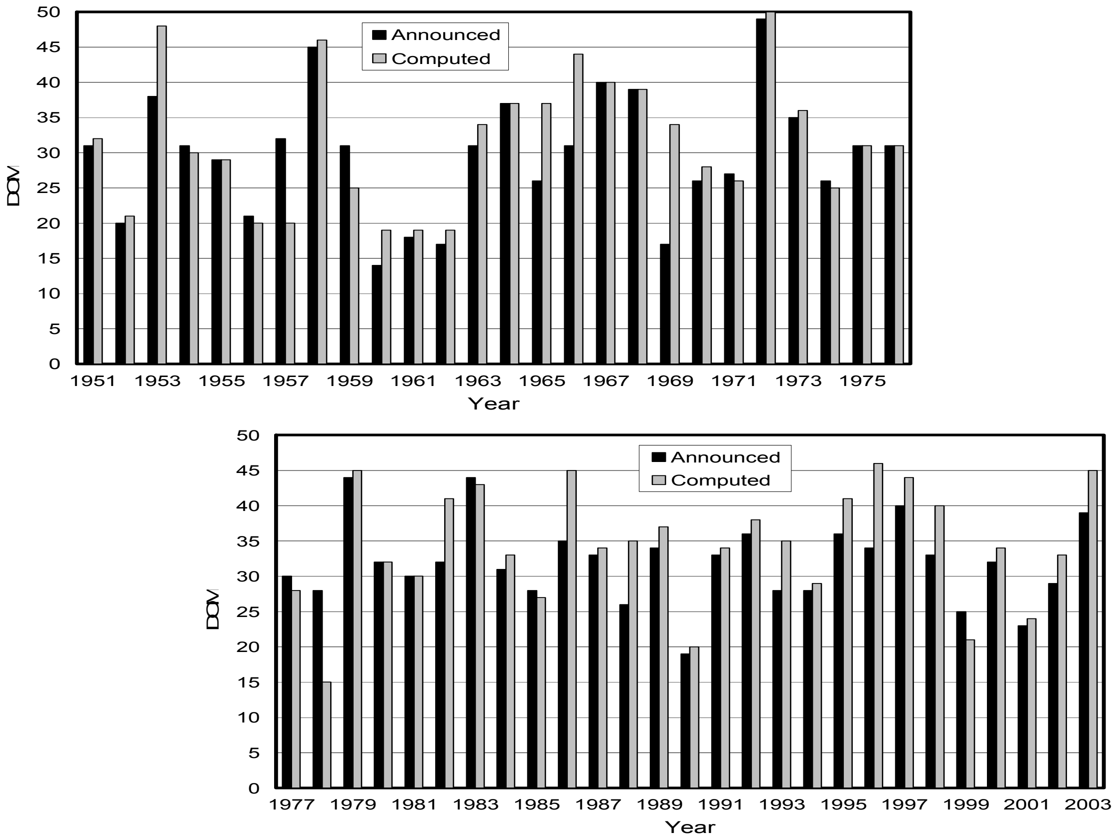
C-MMACS

Objective Determination of the Date of Onset of Monsoon Rainfall over India based on Duration of Persistence



P Goswami and K C Gouda

CSIR Centre for Mathematical Modelling and Computer Simulation
Wind tunnel Road, Bangalore-560 037, India

Research Report RR CM 0711
June 2007

Publication Status: Communicated to Current Science
Distribution: Unrestricted

CSIR Centre for Mathematical Modelling and Computer Simulation
Wind Tunnel Road, Bangalore-560 037, India
Tel No: +91 80 25224667   Fax +91 80 25220392
Website: http://www.cmmacs.ernet.in
E-mail: goswami@cmmacs.ernet.in

Objective Determination of the Date of Onset of Monsoon Rainfall over India based on Duration of Persistence

P Goswami and K C Gouda

Abstract

The onset of summer monsoon marks the beginning of the main rainy season for a large population; accurate determination of the day of the onset of monsoon (DOM) rainfall can be thus a valuable input for many applications. Determination of DOM, however, has been often difficult because of so called false or 'bogus' onset. An objective methodology is presented here to compute DOM from spatial distribution, rather than isolated station observations, of rainfall. We introduce duration of persistence and (% of) spatial coverage of rainfall to avoid false onset. The objective criteria have been calibrated based on daily gridded rainfall data for 53 years (1951-2003) from the India Meteorological Department (IMD). The average error in 53 onsets is 3.8 days, with 83% of the cases having error less than 1 standard deviation and about 15% of the cases with error more than 10 days. The method can also be used to determine application (requirement) based DOM.

Introduction

The onset of the Indian summer monsoon (ISM) over Kerala marks the beginning of the main rainy season for a large population. The onset, however, needs to be carefully distinguished from the synoptic processes that mimic it. The onset of monsoon is a result of a large-scale shift in the regional circulation pattern (e.g., *Ananthakrishnan and Soman,* 1991*; Soman and Kumar,* 1993*; Gadgil,* 2003). In contrast, the so called "false" or "bogus" monsoon onsets are associated with propagating tropical intraseasonal disturbances unrelated to the monsoon onset (*Flatau et al.,* 2001; *Joseph et al.,* 1994). These disturbances are characterized by an enhancement of convection and westerly surface winds similar to the monsoon onset but occurring over a smaller scale and lasting a week or less. The false onsets are often followed by extended periods of heat waves and droughts. An incorrect identification of a bogus onset with DOM can cause considerable economic and agricultural damage, as crops planted in anticipation of the monsoon are likely to fail. The bogus onsets can predate the actual onset by up to a several weeks (*Webster,* 1998); thus one of the biggest challenges in identifying, and predicting, date of onset of monsoon (DOM) is, to avoid these 'bogus' onsets.

A characteristic feature of the dynamics of the onset is that it appears to be primarily driven by large-scale rather than synoptic processes. Several studies have emphasized the roles of large-scale interactions between surface heating and atmospheric dynamic, thermal, and hydrologic processes in monsoon transitions (*Sikka and Gadgil,* 1980; *Webster,* 1983; *Webster et al.*, 1998; *Takagi et al.,* 2000; *Hsu et al.*, 1999; *Kumar* *et al.,* 1997; *Ueda and Yasunari,* 1998; *Wu and Zhang*, 1998). Though a variety of dynamic and thermodynamic precursors have been identified (*Ananthakrishnan and Soman,* 1991; *Murakami et al.,* 1985), one index of the large-scale transition in the regional circulation associated with the onset of monsoon is the characteristic change in the rainfall over Kerala. While there exists no unique definition, at the surface the onset is recognized as a rapid, substantial, and sustained increase in rainfall over a large scale; typically, from below 5 to over 15 mm day^{-1} during onset (*Ananthakrishnan and Soman,* 1988; *Soman and Kumar,* 1993). The suddenness of rainfall fluctuations during the monsoon's transitions has been emphasized in several studies (e.g., *Ramage,* 1971; *Rao,* 1976; *Ananthakrishnan and Soman,* 1991; *Wang and LinHo*, 2002)

Near India, the onset occurs initially across the peninsula's southern tip in late May to early June, progressing north-westward across most of the country in the following month. A feature that characterizes the onset is spatial coherency over a large scale, which is uncharacteristic of synoptic variability. Traditionally, the official announcement of DOM by the India Meteorological Department (IMD) is based on station observations of a number of meteorological variables (*Ananthakrishnan and Soman,* 1988; *Soman and Kumar*, 1993). Given the tremendous spatio-temporal variability of monsoon rainfall, the effectiveness of isolated station observations in capturing the essential characteristics of the onset process is questionable.

Inherent to identification of DOM, of course, is its (objective) definition. A number of techniques have been developed to identify monsoon onset. Two principal methods of identifying DOM are the objective method (Ananthakrishnan and Soman, 1998) and the more subjective declarations of the IMD. However, large disagreements have been noted in objective and subjective assessments of DOM. For example, in 1969 IMD's declaration of onset on 17 May disagrees with the objective classification (*Ananthakrishnan and Soman, 1998*) by 8 days. In 1959, the disagreements are 19 days and in 1943 and 1932 the disagreements are 17 and 19 days, respectively. Moreover, other years, such as 1979 and 1995, are associated with bogus onsets that objective methods can misdiagnose by up to three weeks (*Webster et al.,* 1998). A challenging task is to devise a method to determine and forecast DOM that can discern the occurrence of bogus monsoon onsets.

In this work we show that a set of objective criteria can be formulated to compute the DOM from spatial distribution of (gridded) rainfall. As shown below, these criteria encompass the conventional criteria for DOM based on rainfall, but have extended scope to avoid bogus onset. In particular, our criteria, for the first time, include post-onset persistence of rainfall to enhance reliability of the predicted DOM. Further, these criteria are equally applicable to observed gridded data and model outputs.

Definition and Calibration of Objective Criteria: Post-Onset Persistence

Although the onset manifests itself in various dynamical and thermodynamical variables, these can be expected to be closely interrelated and mutually consistent. We shall therefore consider only one variable, daily rainfall, to examine its potential to identify DOM as announced by India Meteorological Department (IMD). We consider three characteristics of rainfall that distinguish the true onset from a short-lived synoptic event; they are (a) large-scale (spatial coverage), persistence (length of duration) and significance (threshold). While it is possible to consider a duration of persistence (DOP) sufficiently long to announce DOM, such a method would have little practical use for announcing DOM; the announcement must be sufficiently well before the rainy spell of onset is followed by the characteristic lull. It should be noted, however, that the criteria of post-onset persistence, unlike the other three, can not be obtained from observations, and must be derived from model predictions.

Leaving the methodology for obtaining post onset rainfall for a future work, our aim in this work is to formulate and assess such a criterion for determining DOM. The objective criteria adopted by us are thus based on four parameters: pre-onset persistence (PrOP), significance, spatial coverage and post-onset persistence (PoOP). The PrOP and PoOP together ensure that the rainfall observed or predicted is sustained, and not a result of a transitory system. The significance is taken in terms of rainfall above a threshold value, 3 mm/day. The spatial coverage, above a threshold value, ensures the large-nature of the monsoon rainfall. While conventional announcement of DOM is based on station rainfall, this procedure suffers from the intrinsic and significant spatial variability of the distribution of monsoon rainfall from year to year. In particular, it is possible to have relatively high or low rainfall over a few stations without the characteristic large-scale coverage. We have therefore based the measure of large-scale nature of the onset in terms of percentage of spatial coverage of the onset domain; this avoids fixed locations for determining onset and automatically incorporates the spatial variability inherent in monsoon rainfall. However, as the traditional announcements of DOM are based on station observations (in Kerala), an appropriate equivalent value for the spatial coverage has to be determined from observed spatial distribution of rainfall.

The primary requirement for such an analysis, a high-resolution gridded rainfall dataset, was only recently met with the availability of a 53-year (1951-2003) daily rainfall data on a $1^{\circ} \times 1^{\circ}$ grid prepared by the India Meteorological Department (IMD). The IMD dataset (*Rajeevan et al.,* 2006) is based on rainfall records of 1803 stations which had a minimum of 90% data availability during the analysis period (1951–2003). The station rainfall data have been projected into a rectangular grid ($1^{\circ} \times 1^{\circ}$) for each day for the period 1951–2003. In this gridding method, the interpolated values are computed from a weighted sum of the observations. Given a grid point, the search distance is defined as the distance from this point to a given station. The interpolation is restricted to the radius of influence; for search distances equal to or greater than the radius of influence, the grid point value is assigned a missing code when there is no station located within this distance. A predetermined maximum value limits the number of data points used which, in the case of high data density, reduces the effective radius of influence. The starting point of the grid is 6.5°N and 66.5°E. From this point, there are 35 points towards east and 32 points towards north.

Results

As the first necessary step, we calibrate the values of the four parameters so that the statistics of the computed DOM from gridded daily rainfall data match with those of the announced DOM within margins of error inherent in the announced DOM. The calibration, based on multiple choices of each of the four parameters, is presented in Table 1. The results in table 1 are based on a pre-onset persistence of 3 days of daily rainfall of 3 mm/day or more. It may be seen that for post-onset persistence of 3 days and spatial coverage of 20%, (highlighted) the DOM computed from the daily gridded data fit most closely the announced dates of onset. It may be noted that although the case of 0 PoOP produces error statistics similar to those of the case with 3 day PoOP (Table 1), the former has several cases of large errors. In our subsequent analysis we shall therefore use, unless otherwise stated, a spatial coverage of 20% as the criterion for DOM.

The important role of spatial coverage has been already indicated in Table 1. Figure 1 further calibrates and quantifies the role of spatial coverage in terms of average error in computed DOM for different area coverage (%) beginning 10% and going up to 50 %. An important feature in figure 1 is the steep increase in error for coverage (%) beginning 10% and going up to 50 %.

Table 1: Summary of absolute average error and standard deviation between the announced and computed DOM for different persistence cases (PoOP 0,3 and 5 days) and spatial area coverage (10-50%, interval of 10%) with PrOP 3days, and significant rainfall threshold of 3 mm/day. The average announced DOM for period 1951-2003 is May 31 with standard deviation 7.3 days.

PoOP (days)	Area Coverage (%)	Mean Onset Date (day)	Correlation Coefficient	σ (Days)	ē (Days)	No of Years with error <1σ (%)
0	10	May 29 (29)	0.5	8.4	5.1	77
	20	May 31(31)	0.5	8.2	4.3	83
	30	June 3 (34)	0.3	11.3	8.3	56
	40	June 8 (39)	0.57	10.8	9.6	47
	50	June 14 (45)	0.14	13.3	16.1	28
3	10	May 31 (31)	0.6	8.74	4.5	79
	20	June 2 (33)	0.8	8.91	3.8	83
	30	June 3 (34)	0.7	10.7	5.5	77
	40	June 11 (42)	0.42	10.9	11.5	45
	50	June 21(52)	0.25	12.3	19.4	26
5	10	May 31 (31)	0.6	8.8	4.6	77
	20	June 3 (34)	0.75	9.7	4.8	72
	30	June 7 (38)	0.46	11.2	8.6	62
	40	June 12 (43)	0.46	10.2	12.4	36
	50	June 22(53)	-0.05	10.7	21.0	20

An important feature in figure 1 is the steep increase in error for coverage above 20%, while this error increases from 4 to 6 days for PoOP of 3 days, for PoOP of 0 and 5 days the error more than doubles. The large scale nature of monsoon is thus best reflected in an area coverage of 20% of the onset domain. It was found that for area coverage more than 50% a DOM within the May-June period could not be identified for most of the years.

A year-wise comparison of dates of onset from daily gridded data based on the above criteria with the announced dates of monsoon is presented in terms of absolute difference between the two in Figure 2. It can be seen that the case with no PoOP (thick solid line) has an average difference of 4.3 days compared to 3.8 days in the case with PoOP three days (thin solid line). For a PoOP of five days, this difference increases to 4.8 days. In figure 2 the years marked with * signify El-Nino years; as can be seen from the figure the errors in the years with El Nino are not particularly high.

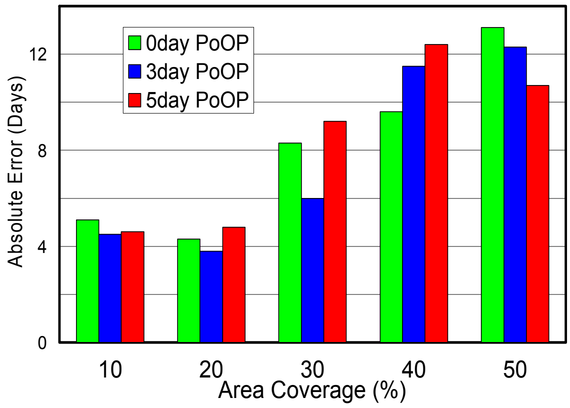


Figure 1: 53-year (1951-2003) average absolute error in computed DOM for different area coverage. The hollow, solid and shaded bars represent, respectively, computed dates based on criteria of 0 (no post-onset persistence), 3 and 5 day post-onset persistence of (threshold) rainfall above 3 mm/day. The Pre-onset persistence is 3 days.

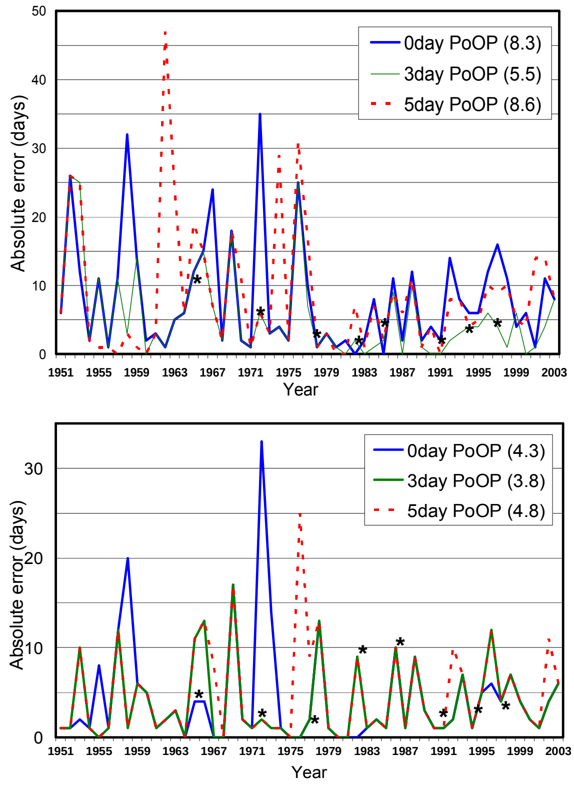


Figure 2: Absolute errors in days in determination of date of onset of ISM computed from gridded daily rainfall data from the India Meteorological Department with respect to announced dates of onset. The thick, thin and the dash line represent, respectively, errors based on criteria of 0 (no post-onset persistence), 3 and 5 day post-onset persistence of rainfall above 3 mm/day and 3 day pre-onset persistence. The * marks represents the false onset years. (a) for area coverage of 20% over the onset domain (b) for area coverage of 30% over the onset domain.

We next evaluate the skill of the methodology in terms of histogram of errors between the predicted dates of onset and the dates of onset computed from the daily gridded rainfall data of IMD following our objective criteria. The hollow, shaded and filled bars in figure 3 represent the (% of) cases in different error bins for PoOP of 0, 3 and 5 days, respectively. As can be seen from figure 3, both for PoOP of 0 and 3 days, 83% of the cases have error less than one standard deviation, while for PoOP of 5 days this number is 72%. The average error between the announced and computed DOM are 4.3 days, 3.8 days and 4.8 days, respectively, for PoOP of 0, 3 and 5 days, as given in the panel.

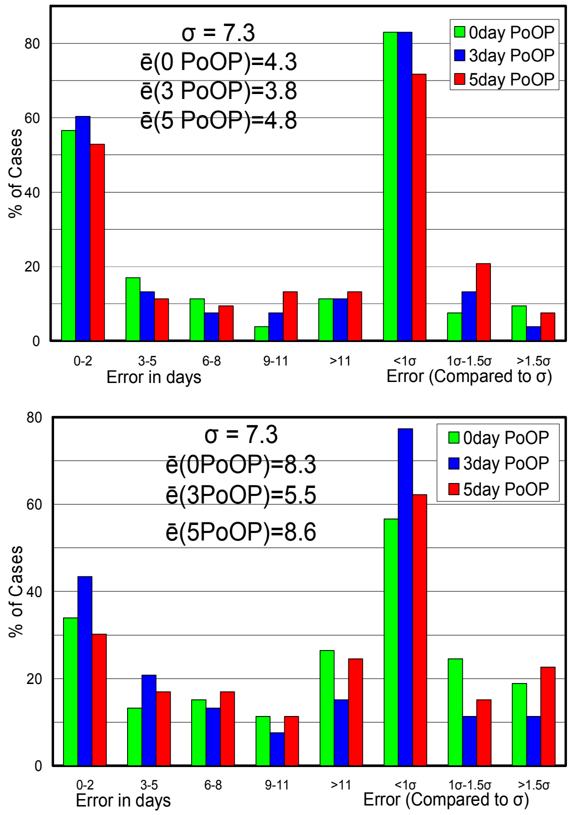


Figure 3: Histogram of errors in computed and announced dates of onset with respect to dates of onset computed from daily rainfall data from the India Meteorological department. Both the computed and the predicted dates are based on the criteria of significant rainfall with Pre-onset persistence of 3 days and threshold of 3 mm/day. The hollow, solid and shaded bars represent, respectively, Post-onset persistence of 0 (no post-onset persistence), 3 and 5 days. (a) for the area coverage of 20% over the onset domain (b) for the area coverage of 30% over the onset domain.

It may appear, based on the analysis so far that inclusion of a PoOP doesn't significantly improve identification of DOM. As mentioned earlier, however, the biggest challenge in identification of DOM is to avoid false onsets. We have therefore compared (Table 2) the success of the methodology for seven years that were characterized by false onsets (*Flatau* et al 2001) within the period 1951-2003. As can be seen from Table 2, the introduction of a PoOP of 3

Table 2 : Error in computation of DOM for seven years of false onset with different PoOP and Coverage

Year	0 day PoOP Coverage			3 day PoOP Coverage			5 day PoOP Coverage		
	10%	20%	30%	10%	20%	30%	10%	20%	30%
1967	3	0	24	3	0	7	3	8	7
1968	0	0	2	0	0	2	0	0	2
1972	33	33	35	33	2	6	33	2	6
1979	0	1	3	0	1	3	0	1	3
1986	6	10	11	6	10	9	6	10	9
1995	5	5	6	5	5	4	5	5	5
1997	4	4	16	4	4	4	4	4	9
Avg	7.2	**7.5**	13.8	7.2	**3.1**	5	7.2	**4.2**	5.8

days considerably reduces error in identification of DOM, the average error for the seven years is only 3.1 days for PoOP of 3 days, as against 7.5 days for no PoOP. Further, with PoOP of 3 days there is only one case of large error (10 days for 1986) as against two large errors (33 days and 10 days for 1972 and 1976, respectively) for no PoOP.

It is, of course, neither meaningful nor necessary to insist on a fixed set of criteria to define DOM beyond ensuring its monsoonal characteristics. The traditional definition of the onset of ISM, whether based on rainfall, dynamical fields or hydrological considerations, uses a single set of criteria. However, once these criteria ensure large-scale and sustained nature (that is monsoonal) rainfall, the parameters defining significance and persistence can be process-specific. In practice, it is necessary to consider onset dates based on multiple sets of criteria, as agro-hydrological requirements are likely to be different for different users (such as crop type and catchment's area). Our methodology allows determination of DOM based on such multiple-criteria, and we have explored it in table 1. Thus, based on different agro-hydrological requirements the date of onset can vary by as much as 11 days. Further, the parameter of emphasis can change from user to user. A reservoir manager over a small area may be more interested in the level and persistence of rainfall, while a state policy planner may give more emphasis on the extent of spatial coverage. For example, for PoOP 3 days, and area coverage 40% (which may be necessary for state wise agricultural planning), the mean date of onset is 11 June. For applications that do not require long persistence (0 PoOP), but with 50% coverage (such as for large-scale water availability) the mean date of onset is June 14. It is also worth noting that changes are comparatively larger for change in the spatial coverage than in the days of persistence.

Concluding Remarks

We have proposed and evaluated a methodology for determining the date of onset of monsoon based on gridded daily rainfall data; our method thus allows accounting for such parameters like spatial coverage (characteristic of ISM) not possible with isolated station data. Further, gridded data allows a uniform procedure for computing DOM both from observations and model simulations. Gridded data from IMD observations of course may not be available on time to be used in determining DOM as preparation of gridded data requires additional analysis. However, use of remotely sensed data with sufficient spatial coverage provides an exciting possibility. Although model forecasts in principle can be downscaled to station scale, the errors involved are often unacceptable. The use of % coverage largely eliminates errors that may arise from using fixed locations.

The separation of rainfall associated with onset to pre-onset and post-onset, may, of course, appear unnecessary; it could be argued that the period of post-onset persistence could be merged with the period of (a longer) pre-onset persistence. However, the combined length of pre-onset and post-onset persistence is about the average length of a monsoon rainy spell. Thus waiting to announce the onset at the end of a pre-onset persistence of 6 days or more will make such an announcement essentially ineffective.

Our work also provides, for the first time, a quantitative measure of large-scale nature of onset that distinguishes it from small-scale synoptic variability. While for spatial coverage 10-20% error are small (figure 1) , such coverage may not be significant for actual application. Beyond 30% coverage, on the other hand, DOM calculated using the present method would differ significantly from DOM announced by IMD. However, such coverage may have to be considered for certain policy decisions and applications.

We note that our method successfully avoids the false onsets; in fact the errors for years like 1959, 1968, 1995 (figure 2) characterized by false onsets are quite low. Another important point to note is that errors for the El-Nino years are not particularly large; this has important implication for forecasting DOM.

The criterion of PoOP of rainfall, of course, can be used effectively only through forecasts (at least short-term) of rainfall over the onset domain. This is a challenging task given the current difficulties in predicting monsoon rainfall. The effectiveness of the criterion of post onset persistence in avoiding bogus onset, however, makes its inclusion highly desirable. Further, it may be noted that such forecasts only need to be accurate in forecasting coverage, persistence and category (significance) rather than precise forecasts of spatio-temporal distribution of rainfall; this will be examined in a subsequent work.

References:

Ananthakrishnan, R., and M. K. Soman (1988), The onset of the southwest monsoon over Kerala: 1901–1980. *J. Clim.,* **8,** 283–296.

Ananthakrishnan, R., and M. K. Soman (1991), The onset of the south-west monsoon in 1990, *Curr. Sci*., **61 (7)** , 447-453.

Flatau, M. K., P. J. Flatau, and D. Rudnick (2001), The dynamics of double monsoon onsets. *J. Climate*, **14**, 4130-4146.

Gadgil, S (2003), The Indian monsoon and its variability, *Annu. Rev. Earth Planet. Sci.*, **31**, 429-467.

Goddard, L., M.J. Zebiak, S.E. Ropelewski, C.F. Basher, and M. A. Cane (2001), Current Approach to seasonal and interannual climate predictions, *Int. Jou. Climatology,* **21***,* 1111-1152.

Hsu, H. H., C. T. Terng., and C. T. Chen (1999), Evolution of large-scale circulation and heating during the first transition of Asian summer monsoon. *J. Clim.*, **12**, 793-810.

Joseph, P. V., J. Eisheid, and R. J. Pyle (1994), Interannual variability of the onset of the Indian summer monsoon and its association with atmospheric features, El Nino, and sea surface temperature anomalies. *J. Clim.*, **7**, 81-105.

Kang, I. S., et al. (2002), Intercomparison of the climatological variation of Asian summer monsoon precipitation simulated by 10 GCMs. *Clim. Dyn.,* **19***,*383-395.

Kumar, K. K., K. R. Kumar, and G. B. Pant (1997), Pre-monsoon maximum and minimum temperatures over India in relation to the summer monsoon rainfall. Int*. J. Clim.,* **17**, 1115-1127.

Lau, K.M., H. T. Wu, and S. Yang (1998), Hydrologic processes associated with the first transition of the Asian summer monsoon: A pilot satellite study. *Bull. Amer*. *Meteor. Soc*., **79**, 1871-1882.

Murakami, T., and T. Nakazawa (1985), Transition from the Southern to Northern Hemisphere summer monsoon*. Mon. Weather Rev*., **113**, 1470-1486.

Pearce, R. P., and U. C. Mohanty (1984), Onsets of the Asian summer monsoon 1979-82. *J. Atmos. Sci*., **41**, 1620-1639.

Rajeevan, M., J. Bhate, J.D. Kale, and B. Lal (2006), High resolution daily gridded rainfall data for the Indian region, *Current Science*, **91** (3), 296-306.

Ramage, C. (1971), *Monsoon Meteorology*. *International Geophysics Series*, Vol. **15**, Academic Press, San Diego.

Rao, Y. P. (1976), Southwest monsoon: Synoptic meteorology. *Meteor. Monogr.*, No.1/1976, India Meteorological Department, 367pp

Sikka, D.R., and S.Gadgil (1980), On the maximum cloud zone and the ITCZ over Indian longitude during the southwest monsoon, *Mon. Weather Rev.,* **108**, 1840-1853.

Soman, M. K., and K. K. Kumar, (1993), Space-time evolution of meteorological features associated with the onset Indian summer monsoon. *Mon. Weather. Rev*., **121**, 1177-1194.

Sperber, K. R., and T.N.Palmer (1996), Interannual Tropical Rainfall Variability in General Circulation Model Simulations Associated with the Atmospheric Model Intercomparison Project. *J. Clim.,* **9***,* 2727-2750.

Takagi, T., F. Kimura, and S. Kono (2000), Diurnal variation of GPS precipitable water at Lhasa in premonsoon and monsoon periods. *J. Meteor. Soc. Japan*, **78**, 175-180.

Tanaka, M., (1992), Intraseasonal oscillation and the onset and retreat dates of the summer monsoon over the east, southeast and western North Pacific region using GMS high cloud amount data*. J. Meteor. Soc. Japan*, **70**, 613-629.

Ueda, H., and T. Yasunari (1998), Role of warming over the Tibetan Plateau in early onset of the summer monsoon over the Bay of Bengal and the South China Sea. *J. Meteor. Soc. Japan*, **76**, 1-12.

Wang B., and LinHo (2002), Rainy season of the Asian-Pacific summer monsoon. *J. Climate*, **15**, 386-398.

Webster, P. J., (1983), Mechanisms of monsoon transition: Surface hydrology effects*. J. Atmos. Sci*., **40**, 2110-2124.

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.

Wu, G. X., and Y. S. Zhang (1998) Tibetan Plateau forcing and the timing of the monsoon onset over South Asia and the South China Sea. *Mon. Weather. Rev*., **126**, 913-927.