

Were the 2010 Pakistan floods predictable?

P. J. Webster,¹ V. E. Toma,¹ and H.-M. Kim¹

Received 30 November 2010; revised 6 January 2011; accepted 10 January 2011; published 22 February 2011.

[1] During July 2010, a series of monsoonal deluges over northern Pakistan resulted in catastrophic flooding, loss of life and property and an agricultural crisis that may last for years. Was the rainfall abnormal compared to previous years? Furthermore, could a high probability of flooding have been predicted? To address these questions, regional precipitation is analyzed using three dataset sets covering the 1981–2010 time period. It is concluded that the 2010 average May to August (MJJA) rainfall for year 2010 is somewhat greater in magnitude than previous years. However, the rainfall rate of the July deluges, especially in North Pakistan was exceptionally rare as deduced from limited data. The location of the deluges over the mountainous northern part of the country lead to the devastating floods. The European Centre for Medium Range Weather Forecasts (ECMWF) 15-day Ensemble Prediction System (EPS) is used to assess whether the rainfall over the flood affected region was predictable. A multi-year analysis shows that Pakistan rainfall is highly predictable out to 6–8 days including rainfall in the summer of 2010. We conclude that if these extended quantitative precipitation forecasts had been available in Pakistan, the high risk of flooding could have been foreseen. If these rainfall forecasts had been coupled to a hydrological model then the high risk of extensive and prolonged flooding could have anticipated and actions taken to mitigate their impact. **Citation:** Webster, P. J., V. E. Toma, and H.-M. Kim (2011), Were the 2010 Pakistan floods predictable?, *Geophys. Res. Lett.*, 38, L04806, doi:10.1029/2010GL046346.

1. Introduction

[2] Two main factors control South Asian rainfall. On 2–5 year time scales, the El Niño–Southern Oscillation (ENSO) phenomena is associated with above average summer precipitation during a La Niña and deficits during an El Niño [Shukla and Paolina, 1983; Kumar *et al.*, 2006]. Far more dramatic and higher amplitude modulations occur on subseasonal time scales. Over much of Asia the summer monsoon is divided into a series of “active” (rainy) and “break” (dry) periods following a roughly 20–40 days cycle [Lawrence and Webster, 2001; Webster and Hoyos, 2004; Hoyos and Webster, 2007] associated with the boreal summer Madden-Julian Oscillation [Madden and Julian, 1972] that produce a northeasterly excursion of large-scale convective anomalies under the action of a strong cross-equatorial pressure gradient [Stephens *et al.*, 2004; Wang *et al.*, 2005,

2006]. The arrival of convection over the Indian subcontinent heralds an active pluvial period. Summer rainfall in Pakistan is also monsoonal and, as such, has active and break periods. However, the total summer rainfall is far less than in the east (Figure 1a) decreasing from the Bay of Bengal (16 mm/day) across the plains of northern India (8–10 mm/day) to values of about 6–8 mm/day in northern Pakistan. Pakistan is at the western edge of the pluvial region of the monsoon.

[3] During the late boreal spring of 2010, the tropical Pacific Ocean entered a La Niña phase and during July 2010 the monsoon over the northern part of the Indian subcontinent was “active” with rainfall extending across the Gangetic Plains between the Bay of Bengal in the east to northern Pakistan in the west (Figure S1 of the auxiliary material).¹ Embedded in this active period were the deluges that caused the devastating floods in Pakistan. In late July, some Pakistan stations recorded rainfall amounts exceeding 300 mm over a four-day period (http://www.pakmet.com.pk/FFD/index_files/rainfalljuly10.htm). During the following days and weeks, flooding extended through the entire Indus Valley eventually reaching the Arabian Sea leaving behind a wake of devastation and destruction. In the end, the death toll was close to 2000 and over 20 million people were affected. An estimated 20,000 cattle were drowned. Power stations and transmission towers were destroyed along with other major infrastructure such as barrages, bridges and roads. Irrigation systems were destroyed and planting of subsequent crops delayed or abandoned with agricultural costs exceeding \$US500M. Overall, estimates of damage exceed \$US40B. In general, it was the poor that suffered the most and many of these will face the prospect of intergenerational poverty as a result of the floods [Webster and Jian, 2011]. Most assessments of the 2010 Pakistan floods have appeared on the internet and in relief organization reports (<http://www.pakistanfloods.pk/>; http://en.wikipedia.org/wiki/2010_Pakistan_floods). Eventually, scholarly articles on the flooding will be forthcoming discussing, in more detail, the climate and meteorological conditions that produced the flooding (e.g., R. A. Houze Jr. *et al.*, Anomalous atmospheric events leading to the summer 2010 floods in Pakistan, submitted to *Bulletin of the American Meteorological Society*, 2010). However, to date there has been an absence of any comment about the predictability of the deluges or the associated risk of floods. Eventually, skill in predicting floods reduces to the predictability of precipitation and the use of an adequately sophisticated hydrological model. Thus, an immediate and critical question is the degree to which rainfall at the western edge of the South Asian monsoon system is predictable on time scales of 1–2 weeks. Is the predictability of precipitation in the western

¹School of Earth and Atmospheric Science, Georgia Institute of Technology, Atlanta, Georgia, USA.

¹Auxiliary materials are available in the HTML. doi:10.1029/2010GL046346.

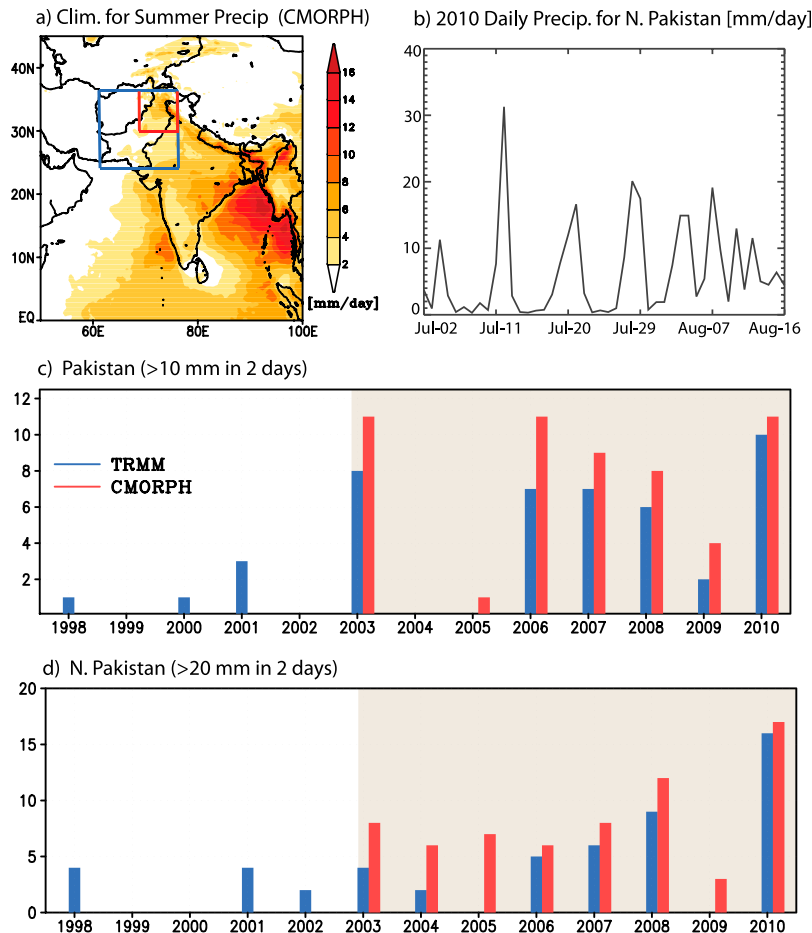


Figure 1. (a) May–August CMORPH precipitation [mm/day] climatology for 2003–2010. (b) The observed CMORPH rainfall averaged for the Northern Pakistan (70°N – 74°N ; 30°E – 36°E , red rectangle). Number of heavy rainfall events over the summer (May–August) in TRMM (blue) and CMORPH (red). The events are defined when two-day accumulated rainfall exceeds (c) 10 mm over entire Pakistan (62°N – 74°N ; 24°E – 36°E , blue rectangle in Figure 1a) and (d) 20 mm over the northern Pakistan (70°E – 74°E , 30°N – 36°N , red rectangle in Figure 1a). Years with available CMORPH data are shaded in gray.

edge of the monsoon comparable to that seen over the Ganges and Brahmaputra basins [Hopson and Webster, 2010; Webster *et al.*, 2010]?

[4] In this study we focus on the predictability of 1–15-day ECMWF EPS forecasts [Buizza *et al.*, 2007] over Pakistan. In the next section details of the observation and numerical model data are introduced. Section 3 discusses the uniqueness of the July–August flooding events and examines the prediction skill of 15-days rainfall forecast followed by conclusions related to the predictability of floods in Pakistan.

2. Data and Analysis

[5] Three precipitation data sets are used to assess the variability of the precipitation over the Pakistan region: the Global Precipitation Climatology Project (GPCP) data [Adler *et al.*, 2003] for the 1981–2009 period, the Tropical Rainfall Measuring Mission (TRMM) [Huffman *et al.*, 2005, 2007] TRMM_3B42 product for 1998–2010, and the NOAA CPC Morphing Technique (CMORPH) Precipitation Product for the 2003–2010 period [Joyce *et al.*, 2004]. GPCP (a merging of rain gauge data with satellite geostationary and low-orbit infrared and passive microwave

information) and TRMM data sets (specifically the TRMM_3B42 set) were chosen for their temporal extension (29 and 13 years, respectively). All of these precipitation products had a $0.25^{\circ} \times 0.25^{\circ}$ horizontal resolution facilitating a comparison with model output. Figure S4 shows time series of monthly rainfall anomalies for each of the data sets.

[6] A comparison of the CMORPH and TRMM data sets (Figures S3) reveals considerable differences in the magnitude of estimates of precipitation during the third precipitation pulse of July 2010 that occurred over the higher terrain of northern Pakistan (Figure S2d). The TRMM rainfall estimate was considerably higher than CMORPH by about a factor of two consistent with the discussion of Gopalan *et al.* [2010] who suggested that TRMM may overestimate precipitation rates over substantial terrain. Comparisons during earlier periods, when the precipitation maxima occurred over the plains of southern Pakistan and northwestern India are more comparable (Figure S2). Consequently, we use CMORPH as the principal data set for determining the sequence of events during 2010 and also as the principal agent for the statistical rendering of the quantified precipitation forecasts.

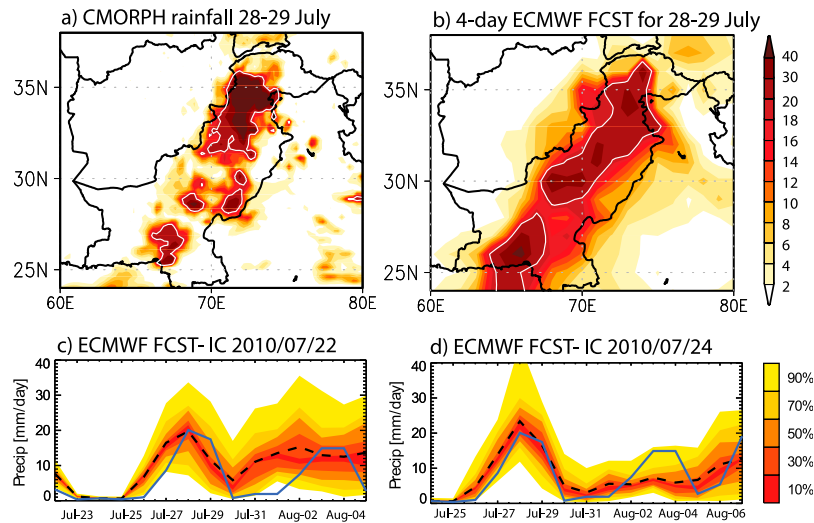


Figure 2. Total precipitation [mm/day] for (a) CMORPH over 28–29 July 2010 and (b) ECMWF ensemble mean of the forecast initialized four days previously (July 24, 2010) for the same time period. White contour shows 20 mm/day. ECMWF 15-day forecast of the precipitation [mm/day] in the red rectangle (Figure 1a) initialized on July (c) 22nd, and (d) 24th, 2010. Black dashed line shows the ensemble mean. Colored shading depicts the probability of precipitation rate based on the 51 ensemble members. Dark blue line represents the observed CMORPH precipitation averaged for the same region.

[7] The ECMWF EPS forecasts consist of 51 ensemble members initialized twice per day (00 and 12 UTC), each ensemble member having a 15-day forecast horizon. The horizontal resolution of the model is 50×50 km from 0 to 10 days and then $80 \text{ km} \times 80 \text{ km}$ from day 10–15 [Buizza *et al.*, 2007]. For this initial study, model forecast precipitation for the months of July and August from 2007 to 2010 was converted into daily cumulative amounts. To minimize systematic model bias differences between the distributions of the ECMWF forecasts and the observed rainfall, a quantile-to-quantile (*q-to-q*) mapping technique was implemented following Hopson and Webster [2010] and Webster *et al.* [2010] (see method description in the auxiliary material). All rainfall forecasts presented here are adjusted using the *q-to-q* technique.

3. Results

[8] Beginning in early July 2010, there were six major pulses of torrential rainfall occurring over Pakistan, each separated by about a week (Figure 1b). One of the most intense periods occurred between July 27–30 over the mountainous regions of the north. Figure S2 shows the distribution of rainfall for the major pulses of monsoon rain. The earlier rainfall events caused flooding in Balochistan in central Pakistan. Flooding followed across northern Pakistan in the Khyber Pakhtunkhwa province with the later July rains extending to the Punjab in late July/early August (<http://www.unitar.org/unosat/node/44/1469>). Here we address the uniqueness and predictability of the floods.

3.1. Uniqueness

[9] There have been 67 reported flooding events in Pakistan occurring since 1900 with a clustering of 52 events of various severity in the last 30–40 years (International Disaster Data Base, <http://www.emdat.be>). Some of these events (e.g., 1950, 1973, 1976, 1977, 1992, 2001, 2007 and

2008) were also accompanied with large loss of life and property. This recent increase is consistent with the increase in intensity of the global monsoon accompanying the last three decades of general global warming (B. Wang *et al.*, Recent intensification of global monsoon and precipitation, submitted to *Nature*, 2011) or perhaps changes in water management strategies, increases of damage due to a rapidly growing population or improved reporting through advances in communication.

[10] Figure S4 shows the temporal variability of seasonal (MJJ) precipitation averaged in Pakistan (62° – 74° E, 24° – 36° N, blue rectangle in Figure 1a) and northern Pakistan (70° E– 74° E, 30° N– 36° N, red rectangle in Figure 1a) relative to the seasonal climatology for each of the data sets: GPCP and CMORPH. While there are amplitude differences between datasets, each shows substantial variability, with seasons of excessive rainfall and drought occurring irregularly over the past 30 years (Figures S3 and S4).

[11] Were the rainfall events of 2010 worse than previous extreme events? Using a 13-year TRMM precipitation record, extreme events can be counted. An extreme event is defined here to occur when the two-days accumulated rainfall exceeds over 10 mm over all Pakistan and 20 mm over the northern Pakistan (Figures 1c and 1d). Note that the chosen thresholds for this analysis are much smaller than maximum daily rainfall measurements at specific stations (see http://www.pakmet.com.pk/FFD/index_files/rainfalljuly10.htm) due to a broader averaging area. Although there is considerable interannual variability, the number of extreme events over entire Pakistan, so defined, is larger in 2010 than in previous years, greater, for example than in 2008. In summary, 2010 stands out as a period of above average rainfall events over northern Pakistan. The number of extreme events over northern Pakistan is far more unique which, based on the very limited TRMM data set would have return periods of >30 years. Long-term variability for extreme events is calculated with GPCP pentad data set from

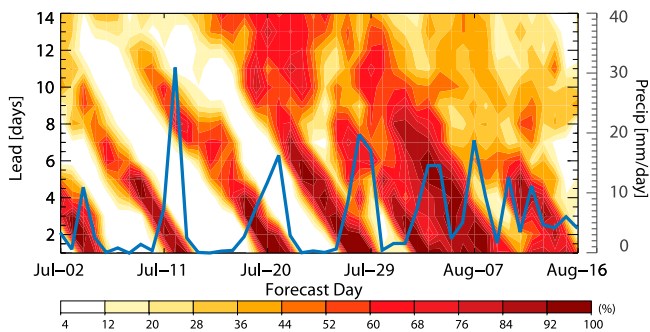


Figure 3. Forecast lead time diagram of the probability that the ECMWF forecast for the red region (Figure 1a) exceeds the observed CMORPH July–August climatology plus 1 standard deviation. The blue line represents the observed CMORPH rainfall [mm/day] averaged for the same region and the same time period (units on the left axis). The July 12 and June 21 events were forecast at 50% probability level to exceed observed climatology plus 1 standard deviation, at least 10 days in advance. An additional 2 days of predictability is evident at the 50% level for the July 28 event. All events were forecast at >70–80% level of probability 6 days in advance. Note that the ECMWF forecast is adjusted using the q -to- q technique.

1981 (Figure S5) to 2007 overlapped with CMORPH pentad from 2003 to 2010. Although, there are differences between data sets, the high occurrence of Northern Pakistan extreme events in 2010 is relatively rare. Rainfall data is not sufficiently reliable prior to 1987 when GPCP data was generated on a daily basis. However, we do have CMORPH and TRMM data for 2008. As shown in Figure S6, the cumulative July–August rainfall for northern Pakistan is larger in 2010 than 2008, with values larger than 0.5 m in several areas.

3.2. Predictability

[12] The next step is to examine the predictability of the rainfall events depicted in Figure 1b. Figure 2a shows the total average precipitation [mm/day] for July 28–29, based on the CMORPH observational dataset and the ECMWF forecast ensemble mean initialized 4 days before the event (Figure 2b). The q -to- q correction was applied to the precipitation forecast data. The forecasts compare well with the observed rainfall with ECMWF slightly underestimating the rainfall intensity in the northern part of the region. The ECMWF forecast showed average precipitation larger than 40 mm/day in some areas which is over 3 times larger than the CMORPH climatological average for the region.

[13] Figures 2c and 2d show the temporal evolution of the ECMWF forecast commencing on 22nd and 24th July, 2010 through August 9, 2010 for the Khyber-Pakhtunkhwa province, located in the north west of the country (red rectangle in Figure 1a). The diagram shows the probability distribution of precipitation based on the 51 ensemble members with the ensemble mean plotted as the black dotted line. The blue line represents the CMORPH observed rainfall. Good predictive skill of the July 28–29 event is found up to 6 days in advance. The same analysis done for various other monsoon pulses have resulted in similar conclusions (Figure 3).

[14] Figure S7 shows an assessment of precipitation predictability in northern Pakistan using all available hindcast data. Predictability is shown as correlations between predicted and observed CMORPH rainfall values as a function of lead time for July based on 2007–2010 period. Note that for 2007, the model prediction extends only up to 10 days but up to 15 days for the 2008–2010 period. Correlations ≥ 0.7 were found for predictions 5 days in advance indicating useful predictive skill. Thus, the quantitative rainfall forecasts could be used as a robust variable in a flood forecasting scheme for Pakistan region.

[15] In order evaluate whether the model can provide useful information with regards to the actual severity of the major rainfall events of July–early August 2010, all ECMWF forecasts made during the period were extracted and bias corrected. Then, the probability that the predicted rainfall would exceed the observed climatological average plus 1 standard deviation was computed. In other words, for each forecast, at each lead time, the percentage of ensemble members exceeding the threshold was computed. The exceedance threshold is calculated using 2003–2010 CMORPH data, with mean and standard deviation based on July–August daily average data. Results are shown in Figure 3 as shaded contours. The blue line represents the observed CMORPH rainfall averaged for the same region and the same time period. For example, the July 28 event was predicted almost 8 days in advance with a probability >60% over the climatological average plus 1 standard deviation (Figure 3). All the other events appear to have similar skill at the 8 to 10 day horizon.

4. Conclusions

[16] From a climatological perspective, July and August precipitation rates were above average in Pakistan although not exceptionally so. However, in terms of rainfall rate, the monsoon pulses were extreme events compared to other years in the period 1998–2010. The devastating flooding occurred from a conspiracy of events. The summer of 2009 was a severe drought period with rainfall well below average (<http://www.pakmet.com.pk/monsoon2009ver.pdf>) so that vegetation may have been sparser during 2010. The region is mountainous with steep valleys and ridges. Furthermore, deforestation in northern Pakistan has been severe [e.g., *Ali et al.*, 2006]. Deforestation and sparse undergrowth would exacerbate runoff through the steep valleys of the heavy rains that occurred during the month of July and early August.

[17] The major result of the study is that the heavy rainfall pulses throughout July and early August were predictable with a high probability 6–8 days in advance. If these forecasts had been available to the regions of northern Pakistan, government institutions and water resource managers could have anticipated rapid filling of dams, releasing water ahead of the deluges. A high probability of flooding could have been anticipated.

[18] Finally, it appears that Pakistan would benefit from a hydrological forecasting scheme similar to that developed for Bangladesh [*Hopson and Webster*, 2010; *Webster et al.*, 2010]. The Bangladesh system incorporates the same form of statistically rendered ensemble precipitation forecasts as discussed above but coupled to a hybrid hydrological model. Working with Government of Bangladesh authorities, these 10-day river forecasts were communicated to the union

(county) and village level allowing time to prepare for the floods for three major Brahmaputra floods during 2007/8 allowing the saving of household and agricultural effects and the successful evacuation of those in peril [Webster et al., 2010; Webster and Jian, 2010] (ADPC, Flood forecasts application for disaster preparedness: Post flood forecasts assessment 2008: Community response to CFAN forecasts, 2009, available at <http://www.adpc.net/v2007/> and <http://pacific.eas.gatech.edu/~pjw/FLOODS>).

[19] **Acknowledgments.** This research has been supported by the Climate Dynamics Division of the National Sciences Foundation under Award NSF-ATM 0965610. Once again, we are indebted to ECMWF for providing data to make this analysis possible. We wish to thank J. A. Curry for interesting discussions as summarized at <http://judithcurry.com/2010/09/20/pakistan-on-my-mind/and-comments-on-the-paper>.

[20] Noah Diffenbaugh thanks the two anonymous reviewers.

References

- Adler, R. F., et al. (2003), The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), *J. Hydrometeorol.*, *4*, 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.
- Ali, Y., B. Shahbaz, and A. Suleri (2006), Analysis of myths and realities of deforestation in northwest Pakistan: Implications for forestry extension, *Int. J. Agric. Biol.*, *8*(1), 107–110.
- Buizza, R., J.-R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart (2007), The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System), *Q. J. R. Meteorol. Soc.*, *133*, 681–695, doi:10.1002/qj.75.
- Gopalan, K., N.-Y. Wang, R. Ferraro, and C. Liu (2010), Status of the TRMM 2A12 land precipitation algorithm, *J. Atmos. Oceanic Technol.*, *27*, 1343–1354, doi:10.1175/2010JTECHA1454.1.
- Hopson, T. M., and P. J. Webster (2010), A 1–10 day ensemble forecasting scheme for the major river basins of Bangladesh: Forecasting severe floods of 2003–2007, *J. Hydrometeorol.*, *11*, 618–641, doi:10.1175/2009JHM1006.1.
- Hoyos, C. D., and P. J. Webster (2007), The role of intraseasonal variability in the nature of Asian monsoon precipitation, *J. Clim.*, *20*, 4402–4424, doi:10.1175/JCLI4252.1.
- Huffman, G. J., R. F. Adler, S. Curtis, D. T. Bolvin, and E. J. Nelkin (2005), Global rainfall analyses at monthly and 3-hr time scales, in *Measuring Precipitation From Space: EURAINSAT and the Future*, edited by V. Levizzani, P. Bauer, and J. F. Turk, pp. 291–306, Springer, Dordrecht, Netherlands.
- Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker (2007), The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined sensor precipitation estimates at fine scales, *J. Hydrometeorol.*, *8*, 38–55, doi:10.1175/JHM560.1.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie (2004), CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, *J. Hydrometeorol.*, *5*, 487–503, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.
- Kumar, K. K., B. Rajagopalan, M. Hoerling, G. Bates, and M. Cane (2006), Unraveling the mystery of Indian monsoon failure during El Niño, *Science*, *314*(5796), 115–119, doi:10.1126/science.1131152.
- Lawrence, D. M., and P. J. Webster (2001), Interannual variations of the intraseasonal oscillation in the South Asian summer monsoon region, *J. Clim.*, *14*, 2910–2922, doi:10.1175/1520-0442(2001)014<2910:IVOTIO>2.0.CO;2.
- Madden, R., and P. Julian (1972), Description of global-scale circulation cells in the tropics with a 40–50 day period, *J. Atmos. Sci.*, *29*, 1109–1123, doi:10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2.
- Shukla, J., and D. A. Paolina (1983), The southern oscillation and long range forecasting of the summer monsoon rainfall over India, *Mon. Weather Rev.*, *111*, 1830–1837, doi:10.1175/1520-0493(1983)111<1830:TSOALR>2.0.CO;2.
- Stephens, G. L., P. J. Webster, R. H. Johnson, R. Engelen, and T. L'Ecuyer (2004), Observational evidence for the mutual regulation of the tropical hydrological cycle and tropical sea surface temperatures, *J. Clim.*, *17*, 2213–2224, doi:10.1175/1520-0442(2004)017<2213:OEFTMR>2.0.CO;2.
- Wang, B., P. J. Webster, and H. Teng (2005), Antecedents and self-induction of active-break South Asian monsoon unraveled by satellites, *Geophys. Res. Lett.*, *32*, L04704, doi:10.1029/2004GL020996.
- Wang, B., P. Webster, K. Kikuchi, T. Yasunari, and Y. Qi (2006), Boreal summer quasi-monthly oscillation in the global tropics, *Clim. Dyn.*, *27*, 661–675, doi:10.1007/s00382-006-0163-3.
- Webster, P. J., and C. D. Hoyos (2004), Prediction of monsoon rainfall and river discharge on 15–30 day time scales, *Bull. Am. Meteorol. Soc.*, *85*, 1745–1765, doi:10.1175/BAMS-85-11-1745.
- Webster, P. J., and J. Jian (2011), Probability, uncertainty and prediction: A pathway towards the alleviation of poverty in the developing world, *Philos. Trans. R. Soc.*, in press.
- Webster, P. J., J. Jian, T. M. Hopson, C. D. Hoyos, P. A. Agudelo, H.-R. Chang, J. A. Curry, R. L. Grossman, T. N. Palmer, and A. R. Subbiah (2010), Extended-range probabilistic forecasts of Ganges and Brahmaputra floods in Bangladesh, *Bull. Am. Meteorol. Soc.*, *91*, 1493–1514, doi:10.1175/2010BAMS2911.1.

H.-M. Kim, V. E. Toma, and P. J. Webster, School of Earth and Atmospheric Sciences, Georgia Institute of Technology, 311 Ferst Dr., Atlanta, GA 30332, USA. (pjw@eas.gatech.edu)