

## RESEARCH LETTER

10.1002/2016GL070711

## Key Points:

- Agronomic onset can precede the regional monsoon flow by over 3 weeks across India
- Nonarbitrary boundaries exist where agronomic onset shows similar interannual variability
- The interannual variability of agronomic onset over Kerala is different to that of its immediate surrounding region

## Supporting Information:

- Supporting Information S1

## Correspondence to:

R. G. J. Fitzpatrick,  
js08rgjf@leeds.ac.uk

## Citation:

Fitzpatrick, R. G. J., D. J. Parker, and P. D. Willetts (2016), Assessing the level of spatial homogeneity of the agronomic Indian monsoon onset, *Geophys. Res. Lett.*, *43*, 11,867–11,874, doi:10.1002/2016GL070711.

Received 3 AUG 2016

Accepted 27 OCT 2016

Accepted article online 29 OCT 2016

Published online 19 NOV 2016

©2016. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Assessing the level of spatial homogeneity of the agronomic Indian monsoon onset

Rory G. J. Fitzpatrick<sup>1</sup>, Douglas J. Parker<sup>1</sup>, and Peter D. Willetts<sup>1</sup>

<sup>1</sup>ICAS, University of Leeds, Leeds, UK

**Abstract** Over monsoon regions, such as the Indian subcontinent, the local onset of persistent rainfall is a crucial event in the annual climate for agricultural planning. Recent work suggested that local onset dates are spatially coherent to a practical level over West Africa; a similar assessment is undertaken here for the Indian subcontinent. Areas of coherent onset, defined as local onset regions or LORs, exist over the studied region. These LORs are significant up to the 95% confidence interval and are primarily clustered around the Arabian Sea (adjacent to and extending over the Western Ghats), the Monsoon Trough (north central India), and the Bay of Bengal. These LORs capture regions where synoptic scale controls of onset may be present and identifiable. In other regions, the absence of LORs is indicative of regions where local and stochastic factors may dominate onset. A potential link between sea surface temperature anomalies and LOR variability is presented. Finally, Kerala, which is often used as a representative onset location, is not contained within an LOR suggesting that variability here may not be representative of wider onset variability.

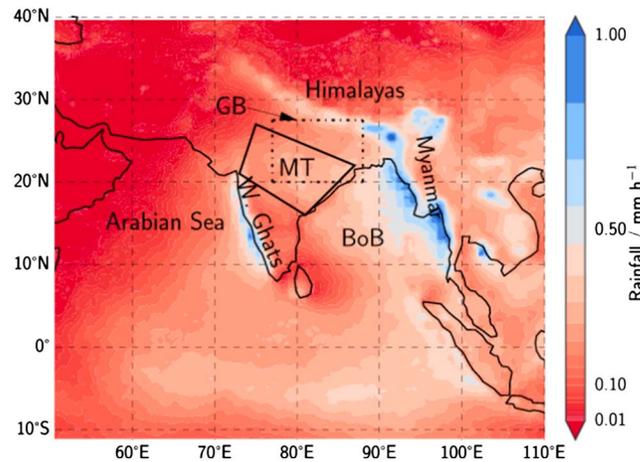
### 1. Introduction

The agronomic onset of the Indian Monsoon marks an important date for local stakeholders (taken here to mean local forecast users, particularly farmers) across the Indian subcontinent. It denotes the optimal planting date for rain-fed crops across the region. In this manuscript, the scale of spatial homogeneity of the agronomic onset is assessed across the Indian Subcontinent.

The onset of the Indian Monsoon rainfall on the regional scale is well documented. Onset first occurs in the extreme southwest over Kerala, over Bangladesh and over Assam, with a mean date of 1 June [Ananthakrishnan and Soman, 1988; Joseph *et al.*, 2006]. Indeed, the onset of the first rains over Kerala is often considered a key onset metric for the whole country [Ananthakrishnan and Soman, 1988; Joseph *et al.*, 2006; Wang *et al.*, 2009; Moron and Robertson, 2014]. After initial onset the northern limit of rainfall progresses north and west across the continent reaching the northwest of India around 15 July. Parker *et al.* [2016] recently explained how the progression of the northern limit of rains is controlled on average by the retreat of the midlevel dry northwesterly winds, which influence both the amount and the type of rainfall. However, as in other monsoon regions (such as West Africa), a regional perspective on monsoon onset is not necessarily consistent with the onset of rains at the local scale.

A prior study over West Africa has shown that local onsets are spatially consistent to a practical level [Fitzpatrick *et al.*, 2016]. By assessing homogeneity of onset over increasing spatial scales at every location, the maximum possible extent over which the interannual variability of onsets is consistent is captured within interannual local onset regions (LORs). Using LORs, one can segment a large monsoon region into subregions where similar drivers of onset (such as sea-surface temperature teleconnections or synoptic scale wave activity) are present. This is in contrast to measuring the total level of onset consistency across a preset domain. In this work, we examine the applicability of the LOR method discussed to the Indian subcontinent.

There are various local, agronomic onset definitions available. Standardly, an agronomic onset definition is presented as a series of threshold conditions that need to be met. Examples include Sivakumar [1988], Omotosho [1990], and Marteau *et al.* [2009] for West Africa, and Ananthakrishnan *et al.* [1967] and Moron and Robertson [2014] (which also includes a wide set of onset definition references) for India. In this work, the agronomic onset definition of Marteau *et al.* [2009], henceforth termed “the local onset,” is used to allow direct comparison with the work of Fitzpatrick *et al.* [2016]. The local onset definition used here was also incorporated as part of the work of Moron and Robertson [2014] and hence has previously been assessed over India. An extension to other onset measures and threshold variations of the local onset is readily possible and left for future work.



**Figure 1.** Mean hourly rainfall over the Indian subcontinent for 1 April to 30 September using TRMM v7 data for 1998–2014. Regions listed are the Arabian Sea, the Western Ghats (W. Ghats), the Ganges-Mahanadi Basin (GB), the Monsoon Trough (MT), the Himalayas, the Bay of Bengal (BoB), and Myanmar.

Figure 1 presents the April to September mean climatological precipitation pattern across the Indian subcontinent. The highest-average rainfall is located over the coastal regions of the Western Ghats and around the east coast of the Bay of Bengal. In addition, although the average rainfall amounts over the Monsoon Trough (defined as a region over India where the majority of rainfall comes from low-pressure systems by *Sikka and Gadgil* [1980]), and Ganges-Mahanadi Basin regions are not striking in Figure 1, they are socioeconomically important regions with large populations involved in farming. It is of interest to see whether the LOR method used captures these

regions as having distinct localized homogeneity of onset or indeed whether local onsets are consistent in any other regions.

In order to validate the use of LORs for sensitivity studies, potential drivers for onset variability across LORs need to be identified. In this work, we follow the work of *Moron and Robertson* [2014] who highlight a potential link between the local onset and El Niño–South Oscillation (ENSO) related anomalies in May across India. While ENSO variability differs through the season and is not expected to be the sole driver of local onset variability, the identification of a link in this work provides emphasis to continue using LORs to capture the causes of onset triggering across the Indian Subcontinent.

Sections 2 and 3 present the data and methods used in this study respectively. Section 4 highlights the mean onset patterns as well as LORs over the study region. Finally, conclusions are presented in Section 5.

## 2. Data

Precipitation data are taken from the Tropical Rainfall Measurement Mission 3B42 v7 product for the period 1 April to 31 July 1998–2014 (TRMM v7, *Huffman et al.*, 2007). Satellite measurements are aggregated to the daily scale and measured at their native quarter-degree resolution. For this work, the spatial limits of the Indian Monsoon region are taken as 60–100°E, 5–40°N. This region covers the entire of India as well as a large portion of the Indian Ocean (through which the zonal band of maximum precipitation moves northward, with the Intertropical Convergence Zone in the boreal spring and early summer). The dates used in this assessment capture the standard monsoon onset period, including the pre-monsoon rains [e.g., *Parker et al.*, 2016].

While other precipitation data can be used for analysis (both rain gauge and satellite based), the choice of TRMM v7 for this work allows for direct comparison between the findings presented here and those of *Fitzpatrick et al.* [2016]. Although data sets exist that have a longer timescale of available precipitation measurements (such as the Global Precipitation Climatology Product [*Adler et al.*, 2003; *Xie et al.*, 2003]), the higher spatial resolution of TRMM v7 allows for more precision when assessing spatial homogeneity without spatial interpolation. Mean onset dates for the local onset of *Marteau et al.* [2009] are spatially consistent between TRMM v7 and the Global Precipitation Climatology Product data set (Figure S1 in the supporting information).

ENSO anomalies are taken from the Climate Prediction Centre analysis of the Oceanic Niño Index across the Niño 3.4 region (5°N–5°S, 120–170°W). Monthly index values are freely available ([http://www.cpc.ncep.noaa.gov/products/analysis\\_monitoring/ensostuff/ensoyears.shtml](http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml)).

### 3. Methods

For this work, the local agronomic onset of *Marteau et al.* [2009] is used. The local onset date, defined at every grid cell for every year is calculated as the first rainy day (with at least 1 mm of precipitation) of two consecutive rainy days (with total precipitation greater than 20 mm) and no 7 day dry spell of precipitation less than 5 mm in the subsequent 20 days. Possible onset dates are calculated between 1 April and 31 July each year.

The inclusion of an extended dry spell test allows the method to distinguish between false and true onsets. It is found that extending the observation window of the dry spell beyond 20 days does not substantially change the timing or interannual variability of onset dates for the studied region (Figure S2) with the caveat that regions of southern India may be of greater risk of false onset compared to the rest of the study region.

The full method for calculating interannual LORs, where the interannual variability of onset dates is consistent, is presented by *Fitzpatrick et al.* [2016]. An abbreviated synopsis of the interannual LOR method is presented here.

For each grid cell, we independently assess the largest possible spatial scale (with a minimum accepted size of a 3-by-3 grid box) over which the onset time series of at least  $n_{crit}\%$  of all grid cells show correlation (at the  $x$  confidence level) with the median onset date time series of the LOR.

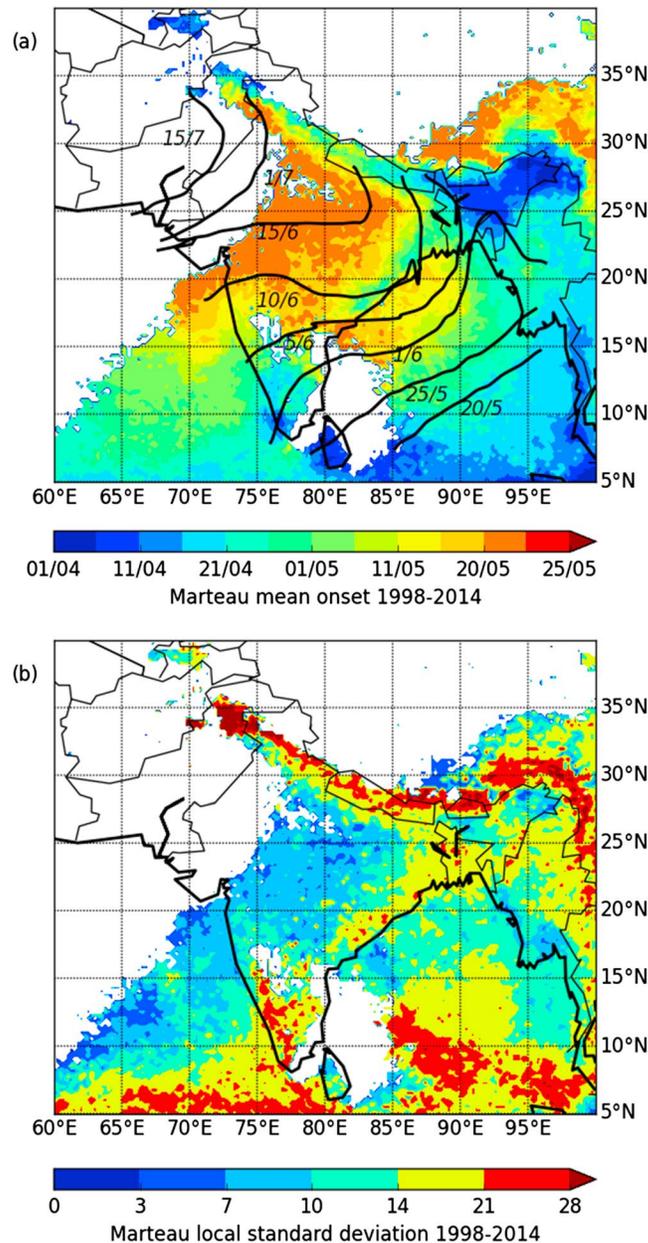
An interannual LOR is therefore a region where at least  $n_{crit}\%$  of grid cells share similar interannual variability. Note that as each grid cell is tested independently, it is possible (and indeed likely) that a grid cell will be contained within many LORs. Given knowledge of the interannual variability of the median onset across an LOR, practical information regarding onset timings can be given to forecast users with a pragmatic level of confidence.

The parameters  $n_{crit}$  and  $x$  are readily modifiable. Here we take  $n_{crit}$  as 80% and  $x$  as the 80%, 90%, and 95% confidence intervals, respectively, consistent with *Fitzpatrick et al.* [2016]. LORs are allowed to extend along any of the four major axes (north, east, south, and west), provided that the criterion listed above can be met. First we attempt to expand an LOR across all four axes. If this potential LOR fails the criterion, we test (in order) potential LORs with an expansion along the north and south axes, the east and west axes, the north axis, the east axis, the south axis, and finally, the west axis. If any of the potential LORs pass the criterion, we repeat the process from the beginning with LOR size recorded once all possible expansion fail. By design, the LOR method therefore is restricted to either square or rectangular LORs; extensions of this method to allow for nonregular-shaped LORs are possible but not performed here. The interannual LOR method employed is simple and repeatable providing relevant results for forecast users and institutional planners by bounding regions where similar monsoon onset sensitivity could be expected.

### 4. Results

Figure 2 shows the mean local [*Marteau et al.*, 2009] onset dates and the localized inter-annual variability across the Indian Subcontinent. The earliest continental onset dates are seen across the Bay of Bengal (taken here to include the coastal regions of West Bengal and Bangladesh) as well as over the southern tip of India and northeast India (90–98°E, 25–30°N) (Figure 2a). Onset dates progress inland in a north-westward direction from the Bay of Bengal and a northward direction from the Western Ghats similar to the findings of Moron and Robertson [2014, Figure 3a]. Over the northwest section of India and Pakistan, there are many years when onset is not triggered according to the local onset definition. Therefore, the local agronomic onset date selected for this work is not applicable over these locations although altered thresholds within the local onset definition may provide different results. A full analysis of the variability of onset triggering and timing due to the modification of onset thresholds is beyond the scope of this work.

It is of note that the local onset date differs substantially from the commonly used regional onset metric presented by the Indian Meteorology Department (IMD; compare shading and contours in Figure 2a). Many local onsets appear to precede the IMD onset by at least 3 weeks. In particular, local onset dates across the southern Western Ghats (south of 15°N) and the Bay of Bengal precede the regional onset of the southwesterly monsoon flow by about a month (solid lines in Figure 2a). This disparity in onset timings is similar to that found over other regions such as West Africa [*Fitzpatrick et al.*, 2015]. It is possible that local



**Figure 2.** (a) Mean local onset and (b) local temporal standard deviation (days) for TRMM v7 data 1998–2014. In Figure 2a, black superimposed lines mark the climatological regional onset dates as defined by the Indian Meteorological Department. White regions denote locations where local onset is found in less than five of the years studied.

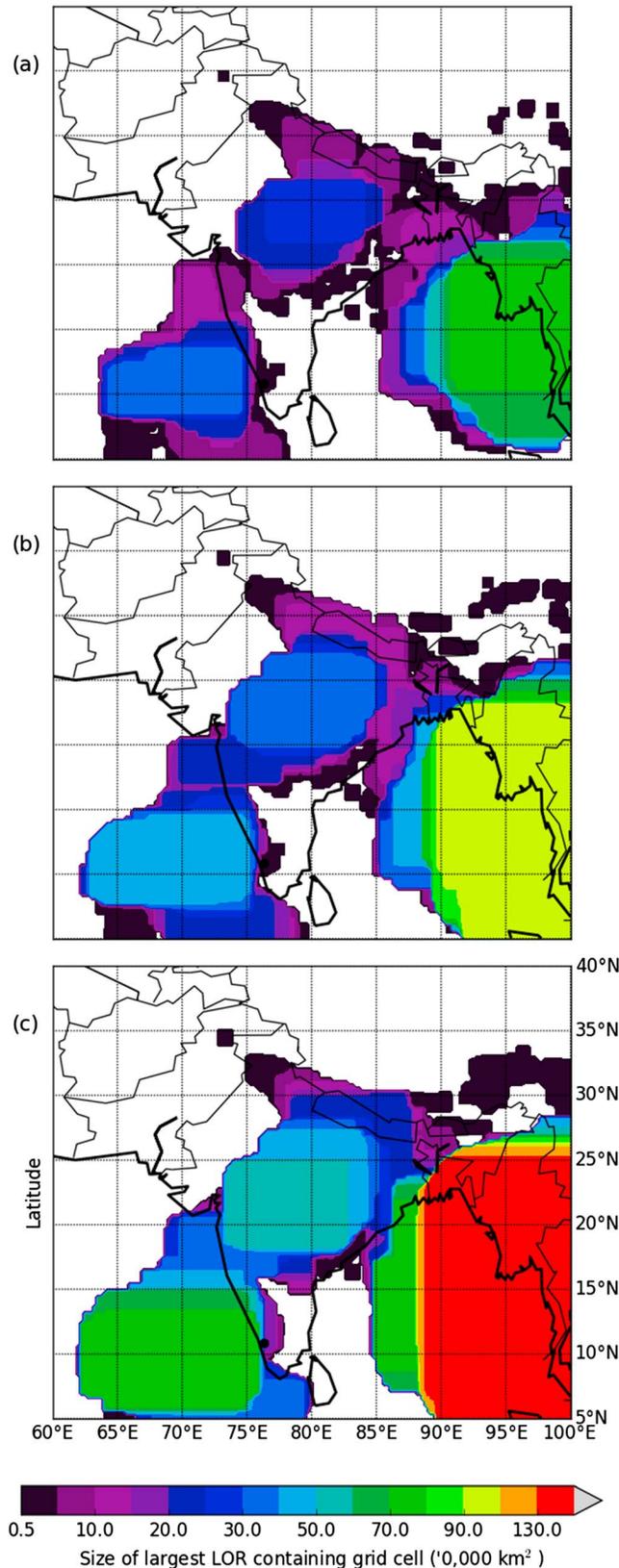
confidence intervals (Figures 3b and 3c), but the same regions are apparent at all confidence intervals. There are also some smaller LORs found over the China and India border at the 90% and 80% confidence interval (clustered around 90–100°E, 30°N in Figure 3c).

Similar to the results found for West Africa, there are regions where local onsets show no spatial homogeneity using this method and onset definition choice (white regions of Figure 3). Most of southeast India, the eastern coast of India and northwest India are not captured within LORs. It is worth noting that the eastern coast of India is within the rain shadow of the Western Ghats during the summer monsoon and has a different seasonal precipitation cycle to much of the rest of the country. Nevertheless, Figure 3 suggests that over substantial areas of India the potential predictability of local onsets may be limited as local variability may dominate onset

onset is triggered by more isolated storms (which are still persistent enough to trigger the local onset) prior to the establishment of the maximum convergence zone across the regions identified. Furthermore, the IMD onset is a subjective date determined using 5 day precipitation averages and therefore is possibly hampered by a lack of objectivity (by contrast the local onset is objective and threshold dependant). The IMD onset has been criticized in prior work and may not be the most relevant onset metric for stakeholders in the region [Wang et al., 2009].

The locations of highest local onset variability (over 3 weeks) are the Western Ghats, the Bay of Bengal, the east coast of the Indian peninsula, and along the Himalayas. By contrast, over central India, the variability of local onset is lower (~10 days). In general, locations with early onset dates have high interannual variability of onset (Figure 2b).

Figure 3 shows the largest LOR each grid cell is contained within at the 95%, 90%, and 80% confidence intervals (Figures 3a–3c, respectively). For completeness, all LORs found including those that only cover ocean are included. At the 95% confidence interval, three distinct clusters of LORs are found. These are termed the Arabian Sea (which includes the Western Ghats), the Monsoon Trough region, and the Bay of Bengal consistent with Figure 1. The largest LORs are found over the Bay of Bengal, followed by the Arabian Sea (Figure 3). As expected, LORs are larger and more numerous for less stringent confidence intervals.

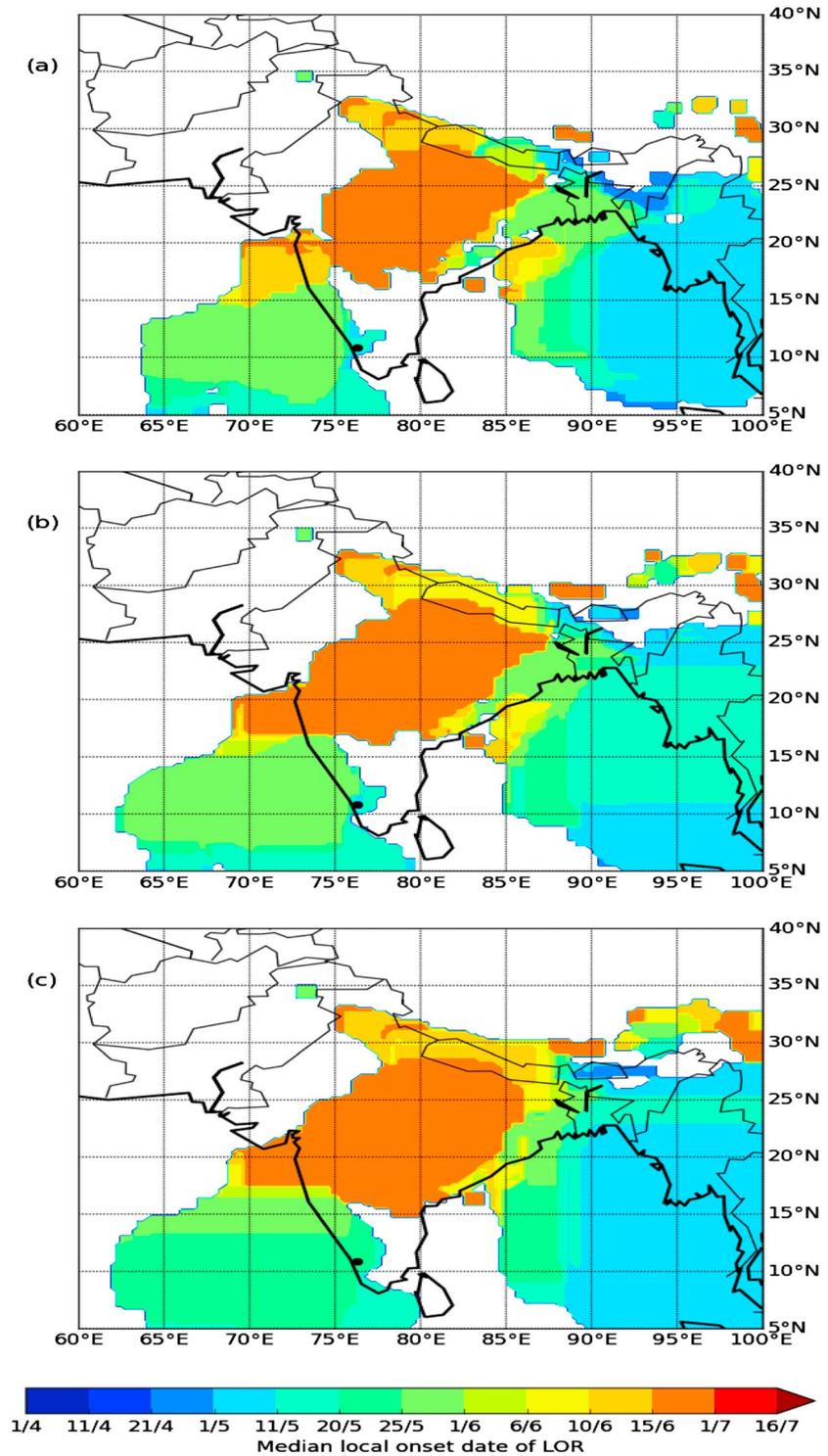


triggering rather than synoptic scale controls. These regions largely coincide with locations where the risk of false onset is potentially quite high (compare white regions in Figures 3 and S2).

Figure 4 shows the median onset dates of the LORs presented in Figure 3 with the LORs of Figures 4a–4c directly relating to Figures 3a–3c, respectively. Consistent with Figure 2a, the earliest cross-LOR onset dates occur across the Bay of Bengal and the Arabian Sea with later dates across the Monsoon Trough and toward the Himalayas. Interestingly, the difference in median LOR onset date between the Bay of Bengal and the Monsoon Trough region matches the step change in both regional and local onsets between these two regions (both the colors and black lines in Figure 2a). Figures 3 and 4 suggest that the LOR method can accurately capture regions of homogeneous local onset variability across the Indian subcontinent and the representative timings of LOR onsets are realistic (when compared to the findings of Wang *et al.* [2009] for example).

It is also of note that Kerala (76.3°E, 10.9°N; black dot in Figures 3 and 4) is at most on the periphery of the LORs found across the Arabian Sea. Figures 3 and 4 suggest that the interannual variability of local onset about Kerala is not consistent with onset variability both in immediately neighboring regions and further afield. The use of onset measurements over Kerala as a representative indicator may give misleading results for the rest of the Indian Subcontinent. Similar results for Kerala using different methods have

**Figure 3.** Size of the largest LOR containing each grid cell using LORs calculated at the (a) 95%, (b) 90%, and (c) 80% confidence interval. White locations denote areas that are not contained within an LOR. Scale of color bar is in the order of 10,000 km<sup>2</sup>.



**Figure 4.** Median onset date of interannual LORs found at different confidence levels. Mean (from 1998 to 2014), median onset dates for each of the LORs of Figures 3a–3c for (a) the 95% confidence interval, (b) the 90% interval, and (c) the 80% interval. White regions denote areas where no LOR is found at the given confidence interval. Note that the same specific location may have a different median date at different confidence intervals depending on the area over which the assessed LOR is present.

been shown previously [cf. *Moron and Robertson, 2014*] and combined these findings raise the question of whether timing of monsoon onset over Kerala should be considered as representative on a large scale.

In summary, local onset dates are consistent across nonarbitrary boundaries over the Indian subcontinent with two of the LOR clusters found covering regions of high local interannual variability of onset. By using LORs, it may be possible to evaluate the causes of interannual local onset variability across regions where high variability is present. This provides the platform for future sensitivity work across the Indian subcontinent.

#### 4.1. Link Between the El Niño–Southern Oscillation and LOR Onset Variability

Previous work has shown a potential link between local onset dates and May ENSO anomalies [*Moron and Robertson, 2014, Figure 8*]. Here we investigate a link between interannual variability of LOR onset dates and ENSO anomalies for the 3 month periods January–March and April–May.

Figures S3 and S4 in the supporting information show LORs that have significant rank correlation (at the 95% confidence interval) with ENSO anomalies in January–March and April–May, respectively. While there appears to be a link between ENSO anomalies at both time periods studied and LOR variability across the Bay of Bengal region, the majority of correlated LORs are found over the ocean and therefore have little agronomic value to stakeholders. Of more interest is the region of correlating LORs seen in Figure S4 around the Monsoon Trough region which may suggest a trigger for onset variability. However, over the Western Ghats region, no correlating LORs are seen in either Figure S3 or S4. This result does not contrast with the findings of *Moron and Robertson [2014]* who find weak correlation between the two metrics in their Figure 8. While ENSO conditions may impact local onset variability across certain parts of India, our work suggests that there are other factors that control relative onset timings over much of the study region.

## 5. Conclusions

Local onset dates provide meaningful information for forecast users in monsoon regions, but their relationship to regional-scale measures of the monsoon can be very weak. Here the spatial homogeneity of local onset dates, taken from *Marteau et al. [2009]*, are calculated to assess the regional coherence of onsets across India.

The local agronomic onset date studied occurs either before or around the average onset associated with the large-scale monsoon flow. Earliest onset dates occur around the Arabian Sea and the Bay of Bengal where local onset can precede the regional monsoon progression by upwards of 3 weeks. Later onset dates are found further inland. The location of earliest local onset dates is similar to the location of high local interannual variability of onset.

Local onset regions (LORs) [*Fitzpatrick et al., 2016*] provide a method to calculate nonarbitrary boundaries over which local onset variability is consistent. Three main clusters of LORs are found in the studied region. These are the Arabian Sea (including the Western Ghats), the Monsoon Trough, and the Bay of Bengal and are roughly similar to currently defined monsoon regions (seen in Figure 1). There is some evidence of ENSO variability in April–June having some level of control over onset variability in the Monsoon Trough and Bay of Bengal regions. However, it is quite likely that the onsets over different regions have different controls or react to the atmospheric conditions differently (for example, due to land-surface interactions or orography). Further work is required to fully understand the cause of onset variability in each region.

There also exist large regions of India where local onset variability is greater than the regional signal, and thus, LORs are not present. Over areas where no large-scale coherency of onset dates is present, local noisiness of precipitation likely dominates the interannual onset variability. For these locations the limit of predictability of onset will be lower than within an LOR.

Larger LORs are present across the Indian subcontinent than are found over West Africa, particularly at the highest-confidence interval tested (compare Figure 3 to Figure 6 of *Fitzpatrick et al. [2016]*). The reason for this difference could be due to more coherent forcing of the large-scale monsoon flow over the Indian subcontinent or more coherent responses to large-scale forcing across the Indian subcontinent. The exact understanding of this result is left for future work.

In summary, understanding the synoptic scale controls on local onset can be performed across the LORs presented. Over other regions, further investigation is required to understand why local variability in onset is so high. Interannual LORs provide the framework on which future research may be performed with relevance to local stakeholders.

#### Acknowledgments

This research has been supported by the NERC/DFID Future Climate for Africa program under the AMMA-2050 project (NE/M020126/1), as well as the joint NERC/MoES INCOMPASS project (NE/L013843/1). The data used for this work have been referenced in the bibliography and are freely available for evaluation and extension to the work. The authors would like to thank Richard Keane from the University of Leeds for his suggestions during the construction of this manuscript as well as Vincent Moron and the anonymous reviewer who provided valuable suggestions to improve the article.

#### References

- Adler, R. F., et al. (2003), The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), *J. Hydrol.*, *4*, 1147.
- Ananthakrishnan, R., and M. K. Soman (1988), The onset of the south-west monsoon over Kerala: 1901–1980, *J. Clim.*, *8*, 283–296.
- Ananthakrishnan, R., U. R. Acharya, and A. R. Ramakrishnan (1967), On the criteria for declaring the onset of the southwest monsoon over Kerala, IMD, Forecasting Manual, No. IV-18.1.
- Fitzpatrick, R. G. J., D. J. Parker, J. H. Marsham, P. Knippertz, and C. L. Bain (2015), The West African monsoon onset—A concise comparison of definitions, *J. Clim.*, *28*, 8673–8694.
- Fitzpatrick, R. G. J., D. J. Parker, J. H. Marsham, P. Knippertz, and C. L. Bain (2016), On what scale can we predict the agronomic onset of the West African monsoon?, *Mon. Weather Rev.*, *144*, 1571–1589.
- Huffman, G. J., R. F. Adler, D. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, T. Hong, E. F. Stocker, and D. B. Wolff (2007), The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale, *J. Hydrometeorol.*, *8*, 38.
- Joseph, P. V., 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.
- Marteau, R., V. Moron, and N. Philippon (2009), Spatial coherence of monsoon onset over western and central Sahel (1950–2000), *J. Clim.*, *22*, 1313–1324.
- Moron, V., and A. W. Robertson (2014), Interannual variability of Indian summer monsoon rainfall onset date at local scale, *Int. J. Climatol.*, *34*(4), 1050–1061.
- Omotosho, J. B. (1990), Onset of thunderstorms and precipitation over northern Nigeria, *Int. J. Climatol.*, *10*, 849–860.
- Parker, D. J., P. Willetts, C. E. Birch, A. G. Turner, J. H. Marsham, C. M. Taylor, S. Kolusu, and G. M. Martin (2016), The interaction of moist convection and mid-level dry air in the advance of the onset of the Indian monsoon, *Q. J. R. Meteorol. Soc.*, *142*, 2256–2272.
- Sikka, D. R., and S. Gadgil (1980), On the maximum cloud zone and the ITCZ over Indian longitudes during the southwest monsoon, *Mon. Weather Rev.*, *108*, 1840–1853.
- Sivakumar, M. V. K. (1988), Predicting rainy season potential from the onset of rains in souther Sahelian and Sudanian climatic zones of West Africa, *Agric. For. Meteorol.*, *42*, 295–305.
- Wang, B., Q. Ding, and P. V. Joseph (2009), Objective definition of the Indian summer monsoon onset, *J. Clim.*, *22*, 3303–3316.
- Xie, P. P., J. Janowiak, P. Arkin, R. F. Adler, A. Gruber, R. Ferraro, G. J. Huffman, and S. Curtis (2003), GPCP Pentad precipitation analyses: An experimental dataset based on gauge observation and satellite estimates, *J. Clim.*, *16*, 2197.