ISM-onset, ISM-peak, and ISM-withdrawal, as well as three dry stages: fall, winter, and spring. The duration of these stages varies from 23 to 90 days. In addition to identifying the stages, we also determined the year-to-year variation of the start and end dates for each stage. Compared with the existing literature using criteria-based methods (Table 1), our SOM-based method delineates the substages reasonably and offers objective determinations of their commencement and withdrawal.
Table 1. Dates of the IM stages derived in this work and in previous studies. Uncertainties are based on one standard deviation if onset dates for individual years are available. Note that the Pearson correlation coefficients between the onset and retreat dates presented in this work (Table S2) and in the previous studies are calculated based on the overlapped period. Statistically significant correlation coefficients at the 95% confidence level are shown next to the dates; otherwise, they are denoted by long dashes (—). OLR = outgoing longwave radiation; $P = $ precipitation.

<table>
<thead>
<tr>
<th>Stage number</th>
<th>References</th>
<th>Methods</th>
<th>Parameters</th>
<th>Dataset</th>
<th>Time</th>
<th>ISM-peak start</th>
<th>ISM-peak end</th>
<th>Fall start</th>
<th>Winter start</th>
<th>Spring start</th>
<th>Spring end</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>This study</td>
<td>SOM-based</td>
<td>850-hPa horizontal winds</td>
<td>ERA5</td>
<td>1979–2020 (42 years)</td>
<td>7 Jun ± 7</td>
<td>5 Sep ± 7</td>
<td>9 Oct ± 8</td>
<td>28 Dec ± 16</td>
<td>19 Mar ± 13</td>
<td>14 May ± 8</td>
</tr>
<tr>
<td>2</td>
<td>Fasullo and Webster (2003)</td>
<td>Vertically integrated moisture transport</td>
<td>Vertically wind shear between 850 and 200 hPa</td>
<td>NCEP–NCAR</td>
<td>1948–2000 (53 years)</td>
<td>4 Jun ± 7 (0.61)</td>
<td>7 Sep ± 11 (0.43)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>Prasad and Hayashi (2005)</td>
<td>Criteria-based</td>
<td>Vertical wind shear between 850 and 200 hPa</td>
<td>ERA–40</td>
<td>1958–2001 (44 years)</td>
<td>5 Jun ± 9 (0.85)</td>
<td>3 Jun ± 7 (0.76)</td>
<td>21 Oct ± 11 (—)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>Wang et al. (2009)</td>
<td></td>
<td></td>
<td>NCEP–NCAR</td>
<td>1948–2007 (60 years)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Estimated from Fig. 4 in Wang et al.’s (2009) paper.*

Our hazard analysis reveals that the subseasonal IM stages exhibit varying risks of natural disasters. As expected, the ISM-peak stage has the highest probability of floods. However, our analysis also highlights the high risks of compound heatwaves and droughts during the spring and ISM-onset stages, particularly in southern and central India. The fall stage stands out as the peak season for tropical storms in the Arabian Sea and the BoB, with more than 50% of tropical storms occurring during this stage. Rather than using October and November as the study season directly, the background circulation conditions are similar within the fall stage, so we may need to consider all the fall months together as the TC study period in further research. Additionally, the winter stage is likely to bring cold waves to northern India and southern Pakistan. These findings underscore the need for improved disaster preparedness and early warning systems during these high-risk subseasonal stages of the IM cycle.

As an interesting side note, there exists an asymmetry in the two transition periods before and after the ISM-peak stage. The ISM-onset (~21 days) tends to be 50% shorter than the ISM-withdrawal (32 days) (Fig. 2). This asymmetry in the IM annual cycle is commensurate with that in the seasonal north–south displacement of the intertropical convergence zone (ITCZ) along the Southeast Asian “land bridge” (He et al. 2007; Hung et al. 2004; Matsumoto 1992), manifested in a sudden poleward advance in spring but a slow retreat to the equator in fall (Lau and Yang 1996; LinHo and Wang 2002; Schneider and Bordoni 2008). The asymmetric response of the coupled atmosphere–land–ocean system to annual solar forcing shapes the rapid northward migration of the ITCZ (monsoon trough) in early summer and its slower return to the equator during September.

Our findings suggest that it may be beneficial to monitor and predict the start of each subseasonal stage. The IM annual cycle plays a critical role in the agrarian societies of the
Indian subcontinent (Jena et al. 2015), where crop production is highly dependent on temperature and the timing of monsoonal rains (Webster et al. 1998). Knowing the onset and retreat dates of each subseasonal stage—not just the ISM-peak stage but also the other stages with varying temperature conditions, precipitation distribution, and hazard risks—can help optimize sowing and planting times and guide crop species choices to improve food security. Moreover, the timing and duration of the subseasonal IM stages can also be useful in subseasonal-to-seasonal prediction, as demonstrated in our recent study on the East Asian monsoon (Dai et al. 2021).

From the numerical modeling perspective, the framework of the IM annual cycle can serve as a benchmark to validate model simulations of subseasonal IM features. Our work also lays the foundation for future studies on the linkage of the IM with other monsoon systems with similar subseasonal features, interannual variabilities of the IM annual cycle, and its projected evolutions in a warmer climate.

Acknowledgments. This research contributes to and is financially supported by the Collaborative Research Fund (CRF) Project C6032-21G by the Research Grant Council of Hong Kong. BW acknowledges the support by NSF/Climate (CRF) Project C6032-21G by the Research Grant Council of Hong Kong. BW acknowledges the support by NSF/Climate (CRF) Project C6032-21G by the Research Grant Council of Hong Kong. BW acknowledges the support by NSF/Climate (CRF) Project C6032-21G by the Research Grant Council of Hong Kong.

Data and software availability statement. The meteorological data are retrieved from the ERA5 by the European Center for Medium-Range Weather Forecast (ECMWF) at https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5. IBTrACS best track data are available at https://doi.org/10.25921/82ty-9e16. The CPC Global Unified Gauge-Based Analysis of Daily Precipitation data are available at https://psl.noaa.gov/data/gridded/data.cpc.globalprecip.html. The SOM training was performed using the R package “kohenon” (Wehrens 2007). Some geographic figures were generated using a Matlab package “M_Map”(Pawlowicz 2020).

REFERENCES


