MAUSAM, **54**, 2 (April 2003), 359-376

551.509.324 : 551.553.21(540)

Verification of model predicted precipitation over India during the summer monsoon of 1997

B. K. BASU

NCMRWF, Lodi Road, New Delhi – 110 003, India

(*Received 18 April 2002, Modified 21 August 2002*)

सार – वर्ष 1997 की मानसून ऋतु के लिए समूचे भारत के राष्ट्रीय मध्यम अवधि मौसम पूर्वानुमान केंद्र के पूर्वानुमान निदर्श के 135 ग्रिड स्थानों के स्थानिक औसतों को तैयार करने के लिए 1333 वर्षा मापी स्टेशनों से प्राप्त किए गए वर्षा के प्रेक्षणों का उपयोग किया गया है। इन स्थानिक औसतों की तुलना माध्य ऋतु, माध्य निरपेक्ष त्रुटि, वर्ग माध्य मूल मान और सहसंबंध गुणांक से आकलित करते हुए वर्षा के निदर्श पूर्वानुमानों के साथ की गई है। वर्षा की मात्रा को मानक श्रेणियों से पृथक करते हुए पूर्वानुमान की निपुणता को भी आकलित किया गया है।

ऋतुनिष्ठ संचयी वर्षा और इनके स्थानिक वितरण को अधिकतम 4 दिन से ही निदर्श पूर्वानुमान में बेहतर रूप से पुनः प्रस्तुत किया गया है। पूर्वानुमान में पाए गए माध्य त्रुटि से पश्चिमी घाट की पहाडियों की दूसरी तरफ हुई वर्षा कब पता चलता है जबकि क्षीण क्षैतिज विभेदन के कारण निदर्श में आए रोध के प्रभाव को वास्तविकता से कम आकलित किया गया है। सहसंबंध गुणांक से विशेषकर अत्याधिक ऋतुनिष्ठ वाले वर्षा के क्षेत्रों में पूर्वानुमान और संभावित वर्षा के मध्य आपस में अच्छे संबंधों का पता चला है। थोड़े से क्षेत्र में इस सहसंबंध के मान 0.6 की अपेक्षा अधिक पाए गए हैं। इस निदर्श में 1.0 इंच तक की दूरी सहित वर्षा की श्रेणियाँ बताने और 2 दिनों का निर्धारित पूर्वानुमान देने की निपुणता है।

वर्ष 1997 में संभावित बाक्स औसत वर्षा में दक्षिण से उत्तर की ओर बढ़ने वाली प्रणाली में किसी प्रकार की निम्न आवर्त्तिता विद्यमान नहीं है।

ABSTRACT . For the monsoon season of 1997 the precipitation observation from 1333 rain gauge stations are used to form spatial averages for 135 grid points of the NCMRWF forecast model covering the whole of India. These spatial averages are compared with the model forecasts of precipitation by computing the mean error, mean absolute error, root mean square error and correlation coefficient. The skill of forecast is also computed by discretizing the precipitation amounts in to standard categories.

The seasonal accumulated precipitation and its spatial distribution are well reproduced in the model predictions up to 4-days in advance. The mean error in prediction shows spread of precipitation across the Western Ghat hills as the barrier effect is under estimated in the model due to poor horizontal resolution. The correlation coefficient shows good in-phase relationship between the predicted and observed precipitation especially over areas of very large seasonal precipitation. Over a small area this coefficient has values higher than 0.6. The model has skill for categories of precipitation with class mark up to 1.0 inch and prediction lead-time up to 2 day.

In 1997 no low frequency south to north propagating mode is present in the observed box average precipitation.

Key words – Model prediction, Summer monsoon rainfall, Global numerical model, NCMRWF.

1. Introduction

Precipitation is the most important parameter for atmospheric prediction in tropics. It is also the most difficult one to predict both in space and time. Most of the atmospheric prediction centres use global numerical models at present to predict the atmospheric variables including precipitation in the medium range (more than 3-days and up to 10 days) and regional models for higher spatial resolution of 50 km or less for short range (1 to 3 days). At the National Centre for Medium Range Weather Forecasting (NCMRWF) at New Delhi, a global spectral model at triangular truncation of 80 waves (T80) is integrated everyday from the initial condition valid at

0000 UTC, to produce forecasts for next 5 days. In this work we propose to verify the precipitation forecasts of this model against the observed precipitation over India during the monsoon season (1 June to 30 September) of 1997.

In the spectral T80 forecast model (Kalnay *et al.*, 1988) all physical parameterisations including those leading to precipitation are invoked on the transform grid used to transfer variables between spectral and physical spaces. The processes parameterised in the model to represent condensation of water vapour are the large-scale process that precipitates excess water vapour in stable regions of super-saturation and the convective process that leads to formation of deep clouds in regions of low-level vertical velocity and thermodynamic instability. The model precipitation is the total amount of condensed water vapour reaching surface after re-evaporation into layers below the cloud bottom.

Precipitation is computed in the atmospheric prediction model at NCMRWF as a by-product of the conservation of water vapour principle. Since, only the vapour phase of the water substances is carried in the model as a prognostic variable, precipitation acts as a sink of water vapour produced by saturation due to convergence of moisture. The model cannot treat the cloud scale explicitly and parameterisation schemes for large-scale condensation in layer clouds (Manabe *et al*. 1965) and for convections of both shallow (Tiedtke, 1983) and deep (Kuo, 1974) types are used to take care of the saturation of water vapour. In the parameterisation of deep convection a fraction of the water vapour converging in the cloud is allowed to mix with the environment to increase the water vapour content of the latter, while the rest is allowed to condense. The condensed water is allowed to evaporate as it falls through unsaturated layers below the cloud base.

During monsoon season, the atmosphere close to the surface is highly moist over the whole of India leading to saturation by lifting at a lower height and has a large amount of convective available potential energy (CAPE). This inevitably leads to the formation of convective clouds over regions where vertical velocity is present. This explains the regions of copious rainfall in the windward side of the Western Ghat and the Khasi and Jayantiya hills. In the real atmosphere a region of rain shadow is formed downstream, where downward velocities prevail in the lee of the hills. Also, moisture stripped in the windward side and adiabatic compression of the descending air lead to low relative humidity. In the model, this sharp demarcation between the windward and leeward sides of a narrow hill/ridge is lost due to the coarse resolution of the model that spreads the topography over a larger area thus smoothing the large gradients. Keeping these observations in view, we proceed to compare the model predicted precipitation with the observed ones.

The standard procedure for model verification (WMO 1992) is to compute the mean error, the standard deviation and the correlation coefficient between predicted and analysed fields valid for the same verification time. These quantities are measures of agreement between the means, the amplitudes and the phases of the predicted and analysed fields. Another quantity used for estimation of prediction error is the root mean square error (rmse) between the predicted and the observed fields at the observation locations. For computing rmse the predicted field is interpolated to the observation points.

For precipitation no routine analysis is available as it is not an analysis variable. Since precipitation is highly variable both in space and time, its interpolation to observation point is not advisable. Therefore, the standard WMO prescribed procedures are not sufficient for the verification of precipitation. In many Numerical Weather Prediction (NWP) centres, the short-range (six-hourly assimilation cycle) model predicted precipitation is taken as the benchmark against which the medium range (3-10 days) prediction of precipitation is compared. The other method (Murphy and Winkler, 1987) employed is to compute statistical parameters related to the skill of forecast based on the model predictions at regular grid points and gauge and radar measurements at irregular observation points.

2. Data and methodology

Data used in this work are the observations of daily rainfall amounts accumulated over the past 24 hours and reported at 0300 UTC of each day. A total of 1333 observations are used. Out of these about 500 are from the regular surface observatories of India Meteorological Department (IMD) and the rest are from the part-time observatories maintained by IMD or from rain/snow gauges maintained by various states and hydrological authorities. Data from all these stations were obtained from the National Data Centre of IMD, located at Pune, where the data are archived after quality control checks. The spatial distribution of data is highly inhomogeneous with larger density over the peninsula and sparse distribution over Rajasthan and Bihar. The precipitation is collected by standard manually operated rain gauges and measured up to the first place of decimal in mm.

For comparison with the regularly distributed grid point predictions from a numerical model, the station data are averaged over an area around each grid point. It is

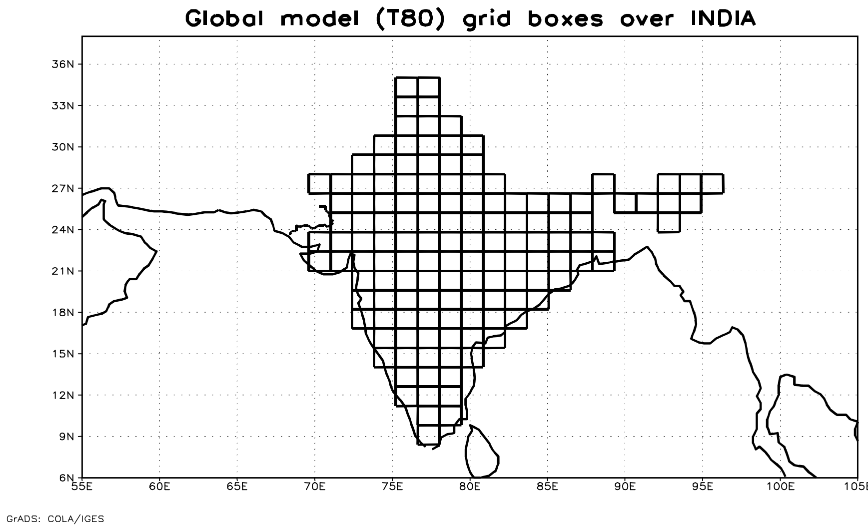


Fig. 1. Distribution of 135 grid boxes over India

assumed that in a model the grid point closest to any area represents the condition over it. Hence, the grid point value is representative of the average condition over the whole area associated with the grid point. For comparison of observed and model predicted values of any parameter, we compute the area average of the field over the area associated with each grid point. In this work the area (henceforth called grid box) associated with each grid point is obtained by drawing the bisectors of lines joining nearest grid points of the model. The 135 grid boxes that approximately cover the land area of India are shown in Fig. 1.

For each of the grid box, the average of the observed precipitation is computed by the Thiessen method (WMO, 1994) in which the area associated with each grid box is divided into a large number of smaller areas and each one of these smaller areas is assigned to the observing station nearest to it. The total area assigned to a station normalized by the grid box area, is the weight of the station in computation of average over that grid box. Since the spatial distribution of the rain gauges is inhomogeneous, the number of observations contributing to each grid box is variable and is listed in Table 1. The grid boxes are numbered from west to east starting with the northernmost row. The number of stations contributing to a grid box varies from 1 for the grid box numbered 13 (west Rajasthan) to 89 for the grid box numbered 134 (south Tamilnadu). The spacing of the NCMRWF forecast model's transform grid, on which the precipitation is computed, is about 150 km and hence each grid box covers an area of approximately 22,500 sq. km. Assuming that each rain gauge represent an area of 2,500 sq. km (a square of side 50 km), 9 or more observations are required for the computation of a reliable grid box average. Out of the 135 grid boxes over India, 72 have contributions from 9 or more rain gauges while 100 have contributions from 7 or more rain gauges.

The Thiessen method gives a higher value for the spatial average compared to that computed by the isohyetal method. For precipitation the implicit assumptions in the latter method are that the rain gauge observations are point values and the sampling includes the local maximum in the field. On the other hand the Thiessen method assumes that observed precipitation is representative of an area around the rain gauge that may contain both higher and lower values than that sampled. Since precipitation is extremely variable in space (and also time), the isohyetal method may not be suitable unless many observations per grid box are available. In this case averages computed by both methods will be close to each other. Also, the isohyetal method involves computing the weights everyday while the weights in the Thiessen method depend on the geographical locations of stations and are fixed as long as the observing network remains the same. The isohyetal method is also difficult to adopt on a computer.

The model predicted grid point values of precipitation are compared with the grid box averages of observed precipitation by computing the parameters: mean error, mean absolute error, root mean square error (rmse) and the correlation coefficient. The mean error gives

TABLE 1

Number of precipitation observation contributing to each grid box over India

Box No.	No. of rain gauges	Box No.	No. of rain gauges	Box No.	No. of rain gauges
1	3	46	18	91	8
2	4	47	10	92	10
3	8	48	18	93	11
4	6	49	20	94	9
5	9	50	4	95	9
6	19	51	5	96	6
7	6	52	6	97	8
8	7	53	7	98	9
9	16	54	8	99	8
10	31	55	7	100	6
11	11	56	9	101	7
12	4	57	7	102	8
13	1	58	4	103	9
14	5	59	11	104	12
15	25	60	19	105	9
16	30	61	20	106	9
17	7	62	27	107	5
18	8	63	55	108	8
19	2	64	14	109	9
20	3	65	10	110	6
21	4	66	9	111	8
22	5	67	14	112	6
23	7	68	10	113	10
24	14	69	5	114	9
25	11	70	7	115	6
26	12	71	6	116	9
27	11	72	6	117	7
28	19	73	14	118	6
29	7	74	10	119	7
30	16	75	43	120	8
31	15	76	32	121	6
32	3	77	40	122	9
33	3	78	18	123	4
34	5	79	12	124	8
35	4	80	14	125	9
36	6	81	14	126	5
37	7	82	10	127	10
38	9	83	6	128	11
39	9	84	7	129	24
40	10	85	10	130	16
41	8	86	7	131	54
42	3	87	8	132	71
43	9	88	7	133	80
44	16	89	29	134	89
45	22	90	10	135	52

TABLE 2

Categories of precipitation based on 24 hr accumulated amounts

Category	IMD limits	Present limits
Trace	0.1 - 2.4 mm	< 0.25 cm
Very light/light	2.5 - 7.5 mm	0.25–1 cm
Moderate	7.6 - 34.9 mm	1–3 cm
Rather heavy	35.0 - 64.9 mm	3–7 cm
Heavy	65.0-124.9 mm	7–13 cm
Very heavy	>125.0 mm	> 13 cm

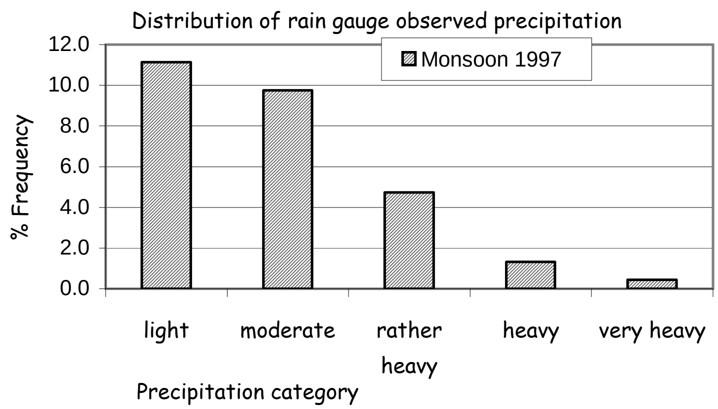


Fig. 2. Percentage frequency distribution of the number of rain gauge-day in different precipitation categories (rainy days only)

a measure of the error in model prediction of the seasonal mean while the mean absolute error and rmse gives the measure of error in prediction of the magnitude of precipitation. The correlation coefficient provides an estimate in the phase error in prediction. In addition, skill of the model in forecasting precipitation is estimated by classifying the quantity of precipitation in various discrete categories.

The model forecasts are initiated from the 0000 UTC analysis of everyday. Since, precipitation observations are accumulated for a 24-hour period ending on 0300 UTC of the reporting day, the model forecasts are accumulated between 3-hr and 27-hr forecasts, between 27-hr and 51-hr forecasts, between 51-hr and 75-hr forecasts and between 75-hr and 99-hr forecasts. These periods are henceforth referred to as day-1, day-2, day-3 and day-4 forecasts.

3. Characteristics of observed precipitation over India

The seasonal precipitation pattern (Rao, 1976) for the monsoon season over India is dominated by a region of large values, exceeding 1000 mm, over the east of the country. Other regions of large seasonal precipitation are along the west coast and over northeast India, where blocking effects of topography induce upward vertical velocity and enhanced precipitation in the windward side and suppress precipitation in the lee side. The rain shadow due to western ghats spreads over a large area of the peninsula and is most prominent in the south of Tamilnadu where seasonal totals are less than 50 mm. Another feature is the presence of an east-west circulation (Das, 1962) with upward limb over the northeast parts of India and downward limb over Rajasthan. Local scale features and passage of synoptic scale systems further modify this gross picture.

The monsoon season of 1997 was a normal one as the spatial average of seasonal total rainfall for the whole of India was 102% of the long term mean (Climate diagnostic bulletin of India 1997). This was in spite of the onset over Kerala being delayed by 8 days. This year was also exceptional as the number of depressions and more intense systems numbered six in the season - the highest in the preceding decade.

Grid box average rain – Monsoon 1997

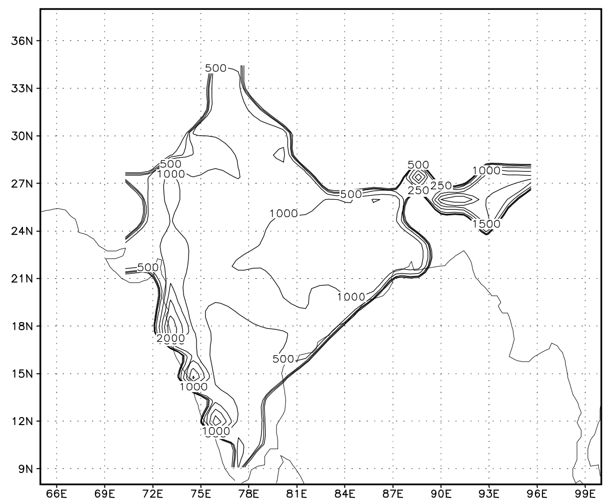


Fig. 3. Observed seasonal precipitation averaged over T80 model grid boxes

For a more detailed study of its distribution, the precipitation amounts are divided into 6 main categories as described in Table 2. The limits of these categories are similar to but slightly different from those defined by IMD. The present limits are used for all subsequent work with both rain gauge observed and grid-box average precipitation.

The frequency distribution of the observed precipitation from all rain gauges during the monsoon of 1997 is shown in Fig 2. As expected, the observations of small amounts of precipitation are more numerous than those of heavier falls. During monsoon 1997, the number of occurrences of trace or no rain constituted 72.6% of all observations while that of heavy and very heavy rainfall constituted only 1.3% and 0.4% respectively. For an individual station, the amount of 8397.4 mm, recorded at Mawsynram is the highest seasonal total precipitation in 1997 with 7636.8 mm, recorded at Cherrapunji, a close second. The lowest cumulative amount recorded during this season is 0.0 mm at Valinockam and Mandapam both located in the district of Ramanathapuram in south Tamilnadu.

In 1997 the heaviest 24-hr precipitation recorded during the monsoon season was 800.0 mm at Barwani/Rajghat in the district of Nimar in Madhya Pradesh on 3 of August. After averaging over a grid box, the magnitudes of heaviest precipitations decrease considerably and also change in both space and time. This is because the areas covering heaviest falls (usually associated with mesoscale convective systems) are much smaller than the grid box area and the horizontal scales of such systems vary considerably from one system to another. The grid box 78, that includes Barwani/ Rajghat recorded 145.8mm after averaging, as the other stations in the box did not record heavy rains. For the monsoon of 1997 the highest grid box average precipitation of 410.6 mm was observed on 29 July for grid box numbered 21 covering parts of northwest Rajasthan. The grid box average precipitation for the 135 grid boxes over India, accumulated over the monsoon season, is shown in Fig. 3. Since, observations for grid boxes covering neighbouring countries like Bangladesh, Nepal, China and Pakistan, and also sea areas, are not available, precipitation values for all grid boxes, excluding the 135 grid boxes covering India, are zeroes. This leads to a spurious strong gradient in the contour values near the outer boundary of the 135 grid boxes used in this work. The three regions of climatologically heavy precipitation namely, the windward sides of Western Ghats extending to the eastern parts of Gujarat, Khasi and Jayantiya hills and the eastern part of the country are well reflected in 1997. In addition an area covering eastern parts of Rajasthan also recorded rainfall exceeding 1000.0 mm, which is much above its climatological normal of 631.0 mm. It may be noted that the sharp gradient in the precipitation amounts between the windward and the leeward sides of the hills are somewhat smoothed in the grid box averaging as the size of the box is much larger than the width of the region of sharp gradient. Another feature of the grid box average precipitation during monsoon 1997 is the appearance of

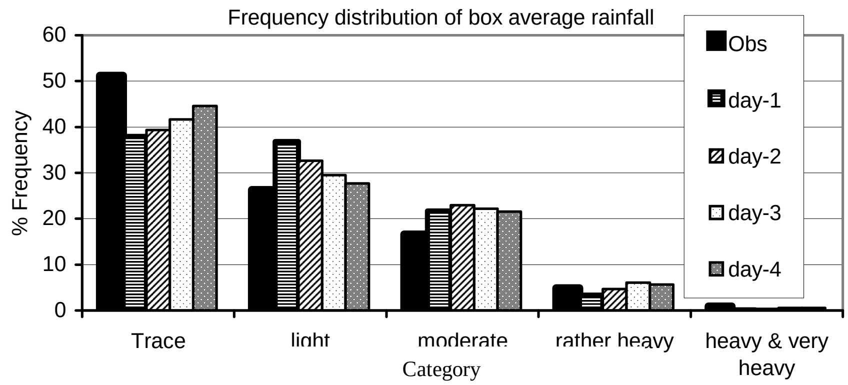


Fig. 4. Percentage frequency distribution of the number of grid box-day in different precipitation categories

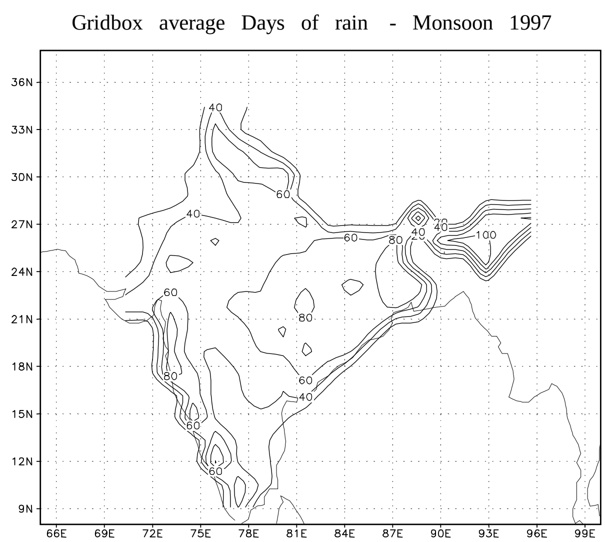


Fig. 5. Observed grid box average rainy days in monsoon 1997

three centres of heavy precipitation along the west coast, instead of a continuous isohyet of 1500 mm covering the whole of west coast as in the long term normal.

The distribution of the percent number of occurrences of box averaged observed and model forecast precipitation is shown in Fig. 4. The frequency of trace/no rain category has reduced from 72.6% to 51.2% by spatial averaging. The frequency of occurrences of heavy and very heavy precipitation also decreases on space averaging and together they contribute only 0.94% instead of 1.7% for rain gauge stations. The spatial averaging tends to cluster the rainfall values in the light and moderate categories.

The spatial distribution of the number of rainy days (days when precipitation exceeds 2.5 mm) in a grid box, as observed in 1997 is shown in Fig. 5. It may be noted that the box numbered 45 (including both Mawsynram and Cherrapunji) has the highest seasonal total precipitation of 3083.1 mm in 1997 with 113 rainy days while the box numbered 108 over southern Konkan recorded a seasonal total precipitation of 3007.5 mm in only 92 rainy days. The lowest number of rainy days in a grid box is 2 for the box numbered 13 (seasonal total precipitation of 35 mm) over northwest Rajasthan and adjoining Haryana.

During every monsoon season there are few spells of enhanced activity, normally associated with some

TABLE 3

List of synoptic scale precipitation giving systems during Monsoon 1997

Number	Intensity	Start date	End date
1	Deep depression	26 June	29 June
2	Deep depression	30 July	2 Aug
3	Deep depression	5 Aug	7 Aug
4	Deep depression	20 Aug	25 Aug
5	Depression	29 Aug	30 Aug

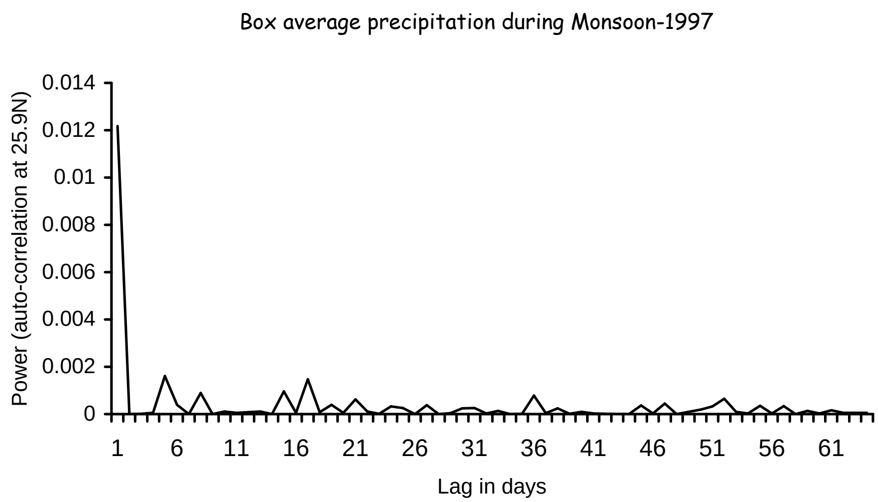


Fig. 6. Power spectrum of adjoining box average precipitation during the monsoon of 1997

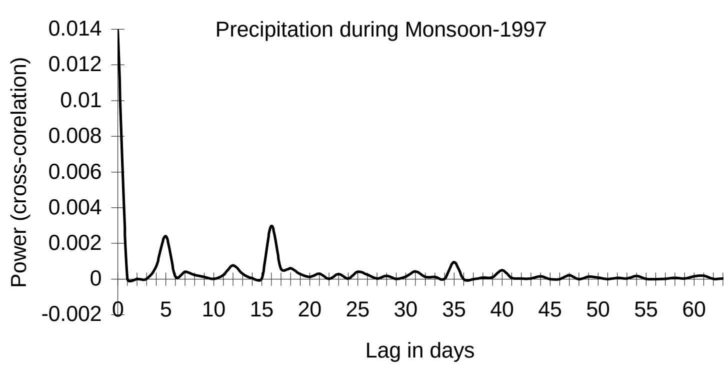


Fig. 7. Power spectrum of cross-correlation between area average precipitation near monsoon trough and over north Andhra Pradesh

Gridbox average total rain – 1997 27-hr forecast

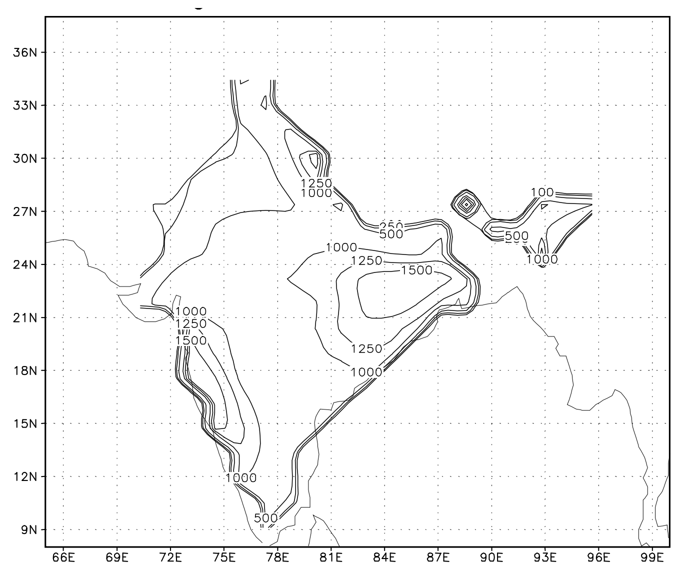


Fig. 8. Model predicted (day 1) seasonal total precipitation in monsoon 1997

well-organized synoptic system moving from east to west. The heavy and very heavy precipitation events usually occur during these active periods. Prediction of such active spells of precipitation is an important component in the verification of model forecasts. During the monsoon of 1997 five depressions/deep depressions occurred between June and August as listed in Table 3. The highest grid box average rain (over box numbered 21) however, was due to a low-pressure area moving from east to west between 21 of July to 29 of July.

In addition to the east-west propagating mode in the rainfall associated with the monsoon systems originating over or near the head Bay, a low frequency south to north mode with a time period of 30 to 40 days was detected [Sikka and Gadgil (1980)] in the pattern of maximum cloudiness during monsoon. To examine whether such a mode exists in the area averaged precipitation in 1997, two regions centred close to 25.9° N and 17.5° N are chosen. Each region consists of two adjoining boxes lying at the same latitude and centred near 78.0° E and 79.5° E meridians. Power spectrum of auto-correlations at both sites and cross-correlations between them does not indicate any significant power near the time period of 30-40 days. In fact for autocorrelations, the only time period where some power is concentrated (Fig. 6), is 1-day. This indicates that near the eastern end of the monsoon trough, the spells of rain most often tend to persist for more than a day. This is substantiated by the large number of rainy days recorded here. The cross-correlation between area average precipitation near 25.9° N and near 17.5° N shows (Fig. 7) that peaks in power spectrum appear for time periods of 0-day, 5-day and 16-day. The first one indicates that precipitation over north Andhra during monsoon is mostly due to large- scale phenomena that give precipitation at the eastern part of monsoon trough and north Andhra simultaneously. The absence of 30-40 day south to north mode, to the north of 17.5° N, in 1997 agrees with the findings of Ramasastry *et al*. (1986).

4. Characteristics of model predicted precipitation

The model predicted day-1 (accumulated between 3 hr and 27 hr forecasts) precipitation for the monsoon season of 1997 (Fig. 8) shows that the three regions of climatologically heavy precipitation are well represented in the seasonal accumulated values. This feature of the model predicted precipitation was also present in the predictions of monsoon 1995 (Basu *et al*., 1999). The area of large precipitation over the east of the country is not reproduced in many atmospheric models in the climate simulation mode (Gadgil *et al*., 1998) or in seasonal simulation mode (Basu, 2001). Unlike the other regions of large precipitation produced by fixed lower boundary

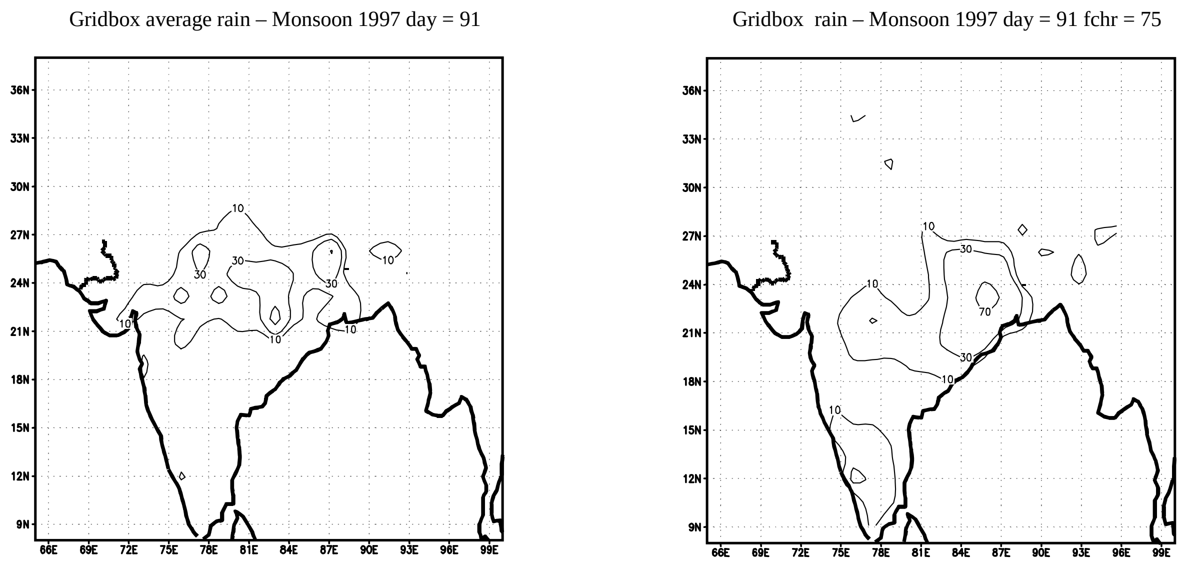


Fig. 9. Precipitation amounts observed (left) and in day-2 model prediction (right) for 30 August 1997

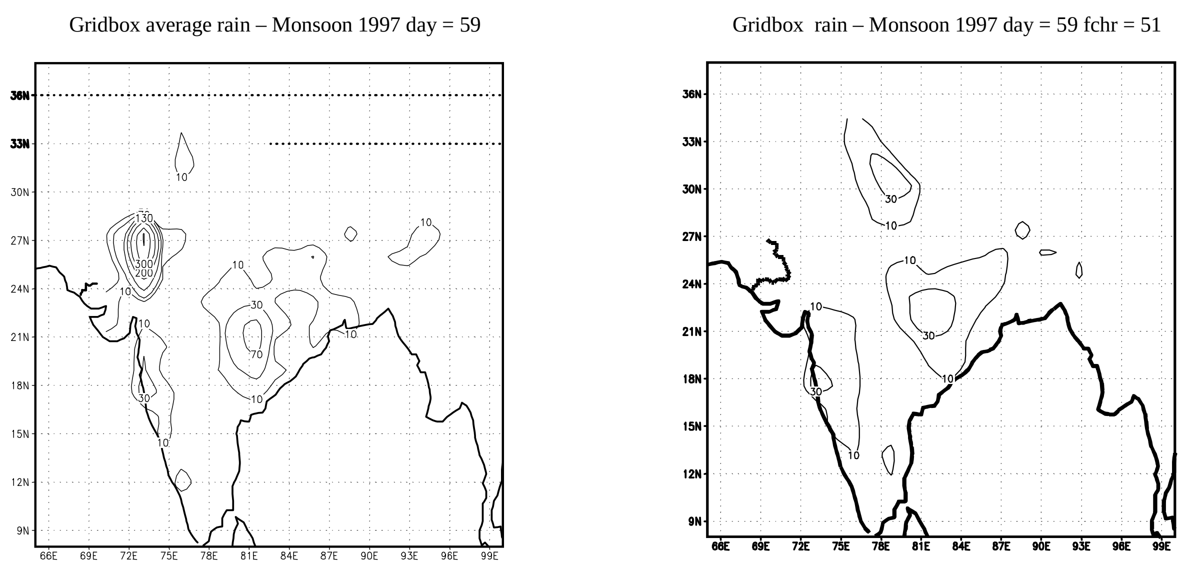


Fig. 10. Observed (left panel) and model predicted day-2 precipitation on 29 July 1997

condition (topography), the area of large precipitation over the eastern part of the country is due to the proper prescription of the atmospheric initial condition. In day-1 forecast the maximum magnitude of precipitation, over all boxes and the whole season, is only 204.4 mm (on 30 June over the grid box numbered 99) that is about half the value of the highest observed grid box average precipitation of 410.6 mm for a day. This, alongwith the frequency of occurrences of precipitation in different categories, implies that the model under-predicts the occurrences of very heavy precipitation. The magnitude of the cumulative precipitation, averaged over all the grid boxes, is 1023.7 mm in the day-1 prediction compared to 937.5mm for the observed. Thus, the model over predicts the accumulated all India total. It can be concluded from the frequency of trace/no rain category in Fig. 4 that the model predicts precipitation more frequently than actually observed. The occurrences of light and moderate precipitation are more numerous in the model prediction while the occurrences of heavy precipitation are considerably reduced in

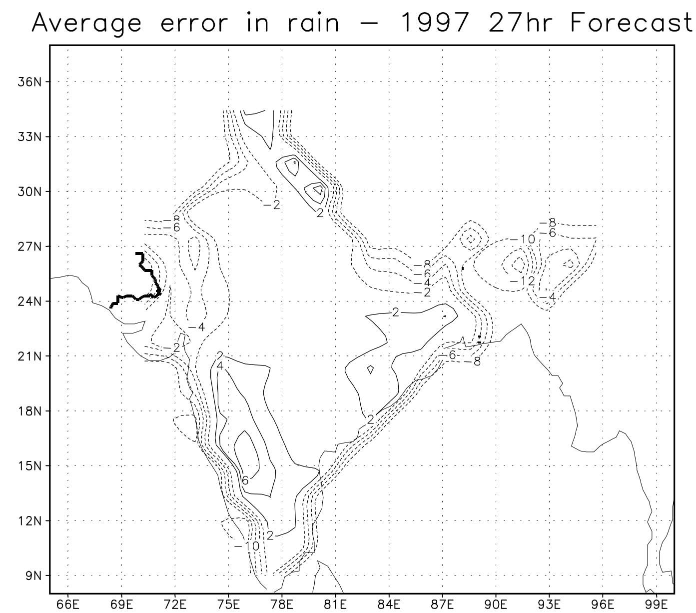


Fig. 11. Seasonal average error in precipitation intensity (mm/day)

number. In the day-2, day-3 and day-4 forecasts, the all grid box average accumulated precipitation during monsoon 1997 were 1065.2 mm, 1004.2 mm and 893.0 mm respectively. The magnitude of highest precipitation decreased to 146.9 mm, 153.3 mm and 113.4 mm respectively for the above three forecast lengths. Thus, the cumulative predicted precipitation, averaged over the whole of India, does not vary much during the evolution of the model forecast at least up to day-3. The maximum 24-hr accumulated precipitation, however, decreases significantly after day-1.

During the days when depressions/deep depressions were present over the Indian land area, the observed grid box average precipitation shows that the model under predicted the precipitation associated with the depressions but often over predicts the distant effect (Mukherjee & Shyamala 1986) of enhanced activity along the west coast and adjoining areas. In 1997 a depression formed close to head Bay on the 29 of August. The spatial distribution of the observed and model predicted precipitation (day-3 forecast) valid for 30 August is shown in Fig. 9. Both precipitation fields include rain over Vidarbha and adjoining areas where no marked synoptic system was present on that day. The model also predicts precipitation along west coast, far to the south.

The system producing the heaviest 24-hour accumulated precipitation (410.6 mm) in a grid box (numbered 21) was a low-pressure area forming over Bihar on 22 July 1997 and arriving over Rajasthan on 29 July. On 29 July, the other regions of heavy rain were located along the west coast and over the area covering northwest Andhra Pradesh and southeast Madhya Pradesh. These areas of rather-heavy/heavy precipitation were predicted well up to a forecast length of day-2 (Fig. 10) but the model shifted the area of phenomenal precipitation over Rajasthan to further northeast, where a small area of observed moderate precipitation was over predicted as a large area of rather heavy precipitation. Thus, the NCMRWF atmospheric forecast model could not predict the incidence of heavy precipitation associated with a synoptic scale system. In this case, the low-pressure area over Rajasthan was not well represented in the initial condition.

5. Comparison of model predicted and observed precipitation

The difference between the model predicted and the observed seasonal total precipitation (Fig. 11) shows that the model over predicts precipitation over an area to the south of the eastern end of the monsoon trough while it

Correlation coefficient in rain – 1997 27 hr Forecast

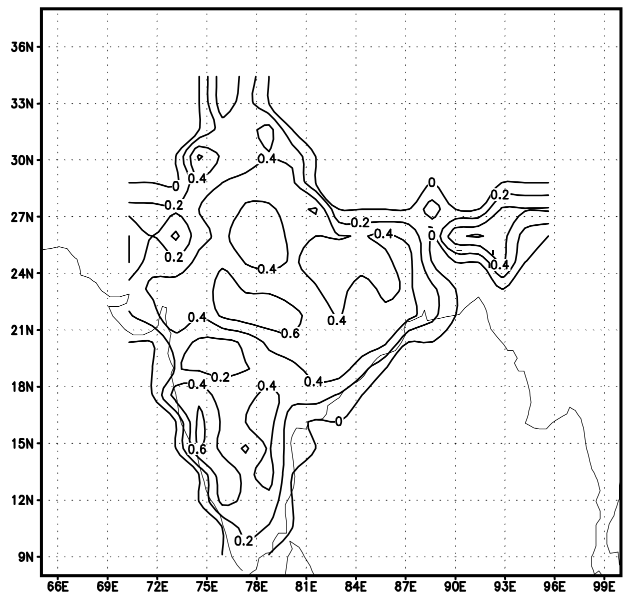


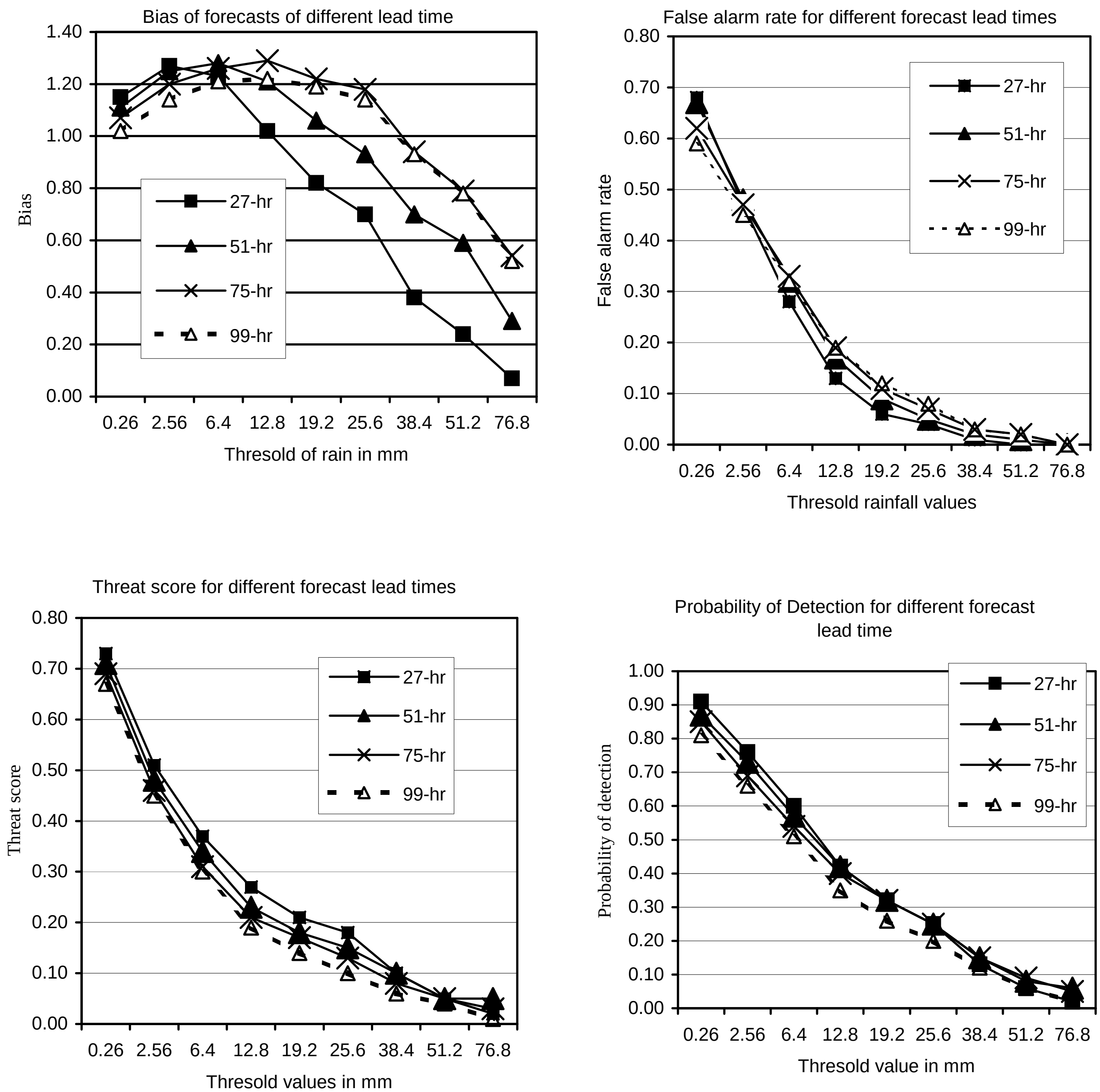
Fig. 12. Correlation between trends in observed and day-1 forecast precipitation

under predicts precipitation over the rest of the country. The predicted seasonal average of precipitation intensity is within 2 mm/day of the observed amounts over most of the monsoon trough region. Near the Western Ghats, the predicted values are smaller than observed (by up to 10mm/day) in the windward side but are higher than observed (by up to 6mm/day) in the rain shadow area (leeward side) to the east. The region of excess precipitation over Rajasthan is not reproduced in the model forecast, as the synoptic systems producing precipitation over this area were not captured by the assimilation procedure even though the systems were over land.

The rmse of the day-1 predicted precipitation (not shown) has magnitude less than 15 mm/day over most of the country. This indicates that the forecast error in precipitation is not purely random but has some systematic component. The magnitude of rmse increases slowly with the forecast length and by day-4 contour values increase to 20 mm/day.

The correlation coefficient (CC) between the observed and the forecast precipitation at day-1 is shown in Fig. 12. Over most of the central India, the magnitude of CC exceeds 0.4, which is considered to be good for precipitation. A small area bordering Maharashtra and Madhya Pradesh has a magnitude of CC exceeding 0.6. The heavy precipitation areas in the windward sides of hills also have small regions of CC exceeding 0.6. This indicates that the day-1 predicted precipitation trends are in phase with that observed over the regions of heavy precipitation. The magnitude of CC decreases with forecast length and by day-4, CC values over most of India are less than 0.2 except in a few pockets near the Western Ghat where the CC values still exceed 0.4.

As discussed earlier, the standard WMO method of verification of numerical weather prediction model outputs by computing the mean error, rmse and CC is not suitable for precipitation due to its great temporal and spatial variability. The statistical parameters based on the frequency of occurrences in various classes are more suitable for determining the skill of a model in predicting precipitation. In Fig. 13, the bias, false alarm rate, threat score and probability of detection for the classes with class mark as 0.256 mm, 2.56 mm, 6.4 mm, 12.8 mm, 19.2 mm, 25.6 mm, 38.4 mm, 51.2 mm, and 76.8 mm corresponding to 0.01 inch, 0.10 inch, 0.25 inch, 0.50 inch, 0.75 inch, 1.00 inch, 1.50 inch, 2.00 inch and 3.00 inch respectively are presented. For same class marks the precipitation forecasts from the global spectral model at the National Centers for Environmental Prediction (NCEP) of USA, have similar values for the bias and the equitable threat scores during June to September of 1997 (available at the NCEP website).



Figs. 13. Variation of statistical parameters of model predicted precipitation for different forecast length

The bias of a model forecast is a comparison of the model predicted number of occurrences of an event with that actually realised in nature. In the present case, the model over predicts rainfall up to 12.8 mm (0.5 inch) in the 24-hr forecast while it under predicts events of higher magnitude. The NCEP model also over predicts events of lower magnitude but the cross over to under prediction occurs at a higher value close to 1.5 inch. Also for the NCEP model the value of bias increases initially to exceed 1.4 for the class mark 0.5 inch while the bias never exceeds 1.3 for the present model. For increasing forecast length the value of the class mark at the cross over point (from over prediction to under prediction) also increases. For day-3 and day-4 predictions this value is close to 1.25 inch.

The threat score is the ratio of the number of successful model prediction of an event to the number

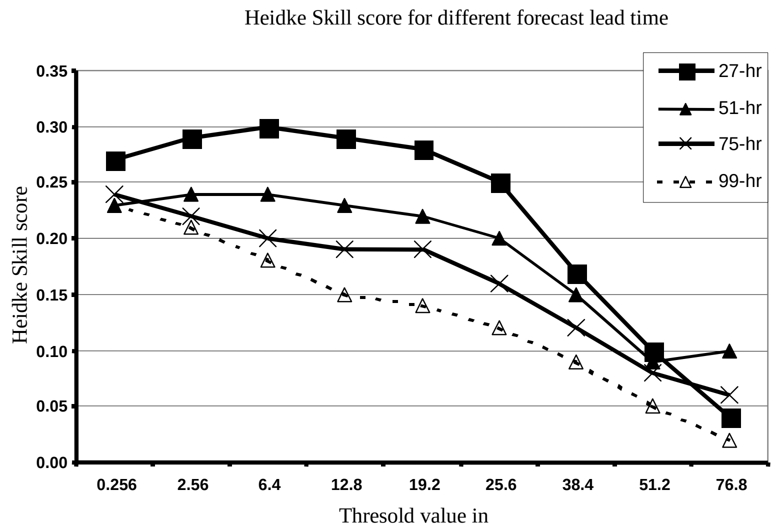


Fig. 14. The Heidke skill score for monsoon 1997 for different prediction length

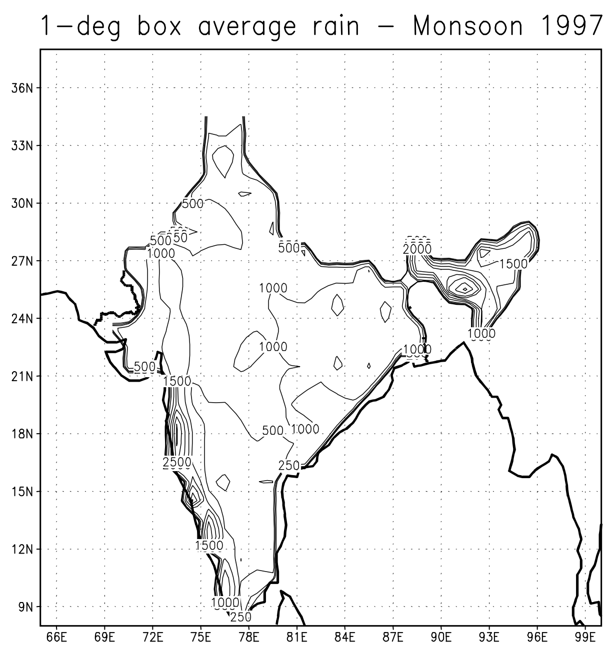


Fig. 15. Observed seasonal total of monsoon 1997 : 1-degree square average

of all such events in both observed and predicted. Higher value of threat score indicates better prediction with a theoretical limit of 1.0 for a perfect model. The average threat score for the summer monsoon of 1997 over the Indian region starts close to 0.7 and then decreases to 0.3 near the 0.25 inch mark. This is better than the prediction of NCEP models over the USA. The latter never exceeds a value of 0.3 for the same period of time.

The false alarm rate (FAR) is the fraction of wrong prediction out of the total number of non-occurrences of the event. For perfect prediction value this parameter should be 0.0. In the present case FAR is large for classes with small class mark but decreases markedly with increase in class mark and is practically zero for class marks above 1.0 inch. The high values for both threat score and FAR for lower class marks indicate that predicted occurrences of precipitation in these classes far exceed the observed frequencies.

The probability of detection (POD) is the fraction of correct prediction out of all predictions of the occurrence

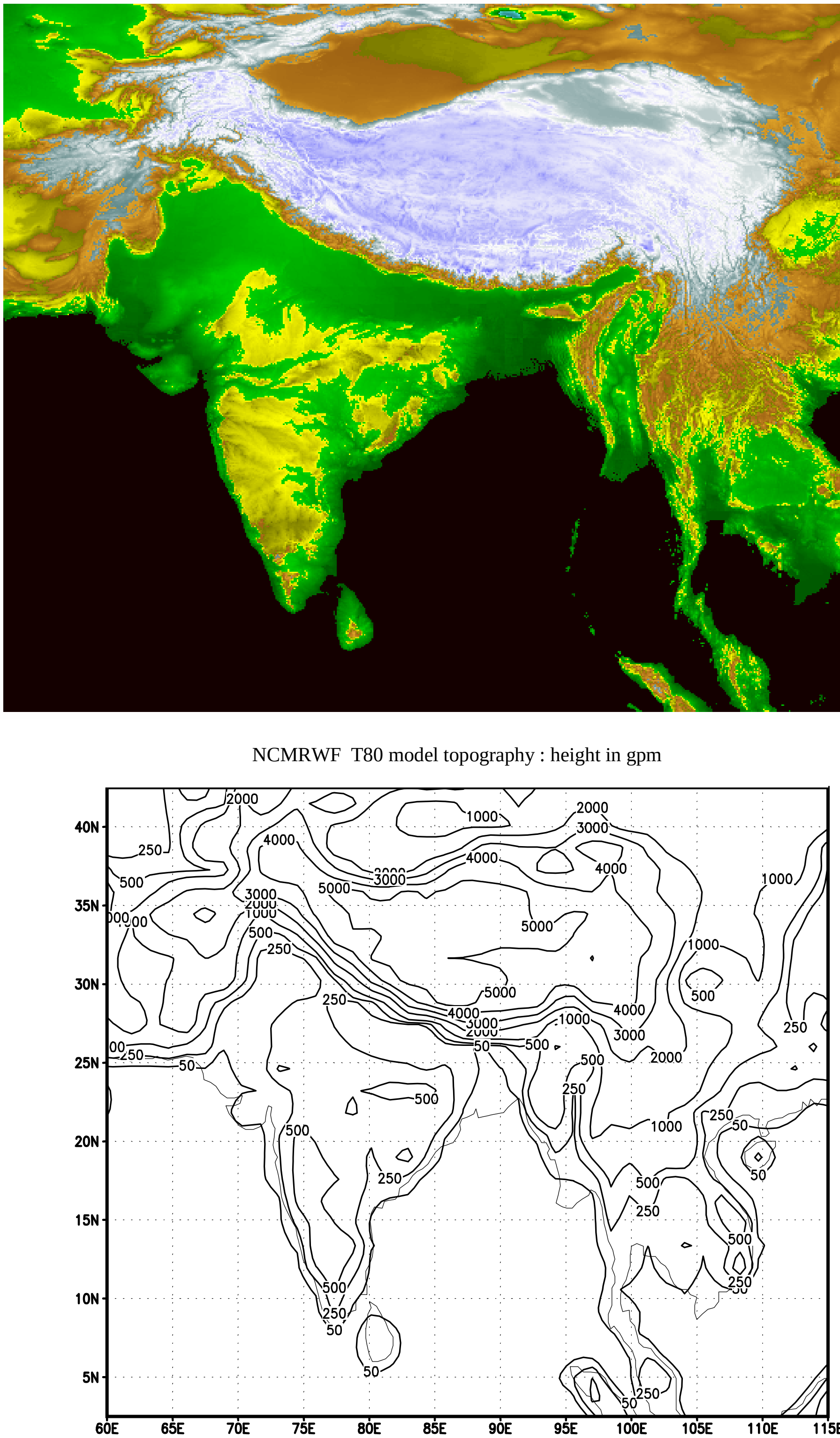


Fig. 16. Actual (1km resolution GLOBE data) and NCMRWF T80 model (lower panel) topography

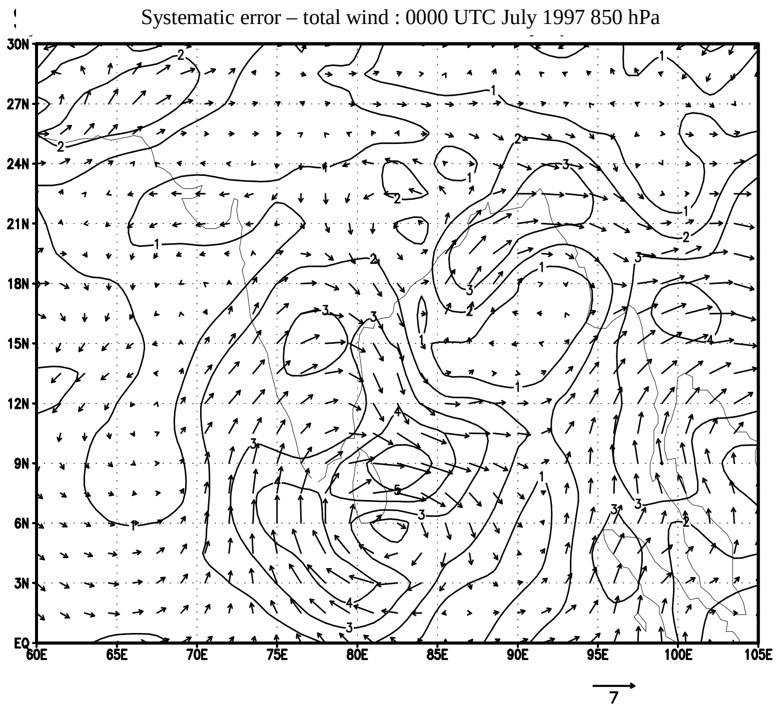


Fig. 17. Systematic wind error at 850 hPa during July 1997 for 24-hr prediction

of an event and a higher value for the quantity indicates the ability of the prediction model to capture the occurrence of desired events. In the present case, the probability of detection is more than 50 percent for class marks below 0.5 inch. The bias, threat score, FAR and POD together provide an estimate of the quality of model prediction. A comparison with the values of bias and threat score of the NCEP model predictions over the USA, suggests that the quality of precipitation forecast over India during the monsoon of 1997 is similar to that of the NCEP model over USA during the same year.

Another statistical parameter that directly provides an estimate of the usefulness of the model prediction is the skill score, which is the ratio of the number of correct predictions by the model above those obtained by chance to the number of total predictions above those obtained by chance. Large positive value of skill score implies that the model predictions are useful. The skill score in the present case is shown in Fig. 14. Both day-1 and day-2 predictions of precipitation have significant skill scores for categories of precipitation with class mark up to 1.0 inch. The skill falls off rapidly for larger precipitation amounts and also for longer prediction range.

6. Discussion & conclusion

Quantitative precipitation forecast is one of the most difficult tasks of numerical weather forecasting, especially as liquid water content is not a prognostic variable in most of the atmospheric models. In most of the atmospheric models precipitation is an ad hoc fraction of the near ground moisture condensed by lifting. The verification of the model predicted precipitation is also a difficult task as routine analysis is not available for comparison. Because of this no routine verification of precipitation, produced by the model, is done at NCMRWF or other numerical prediction centres. At NCEP an estimate of the skill of precipitation is obtained by computing a few statistical parameters using the model forecast precipitation at regular grid points and observed precipitation at irregular rain gauge sites. At NCMRWF location specific forecasts of precipitation and other atmospheric variables are derived from the model predictions by applying statistical and synoptic methods. These location specific forecasts are routinely verified against observations at the specific locations but are not the same as model predicted precipitation. To the best of knowledge of the authors, the present method of spatial averaging of observations for verification of model predicted precipitation is being attempted for the first time.

Spatial averaging over a number of stations has an effect of reducing the extreme values and smoothing out the sharp gradients. These effects are heightened for variables like precipitation that have large variations both in space and time. An averaging of the same data over 1-degree square boxes produces a highest value of seasonal total of 5308.1 mm over northeast India instead of 3083.1mm in the T80 grid box (approximately 1.5-degree square) averaging. Over the Western Ghat, the highest value of the box average is not changed much due to the decrease in the area of averaging but a new centre of heavy precipitation appears in the contour plot of the seasonal precipitation (Fig. 15). Thus, the precipitation field over northeast India is more localised in nature than that over the Western Ghat. The rain shadow effect is also more enhanced over the Khasi and Jayantiya hills where observed seasonal total precipitation amounts drop from 8397.4 mm over Mawsynram and 7636.8 mm over Cherrapunji to 426.9mm over upper Shillong, within a distance of about 50 km. The all India average of all (264) 1-degree squares, however, has a value of 933.4 mm that is almost same as the value of 937.5 mm obtained by averaging over all (135) T80 grid boxes. Over the monsoon trough region and also the rest of India, the area of spatial averaging has little effect indicating that the seasonal total is more homogeneous over these regions. Thus conclusions based on T80 grid box averaging is not unduly affected by the size of the grid box over which averaging is done.

The dipole pattern in the seasonal precipitation forecast error (Fig. 11) across the Western Ghats is due to the smearing of the narrow mountain ridge by the low-resolution forecast model. A comparison of the observed 1-km resolution topography [The Global Land One-km Base Elevation (GLOBE) Project] and that used in the NCMRWF model (Fig. 16) clearly shows the shift of the ridge away from the west coast of India and also the flattening of the steep gradient in the windward side of the plateau. This explains the decrease in the intensity of precipitation in the model over the windward side of the Western Ghats and increase over the rain shadow area.

The excess precipitation near the eastern end of the monsoon trough is related to the systematic error in the wind field. For the representative month of July, the monthly mean systematic error of 24-hr prediction at the 850 hPa level shows (Fig. 17) cyclonic turning of wind and convergence over the east coast adjoining the head Bay. The anticyclone that dominates the error flow over the peninsula, also contributes to the excess precipitation in the lee of the western Ghat by pumping moisture from south. The anticyclone, in general, induces subsidence and decrease in precipitation.

From the result presented above, it can be concluded that, in 1997, the rainfall predictions from the NCMRWF global model were directly usable for quantitative precipitation forecast up to a forecast length of at least day-2. For longer forecast periods, rainfall values derived statistically from model predicted circulation features might be more useful than the direct model output of precipitation. The mean forecast error is large across the Western Ghat, whose role as a barrier to the atmospheric flow is reduced in the model due to its low horizontal resolution. Over most of the monsoon trough region the magnitude of error in the intensity of precipitation is within 2 mm/day. The predicted precipitations are mostly in phase with the observed trends as seen from the magnitude of the CC between the predicted and observed precipitation fields.

The model has a bias towards over predicting the occurrences of light to moderate precipitation events while it under predicts the events in heavier categories. The skill of the forecast is more than 0.20 for categories of precipitation with class mark up to 1.0 inch.

Acknowledgements

The Director General of Meteorology (IMD) has kindly provided the observations of precipitation used in this work. The model used at NCMRWF for medium range forecasting is a modification of an older version of the forecast model of the NCEP.

The author is thankful to Dr. U. S. De, the Additional Director General of Meteorology (Research) for his help and advice regarding the use of observed precipitation.

References

Basu, B. K., 2001, "Simulation of the summer monsoon over India in the ensemble of seasonal simulation from the ECMWF reanalysed data", *J. of Climate*, **14**, 1440-1449.

Basu, B. K., Ramesh, K. J. and Harathi, P. A., 1999, "Intercomparison of the characteristic features of the southwest monsoon over India as reproduced in the mean monthly analyses and forecasts of some operational NWP centres during 1995", *Meteor. and Atmos. Phys.*, **69**, 157-178.

Das, P. K., 1962, "Mean vertical motion and non-adiabatic heat source over India during the monsoon", *Tellus*, **14**, 212-220.

Gadgil, S., Sajni, S. and participating AMIP Modelling groups, 1998, "Monsoon precipitation in the AMIP runs", WMO/TD No. 837, pp28+figures.

India Meteorological Department, 1997, "Climate diagnostic bulletin of India", Monsoon season, Office of the ADGM (Research), Pune, p22.

Kalnay, E., Sela, J. G., Campana, K. A., Basu, B. K., Schwarzkopf, M. D., Long, P. E., Caplan, P.M. and Alpert, J. C., 1988, "Documentation of research version of the NMC medium range forecasting model".

Kuo, H. L., 1974, "Further studies of the parameterization of the influence of cumulus convection on large-scale flow", *J. Atmos. Sci*., **31**, 1232-1240.

Manabe, S., Smagorinsky, J. and Strickler, R., 1965, "Simulated climatology of a general circulation model with a hydrological cycle", *Mon. Wea. Rev*., **93**, 769-798.

Mukherjee, A. K. and Shyamala, B., 1986, "Distant effect of monsoon depressions: part 1 – Effect of formation of depressions on rainfall activity along the west coast of India", *Mausam*, **36**, 429-434.

Murphy, A. H. and Winkler, R. L., 1987, "A general framework for forecast verification", *Mon. Wea. Rev*., **115**, 1330-1338.

Ramasastry, A. A., De, U. S., Vaidya, D. V. and Sundari, G., 1986, "Forty-day mode and medium range forecasting", *Mausam*, **37**, 305-312.

Rao, Y. P., 1976, "Southwest monsoon", India Meteorological Department, Meteorological Monograph on Synoptic Meteorology, No. 1/1976, Delhi, p337.

Sikka, D. R. and Gadgil, Sulochana, 1980, "On the maximum cloud zone and the ITCZ over Indian longitude during the southwest monsoon", *Mon. Wea. Rev*., **108**, 1840-1853.

Tiedtke, M., 1983, "The sensitivity of the time-mean large scale flow to cumulus convection in the ECMWF model", Proc. of the ECMWF workshop on convection in large scale numerical models, Reading, UK, 297-316.

WMO No. 485, 1992, "Manual on the global data processing system", Attachment II-7.

WMO No. 168, 1994, "Guide to hydrological practices".