

Seasonal simulation of the monsoon with the NCMRWF model

Ravi S. Nanjundiah

Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, Bangalore 560 012, India

The development of a climate version of the National Centre for Medium Range Weather Forecasting (NCMRWF) model and its implementation on a message passing platform are discussed. The ability of the model to simulate the Indian Summer Monsoon is studied. It is found that the model realistically simulates most of the major features of the Indian Summer Monsoon. The intraseasonal variations of the Indian monsoon are also studied.

1. Introduction

It has long been recognized that understanding the monsoon and its variability is not only a problem of great academic interest but of applied importance as well. Many attempts have been made to understand/predict monsoons using statistical models. With rapid strides in the field of numerical weather prediction and computer technology, attempts are now being made to predict monsoons using atmospheric General Circulation Models (GCMs). However, before such attempts are undertaken it is necessary to understand the capability of the model in simulating the monsoons and its variability. The simulation of the monsoon rainfall over India is considered as a demanding test for atmospheric GCMs^{1,2}. Attempts to obtain a realistic simulation of this feature have continued for over two decades.

The earliest attempts to simulate the mean monsoon circulation was by Murakami, Godbole and Kelkar³. They were successful in simulating some of the large-scale features of the monsoon circulation. Subsequently, there have been a large number of modelling studies. While some important features related to monsoons such as the tropical easterly jet, Somali jet, low level westerly flow over India, etc. have been reasonably well simulated (with varying degrees of success), the simulation of the mean rainfall pattern has proved to be a challenging problem. Washington⁴ has shown that 'the precipitation amounts (and patterns) in the model varied greatly for each other and from those observed.'

Recently the first phase of the massive effort – the Atmospheric Model Intercomparison Project (AMIP, Gates⁵) has been completed and results of simulations of

all the major atmospheric models run with specified observed Sea Surface Temperatures (SSTs) for the period 1979–1988 are available. The analysis of AMIP simulations has clearly shown that a robust simulation (i.e. a simulation which is not very sensitive to changes in the model) of the monsoon rainfall and its variability is still a major problem with almost all the GCMs⁶.

Realizing the importance of weather forecasting for the Indian region, the Government of India set up the National Centre for Medium Range Weather Forecasting (NCMRWF). During the last decade this centre has developed considerable expertise in generating numerical weather forecasts for the medium range (3–5 days). It is now planned to conduct extended range forecasts (up to 15 days) and seasonal scale forecasts (i.e. forecasts for the entire season) using high resolution GCMs.

Before such attempts are made for extended range and seasonal scale forecasts it is necessary to study the characteristic of the GCM and its capacity to simulate the seasonal pattern of precipitation with specific reference to its capability in simulating the Indian Summer Monsoon.

Such studies could not be easily conducted, except at a few centres world-wide because of the large computational resources required for this. Now, with the advent of cost-effective parallel computers and the development of skills to use this effectively, such studies can be routinely undertaken within the country. In fact the first attempt to apply parallel computing to the study of the monsoon was made in India using a simple climate model and the Flosolver parallel computer (parallel computers designed and developed at NAL, Nanjundiah⁷, Sinha *et al.*⁸). The present work is the first effort at studying the seasonal pattern as simulated by the NCMRWF GCM. A multi-year simulation with such a high resolution GCM would not have been possible (especially at our centre) but for the use of parallel computing techniques.

This paper deals with the development of the climate version of the NCMRWF model, its implementation on a message passing parallel computer and the comparison of the simulated monsoon rainfall with observed features.

2. The NCMRWF model

The present study uses a version of the NCMRWF model. This is a global model (i.e. its domain of integration is

*e-mail: ravi@caos.iisc.ernet.in

global in the horizontal direction) in the vertical, its domain extends from the surface (nominally at a pressure of 1000 hPa) to the top level at about 21 hPa. The main variables predicted by this model are vorticity, divergence, temperature and specific humidity. Winds are computed from vorticity and divergence. Other prognostic variables are surface temperature, precipitation (both snow and rainfall) and surface wetness.

The model is spectral, i.e. time integration is done in the wave number space. However, some of the computations for the model need to be done in the physical (latitude–longitude) space, e.g. computation of radiation, surface effects, clouding and precipitation, evaporation, etc. (generally called as the physics of the model). The processes of advection, horizontal diffusion and time incrementation are generally referred to as the model's dynamics and these computations are conducted in the spectral space.

The tendencies calculated in the physical space need to be transformed to the spectral space and variables in the spectral space need to be transformed to the physical space (for the purpose of conducting such computation as radiation, cumulus convection, boundary layer and surface processes). The model's initial state is in the spectral space.

The model uses a semi-implicit scheme (with leap-frog technique) for temporal integration. It has a horizontal resolution of T80 (i.e. 80 waves are considered to represent a variable). In the latitudinal direction (i.e. north–south direction) it has 128 Gaussian latitudinal circles. On each of these circles it has 256 grid points in the zonal direction (east–west direction). A comprehensive documentation of the model's dynamics is given in Kalnay *et al.*⁹.

2.1 Development of the climate version

The original model as used at NCMRWF was a forecast model developed for a vector supercomputer such as the CRAY-XMP-216. The forecast models are usually run for short periods such as 3–5 days. The present version used for our studies is different from the operational version of the NCMRWF model, the most significant difference being the use of Simplified Arakawa–Schubert (SAS) parameterization scheme of Grell¹⁰ (when compared to the Kuo Cumulus parameterization scheme¹¹ used by the present operational version). The use of SAS in our study is of greater relevance as the newer versions of the operational model are likely to use this scheme. The major difference between the two cumulus schemes is that while the SAS considers the instability of the atmosphere as the dominant factor for determining the presence of clouds, the Kuo scheme considers large-scale low-level convergence of moisture as the major contributing factor for the presence of clouds and precipitation.

Climate models need to be run for longer periods (typically a few years and up to thousands of years if very long

range effects need to be studied). While the physics (i.e. parts of the model pertaining to radiation, clouds, precipitation, surface fluxes, etc.) and dynamics of the model (pertaining to numerics, advection and horizontal diffusion) could be the same, certain modifications are necessary to suit the specific needs of a climate model.

To suit the purpose of climate studies at the Centre for Atmospheric and Oceanic Sciences (CAOS), the model was modified as follows:

1. Modification of the code from the CRAY architecture to a generic version capable of running on any workstation. It is noteworthy that a version of the NCMRWF model has been completely re-engineered, simplified and reduced considerably in length. The re-engineered version is now capable of running on a variety of platforms including a PC¹².
2. Modifying the version obtained in the previous study to a message-passing architecture using message passing constructs such as the MPL/MPI for implementation on the IBM-SP2 architecture.
3. Evolution of the above forecast model into a climate model. This needed the introduction of the following capabilities:

- (i) To update the boundary conditions over the oceans (by specifying the SSTs, Shea, Trenberth and Reynolds¹³).
- (ii) Storing the history state of the model (to study the evolution of the simulated state) in a format to facilitate further post-processing.
- (iii) Modification of the radiation grid. The original forecast model conducted radiation computation on a reduced grid (81 × 162) and these computations were conducted once a day. However, for a climate study, the interaction between radiation (especially the diurnal variation) and cloudiness is an important factor; thus the radiation needs to be invoked more frequently (invoked every three hours in the climate run). Here we also conduct the radiation computation on the full physical grid. Computations on the reduced grid required global communication of the gather–scatter variety whose overhead was considerable (as a large array of data had to be moved between processors) and upon using a larger number of processors the increased throughput due to the reduced grid was largely negated.

2.2 Initial and boundary conditions

The model's integration was begun from an initial data set corresponding to 4 February 1993. For atmospheric computations a spin-up time (i.e. time for the initial transients to die down) of about sixty days is considered sufficient^{14,15}. For this purpose, we have not analysed the first three months of the simulations.

The temperature over the ocean was prescribed from the monthly mean climatological values of Shea,

Trenberth and Reynolds¹³. A linear interpolation technique was used to update SSTs from the monthly mean climatological values. The mean climatological values were considered to correspond to the middle of every month and for any calendar date a linear interpolation between the preceding and succeeding months was used. This is a standard technique used for prescribing SSTs in most atmospheric GCMs. The reason that SSTs are prescribed rather than calculated is that computing oceanic temperatures would require the use of an Oceanic General Circulation Model which would not only be computationally expensive (with spin-up times of the order of hundred simulated years) but it has also been found that simulations are less realistic when the oceanic and atmospheric components are allowed to evolve interactively, e.g. a simulation with Community Climate Model Version 3 (CCM3) with prescribed SSTs has a realistic simulation of the Indian Summer Monsoon while a coupled ocean-atmosphere system using CCM3 as the atmospheric component has a very unrealistic simulation of the Indian Summer Monsoon¹⁶. Almost all weather forecasting models use prescribed SST for specifying the oceanic conditions.

2.3 Parallel implementation of the model

We have used the method of domain decomposition to distribute load between processors. The global domain was split into smaller sub-domains, i.e. slices of latitude-circles assigned to a processor, (in a conventional computer, the entire domain would be assigned to a single processor, here the domain is split and each resultant sub-domain is assigned to processor). Symmetry about the equator is exploited in computing the spectral coefficients, with pairs of latitude-circles assigned to a processor (i.e. a processor computing for 30°N would also compute 30°S). This assignment of co-latitudes has the additional advantage of load-balancing in radiation computations. The pole that is in the summer hemisphere needs additional computing for radiation (due to longer days over this region) while the pole in the winter hemisphere needs less computing for radiation. Since the two poles would be assigned to the same processor, this computational imbalance is largely removed.

The method of latitudinal decomposition requires communication to complete the summation of spectral coefficients (Legendre transform)^{8,17}. The communications are global in nature, i.e. calculation of spectral coefficients requires data from every point on the sphere. This stencil of communication greatly constrains the scalability (i.e. decrease in computational time with increase in the number of processors). We have used the method of binary tree summation to partially reduce this bottleneck. Work is presently underway on more refined models of communication to improve scalability. These include:

- Overlaying of computations (to reduce wait time of processors).
- Communications in the Fourier space to parallelize computations in the wave number space.
- Load balancing – We found that there exists a continuous increase in computational requirements from the pole to the equator. This is due to the increased computations for clouds and convection in the tropical regions *vis-à-vis* the extra-tropics. Therefore instead of allocating contiguous latitude-circles to a processor (as its sub-domain), the processor's computational sub-domain is a mix of latitude-circles from the extra tropics and the tropics. This has considerably reduced the imbalance in the computations.
- Hybrid programming model – On shared memory machines, parallelization in the spectral space can be done using the shared memory constructs while the summation for spectral coefficients can be done using message-passing techniques. This increases the scalability considerably¹⁸.

It typically takes about thirty-five minutes to simulate one day on a Pentium II PC.

2.4 Validation of the parallel model

A lingering doubt always persists whether the massive changes done to the model's code have incorporated errors into the model or if the identified communication stencil is incorrect. Thus an important step in the development of a parallel version of the climate model is its validation. Presently available literature has little information on a commonly accepted yardstick by which transmigration of a GCM can be considered to be a valid one. Jones, Kerr and Hemler¹⁹ discussing the validation of a port, the SKYHI GCM on the CM-5, also highlight the fact that 'small differences in initial conditions and numerical accuracy of computation can have a significant effect on the statistics of the final solution. There are also systems that exhibit multi-decadal transient behaviour. It is not known whether the class of systems that encompasses general circulation models exhibit this behaviour. With constraints on the computational resources available, we can only validate and test the model for the periods of time SKYHI will be used to simulate. For SKYHI, these timescales are on the order of months or decades.' They consider a transmigration to be valid if the diagnostics were less than 10^{-12} after one time-step and less than 1% after one day for the most sensitive diagnostics. Using their approach we have examined the global rms errors for temperature, divergence, vorticity and moisture and found them to be within the limits suggested by Jones, Kerr and Hemler¹⁹. A detailed description of the model's validation is given in ref. 20.

3. The problem: The mean monsoon pattern

The validated parallel climate model has been used to study the simulation of the Indian Summer Monsoon. In this paper we present the average of four summer seasons for which simulations had been conducted. Further integration of the model is currently underway. We compare this with

the long-term observed climatology of the station data (Figure 1 *a*) obtained from India Meteorological Department (IMD) and with the climatology of the National Centres for Environmental Prediction (NCEP) data set obtained by merging the station data over land with Microwave Sounding Unit (MSU) data over the oceans (Schemm *et al.*²¹, we shall refer to this as NCEP-merged data, Figure 1 *b*).

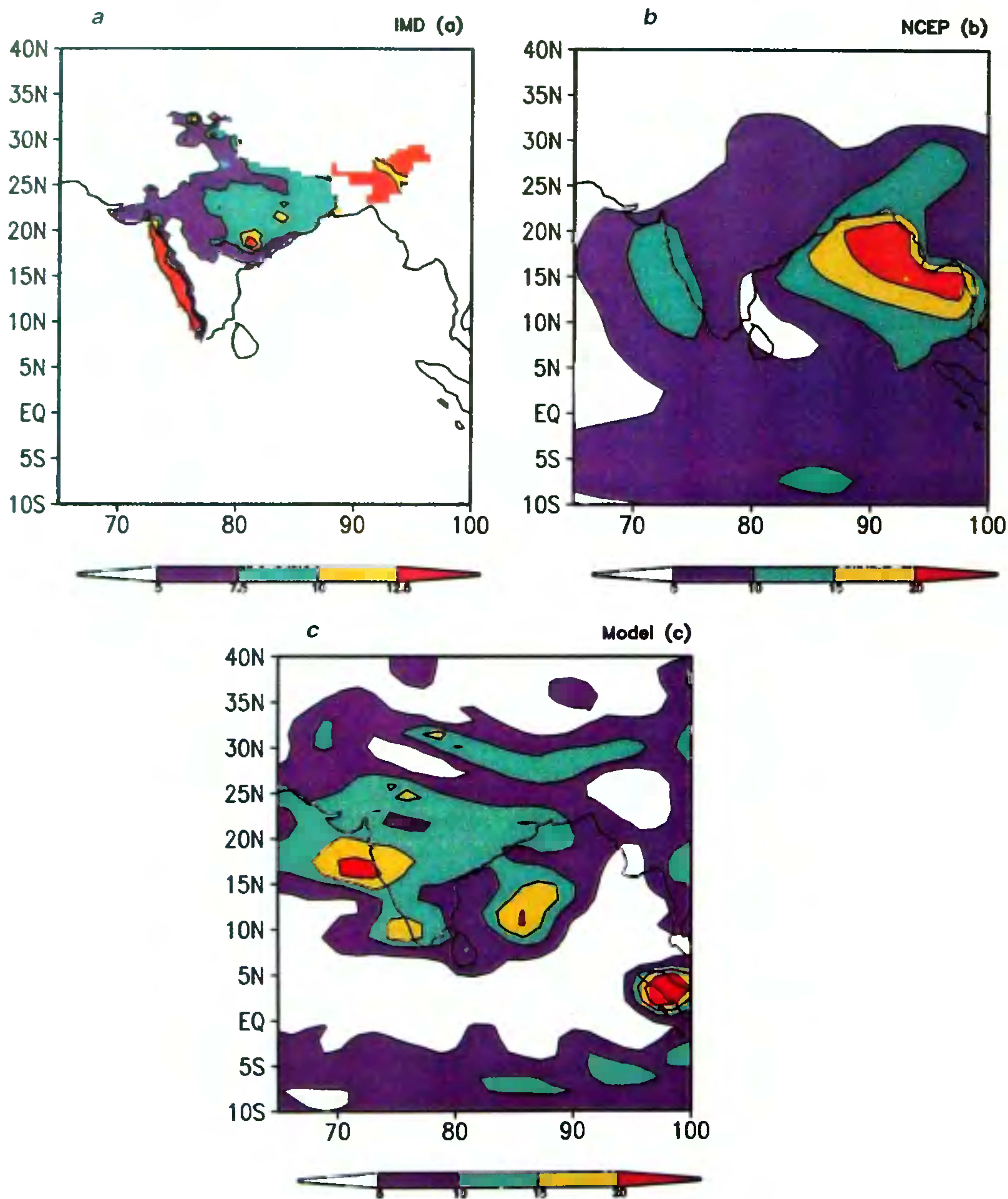


Figure 1. June–September seasonal mean rainfall (mm/day) in *a*, station data from IMD; *b*, NCEP merged dataset; *c*, NCMRWF model simulation.

Most of the rainfall over the Indian subcontinent occurs during the summer monsoon season, i.e. June–September²². We present a comparison for the entire summer monsoon period, comparing the climatology of the four summer seasons with the observed climatological mean.

The summer monsoon rainfall is associated with a rainbelt extending from the Head Bay (of Bengal), west north-westward across the Indo-Gangetic plain (Figure 1 *a* and *b*). This rainbelt fluctuates within the monsoon zone, north of about 15°N and its variation in location and intensity gives rise to the variation in the large-scale monsoon rainfall. This rainbelt is associated with the low pressure zone known as the monsoon trough (we shall refer to this as the monsoon rainbelt). Another region of intense convection occurs over the waters of equatorial Indian Ocean (Figure 1 *b*) intermittently throughout the summer season²³. Intense precipitation also occurs over areas of elevated orography such as the Himalayas and the Western Ghats. The fact that these two rain bands are not

as prominent in Outgoing Longwave Radiation (OLR) (indicators of organized large-scale convection) as in station data possibly suggests that these are small-scale features which are not resolved by the coarser resolution data of OLR. It is worth pointing out that the NCMRWF T80 model has a resolution of about 130 km in the Indian region, much finer than the OLR data set (which have a typical resolution of ~ 250 km) but much coarser than the station network observations (which are irregularly spaced). We also would like to point out that very few climate studies are presently done at such a high resolution, e.g. a majority of the simulation for AMIP was done at T42 with only NCEP conducting a rerun at T63.

4. The mean simulated pattern

In this section we compare the simulated patterns of precipitation during the summer monsoon period (averaged for the four summers of simulation) with climatology. The

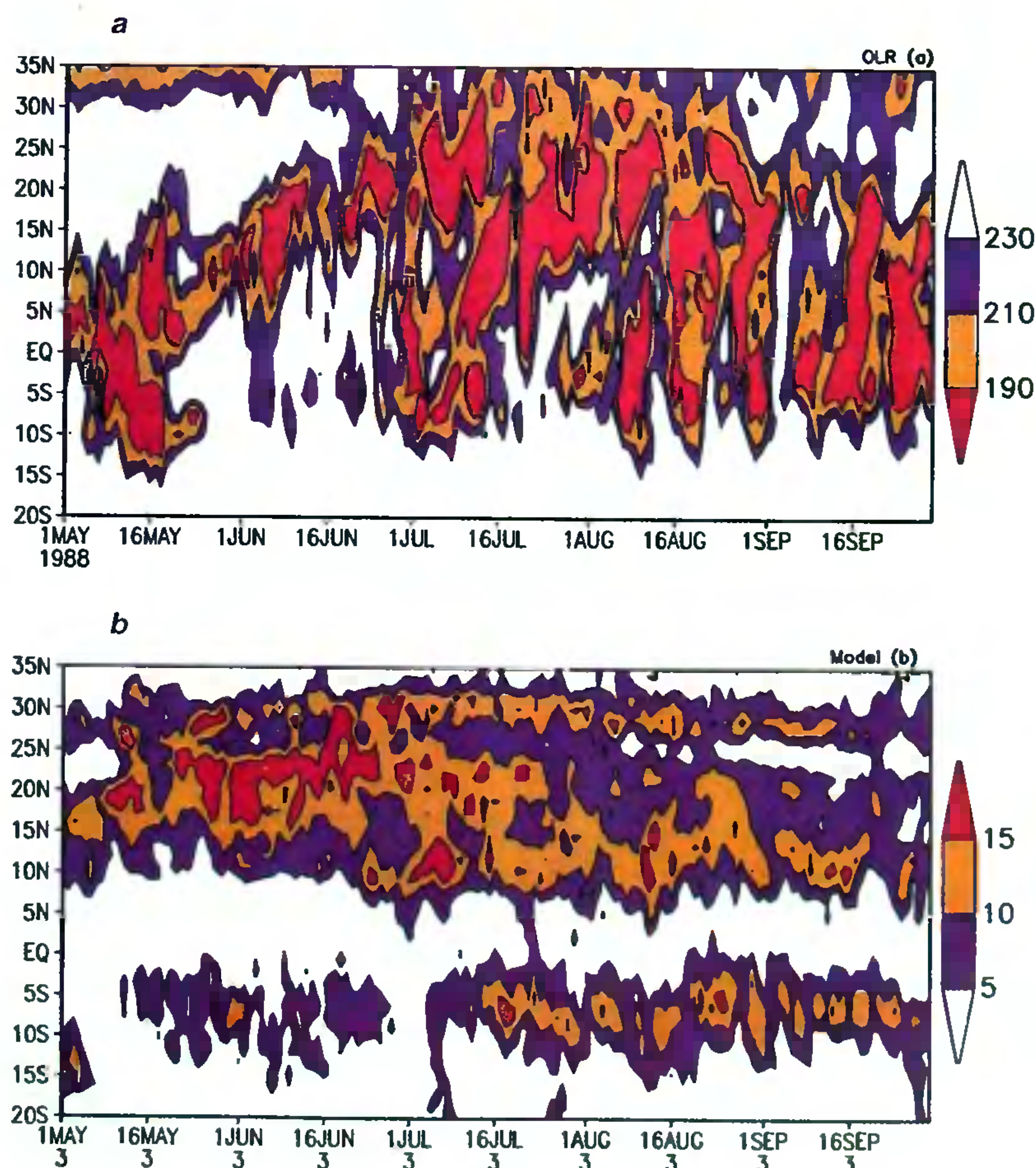


Figure 2. Intraseasonal variation in (a) observations of Outgoing Longwave Radiation (W/m^2), and (b) model simulated rainfall (mm/day).

simulation for each of the individual summers was qualitatively similar to the four-year mean.

Here we compare the summer (June–September) mean rainfall with observed climatology. Table 1 shows the comparative figures for the Indian region. This region has been subdivided into three regions: The monsoon zone (15°N–25°N), the peninsular region (7.5°N–15°N) and the Indian ocean region (5°S–5°N). We note that for the summer season, the rainfall in the monsoon zone is considerably over-estimated. The rainfall over peninsular region is much higher at 11.1 mm/day in comparison to the observed value of 5.4 mm/day. Thus while the observations show a ratio of 0.8 (peninsular/monsoon zone), the model simulates a ratio of 1.1. Interestingly, the equatorial oceanic rainfall is underestimated at 4.4 mm/day while the observations show this to be 6.1 mm/day. However it is necessary to point out that oceanic rainfall has been mainly obtained from MSU estimates which are known to overestimate the precipitation.

Comparing the simulated mean monsoon pattern (Figure 1 c) with those from the observation station (Figure 1 a) and the NCEP merged data set (Figure 1 b), we note the following:

1. A region of high rainfall is simulated in the monsoon trough region which coincides well with the region of high rainfall in IMD station data.
2. A region of intense convection over the equatorial Indian Ocean is simulated. This agrees well with the equatorial rainbelt in the NCEP merged data (Figure 1 b).
3. The rainfall on the western coast of India with peaks around Kerala coast and Konkan corresponding to regions of high orography. This feature corresponds qualitatively well with the station observations.
4. The rainfall over the Himalayan region is underestimated. The model's simulation underestimates the Himalayan rainfall though some orography-related precipitation can be seen at 90°E and 30°N.
5. The rainfall over Rajasthan is overestimated.

It can be said that the model captures qualitatively at least to a first order, most of the features of the mean summer monsoon rainfall.

Table 1. Observed and simulated June–September mean rainfall (mm/day) over the monsoon zone

Model	Mean	P (P/MZ)	IO(IO/MZ)
Obs (NCEP)	6.9	5.4 (0.8)	6.1 (0.9)
NCMRWF	10.4	11.1 (1.1)	4.4 (0.4)

(MZ, 15°–25°N); P = mean rainfall over the peninsula (7.5°–15°N); IO = mean rainfall over the Indian Ocean (5°S–5°N). Values in parenthesis indicate the ratio of rainfall in that region to the mean rainfall in the monsoon zone.

5. Intraseasonal variations

Rain does not fall continuously during the entire monsoon season but in spells of copious precipitation (called the active periods) punctuated with periods when the rainfall is low (called the break periods). We term this variation within a season as the intraseasonal variation. The inter-annual variability of the Indian Summer Monsoon is closely linked to the intraseasonal variation²⁴, i.e. the variation of monsoon from year to year is related to the length of active and break spells within a season; during weak monsoon years breaks are likely to be more frequent or of longer durations, while during strong monsoon years the active spells might be of longer durations. Hence if a model is to be used for seasonal scale forecasts, then its ability to simulate intraseasonal variations needs to be studied. We compare the model's third summer season with observed data. The intraseasonal structure of the model's simulation is largely similar for other years also. We compare this with the summer of 1988, a year in which the monsoon was very active and vigorous. However, similar patterns are seen to occur in almost all the years. The robustness of the intraseasonal structure of the Indian Summer Monsoon rainfall is a unique feature seldom observed in other monsoonal regions²⁴.

We use OLR as a proxy for rainfall; in the tropics, regions of low OLR ($< 240 \text{ W/m}^2$) correspond to regions of intense cloudiness and high rainfall (Figure 2 a). The salient features of intraseasonal variation²³ are:

- (i) Northward movement of the rainbelt from the equatorial Indian Ocean onto the continent during the onset phase of the monsoon, during late May or early June.
- (ii) Persistence of the rainbelt in the northern monsoon zone ($\sim 25^\circ\text{N}$).
- (iii) Disappearance of the rainbelt from the northern monsoon zone for periods ranging from a few days to a few weeks.
- (iv) Revival of the rainbelt from (iii) either by genesis within the monsoon zone or by northward propagation of the oceanic rainbelt. The mean period between northward propagations is about forty days.

The intraseasonal variation within a single seasonal is an envelope of a repetition of events (ii)–(iv).

We find that the model realistically simulates (Figure 2 b, averaged over the Indian longitudes, 75°–95°E) two regions of high rainfall, viz. the equatorial Indian Ocean and the northern monsoon zone. However, the observed poleward propagations from the equatorial region to the monsoon zone are absent. We do find some poleward movement within the monsoon zone itself. The model also simulates an early onset during May (compared to late May/early June in observations). The simulated withdrawal is also early occurring in late August/early September, while the observed withdrawal occurs much later during late September/early October.

6. Conclusion

In this paper we have described the development of a climate version of the NCMRWF model suitable for message passing architecture and a discussion of the mean simulated pattern of the Indian Summer Monsoon (based on a four-year run of the model). Initial results with the model are very encouraging. It appears to capture the monsoonal rainbelt, the rainbelt over the warm equatorial Indian Ocean and some of the orographic rainfall realistically. The area averaged rainfall is overestimated at 10.5 mm/day (compared to the observed value of 7.5 mm/day for the July–August season). However, the simulation of intraseasonal variation by the model is not very satisfactory. The model simulates early onset and withdrawal. Poleward propagations from the oceanic region onto the monsoon zone are also absent.

An indepth systematic analysis needs to be conducted to study the robustness of the simulated monsoon to various aspects of modelling. Perhaps it is necessary to conduct a systematic study to understand the model's sensitivity to the important sub-models such as cumulus convection and boundary layer parameterization to understand the cause of these shortcomings.

1. WCRP, Report of Workshop, Boulder, Colorado, USA, 21–24 October 1991, WCRP-68, WMO/TD-No. 470, Geneva, Switzerland, 1992.
2. WCRP, WCRP-80, WMO/TD-No. 546, Geneva, Switzerland, 1993.
3. Murakami, T., Godbole, R. V. and Kelkar, R. R., Proc. Conf. Summer Monsoon of Southeast Asia, Navy Weather Research Facility, Norfolk, Virginia, 1970, pp. 39–51.
4. Washington, W. M., in *Monsoon Dynamics* (eds Lighthill, J. and Pearce, R. P.), Oxford University Press, 1981, pp. 111–130.
5. Gates, W. L., *Bull. Am. Meteorol. Soc.*, 1992, **73**, 1962–1970.
6. Gadgil, S. and Sajani, S., *Clim. Dyn.*, 1998, **14**, 659–689.

7. Nanjundiah, R. S., NAL Report PD-AE-8813, 1988.
8. Sinha, U. N., Sarasamma, V. R., Rajalakshmy, S., Subramanian, K. R., Bharadwaj, P. V. R., Chandrasekhar, C. S., Venkatesh, T. N., Sundar, R., Basu, B. K., Gadgil, S. and Raju, A., *Curr. Sci.*, 1994, **67**, 178–184.
9. Kalnay, E., Sellers, J. G., Campana, K. A., Basu, B. K., Schwarzkopf, M. D., Long, P. E., Caplan, P. M. and Alpert, J. C., Documentation of Research Version of the NMC Medium Range Forecasting Model, 1988.
10. Grell, G. A., *Mon. Weather Rev.*, 1993, **121**, 764–787.
11. Kuo, H. L., *J. Atmos. Sci.*, 1965, **39**, 1290–1306.
12. Nanjundiah, R. S. and Sinha, U. N., *Curr. Sci.*, 1999, **76**, 1114–1116.
13. Shea, D. J., Trenberth, K. E. and Reynolds, R. W., *J. Clim.*, 1992, **5**, 987–1001.
14. Molod, A., Helfand, H. M. and Takacs, L. L., *J. Climatol.*, 1996, **9**, 764–785.
15. Nanjundiah, R. S., *Clim. Dyn.*, 2000 (in press).
16. Meehl, G. A. and Arblaster, J. M., *J. Climatol.*, 1998, **11**, 1356–1385.
17. Sinha, U. N. and Nanjundiah, R. S., in 6th Symposium on use of MPPs for Weather Forecasting, ECMWF, Reading, UK, 1996.
18. Venkatesh, T. N., Sivaramkrishnan, R. and Sarasamma, V. R., *Curr. Sci.*, 1998, **75**, 709–712.
19. Jones, P. W., Kerr, C. L. and Hemler, R. S., *Parallel Comput.*, 1995, **21**, 1677–1694.
20. Nanjundiah, R. S., Raju, A., and Sinha, U. N., CAOS, IISc Report 96AS3, 1996.
21. Schemm, J., Schubert, S., Terry, J. and Bloom, S., NASA Tech. Mem. 104571, Goddards Space Flight Center, Greenbelt, MD 20771, October 1992.
22. Rao, Y. P., Southwest Monsoon India Meteorological Department. Meteorological Monograph Synoptic Meteorology, No. 1/1976, Delhi, 1976, p. 367.
23. Sikka, D. R. and Gadgil, S., *Mon. Weather Rev.*, 1980, **108**, 1840–1850.
24. Gadgil, S. and Asha, G., *J. Meteorol. Soc. Jpn.*, 1992, **70**, 517–527.

ACKNOWLEDGEMENTS. I thank the Chairman SERC, IISc for computational support provided on the SP2 parallel computer, and NCMRWF for providing the model. The parallel implementation of the NCMRWF model was partially supported by IBM under the joint study agreement #5385 and the same is gratefully acknowledged.