

*Will the Weather Gods Smile or Frown? Evaluating Monsoon Forecasts**

The India Meteorological Department (IMD) has predicted normal and well distributed South West Monsoon rainfall for 2019, which is at odds with the forecasts of private and international agencies. This article drills analytically into past patterns of prediction-outcome performances of various agencies through multiple statistical measures to evaluate their forecast accuracy. The comparative assessment suggests that for generating macroeconomic forecasts, the use of IMD's second stage long range forecast and the predictions of international agencies in conjunction may be appropriate as the preliminary forecasts of IMD and Skymet released in April appear to be noisy.

Introduction

India receives 75 per cent of its annual rainfall during the South West Monsoon (SWM) season, which is spread over June to September. With 65 per cent of Gross Cropped Area (GCA) in the country not under any irrigation cover and even different sources of irrigation depending on rainfall during the SWM season to build up storage levels (GOI, 2018), the SWM remains the life force of India's agriculture and for the broader economy¹ (Gulati, *et al.*, 2013). Furthermore, the SWM influences both *Kharif* (June-September) production through its temporal and spatial distribution, and *Rabi* (October-March) production through its impact on soil moisture conditions and storage levels in reservoirs. Consequently, SWM forecasts remain critical to every assessment of India's macroeconomic outlook.

* This article is prepared by Priyanka Bajaj, D. Suganthi, Rishabh Kumar and Atri Mukherjee of the Division of Rural Economics, Department of Economic and Policy Research, Reserve Bank of India. The views expressed in this article are those of the authors and do not represent the views of the Reserve Bank of India.

¹Share of Agriculture and allied sector in India's total Gross Value Added is 15.3 percent (average of last five years).

At present, the official forecast of the SWM is given by the IMD, first in April - the First Stage Long-Range Forecast (FSLRF) - and again in May/June - Second Stage Long-Range Forecast (SSLRF). Skymet, a private forecaster, releases its preliminary forecast in April and a revision in May. Apart from these Indian agencies, the forecasts of international meteorological organisations, *viz.*, USA's National Oceanic and Atmospheric Administration (NOAA) and Australia's Bureau of Meteorology (BOM) on *El Nino/La Nina* and dipole conditions in the Pacific Ocean and Indian Ocean, respectively, are also used to predict the monsoon conditions in India.

Weather forecasters contend with enormous complexities in no small measure, which has catalysed extraordinary progress in the field with supercomputers and satellite images enabling high frequency precise measurement of temperature, air pressure, humidity, wind speed and direction (Silver, 2012). India received a significant boost to its weather forecasting capacity in 2018 with the Pratyush and Mihir supercomputers, elevating it next only to Japan, the UK and the US in terms of dedicated high capacity computing. The IMD's precise prediction of the cyclone *Fani* in Odisha in 2019 and resultant disaster risk reduction received global acclaim, including from the United Nations.

Evaluating forecast performance in respect of the SWM is the key motivation of the article. At the outset itself, it is sensitive to the reality that the science of weather has enabled enormous progress even on long-range forecasts, but no forecast is free of errors. Recognising this caveat, the purpose of the article is to drill analytically into past patterns of prediction-outcome performance in order to draw some meaningful insights from the IMD's forecasts. Both FSLRF and SSLRF released in April and May, have predicted rainfall during June-September 2019 to be normal at 96 per cent of the Long Period Average

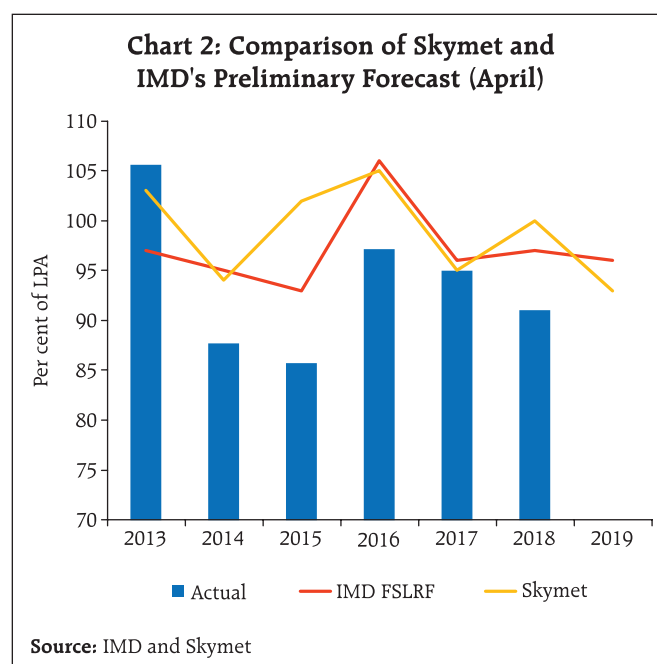
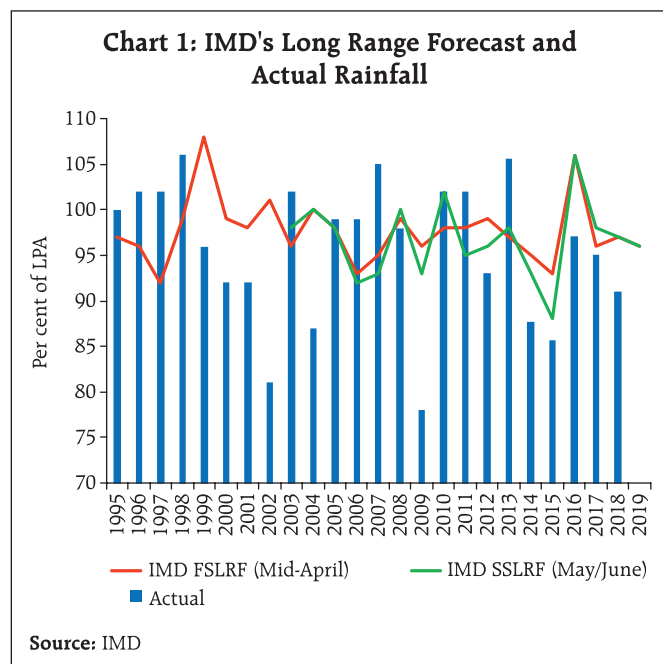
(LPA) with a model error of ± 5 and ± 4 per cent, respectively.² The IMD has also predicted the rainfall to be well distributed across the states. This, however, is at odds with forecasts of Skymet³, NOAA and BOM⁴, which predict a weak monsoon this year. The rest of the paper is structured as follows. Section II presents some stylised facts relating to monsoon forecasts by different agencies. Data, methodology and empirical results are discussed in Section III. Concluding remarks and some policy perspectives are set out in Section IV.

II. Some Stylised Facts

The IMD introduced SSLRF in 2003, after the failure of FSLRF to predict the massive drought of 2002. Out of the 24 years since 1995, FSLRF over-predicted rainfall in 13 years and under-predicted it in 11 years. In the case of the drought years (2002, 2009, 2014 and

2015)⁵ and deficient rainfall year (2004), forecast errors (actual minus predicted) for FSLRF were negative and large in value (Chart 1). SSLRF is more broad-based than FSLRF in that it provides rainfall forecasts for the four broad geographical regions of India as well as for the country as a whole. The assumed model error for the four regional forecasts is ± 8 per cent of the LPA. The SSLRF has out-performed the FSLRF: it could successfully predict the drought of 2015, which was missed out by the FSLRF.

A comparison of IMD's FSLRF with the predictions of Skymet⁶ (both are released in April) reveals that both the agencies have missed the target in 5 out of the last 6 years. The year 2017 was an exception - the forecast was exactly the same as actual for Skymet, and very close to actual in the case of the IMD's FSLRF (Chart 2).



²The LPA (1951-2000) for the country is 89 cm.

³Skymet has predicted a below normal monsoon of 93 per cent of LPA with an error margin of ± 5 per cent. Spatially, it expects less than normal rainfall for all the four regions of the country.

⁴Climate Diagnostics Bulletin (various issues), National Weather Service Climate Prediction Centre, National Oceanic and Atmospheric Administration, United States Department of Commerce. (Retrieved from https://www.cpc.ncep.noaa.gov/products/CDB/CDB_Archive_html/CDB_archive.shtml)

⁵All-India drought years are announced when the rainfall deficiency is more than 10 per cent of LPA and 20-40 per cent area of the country is under drought conditions.

⁶Performance of the revised outlook of Skymet could not be evaluated due to paucity of data as it was introduced only in 2017. Incidentally, it did not revise its prediction for 2017 and 2018.

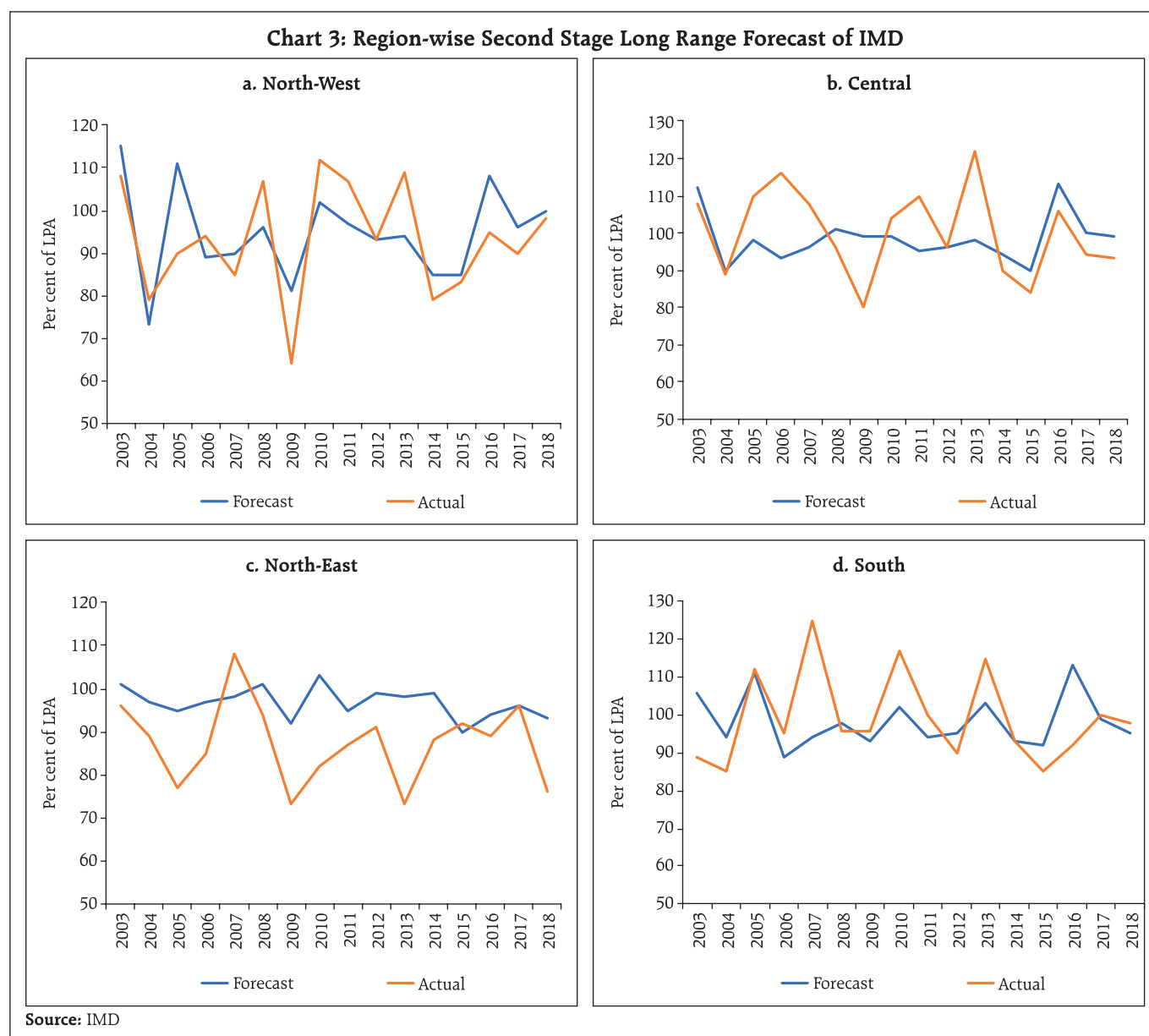
Regional Forecast Performance

Often, the all-India forecast masks the details - spatial distribution of rainfall - which assumes crucial importance for the outlook on agricultural production. For instance, in the year 2007, the overall rainfall in India was 105 per cent of LPA while the North West region received only 85 per cent of LPA. The largest deviations between the IMD's SSLRF and actual rainfall are in relation to the North-Eastern states of India than for the other regions (Chart 3)⁷. For

Skymet, which started its regional forecast since 2018, deviation of actual rainfall from forecast is lower for North West and Southern regions than for Central and North Eastern regions.

Forecast Performance of International Agencies

The SWM is significantly influenced by *El Nino* or *La Nina* conditions. During *El Nino*, sea level pressure tends to be lower in the Eastern Pacific and higher in the Western Pacific. This implies warmer ocean waters in the East and colder ocean waters in the



⁷The region-wise forecast under SSLRF is available only from 2003 onwards.

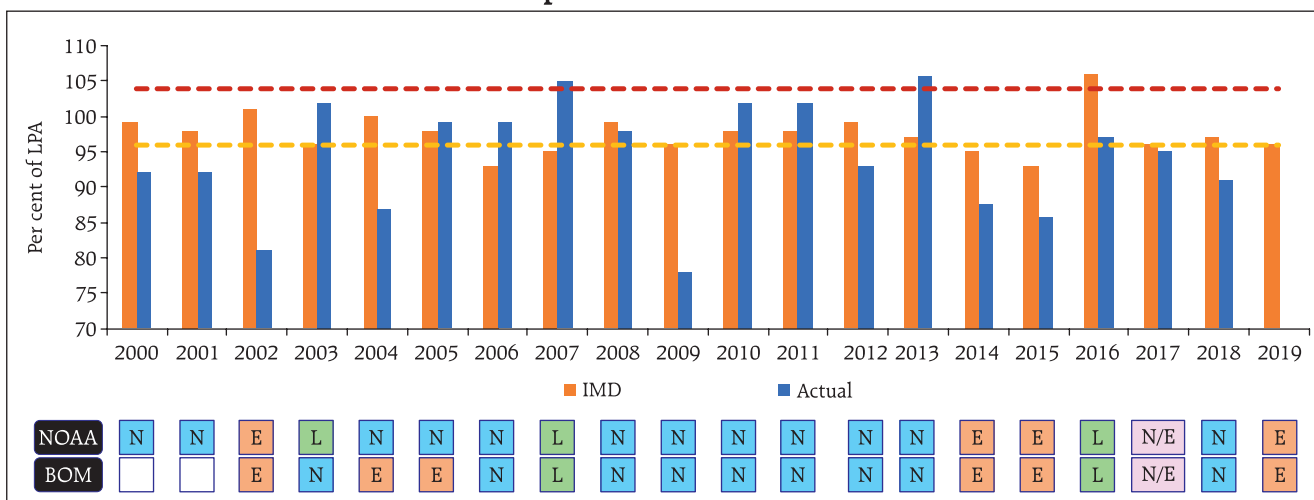
West. The opposite condition prevails during *La Nina*. The warming of ocean waters triggers a see-saw in atmospheric pressure between the Eastern and Western Tropical Pacific, known as Southern Oscillation (SO). Since *El Nino* and SO are related, the two terms are often combined into a single term '*El Nino* Southern Oscillation' or 'ENSO'. A warm ENSO phase signals *El Nino* and a cold ENSO phase is related with *La Nina*. *El Nino* is generally associated with deficient rainfall in India, whereas the development of *La Nina* tends to bring above normal monsoon (Rajeevan and Pai, 2006).

In addition to these developments in the Pacific Ocean, the SWM also depends on the developments which take place nearby in the Indian Ocean, known as Indian Ocean Dipole (IOD) conditions. A positive IOD occurs when the western basin of the Indian Ocean warms up relative to the tropical eastern basin. A negative IOD occurs when the reverse takes place. The occurrence of a positive IOD can reduce the impact of *El Nino*, which takes place in the far away Pacific Ocean and can bring in better rains for India. It is believed that positive IOD had facilitated normal rainfall in

1997 despite it being a strong *El Nino* year, while in 2014, a negative IOD and *El Nino* had worked together to produce deficient rainfall (Karumari, *et al.*, 2001). In 2015, however, a positive IOD could not counteract a very strong *El Nino* and the SWM eventually turned out to be deficient.

The predictions of ENSO and IOD are released by NOAA and BOM on a monthly and fortnightly basis, respectively.⁸ It is observed that the NOAA's forecasts of *El Nino/La Nina* through the spring (March-April) tend to be less reliable, while the outlook released in June seems to perform better. For instance, in 2017, the NOAA's April outlook predicted a 50 per cent probability for *El Nino*; the prediction status was changed to neutral in the June outlook and eventually it turned out as a year of normal rainfall. Similarly, in the severe drought year 2009, the April outlook predicted neutral conditions, which was changed to *El Nino* in the June outlook by both NOAA and BOM (with more than 50 per cent probability). In the drought years 2002, 2014 and 2015, both NOAA and BOM predicted *El Nino* in April and also in June, which turned out to be correct (Chart 4a and 4b).

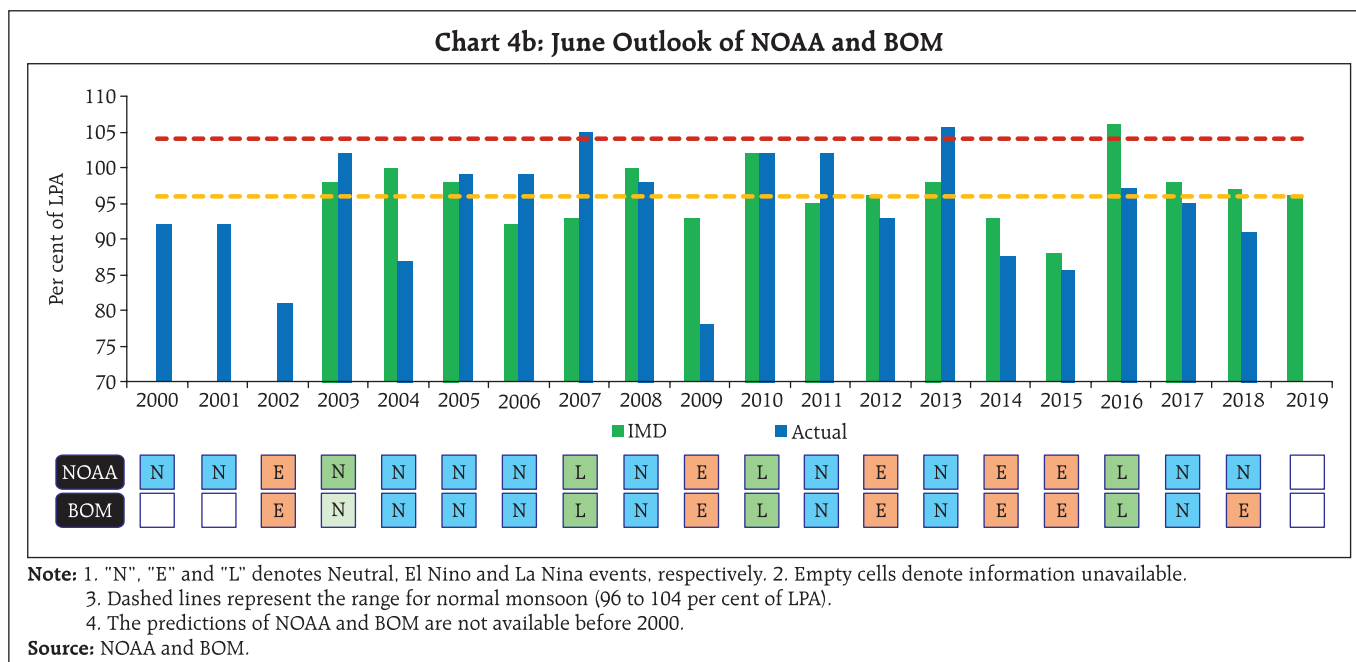
Chart 4a: April Outlook of NOAA and BOM



Note: 1. "N", "E" and "L" denotes Neutral, El Nino and La Nina events, respectively. 2. Empty cells denote information unavailable.
 3. Dashed lines represent the range for normal monsoon (96 to 104 per cent of LPA).
 4. The predictions of NOAA and BOM are not available before 2000.

Source: NOAA and BOM.

⁸ ENSO Wrap-Up: Current State of the Pacific and Indian Ocean (various issues), Australia Government Bureau of Meteorology. (Retrieved from <http://www.bom.gov.au/climate/ens/wrap-up/archive.shtml>)



For the 2019 SWM, both NOAA and BOM had predicted the chances of formation of *El Niño* to be high in their April outlook, given the above average warm temperature conditions in the Eastern Pacific Ocean. While the BOM had predicted 70 per cent chance of an *El Niño* developing this year, the NOAA had predicted occurrence of *El Niño* with 60 per cent probability during the June-August period. In the May outlook, the BOM has reduced the chances of *El Niño* developing in 2019 to 50 per cent, whereas the NOAA has revised it upwards to 70 per cent.

It has been observed that most of the severe droughts in India were influenced by *El Niño*. In the period between 1995 and 2018, there have been five extreme deficient rainfall years, four all-India drought years (2002, 2009, 2014, and 2015) and one deficient rainfall year (2004). All these five years were associated with *El Niño*, though the intensity of *El Niño* differed across the years. However, development of *El Niño* is not a sufficient condition for drought/deficient rainfall. Less than half of *El Niño* events are associated with deficient rainfall over India (Rajeevan and Pai, 2006). For instance, despite being a strong *El Niño*

year, near normal rainfall was received in 1997. On the other hand, only a moderate *El Niño* was observed in 2002, which resulted in one of the worst droughts (Table 1).

III. Evaluating SWM Forecasts

The forecast performance of various models of IMD has been examined in the literature (Rajeevan, *et al.*, 2004; Kar, *et al.*, 2012; Pandey, *et al.*, 2016; Prasad, *et al.*, 2010; Stern, 2008), including the accuracy of the forecast *vis-à-vis* actual rainfall either for the drought years or for a particular spatial area (Joseph, *et al.*, 2017; Sagar, *et al.*, 2017). However, comparisons of forecast accuracy of IMD *vis-à-vis* other private and international agencies has remained a gap.

Accordingly, the forecast accuracy of different agencies is evaluated through (i) Pearson's Correlation Coefficient (PCC); (ii) Root Mean Square Error (RMSE); and (iii) success score of predicting extreme events⁹.

⁹The success score is defined as the proportion of correct forecasts of extreme events to the total actual occurrence of extreme events expressed in percentage terms. This includes drought years (2002, 2009, 2014 and 2015), deficient rainfall year (2004) and above normal rainfall years (1998, 2007, 2013). The success score value ranges between 0 and 100 per cent where higher values denote better forecasting.

Table.1 Vulnerability of Monsoon During *El Nino* Years

Year	Occurrence	Indian Ocean Dipole (IOD)	Actual performance compared to normal	Monsoon (per cent departure from LPA)
1995	Moderate <i>La Nina</i>	Neutral	Near Normal	100
1996	Neutral	Negative	Near Normal	102
1997	Very Strong <i>El Nino</i>	Positive	Near Normal	102
1998	Strong <i>La Nina</i>	Negative	Above Normal	106
1999	Strong <i>La Nina</i>	Neutral	Near Normal	96
2000	<i>La Nina</i>	Neutral	Below Normal	92
2001	Neutral	Neutral	Below Normal	92
2002	Moderate <i>El Nino</i>	Neutral	Deficient	81
2003	Neutral	Neutral	Near Normal	102
2004	Weak <i>El Nino</i>	Neutral	Deficient	87
2005	Neutral	Neutral	Near Normal	99
2006	Weak <i>El Nino</i>	Positive	Near Normal	99
2007	Strong <i>La Nina</i>	Neutral	Above Normal	105
2008	<i>La Nina</i>	Neutral	Near Normal	98
2009	Moderate <i>El Nino</i>	Neutral	Severe Drought	78
2010	Strong <i>La Nina</i>	Negative	Near Normal	102
2011	Moderate <i>La Nina</i>	Neutral	Near Normal	102
2012	Neutral	Positive	Below Normal	93
2013	Neutral	Neutral	Above Normal	106
2014	Weak <i>El Nino</i>	Negative	Deficient	88
2015	Very Strong <i>El Nino</i>	Positive	Deficient	86
2016	<i>La Nina</i>	Negative	Near Normal	97
2017	Neutral	Neutral	Below Normal	95
2018	Neutral	Neutral	Below Normal	91

Source: NOAA, BOM and IMD

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=0}^n (\hat{y}_i - y_i)^2}$$

where, \hat{y}_i and y_i are the forecast and actual values, respectively, and n is the number of observations.

PCC and RMSE have been used to examine the forecast accuracy of IMD and Skymet as the predictions and actual values are reported as percentage of LPA. Success score has been used to assess the forecast

accuracy of all the four agencies (IMD, Skymet, NOAA and BOM) as the predictions are reported in terms of probability of occurrence by BOM and NOAA. For PCC and RMSE, the period of study is from 1995 to 2018, which is characterised by three distinct phases – Period 1 (1995 to 2002) with only FSLRF; Period 2 (2003 to 2012) with both FSLRF and SSLRF and Period 3 (2013 to 2018) with FSLRF, SSLRF and Skymet forecasts.

Analysis of Results

The PCC results reveal that the correlation coefficient between the IMD's FSLRF and actual rainfall for Period 1 is negative (-0.36) and insignificant (Table 2). For Period 2, the correlation coefficient of the FSLRF and SSLRF with actual rainfall is negative (-0.21) and positive (0.12), respectively, and both turn out to be insignificant. In Period 3, the correlation coefficient between the FSLRF and actual rainfall improves to positive and is insignificant (0.45). For the same period, the correlation coefficient of SSLRF improves substantially to 0.64 though it remains insignificant. The correlation coefficient of Skymet prognosis and actual rainfall is positive and insignificant (0.41). Clearly, none of the forecasts are significantly correlated with the actual rainfall data across the entire period of study to arrive at a conclusive inference.¹⁰

The RMSE results are similar to the correlation results. The RMSE of the IMD's FSLRF is higher than the SSLRF across time periods. In Period 3, the RMSE of both FSLRF and SSLRF improves though it still remains above the model error of 5 and 4 per cent, respectively. The RMSE of Skymet is above both the IMD's FSLRF and SSLRF (Table 3).

These results are reflected in the prognosis of extreme events like droughts and above normal rainfall. For instance, the IMD's FSLRF failed to

Table 2: Comparison of PCC Between Rainfall Forecasts and Actual

Time period/ Indicators	FSLRF and actual (pval)	SSLRF and actual (pval)	Skymet and actual (pval)
Period 1 (1995 to 2002)	-0.36 (0.37)	-	-
Period 2 (2003 to 2012)	-0.21 (0.57)	0.12 (0.74)	-
Period 3 (2013 to 2018)	0.45 (0.37)	0.64 (0.17)	0.41 (0.42)

Source: Authors' estimates

¹⁰It is recognised that the period wise sample size is low, due to the paucity of data for various agencies. Nevertheless, even for the entire period from 1995 to 2018, the PCC of FSRLF (-0.07) is negative and insignificant. Similarly, the PCC of SSLRF (2003-2018) is positive and insignificant (0.34). These results are further corroborated by RMSE results and Charts 1 & 2 of the Stylised facts.

Table 3: Comparison of RMSE of Rainfall Forecasts and Actual

Root Mean Square Error	FSLRF and actual	SSLRF and actual	Skymet and actual
Period 1 (1995 to 2002)	10.14	-	-
Period 2 (2003 to 2012)	8.57	8.16	-
Period 3 (2013 to 2018)	7.03	5.99	8.71

Source: Authors' estimates

predict all-India droughts (2002, 2009, 2014 and 2015), deficient rainfall (2004) and above normal rainfall (1998, 2007 and 2013). Incidentally, it got predictions correct for near normal monsoon only 38 per cent of the time. Similarly, Skymet also failed to forecast extreme events. Nevertheless, the SSLRF nailed down the 2015 drought and its probability of predicting near normal monsoon is over 44 per cent. Contrastingly, international agencies NOAA and BOM have been relatively more successful in forecasting extreme rainfall years, which generally coincide with *El Nino* and *La Nina* conditions. This is corroborated by higher values of success score for international agencies compared to IMD and Skymet (Table 4). Among national forecasts, the success score is higher for the IMD's SSLRF than FSLRF and Skymet.

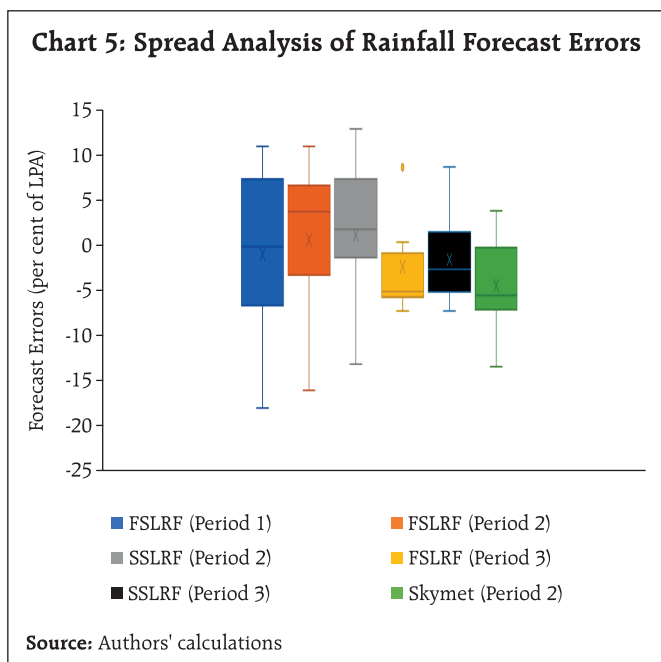
The box plot of the spread of forecast errors (actual minus forecast) of IMD (FSLRF and SSLRF) and Skymet shows interesting variations (Chart 5). The spread is higher for the IMD's FSLRF in Period 1 which has gradually reduced in subsequent periods. In Period 2, as expected, the error spread of the SSLRF is lower than that of the FSLRF. In Period 3, although the spread is higher for SSLRF as compared

Table 4: Success Score of Various Forecasting Agencies During Extreme Rainfall Years

Agencies/Month of forecast	March	April	May	June
NOAA	57.14	57.14	71.43	57.14
BOM	71.43	71.43	71.43	71.43
IMD	-	0.00	-	16.67
Skymet	-	0.00	-	-

Note: '-' represents not applicable.

Source: NOAA, BOM, IMD and Skymet and author's estimate.

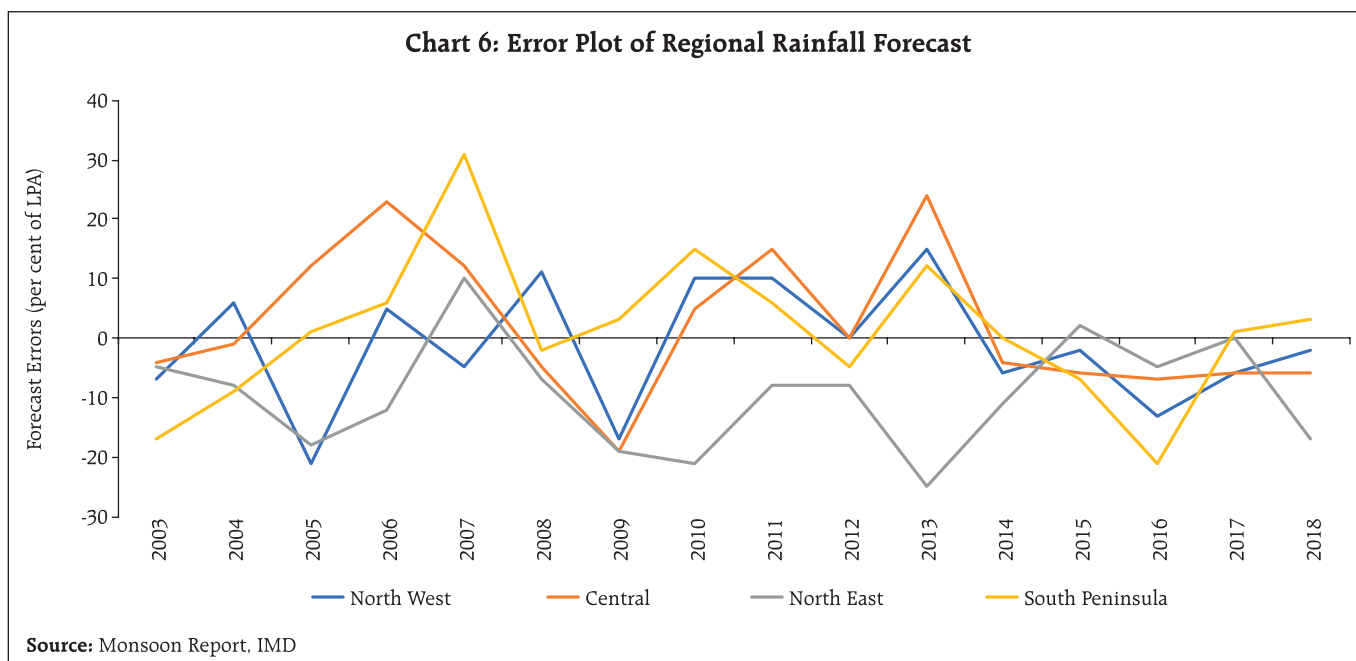


to the FSLRF, the latter has an outlier which makes it less credible. Likewise, the error spread of Skymet forecast is higher than the IMD's SSLRF. Clearly, in Period 3 the error spread for SSLRF has reduced from Period 2; however, since the mean is above the median in the former, SSLRF seems to over-predict the rainfall.

As mentioned earlier, the IMD also provides forecast of regional distribution of rainfall (North West, Central, East and North East and Southern Peninsula) in the SSLRF since 2003. The error plots (difference between actual and prediction) show that there is no systematic pattern of prediction errors across the regions (Chart 6). However, the RMSE is highest for the Southern Peninsula (13.24), followed by East and North East (12.54), Central India (11.18) and North West (9.42). The forecasts tend to over-project rainfall in the East and North East and Southern peninsula in 77 per cent and 50 per cent of the time, respectively. The forecasts have been correct for North West nearly 50 per cent of the time and, hence, the low RMSE.

IV. Conclusion

There is no significant correlation between the projected rainfall (IMD and Skymet) and actual rainfall in India. While none of the forecasts are close to the actual, the performance of the IMD's SSLRF is better than FSLRF and Skymet. Both IMD and Skymet have failed to predict drought and excess rainfall in most of the cases. Nevertheless, the SSLRF nailed down the



2015 drought and its probability of predicting near-normal monsoon has been reasonable and higher than FSLRF and Skymet. In contrast, the predictive power of the international agencies, viz., BOM and NOAA in forecasting extreme rainfall (which generally coincides with the *El Nino* and *La Nina* conditions) is much better than that of the IMD. The comparative assessment of all forecasts suggests that for generating macroeconomic forecasts, the use of IMD's SSLRF and the predictions of international agencies like NOAA and BOM in conjunction may be appropriate as the preliminary forecasts of IMD (FSLRF) and Skymet released in April appear to be noisy.

References

- Government of India (2018), Agriculture at a Glance, Directorate of Economics and Statistics, Ministry of Agriculture and Farmers Welfare, New Delhi.
- Gulati, A., Saini, S., and Jain, S. (2013), 'Monsoon 2013: Estimating the Impact on Agriculture', *Working Paper No.269*, Indian Council For Research on International Economic Relations.
- Joseph, M., Murugan, E. and Hemalatha (2017), 'Forecast Verification Analysis of Rainfall for Southern Districts of Tamil Nadu, India', *International Journal of Current Microbiology and Applied Sciences*, 6: 299-306.
- Kar, S.C., Acharya, N., Mohanty, U.C., and Kulkarni, M.A. (2012), 'Skill of Monthly Rainfall Forecasts over India using Multi-Model Ensemble Schemes', *International Journal of Climatology*, Vol 32, 1271-1286.
- Karumuri, A., Guan, Z. and Toshio, Y. (2001), 'Impact of the Indian Ocean Dipole on the relationship between the Indian Monsoon Rainfall and ENSO', *Geophysical Research Letters*, 28.
- Monsoon Report (various issues), Indian Meteorological Department (IMD), Government of India.
- Pandey, D.K., Rai, S., Sahai, A.K., Abhilash, S. and Shahi, N.K. (2016), 'Prediction and Error Growth in the Daily Forecast of Precipitation from the NCEP CFSv2 over the Subdivisions of Indian Subcontinent', *Journal of Earth System Science*, Vol 125(1), 29-45.
- Prasad, K., Dash, S.K., and Mohanty, U.C. (2010), 'A Logistic Regression Approach for Monthly Rainfall Forecasts in Meteorological Subdivisions of India based on DEMETER Retrospective Forecasts', *International Journal of Climatology*, Vol 30, 1577-1588.
- Rajeevan, M. and Pai, D.S. (2006), 'On the *El Nino*-Indian Monsoon predictive relationships', *Geophysical Research Letters*, Vol. 34.
- Rajeevan, M., Pai, D.S., Dikshit, S.K., and Kelkar, R.R. (2004), 'IMD's New Operational Models for Long-Range Forecast of Southwest Monsoon Rainfall over India and their Verification for 2003', *Current Science*, Vol 86(3), 422-431.
- Sagar, K.S., Mrudula, G., Kumari, K.V., and Rao, S.V.B (2017), 'Verification of VARSHA Rainfall Forecasts for Summer Monsoon Seasons of 2009 and 2010', *International Journal of Current Research Review*, Vol 9(21), 24-36.
- Silver, N. (2012), 'The Weatherman is not a Moron' New York Times dated September 7. (Retrieved from <https://www.nytimes.com/2012/09/09/magazine/the-weatherman-is-not-a-moron.html>)
- Stern, H. (2008), 'The Accuracy of Weather Forecasts for Melbourne, Australia', *Meteorological Applications*, 15: 65-71.