

22. Signoret, P. A., Bernaux, P. C. and Poinso, B., Soybean diseases in France in 1974. *Plant Dis. Rep.*, 1975, **59**, 616–617.
23. Zad, J., Mycoflora of soybean seeds. *Iran. J. Plant Pathol.*, 1979, **15**, 32–33.
24. Rossi, R. L., First report of *Phakopsora pachyrhizi*, the causal organism of soybean rust in province of Misiones, Argentina. *Plant Dis.*, 2003, **87**, 102.
25. Kurlle, J. E., Gould, S. L., Lewandowski, S. M., Li, S. and Yang, X. B., First report of sudden death syndrome (*Fusarium solani* f. sp. *glycine*) of soybean in Minnesota. *Plant Dis.*, 2003, **87**, 449.
26. Sikora, E. J. and Murphy, J. F., First report of bean pod mottle virus in soybean in Alabama. *Plant Dis.*, 2005, **89**, 108.
27. Phibbs, A., Barta, A. and Domier, L. L., First report of soybean dwarf virus on soybean in Wisconsin. *Plant Dis.*, 2004, **88**, 1285.
28. Mengistu, A., Kurtzweil, N. C. and Grau, C. R., First report of frogeye leaf spot (*Cercospora sojina*) in Wisconsin. *Plant Dis.*, 2002, **86**, 1272.
29. Gupta, V. P. and Kaur, A., *Phakopsora pachyrhizi* – soybean rust pathogen new to Rajasthan. *J. Mycol. Plant Pathol.*, 2004, **34**, 151.
30. Verma, K. P., Thakur, M. P., Agarwal, K. C. and Khare, N., Occurrence of soybean rust: some studies in Chhattisgarh state. *J. Mycol. Plant Pathol.*, 2004, **34**, 24–27.

ACKNOWLEDGEMENTS. We thank all the Directors of NBPGR, New Delhi during 1976–2005 for encouragement and facilities. Thanks are also due to Drs R. C. Agrawal and Hanuman Lal Raiger, NBPGR for statistical analysis.

Received 25 July 2005; revised accepted 20 March 2006

Forecasting of seasonal monsoon rainfall at subdivision level

R. N. Iyengar^{1*} and S. T. G. Raghu Kanth²

¹Department of Civil Engineering, Indian Institute of Science, Bangalore 560 012, India

²Department of Civil Engineering, Indian Institute of Technology, Guwahati 781 039, India

It is shown that time series data of monsoon seasonal rainfall at subdivision level is decomposable into six uncorrelated components. These narrowband processes called intrinsic mode functions, in decreasing order of importance, reflect the influence of ENSO, sunspot activity and tidal cycle on inter annual rainfall variability. The decomposition helps in proposing a statistical method to forecast monsoon rainfall in the three subdivisions of Karnataka.

Keywords: Forecasting, intrinsic mode functions, rainfall, subdivisions.

INDIAN rainfall data are available at two spatial scales in the archives of IITM, Pune (www.tropmet.res.in). The country

is considered to be consisting of 33 subdivisions for reporting the monthly rainfall data for the period 1871–2004. This is the smaller spatial scale database. Another database, consisting of the monthly rainfall at the larger spatial scale of eight regions including one at all-India level, is also made available. Recently, we have proposed¹ an approach to analyse and forecast monsoon rainfall data of the regional and all-India time series. The method is based on decomposing the data into empirical modes called Intrinsic Mode Functions (IMFs). The regional-level data have a coefficient of variation defined as the ratio of standard deviation to climatic average (σ/m), ranging from 10 to 18%. The efficiency of the IMF model in one-step-ahead forecasting is about 80%. Thus, the model is able to capture the most important inter-annual variability signatures on the larger spatial scale. A known property of rainfall data is that on larger scales, the variability tends to decrease due to smoothening or averaging effects. Thus, subdivision-level data will show higher variability in comparison with regional data. Since IMF model decomposes the time series into basic uncorrelated empirical modes, one would expect the approach to be qualitatively valid at any scale. However, the forecasting skill will depend on how best the temporal patterns of the signatures are translated into the decomposed modes. The present investigation is aimed at studying the basic modes present in subdivision-level rainfall data, with a view towards forecasting the amount of rainfall.

Rainfall data are available for 33 subdivisions (SD), which make up the geographical extent of the country. The data have been extensively studied for understanding spatial and inter-annual variability (IAV) of the monsoon^{2–5}. Also, the data have been used to study the relationship between rainfall and other atmospheric processes such as quasi-biennial oscillation (QBO)⁶, Southern oscillation index (SOI)⁷ and sunspot index⁸. In spite of the existence of long-term variability or memory signatures, quantitative forecasting of rainfall has remained a daunting task⁹. Here, seasonal (June–September) data of three subdivisions are selected for further study. These are SDs 31, 32 and 33 covering the State of Karnataka (Figure 1). This selection is based on previous studies¹⁰ on the variability structure of station-level rainfall in Karnataka. It was found that broadly the State comprises of three homogeneous regions nearly overlapping with the three IMD subdivisions. It was also found that coastal Karnataka (SD-31) has a significant transition probability structure from June to July. There is a tendency for below-normal June rainfall to be compensated by above-normal rainfall in July. Parthasarathy and Pant⁷ have shown that rainfall in the above subdivisions is well correlated with QBO and that the data are significantly correlated at 14-year lag. The basic statistics of the data considered here is given in Table 1.

Monsoon rainfall evolves in a random fashion around a few central periods. This can be seen by spectral analysis, wavelet or principal component analysis. Recently, it has

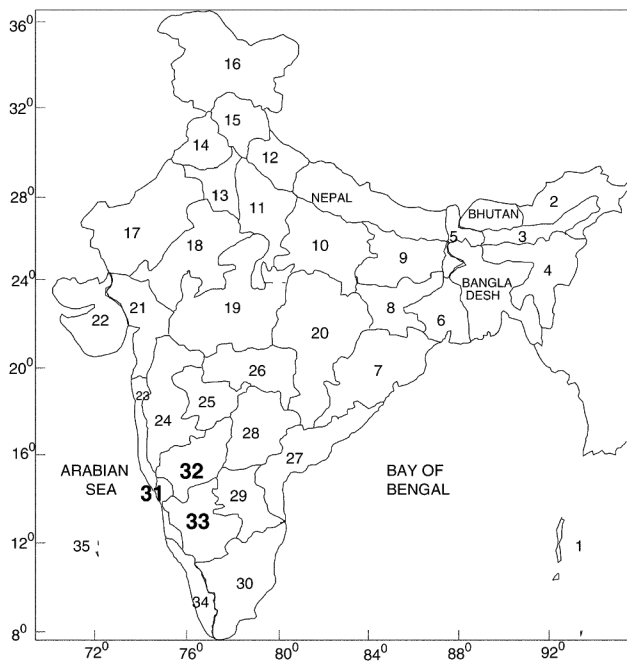
*For correspondence. (e-mail: rni@civil.iisc.ernet.in)

Table 1. Subdivision (IITM data 1871–1990)

| Sub-division | Name | Mean m_R Cm. | SD σ_R Cm. | COV (σ_R/m_R) % | Skewness | Kurtosis |
|--------------|-------|-------------------|----------------------|-----------------------------|----------|----------|
| 31 | COKNT | 285.22 | 50.94 | 15.42 | 0.624 | 4.310 |
| 32 | NIKNT | 60.09 | 11.96 | 19.90 | 0.185 | 2.847 |
| 33 | SIKNT | 50.33 | 10.19 | 20.25 | 0.334 | 2.752 |

Table 2. Central period of the IMFs in years and % variance contributed to IAV

| Region | IMF ₁ | | IMF ₂ | | IMF ₃ | | IMF ₄ | | IMF ₅ | |
|--------|------------------|------|------------------|------|------------------|------|------------------|------|------------------|------|
| | <i>T</i> | IAV% | <i>T</i> | IAV% | <i>T</i> | IAV% | <i>T</i> | IAV% | <i>T</i> | IAV% |
| COKNT | 2.55 | 60.7 | 5.71 | 31.7 | 12 | 10.5 | 18 | 4.3 | 40 | 1.0 |
| NIKNT | 3 | 62.4 | 5.71 | 17.6 | 12 | 11.9 | 20 | 6.6 | 60 | 4.1 |
| SIKNT | 2.86 | 69.7 | 5.45 | 20.5 | 10.9 | 4.5 | 24 | 1.6 | 60 | 3.1 |

**Figure 1.** Meteorological subdivisions of India and Karnataka.

been shown that the empirical mode decomposition proposed by Huang *et al.*¹¹ has distinct advantages over other methods in identifying the dominant periods and their amplitudes. This method decomposes the data series into finite number of empirical modes called IMFs. These are uncorrelated with each other at zero lag, but correlated with the original data in a decreasing order of importance. IMF is a data-derived function such that in its interval of definition, the number of zeros and extrema is equal or differs as at most by one. Each IMF is a narrow band process with an identifiable central period. In Figure 2 a–c, the present data and their IMFs are shown. The sum of the IMFs will be equal to the original data at every time instant, that is $R(t) = \sum \text{IMF}_i(t)$. In each of the figures, data variance

and variance of each mode are also presented. It is observed that IMFs can be organized with decreasing level of importance. The sum of variances of IMFs should be ideally equal to the data variance. However, it is observed that the last two IMFs get correlated due to round-off errors and hence the precise decomposition of variance may not always be achieved.

In Table 2, the central period of the IMFs and the percentage of variability explained by each IMF are given. The last IMF in all the three cases is the climatic normal, which can be taken to be deterministic. Table 3 shows the correlation matrix of the IMFs. IMF₁ is the predominant mode, with an average period of about 2.7 years, contributing to more than 60% of IAV. It is also the mode maximally correlated with the basic data. Like with the regional-scale data studied previously¹, here also IMF₁ and IMF₂ are connected with QBO and ENSO, which show quasi-periodic behaviour with a central period of 2–5 years. IMF₃ can be associated with the 11-year sun-spot cycle. IMF₄ most probably reflects tidal forcing linked to the Metonic cycle of 18–19 years. It is important to verify whether the extracted IMFs are spurious signatures of an originally uncorrelated random noise data sample. In Figure 3, the white noise test developed by Wu and Huang¹² is applied on the IMFs of the three subdivisions. For a strict white noise, the variance of the IMFs and their respective central periods varies linearly on a double log plot. Thus for the data to be accepted as pure noise, all the variance values have to lie within the 99% confidence band of acceptance. It is observed that the null hypothesis that original data are white noise with no patterns gets rejected. This indicates the possibility of statistical forecasting of the data series through modelling and forecasting of the IMFs.

A model is a mathematical equation or an algorithm that can closely replicate the data of a particular length with minimum error. Since such a model is not unique, the efficiency of any particular model can only be verified by comparing it with other claimants in a specified period of

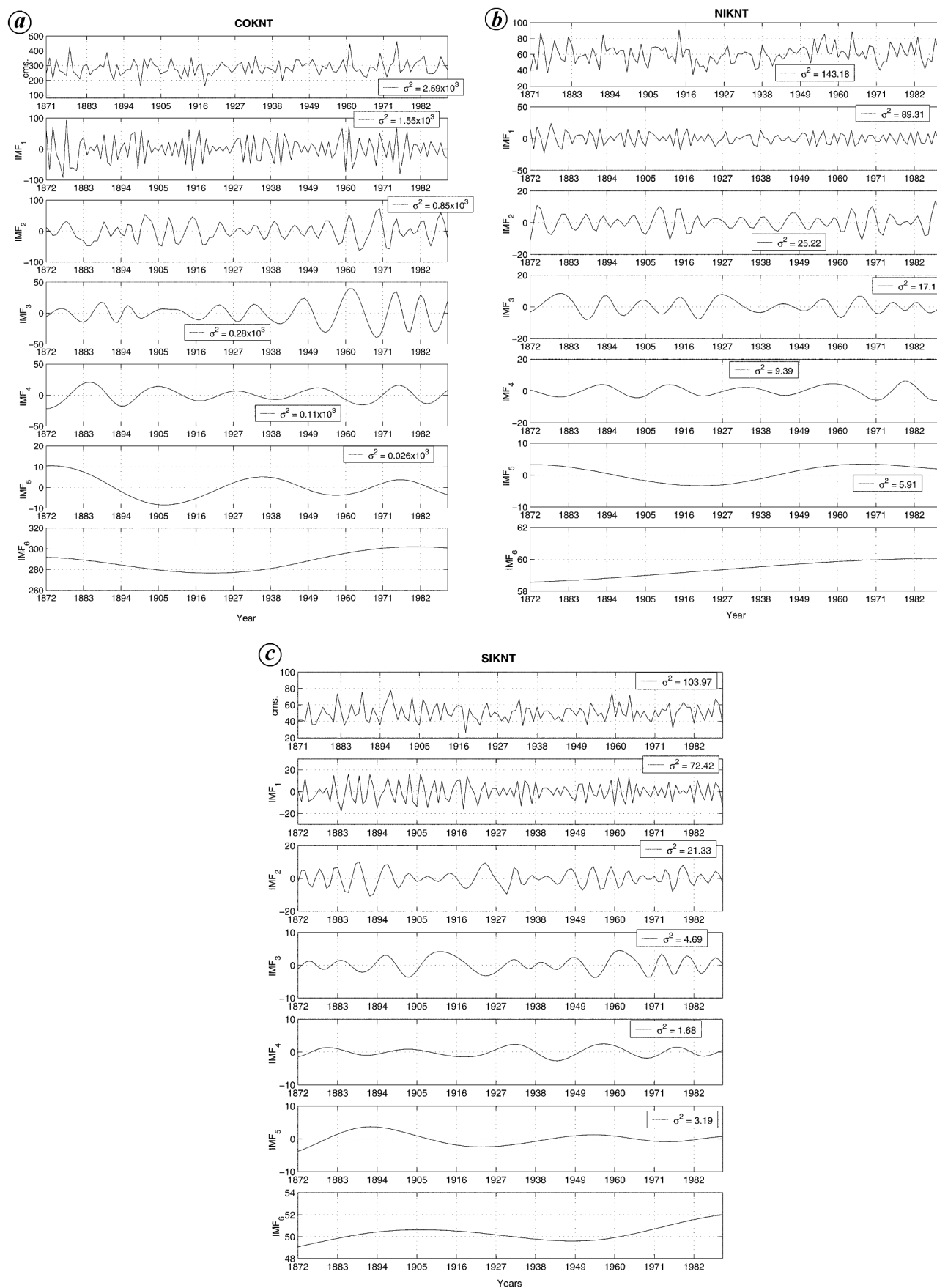
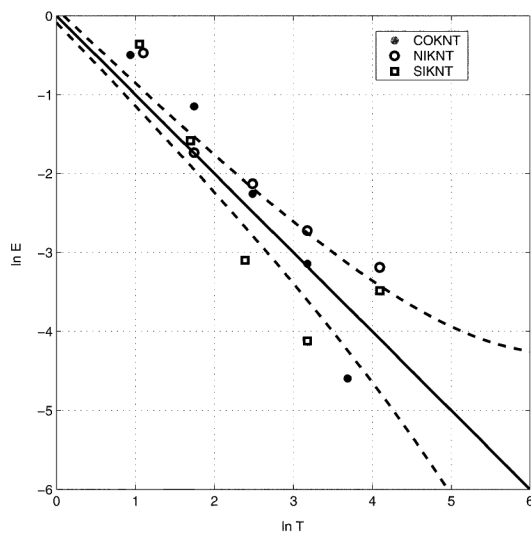


Figure 2. IMFs of COKNT rainfall (a), NIKNT rainfall (b) and SIKNT rainfall (c).

Table 3. Correlation matrix

| | Data | IMF ₁ | IMF ₂ | IMF ₃ | IMF ₄ | IMF ₅ |
|------------------|--------|------------------|------------------|------------------|------------------|------------------|
| COKNT | | | | | | |
| Data | 1.0000 | 0.7452 | 0.5126 | 0.2340 | 0.1199 | 0.0279 |
| IMF ₁ | | 1.0000 | -0.0075 | -0.0386 | -0.0105 | -0.0758 |
| IMF ₂ | | | 1.0000 | -0.0880 | -0.0958 | -0.0356 |
| IMF ₃ | | | | 1.0000 | 0.0025 | -0.0517 |
| IMF ₄ | | | | | 1.0000 | -0.0929 |
| IMF ₅ | | | | | | 1.0000 |
| NIKNT | | | | | | |
| Data | 1.0000 | 0.7615 | 0.4192 | 0.2370 | 0.2417 | 0.1838 |
| IMF ₁ | | 1.0000 | -0.0020 | -0.0734 | -0.0016 | 0.0033 |
| IMF ₂ | | | 1.0000 | -0.0029 | -0.0487 | -0.0186 |
| IMF ₃ | | | | 1.0000 | -0.0342 | 0.0048 |
| IMF ₄ | | | | | 1.0000 | -0.0378 |
| IMF ₅ | | | | | | 1.0000 |
| SIKNT | | | | | | |
| Data | 1.0000 | 0.8260 | 0.4606 | 0.2406 | 0.1123 | 0.1677 |
| IMF ₁ | | 1.0000 | -0.0011 | 0.0055 | -0.0243 | -0.0340 |
| IMF ₂ | | | 1.0000 | 0.0369 | -0.0141 | -0.0041 |
| IMF ₃ | | | | 1.0000 | -0.0020 | 0.0158 |
| IMF ₄ | | | | | 1.0000 | 0.0931 |
| IMF ₅ | | | | | | 1.0000 |

**Figure 3.** White noise test for rainfall of three subdivisions. *IMF, — Expected for white noise. ---- 99% confidence bands.

time. It is useful to benchmark a model with respect to the climatic variation explained by it in the modelling period. The previous model of the authors¹ for regional-scale rainfall had two parts denoted as nonlinear and linear. The former, which was IMF₁, was shown to be amenable for modelling through artificial neural network (ANN) techniques. The remainder was a stationary random process modelled with a simple linear representation regressed on the antecedent five years data. However, in the present case ($R_i - \text{IMF}_{1i}$) at subdivisional level is found to be non-stationary, as verified through the standard run test. Fur-

ther, it is seen that IMF₂ cannot be treated as a Gaussian process, unlike in the case of regional data. On smaller spatial scales, not only does the variability increase, but also the non-Gaussian and hence the nonlinear character of rainfall time series gets accentuated. This necessitates modelling the first two IMFs individually through ANN approaches. The remaining part $y_i = (R_i - \text{IMF}_{1i} - \text{IMF}_{2i})$ is nearly Gaussian and can be modelled as a linear process separately. It can be observed that as one goes to higher empirical modes, computing IMFs at the end-points becomes difficult. Since the data are available for $(i = 1, 2, 3, \dots, n)$, IMFs can be found only for $(i = 2, \dots, n - 1)$. This makes the estimation of linear and nonlinear parts of the data for the last point difficult. This is precisely what would be required in a forecasting exercise that makes use of the above type of decomposition. This difficulty is overcome here by making y_j depend on the known current rainfall R_j , the previous year value R_{j-1} and three past values of y_i . This model has been arrived at based on several trials. The coefficients in the equation

$$y_{n+1} = C_1 R_n + C_2 R_{n-1} + C_3 y_{n-2} + C_4 y_{n-3} + C_5 y_{n-4} + C_6 + \epsilon, \quad (1)$$

are found by minimizing the mean square error between the model and the data in the modelling period (1871–1990). These are shown in Table 4. The modelling of IMF_{1i} and IMF_{2i} is carried out using ANN techniques. The architecture of the neural network model is shown in Figure 4. This consists of one hidden layer with five nodes, dependent on five past values. The model needs 36 parameters, which can be found with the help of MATLAB software using

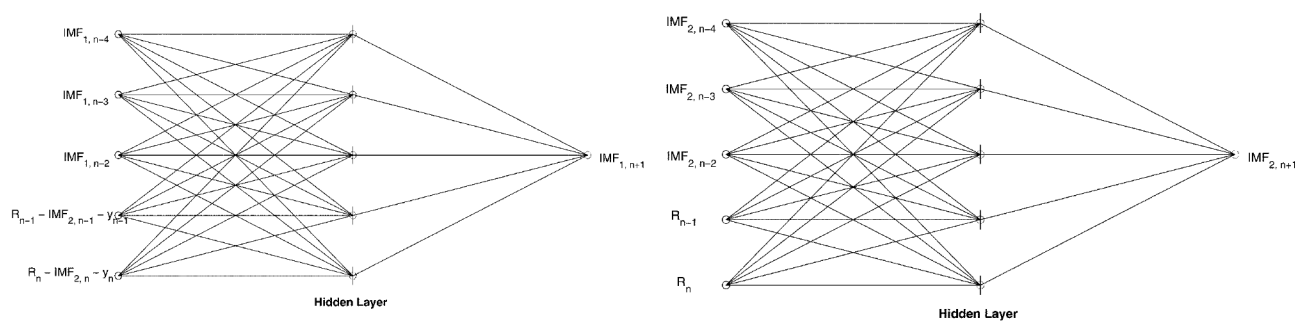


Figure 4. ANN architecture for modelling IMF₁ and IMF₂.

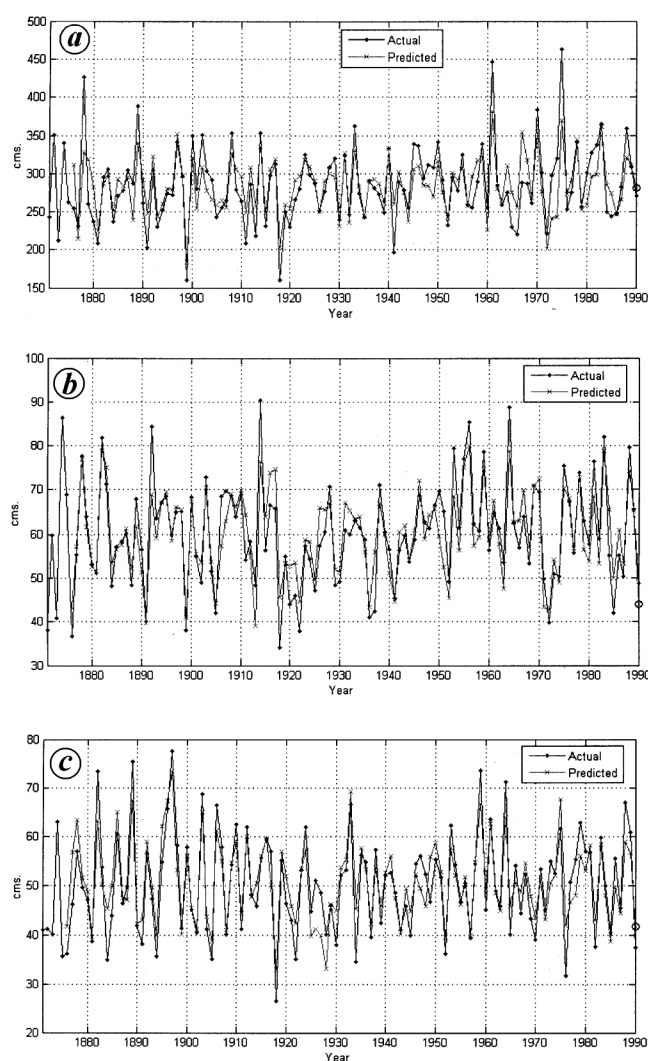


Figure 5. Comparison between actual data and hindcasts. *a*, COKNT; *b*, NIKNT and *c*, SIKNT.

the backpropagation algorithm. At any stage, the model for R_i is given by $(IMF_{1i} + IMF_{2i} + y_i)$. The efficiency of this model has been first verified by hindcasting the data from the model algorithm. In Table 5, the goodness-of-fit

is demonstrated by presenting three statistics. These are (i) standard deviation (σ_m) of the error ϵ in the model fit, (ii) correlation between the data and the model hindcast, and (iii) the performance parameter defined as $PP_m = 1 - \sigma_m^2 / \sigma_d^2$, where σ_m^2 is the mean square error and σ_d^2 is the actual data variance. In a perfect model, $\sigma_m(\epsilon)$ will be zero and both CC_m and PP_m will tend towards unity. It is observed that the present model consisting of eq. (1) and IMF₁ and IMF₂ represented through the ANN architecture of Figure 4, explains 65–70% of the variance of rainfall over the subdivisions considered here. In Figure 5, the actual data and analytical hindcasts obtained from the present model are shown for a visual comparison of the goodness-of-fit.

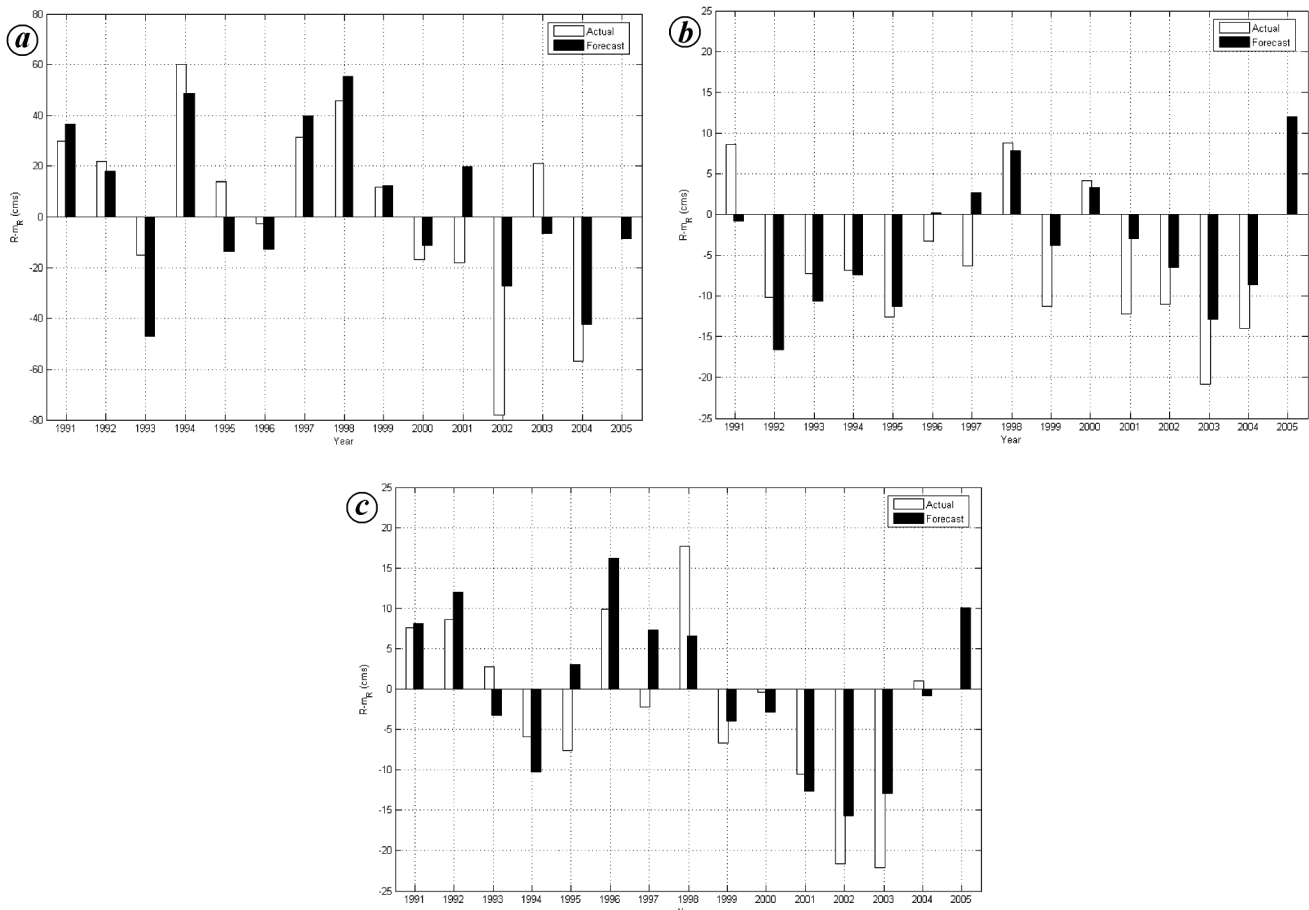
Estimation of rainfall for year $(j+1)$ based on knowledge of data of the current year j and past values of years $(j-1, j-2, \dots, 3, 2, 1)$ is defined as a forecast. If the process were to be stationary, constants found previously could be used in eq. (1) for finding y_{j+1} and further IMF_{1,j+1} and IMF_{2,j+1}. In the present case, since statistical tests indicate the data to be non-stationary, all past data are used in every year j to find R_{j+1} . The modelling results are used in a qualitative sense retaining the same ANN architecture and eq. (1) for updating the model constants at each step. In Figure 6 *a–c* year-by-year forecasts along with observed values are shown for the three subdivisions of Karnataka. Forecasts are desired as unique numbers or as point estimates. However, the procedure leading to the forecast is statistical and hence the prediction has to be interpreted as a random variable with a definite probability distribution. The skill of such a forecast has to be evaluated, necessarily again in a statistical sense by carrying out the exercise on an independent sample. Here this has been carried out for the period 1991–2004, as shown in Figure 6 *a–c*. In Table 5, the forecast skill measures, namely $\sigma_f(\epsilon)$, CC_f and PP_f are presented. With a length of 14 years, correlation between observed and forecast values can be taken to be significant if it is higher than about 0.55. It is seen that the forecast skill of the present methodology is much above the threshold of significance. The performance parameters are quantitatively less than the ones during the modelling period. This is clearly attribut-

Table 4. Coefficients of eq. (1)

| Region | C_1 | C_2 | C_3 | C_4 | C_5 | C_6 | $\sigma_y(\epsilon)$ | CC |
|--------|--------|---------|--------|---------|--------|----------|----------------------|--------|
| COKNT | 0.0793 | 0.0246 | 1.8409 | -2.8216 | 1.3164 | 160.1436 | 16.7992 | 0.6380 |
| NIKNT | 0.0304 | -0.0053 | 3.6532 | -5.2693 | 2.2582 | 20.0185 | 3.4299 | 0.8128 |
| SIKNT | 0.0023 | -0.0119 | 3.6290 | -5.1141 | 2.1857 | 15.7583 | 1.8636 | 0.8004 |

Table 5. Performance index

| Region | Modelling period (1872–1990) | | | Forecasting period (1991–2004) | | |
|--------|------------------------------|--------|--------|--------------------------------|--------|--------|
| | $\sigma_m(\epsilon)$ | CC_m | PP_m | $\sigma_f(\epsilon)$ | CC_f | PP_f |
| COKNT | 30.13 | 0.81 | 0.65 | 24.18 | 0.77 | 0.53 |
| SIKNT | 5.21 | 0.85 | 0.72 | 6.65 | 0.83 | 0.68 |
| NIKNT | 6.31 | 0.86 | 0.71 | 5.95 | 0.69 | 0.53 |

**Figure 6.** Independent test forecasting. *a*, COKNT; *b*, NIKNT and *c*, SIKNT.

able to the short length of 14 years of forecasting regime. As the forecast exercise length increases, the performance parameter PP_f would approach PP_m .

A novel statistical approach for forecasting monsoon rainfall at subdivision level has been proposed here. This is an extension of our previous work¹ on forecasting all-India and regional rainfall using empirical mode decom-

position. It is recognized that seasonal monsoon rainfall on a given space regime exhibits specific patterns on a few preferred timescales. The subdivision data studied here show higher coefficients of variation than the larger-scale regional data series. Like the regional data, the present data also get decomposed into six uncorrelated IMFs. These can again be interpreted in decreasing order of im-

portance, as being related to ENSO, sunspot and tidal phenomena. The modelling and forecasting skill of the proposed method has been demonstrated to be statistically significant. Apart from the statistics reported in Table 5, results of Figure 6 *a–c* are interesting. It is observed that the nature of departure from long-term average (normal) rainfall has been foreshadowed correctly in eleven out of fourteen years. Even in years where the forecast appears to be poor, the value is within a known error band. Persistence of drought-like conditions in SIKNT and NIKNT during 1999–2003 has also been captured by the present model in a forecast mode. In comparison with regional-scale rainfall, the present data show lower levels of modelling and forecasting efficiency measured in terms of PP_m and PP_f . This is attributable to the higher coefficient of variation and lack of stationarity property with the present data series. The efficiency of the present method for modelling other subdivisions which have still higher levels of (σ/m) value is yet to be investigated.

1. Iyengar, R. N. and Raghu Kanth, S. T. G., Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall. *Meteorol. Atmos. Phys.*, 2005, **90**, 17–36 (on-line 2004).
2. Iyengar, R. N. and Basak, P., Regionalization of Indian monsoon rainfall and long-term variability signals. *Int. J. Climatol.*, 1994, **14**, 1095–1114.
3. Mooley, D. A. and Parthasarathy, B., Fluctuations in all-India summer monsoon rainfall during 1871–1978. *Climatic Change*, 1984, **6**, 287–301.
4. Gregory, S., Macro-regional definition and characteristics of Indian summer monsoon rainfall, 1871–1985. *Int. J. Climatol.*, 1989, **9**, 465–483.
5. Rupa Kumar, K. *et al.*, Spatial and subseasonal patterns of the long-term trends of Indian summer monsoon rainfall. *Int. J. Climatol.*, 1992, **12**, 257–268.
6. Raja Rao, K. S. and Lakhole, N. J., Quasi-biennial oscillation and summer southwest monsoon. *Indian J. Meteorol. Hydrol. Geophys.*, 1978, **29**, 403–411.
7. Parthasarathy, B. and Pant, G. B., The spatial and temporal relationships between the Indian summer monsoon rainfall and the Southern Oscillation. *Tellus*, 1984, **A36**, 269–277.
8. Bhalme, H. N. and Jadhav, S. K., The double (Hale) sunspot cycle and floods and droughts in India. *Weather*, 1984, **39**, 112.
9. Gadgil, S. *et al.*, Monsoon prediction – why yet another failure? *Curr. Sci.*, 2005, **88**, 1389–1400.
10. Iyengar, R. N., Variability of rainfall through principal component analysis. *J. Earth Planet. Sci.*, 1991, **100**, 105–126.
11. Huang N. E. *et al.*, The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc. London, Ser. A*, 1998, **454**, 903–995.
12. Wu, Z. and Huang, N. E., A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. R. Soc. London, Ser. A*, 2003, **460**, 1597–1611.

Received 3 November 2005; revised accepted 19 April 2006

Quick look isoseismal map of 8 October 2005 Kashmir earthquake

A. K. Mahajan*, Naresh Kumar and B. R. Arora

Wadia Institute of Himalayan Geology, 33 GMS Road, Dehradun 248 001, India

The isoseismal map for the devastating M 7.6 Kashmir earthquake of 8 October 2005 is constructed based on the immediate damage scenario provided by the teachers trained under the Himalayan School Earthquake Laboratory Programme as well as that reported in electronic and print media. The nature of the damage pattern imprinted on different vulnerable classes of buildings at some 80 sites enabled to map out intensity distribution in earthquake-affected region to a value above IV on the European Macroseismic Scale (EMS-98). The isoseismal map provides a fair picture of the distribution of ground-shaking effects to distant places. This would serve as a useful guide in future earthquake hazard assessment in the region. The Kashmir valley was widely affected and the meizoseismal zone encompassing the township of Balakot and Muzaffabad experienced a maximum intensity of XI on the EMS-98 scale. The use of this maximum intensity and the dimension of the area covered by isoseismal VI in the well-established intensity–focal depth and intensity–moment relations respectively, allowed for estimating the focal depth and the magnitude. Since in the present approach, the map is prepared based on the damage scenario immediately after the main shock, it will be free from biases due the subsequent damages caused by aftershocks that advertently tend to contaminate the maps prepared by conventional field surveys.

Keywords: Focal depth, Kashmir earthquake, intensity distribution, isoseismal map, magnitude.

THE M_w 7.6 worst ever earthquake shook the Kashmir valley on 8 October 2005 at 03:52 UT (09:22 IST). The shallow focus earthquake (depth 10 km) with its epicentre (34.432°N, 73.537°E, USGS), ~124 km to the west of Srinagar, caused widespread destruction and casualties (>50,000) in the region. In the west, the event was widely felt in Pakistan and Afghanistan and in the east shaking of the earth was felt as far as Himachal, Punjab, Haryana, Uttaranchal, Delhi, Rajasthan, Gujarat and western Uttar Pradesh. Earlier also this region experienced a number of moderate and major earthquakes. Among those the most recent ones are the Northwest Kashmir earthquake of 2002 (M 6.4) and Pattan earthquake of 1974 ($M_w = 7.4$)¹. Previous destructive earthquakes in the Kashmir valley that occurred in 1555 (magnitude not known), 1885 (M_w 7.5), 1842 (M_w 7.5 Kinnuar) and the Kangra earthquake of 1905 (M_w 7.8) are reported at http://asc-india.org/events/051008_pak.htm.

*For correspondence. (e-mail: mahajan@wihg.res.in)