

The Indian Monsoon

5. Prediction of the Monsoon

Sulochana Gadgil and M Rajeevan

In this article we first consider the importance of prediction of the monsoon, and events such as the intense rainfall event over Mumbai in July 2005. We then discuss how meteorologists make short-, medium-, and long-range forecasts and the concept of the limit of predictability in a chaotic system such as the atmosphere. Problems and prospects of prediction on different time-scales are discussed by using one example of short-range forecasts and the prediction of the monsoon by dynamical and statistical methods. Finally we consider measures of the skill of a forecast and how high the skill has to be for it to be useful for applications.

1. Introduction

From the beginning of this year there has been a steady increase in the inflation rate due to the dramatic rise in the price of rice, other food grains and oil. By July the inflation rate had risen above 11%, its highest in 13 years. As efforts were made to arrest it, the question arose as to when it would actually start decreasing. The response of one of the experts (Deputy Chairman of the Planning Commission) was very interesting. He suggested that the situation would improve if the monsoon turned out to be normal, as predicted¹. Thus the impact of the vagaries of the monsoon on critical facets of our economy is perceived to be very significant. Not surprisingly, then, we never take the monsoon for granted. Every year, as the heat scorches the countryside in May, we start worrying about the upcoming monsoon and the media becomes obsessed with predictions about the monsoon rainfall.

Yet the Indian summer monsoon rainfall is one of the most reliable events in the tropical calendar. The typical year to year variation (i.e., the standard deviation) of the all-India Summer



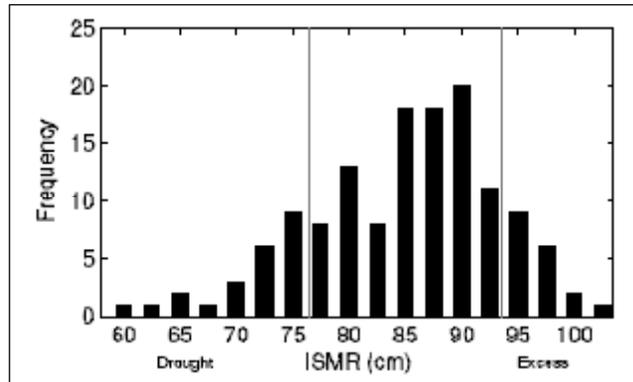
(left) Sulochana Gadgil is an honorary Professor at the Centre for Atmospheric and Oceanic Sciences at the Indian Institute of Science. Her main research interests are monsoon dynamics, the coupling of the tropical cloud systems to the oceans. She has worked on mathematical models of interactive populations.

(right) Madhavan Nair Rajeevan is a Senior Scientist at the National Atmospheric Research Laboratory, Tirupati. Earlier, he was at the India Meteorological Department. His research interests are monsoon variability and prediction and radiation studies.

¹ In fact, the monsoon did turn out to be normal and the inflation rate did decrease in October. However, it is likely that the sharp decrease in oil prices associated with the economic meltdown was a far more important factor than the monsoon.



Figure 1. Number of years with monsoon rainfall in the range specified on the x-axis.



Monsoon Rainfall (ISMR) is only about 10% of the average rainfall of about 85 cm. The frequency distribution of ISMR over the period 1876–2007 is shown in *Figure 1*. Droughts are characterized by the ISMR anomaly (difference between the ISMR of a specific year and the long term average) being negative and of magnitude larger than 10% of the long term average. On the other hand, excess rainfall years are characterized by positive ISMR anomalies of magnitude larger than 10% of the average. Normal monsoon years are characterized by the magnitude of the ISMR anomaly being less than 10% of the average. The distribution is not symmetric and is characterized by a longer tail with negative anomalies than that with positive anomalies. Over the 132-year period there have been 23 droughts and 19 excess rainfall years. Thus historical records show that the chance of the so-called normal monsoon is a little over 68%, of droughts around 17%, and of excess rainfall about 14%. While for the worst drought (1877) the ISMR deficit was 25 cm, for the season with maximum rainfall (1961) the ISMR anomaly was 17cm. It is seen that the most likely value of ISMR (the mode) is around 90 cm, i.e., higher than the average and the rainfall is in the range of 83.75–91.25cm in 44% of the years.

Why, then, are we so anxious about the monsoon? It turns out that although the amplitude of the variation of ISMR from year-to-year is not large, it has a substantial impact on the agricultural production in the country [1]. Before independence, this also implied a large impact on the economy of the country since the

Keywords

Monsoon, prediction, droughts, predictability,chaos.

Part 1. Variations in Space and Time, *Resonance*, Vol 11, No.8, pp.8–21, 2006.

Part 2. How Do We Get Rain?, *Resonance*, Vol.11, No.11, pp.8–21, 2006.

Part 3. Physics of the Monsoon, *Resonance*, Vol 12, No.5, pp.4–20, 2007.

Part 4. Links to Cloud Systems over the Tropical Oceans, *Resonance*, Vol.13, No.3, pp.218–235, 2008.



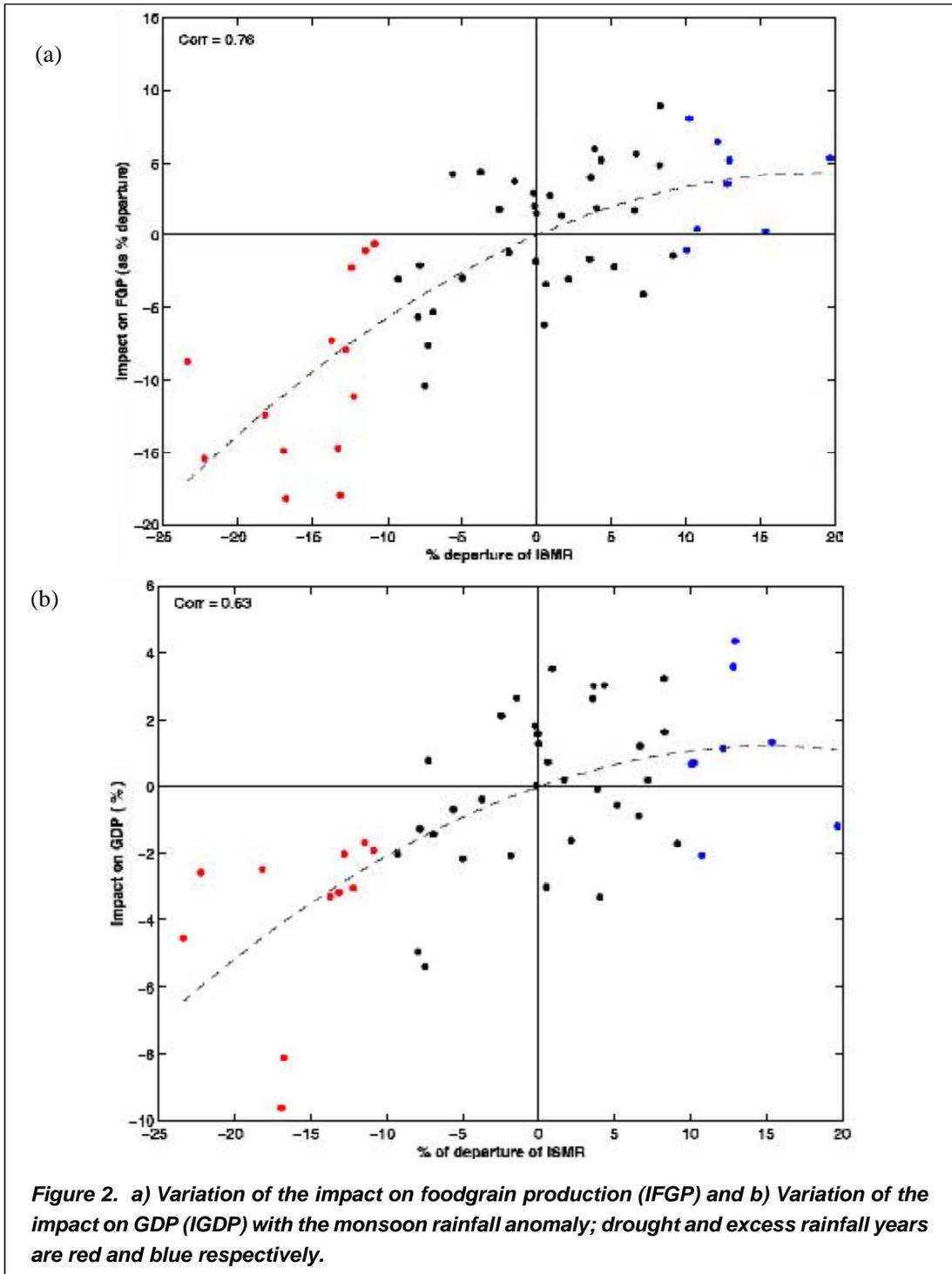
economy was primarily dependent on agriculture. With planned development since independence, the contribution of agriculture to the Gross Domestic Product (GDP) decreased substantially and led to the expectation that the impact of the monsoon on the economy would have also decreased. However, a recent analysis of the variation of the GDP and the monsoon has revealed that the impact of severe droughts on GDP has remained between 2 to 5% of GDP throughout [2]. The large impact of droughts on GDP (despite the substantial decrease in the contribution of agriculture to GDP) can be attributed to the indirect impact on the purchasing power of the large fraction of the population dependent on agriculture. Hence if the monsoon turns out to be a 'normal' monsoon, as in this year, the nation heaves a sigh of relief and carries on with business as usual. If it turns out to be a drought, there is a significant impact on agriculture and the economy, and major drought relief programmes are launched.

If the monsoon turns out to be a 'normal' monsoon, as in this year, the nation heaves a sigh of relief and carries on with business as usual.

It has been shown [2] that while the magnitude of the adverse impact on food-grain production (IFGP) and the GDP (IGDP) of deficit rainfall is large, the positive impact of surplus rainfall is not large (*Figure 2*). In other words, there is an asymmetry in the response of foodgrain production and GDP to the variation of the monsoon. It has been suggested that a possible reason for the relatively large asymmetry in the response of the foodgrain production after 1980, is that the strategies that would allow farmers to reap benefits of the good rainfall years (such as adequate investments in fertilizers and pesticides for rain-fed areas) are not economically viable in the current milieu. Such strategies would become economically viable if reliable predictions for 'no droughts' could be generated. Thus prediction of the interannual variation of ISMR and particularly for the occurrence or non-occurrence of the extremes continues to be extremely important.

In addition to prediction of the monsoon rainfall over the country as a whole, there is demand for prediction of some events such as the intense rainfall event on 26 July 2005 when Mumbai received 94.4 cm of rainfall on a single day, or of the severe cyclone that





devastated Orissa in 1999, because of the enormous impact they have on a large number of people. There is also a need for several user-specific predictions such as prediction of low-level wind for sailors and for paragliding enthusiasts, quantitative precipitation forecasts for reservoir and flood management. The time-scales of the events for which prediction is required also varies with the application. Thus while some farmers need prediction for occurrence of a dry spell of duration of a week or more, for managers of reservoirs, prediction of the total rainfall in a month or a season is often adequate.

Meteorological forecasts are generated for different time-scales.

Meteorological forecasts are generated for different time-scales. Forecasts of daily weather with a lead time of 1–3 days are short-range forecasts and with a lead time of 3–10 days are called medium-range forecasts. Forecasts for monthly or seasonal rainfall come under the category of long-range forecasts. The official government agency which has the responsibility of disseminating short-, medium- and long-range forecasts in our country is the India Meteorological Department (IMD). At present, efforts are being made the world over to generate predictions over an intermediate time-scale, the so-called extended-range prediction with a lead time of 10 days to a month for rainfall, temperature, etc., averaged over about 5 days. Since spatial and temporal scales are inexorably linked, short-range forecasts are generated for the meteorological subdivisions of India (shown in *Figure 3* in which the rainfall anomalies for the drought of 2002 are depicted for each subdivision) and for smaller spatial scales such as district level; whereas long-range forecasts are made for larger regions such as the all-India scale or for 3–4 sub-regions of the country.

In this article we first consider how the forecasts over short, medium and long range are generated (Section 2), mention an example of short range forecast (Section 3), and then focus on the problems and prospects of predicting the monsoon rainfall over India (Section 4). Finally we discuss measures of the skill of a forecast and the minimum skill which has to be attained by a forecast before it can be useful for decision making (Section 5).



interaction between different spatial scales from 3–4 kilometres (as in a single cloud) to hundreds of kilometres (as in a monsoon depression or a hurricane).

Let us first try to understand how predictions are generated. The state of the atmosphere at any point of time (in terms of temperature, wind, rainfall, etc. as a function of space) evolves according to Newton's laws as applied to a compressible fluid in a rotating system. Hence the logical way of predicting the future state of the atmosphere (say 24 or 48 hours ahead) is to integrate the governing equations, starting with the observed state of the atmosphere at the initial instant as the initial condition, and the observed conditions at the surface of land or ocean as the boundary condition, for 24/48 hours. The errors in the short-range forecasts occur because (i) the models are not perfect (involving many assumptions like how sub-grid scale processes such as clouds affect the heating), and (ii) there are errors and gaps in the observations of the initial state.

An important question is: Even with a perfect model and high resolution observations, can we predict a week, month or a season ahead, the weather at a particular place at a specific instant viz. state of the atmosphere at that instant, at that point in space? In fact, even with a perfect model, it will never be possible to predict 'weather' more than about seven days ahead. This is because there is an inherent limit to predictability of weather. In a pioneering study Lorenz [3] showed that if we start integrating the governing equations from two very similar initial conditions (i.e., two similar states of the atmosphere), as they evolve, because of the instabilities in the atmosphere, the two solutions start diverging with time, i.e., the difference in the predicted states increases with time. By about seven days, the initial condition appears to be forgotten. The difference between the two states then becomes comparable to the difference between two states evolving from two randomly chosen initial conditions (not arbitrarily close ones as assumed earlier). Lorenz's study introduced the concept of chaos and the atmosphere became the first known example of a chaotic system.

Seasonal forecasting is primarily a boundary value problem, while short- or medium-range weather forecasting is primarily an initial value problem.

However, fortunately, every facet of the atmosphere is not chaotic on all time-scales. In fact, the variation of climatic elements averaged over different spatial and temporal scales (e.g., the interannual variation of the seasonal rainfall over the Indian region) arises partly from the variation of the conditions at the lower boundary of the atmosphere such as the sea surface temperature (SST) or snow cover over Eurasia. Hence such variables can be used as predictors for this time-scale. *Thus seasonal forecasting is primarily a boundary value problem, while short- or medium-range weather forecasting is primarily an initial value problem. Extended range prediction will depend on the initial as well as boundary conditions.*

The first short-range weather forecasts were made by meteorologists with empirical knowledge of how weather maps evolved from day-to-day. By the 1950s, development of physical models of the atmosphere on the one hand and detailed observations of the system on the other, led to insights into the physics of the variation on the scale of a few days. With the advent of satellites the density of observations increased enormously and with the phenomenal increase in computing power, complex models of the atmosphere that could simulate the short- and medium-range variation realistically, were developed by the 1980s. Now, the integration of such models with initial conditions obtained from the worldwide observation network, is a major input for weather prediction on these time-scales. Atmospheric models are run regularly for this purpose at IMD and the National Centre for Medium Range Weather Forecasting (NCMRWF).

Predictions on the seasonal to interannual scale can be generated by using ensemble runs of atmospheric models with specified boundary conditions and varying initial conditions. Since oceans evolve more slowly than the atmosphere, the conditions at the surface of the ocean could be specified for these runs. Operationally at many centres in the world, long-range predictions are generated by running atmospheric models with specified boundary conditions or by running coupled models in which the oceans also evolve. For long-range predictions, an alternative approach



is the traditional one, in which statistical models are used for prediction. These models are generally based on the links of the predictand (in our case rainfall) with prior values of that variable and/or other variables (such as pressure, temperature of the atmosphere or ocean, over the same or different regions of the atmosphere/ocean) discovered by analysis of large number of data sets.

3. Short Range Forecasts

We consider one example of a short-range forecast here. On 26/27 July 2005, Mumbai received unprecedented heavy rainfall, with its suburb Santa Cruz recording 94.4 cm of rainfall in 24 hours. There were reports of even heavier rainfall of 104.5 cm near Vihar lake. It disrupted life in the metropolis and led to a large number of deaths. The intensity of this event was not predicted either by IMD or by other operational forecasts generated by major weather prediction groups like UK Met office and US weather service. IMD's prediction made 24 hours ahead, suggested a high probability of heavy rainfall (rainfall exceeding 12.5 cm) over the region. However, while rainfall at Mumbai exceeding 12.5 cm in a day is a very common event in the rainy season, rainfall over 90cm in a day had never been experienced before. Had the forecast been more specific in terms of the probable intensity, the damage could have been reduced to some extent and a number of lives could have been saved.

A post facto analysis of the prediction of the Mumbai event [4] suggests that it would have been possible to predict the intensity this event with reasonable accuracy, with high resolution atmospheric models, provided high resolution data (particularly on clouds organized over meso-scale and higher scales) available from satellites and quality-controlled local meteorological data was used in specifying the initial condition. Once the system which can assimilate relevant data from Doppler radars, from satellites, the high density meteorological observations in the metropolis as well as high resolution data on the terrain and land surface conditions is in place, it should be possible to generate



reliable predictions of such events using the high resolution models available in the country.

4. Predicting the Indian Summer Monsoon Rainfall

4.1 Predictions with Dynamical Models

Consider first the prediction of the monsoon rainfall by integration of complex models of the atmosphere or the coupled atmosphere-ocean system based on equations governing fluids in a rotating system. It is important to note that the breakthroughs in seasonal forecasting over the tropics have come from the phenomenal progress since the 80s in the understanding of the physics of El Nino–Southern Oscillation (ENSO) (see Parts 1 and 4), the dominant signal of the interannual variation of the coupled atmosphere-ocean system over the Pacific. The elucidation of the nature of ENSO, and unravelling of the underlying mechanisms led to development of models to a level at which they could realistically simulate the phenomenon and its impact on the climate of different regions.

It is important to note that the breakthroughs in seasonal forecasting over the tropics have come from the phenomenal progress since the 80s in the understanding of the physics of El Nino–Southern Oscillation (ENSO).

Given the links between the Indian monsoon and ENSO, it was expected that it would be possible to simulate the interannual variation of the ISMR with atmospheric general circulation models when the observed SSTs are specified as a boundary condition. However, the results of several such studies suggest that the problem remains a challenging one. Analysis of the simulations for the years 1979–95 by 20 state-of-the-art atmospheric general circulation models showed that while almost all models simulated the correct sign of the ISMR anomaly in 1988, a vast majority of the models failed to capture the anomaly for the excess monsoon season of 1994. Thus, the skill of the models in simulating the sign of the anomalies is not the same for all the droughts or excess rainfall years. During the excess monsoon season of 1988, ENSO was favourable while during 1994, ENSO was unfavorable but EQUINOO was favourable and excess rainfall occurred. On the whole, the skill of the models in simulating the sign of the anomaly for extreme ISMR seasons is higher for



the extreme seasons which are associated with ENSO. Clearly more research and development effort is required to develop models which are capable of a realistic simulation of the links of the monsoon to EQUINOO.

The success of the atmospheric models in simulating the extremes of ISMR when they are linked to ENSO was achieved by concerted efforts under an international programme MONEG in the 90s under which the cases of 1987 and 1988 were studied with a slew of models. We now need research and development of atmospheric and coupled models to a level at which they can simulate realistically the response of the Indian monsoon to EQUINOO as well as ENSO. Once this is achieved it may be possible to generate reasonable predictions of the ISMR with dynamical models.

Predicting ISMR is one of the important mandates of the India Meteorological Department IMD. Hence until the skill of the atmospheric and coupled models in simulating and predicting the interannual variation of the Indian monsoon improves, the empirical approach has to be adopted for operational forecasts.

4.2 Statistical Models

Forecasting of monsoon rainfall has been attempted for over a hundred years in India. In 1871 the Madras famine commission recommended that, "so far as it may be possible, with the advance of knowledge to form a forecast of the future, such aids should be made use of, though with due caution". A major drought and famine occurred in India in 1877 soon after the IMD was established. The first long-range prediction in the world was made in 1886 by Blanford, who was the Chief Reporter of IMD, at the request of the colonial government in the wake of this drought. The prediction was based on the relationship between Himalayan snow cover and monsoon rainfall, discovered by Blanford in 1884. IMD has always been the responsible agency for the operational long-range forecasts of monsoon rainfall, which until recently have been based only on empirical models such as

Forecasting of monsoon rainfall has been attempted for over a hundred years in India.



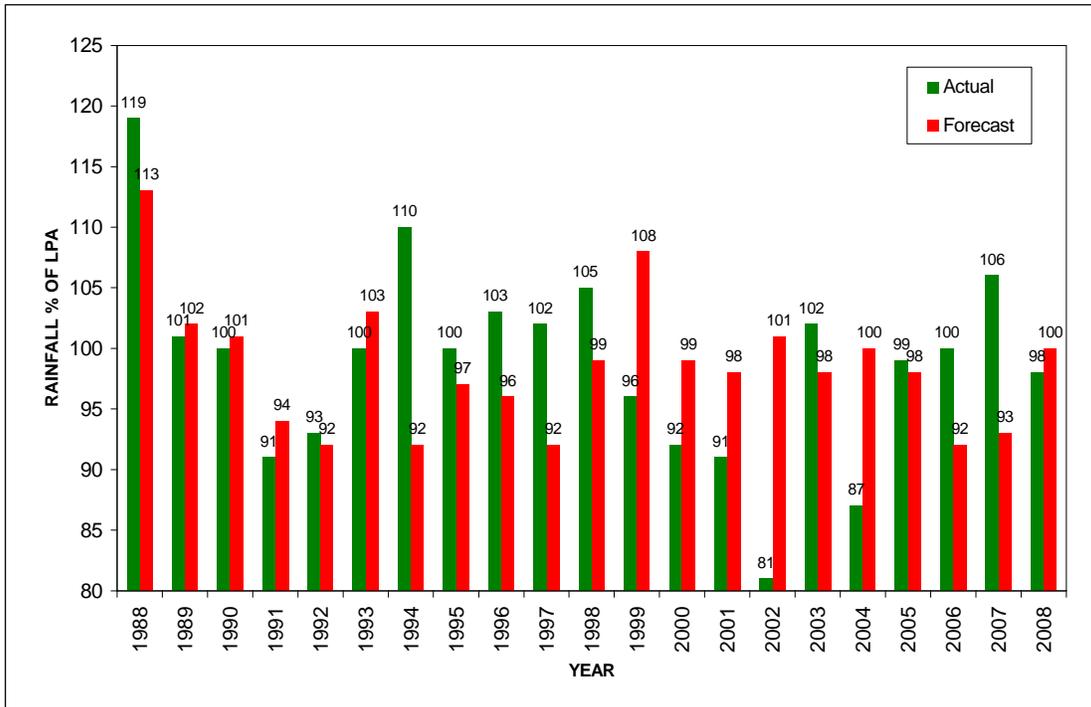
³ Part 1. Variations in space and time, *Resonance*, Vol 11, No.8, pp.8–21, 2006.

Blanford's. Forecasts during the initial years were subjective and qualitative. In the early part of the last century, Sir Gilbert Walker³ initiated extensive studies of the worldwide variation of weather elements (e.g., pressure, temperature, etc.) to develop models for monsoon prediction. In 1909 Walker introduced an objective technique based on correlation and regression analysis. The first model used by Walker in 1909 for prediction of ISMR was a linear regression model based on four predictors (Himalayan snow accumulation at the end of May, South American pressure during March–May, Mauritius pressure in May and Zanzibar rain in April and May). However, assessment of the predictions by this model up to 1936 showed that, in spite of its early encouraging performance, the formula had broken down completely in the 15 years from 1921. While investigating the links of the Indian monsoon with atmospheric conditions over the rest of the globe, Walker discovered the Southern Oscillation, which is a see-saw of pressure between Darwin, Australia and Tahiti in the Pacific Ocean. This discovery was to play a major role in the phenomenal advances in the understanding and prediction of the interannual variability of the tropical ocean–atmosphere system witnessed over the last decade.

⁴ Part 4. Links to Cloud Systems over the Tropical Oceans, *Resonance*, Vol.13, No.3, pp.218–235, 2008.

After the discovery of strong links between the El Nino and the Indian monsoon⁴, the empirical models for monsoon prediction have developed rapidly. In the tradition of Walker, a large number of potential predictors have been identified by analysis of the ever-increasing data from conventional and satellite observations on many atmospheric and oceanic variables, and their lag correlation with the ISMR. Some of these parameters are related to El Nino and Southern Oscillation, others to snow over the Himalayas and Eurasia, and some to global and regional conditions on spatial scales ranging from one station (e.g., surface temperature at De Bilt in Holland) to hemispheric (e.g., northern hemispheric surface air temperature in January and February). In fact, as the sample of years increased with time, the correlation coefficient with several parameters became poor and for some of them even changed sign; hence many revisions were made on the





model by changing the predictors. In 1988, IMD introduced the 16-parameter power regression and parametric models which were used operationally during the period 1988–2002. However there are very large errors in 1994, 1997, 2002 (*Figure 4*). In fact our analysis of the predictions generated by the empirical models used operationally by IMD during 1932–2002, suggests that the performance of these models, based on the relationship of the monsoon rainfall to atmospheric/oceanic conditions over different parts of the globe, has not been satisfactory [5].

Figure 4. Performance of Operational Forecast (1988–2008).

After the failure of forecast in 2002, IMD introduced a new two-stage forecast strategy in 2003, according to which the first-stage forecast for the summer monsoon rainfall over the country as a whole is issued in April and the update is issued in June. Along with the update forecast, separate forecast for seasonal rainfall over broad homogeneous rainfall regions of India and July rainfall over country as a whole are also issued. In 2007, IMD introduced a new statistical forecasting system based on



ensemble technique for the summer monsoon season using 8 predictors. In the ensemble method, instead of relying on a single model, all possible models based on all the combination of predictors are considered. Out of all the possible models, the best few models are selected based on their skill in predicting monsoon rainfall during a common period. The forecast is then generated from the weighted average of the forecast from the selected models. In fact the forecast for 2008 turned out to be rather accurate (*Figure 4*). But 2008 was a normal monsoon. Whether the model can forecast a drought or an excess rainfall season will be tested only when an extreme monsoon season does occur.

Our experience of the monsoon of 2006 suggests that incorporation of predictors associated with EQUINOO along with those associated with ENSO may improve predictions. Based on the analysis of predictors, IMD issued a long-range forecast for the 2006 monsoon season rainfall as 93% of long-period average. This inference of below normal was drawn based primarily on the warming tendency of SST anomalies over the equatorial Pacific which suggested the development of an El Nino. The monsoon rainfall performance, in fact, was alarming till the third week of July with all-India cumulative rainfall departure being 13% below normal. However, rainfall activity revived by the third week of July and good rainfall activity extended almost unabated till the middle of September, thus improving the rainfall situation in the country. At the end of the monsoon season, seasonal rainfall was 100% of its long-period average. During August–September a positive phase of EQUINOO had developed with enhanced convection over the western equatorial Indian Ocean and suppressed convection over the eastern part. The enhanced rainfall during the second half of the monsoon season could be attributed to this. Had the development of positive EQUINOO phase by August been predicted, it might have been possible to predict that the deficit would certainly not be as large as expected from ENSO alone.



5. Assessment of Skill and Minimum Level Required for Applications

Obviously forecasts will be useful only if they are reliable. Hence it is important to assess the skill of forecasts generated by the different models used. For assessment of the skill of prediction of an event (such as rainfall at Mumbai greater than 70cm in a day), a large number of predictions generated for an event have to be compared with the observations. Let the number of occasions on which the event was predicted and observed to occur be a ; in which the event was predicted but did not occur be b ; the number of occasions on which it was predicted not to occur, but occurred c , and those for which it was predicted not to occur and did not d . The skill of the prediction is assessed by how large the hit rate (i.e., probability of correct forecasts = $(a + d)/(a + b + c + d)$) is *vis a vis* the false alarm rate (i.e., fraction of the times it was predicted but did not occur = $b/(a + b)$). The forecast is said to have a reasonable skill only when the hit rate is larger than the false alarm rate. For rare events, the threat score ($a/(a + b + c)$) is a more appropriate measure than the hit rate since d is much larger than a in this case.

In fact, the level of skill (i.e., the probability of correct forecast) has to be sufficiently high for a prediction to be useful for an application. How good is good enough, depends on the cost of adopting a strategy which is appropriate for the prediction and the expected benefit from such a change in strategy. For example, suppose that after an attack by insect pests, a farmer has to choose between two management strategies (e.g., to spray pesticide or not). Let the cost incurred in spraying the pesticide be C . The benefit B in terms of enhancement of the yield due to the spraying of pesticides (which is necessarily greater than the cost C for spraying to be considered at all) will be realized only if it is not immediately followed by a wet spell. Thus, if it is predicted that a wet spell will not occur, the appropriate strategy would be to spray. However, it has been shown [6] that such a strategy will in fact be beneficial only if the probability that this prediction is correct is greater than C/B .



At present in addition to the operational forecasts by IMD, predictions are also available from different meteorological centres of the world and some predictions are also generated by different groups in the country. However, a quantitative assessment of all the available predictions to determine their skill for different events of importance is yet to be made. Such an exercise has to be done on a continuing basis as and when new models are developed here or abroad, to ensure that predictions generated with models or a combination of models with the maximum possible skill are disseminated to users.

Suggested Reading

- [1] Sulochana Gadgil, Y P Abrol and P R Seshagiri Rao, On growth and fluctuation of Indian foodgrain production, *Current Science*, No.4, Vol.76, pp. 48–56, 1999.
- [2] Sulochana Gadgil and Siddhartha Gadgil, The Indian Monsoon, GDP and agriculture, *Economic and Political Weekly*, XLI, pp.4887–4895, 2006.
- [3] E N Lorenz, The predictability of a flow which possesses many scales of motion. *Tellus*, Vol 21, pp.289–307, 1969.
- [4] A K Bohra, Swati Basu, E N Rajagopal, G R Iyengar, M Das Gupta, R Ashrit and B Athiyaman, Heavy rainfall episode over Mumbai on 26 July 2005: Assessment of NWP guidance, *Current Science*, Vol.90, pp.1188–1194, 2006.
- [5] Sulochana Gadgil, M Rajeevan, and R Nanjundiah, Monsoon Prediction: Why yet another failure?, *Current Science*, Vol.88, pp.1389–1400, 2005.
- [6] Sulochana Gadgil, P R Sheshagiri Rao, N V Joshi, S Sridhar, Forecasting rain for groundnut farmers - How good is good enough? *Current Science*, Vol.68, pp.301–309, 1995.

Address for Correspondence

Sulochana Gadgil¹ and M Rajeevan²

¹Centre for Atmospheric and Oceanic Sciences
Indian Institute of Science
Bangalore 560 012, India.
Email: sulo@caos.iisc.ernet.in

²National Atmospheric Research Laboratory
PB NO 123, SVU Campus
Tirupati 517 502, India.
Email: rajeevan@narl.gov.in

