Journal of Hydrology 222 (1999) 1–17

Stochastic modelling of daily rainfall in Nigeria: intra-annual variation of model parameters

O.D. Jimoh[a,*], P. Webster[b]

[a]Department of Civil Engineering, Federal University of Technology, Minna, Nigeria
[b]School of Civil Engineering, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received 9 October 1998; received in revised form 8 June 1999; accepted 9 June 1999

Abstract

A Markov model of order 1 may be used to describe the occurrence of wet and dry days in Nigeria. Such models feature two parameter sets; P_{01} to characterise the probability of a wet day following a dry day and P_{11} to characterise the probability of a wet day following a wet day. The model parameter sets, when estimated from historical records, are characterised by a distinctive seasonal behaviour. However, the comparison of this seasonal behaviour between rainfall stations is hampered by the noise reflecting the high variability of parameters on successive days. The first part of this article is concerned with methods for smoothing these inherently noisy parameter sets. Smoothing has been approached using Fourier series, averaging techniques, or a combination thereof. It has been found that different methods generally perform well with respect to estimation of the average number of wet events and the frequency duration curves of wet and dry events. Parameterisation of the P_{01} parameter set is more successful than the P_{11} in view of the relatively small number of wet events lasting two or more days. The second part of the article is concerned with describing the regional variation in smoothed parameter sets. There is a systematic variation in the P_{01} parameter set as one moves northwards. In contrast, there is limited regional variation in the P_{11} set. Although this regional variation in P_{01} appears to be related to the gradual movement of the Inter Tropical Convergence Zone, the contrasting behaviour of the two parameter sets is difficult to explain on physical grounds. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Daily rainfall; Stochastic modelling; Markov chain; Inter-tropical convergence zone

1. Introduction

Time series of daily rainfall records are often required as input for water resources projects. The availability of such records is often constrained by economic, technical and personnel reasons. As an alternative, the rainfall records can be simulated using stochastic rainfall models (Haan et al., 1976). This involves using the historical rainfall records to estimate the model parameters of an appropriate model, which may then be used to simulate the desired length of rainfall series. Stern (1980a,b), Garbutt et al. (1981) and Jackson (1981) reported that Markov chain models of various orders are adequate for describing the occurrence of daily rainfall in Nigeria. Jimoh and Webster (1996), however, showed that the order 1 Markov model is sufficient for representing the occurrence of daily rainfall in the country. This observation was based upon the ability of the model to reproduce the characteristics of the observed series, rather than formalised statistical tests. They estimated the parameter sets of the order 1 model, P_{01} and P_{11}, using 30

* Corresponding author.

0022-1694/99/$ - see front matter © 1999 Elsevier Science B.V. All rights reserved.
PII: S0022-1694(99)00088-8

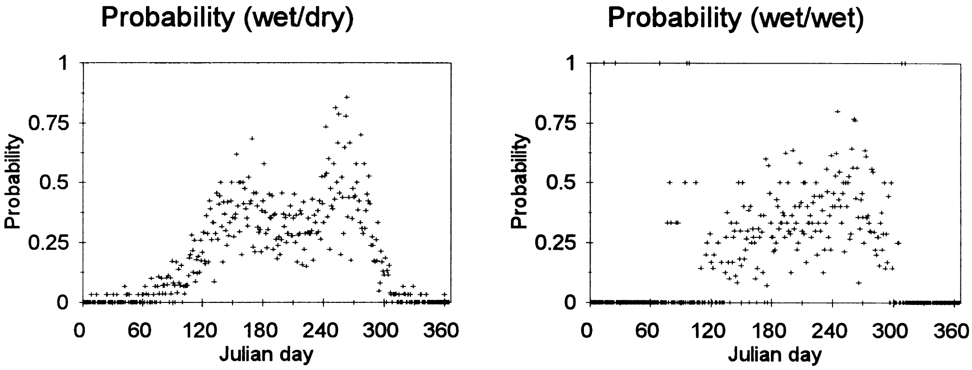

Fig. 1. Model parameters at Bida (*H*1-0, estimated using 1931-1960 records).

years of daily rainfall records for a period of stationary record from 1931 to 1960. An example of the parameters obtained for Bida in the Midland region is shown in Fig. 1.

The seasonal behaviour of the parameter sets is evident in Fig. 1, with a bi-modal pattern for the P_{01} set and an uni-modal pattern for the P_{11} set. However, the figure also illustrates the considerable noise that characterises the parameter sets. Its presence limits the ability to compare parameter sets either from

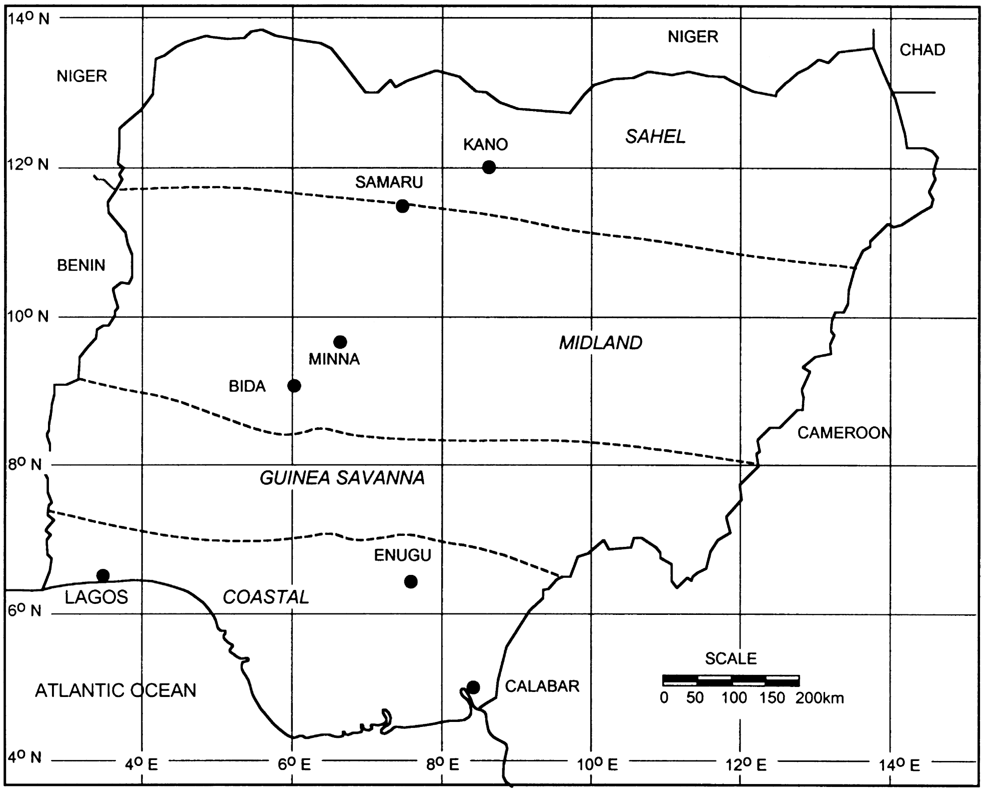

Fig. 2. Map of Nigeria showing the geographical regions and the selected seven synoptic stations.

Table 1
Summary of daily rainfall records collected for this study

Region	Climatic mean (mm)	Mean no of wet days	Station	Period of record 1931–1990	Latitude (N) Longitude (E)
Sahel	875	49	Kano	ok	12.05° 8.53°
Midland	1050	70	Samaru	ok	11.18° 7.63°
	1300	81	Minna	ok	09.62° 6.53°
	1200	70	Bida	ok	09.10° 6.02°
Coastal	2000	94	Enugu	1966–68 & 70 missing	06.47° 7.55
	2000	88	Lagos	1936 missing	06.60° 3.40°
	3000	147	Calabar	1967 missing	04.97° 8.35°

different periods of record for the same gauge, or between gauges for equivalent periods of record. It is essential that some form of smoothing is carried out before such comparisons can be made. This paper describes different approaches that have been adopted for smoothing the parameter sets obtained from seven gauges in Nigeria. Their locations are shown in Fig. 2, and other station details are shown in Table 1. The approaches are compared in terms of the ability of the smoothed sets to reproduce the characteristics of the observed series. Once smoothed, the parameter sets could then be more readily compared between gauges and between periods of record. Prior to analysis, the records were subjected to some quality control checks including double mass analysis and were found to be of acceptable quality.

2. Smoothing model parameters

Smoothing of the model parameters can be achieved using mathematical functions. For example, Coe and Stern (1982) and Zucchini and Adamson (1984) used Fourier functions to smooth the model parameters at some stations in Africa. In addition, Woolhiser and Pegram (1979) averaged the model parameters over an interval of 15 days for stations in the US, and then used Fourier series to smoothen the variation in the parameters. For these studies, the performance of the Fourier series in describing the seasonal behaviour was described in terms of statistical tests such as log-likelihood function or deviance of the estimation. It is perhaps of more direct relevance to know the extent to which the Fourier fitted model parameters are able to reproduce the characteristics of the observed sequence of wet and dry days. Although Jimoh and Webster (1996) reported that the sequences of wet and dry days generated with the unsmoothed model parameters are similar to the observed sequences, the ability of the Fourier fitted parameters to reproduce the characteristics of the historical sequence has not previously been reported. This is clearly of importance in the use of models where limited observed data are available for parameter identification.

In this analysis a day is defined as a wet day if rainfall depth is equal to, or exceeds a threshold value (2 mm for this study); otherwise the day is referred to as a dry day. A wet (or dry) event refers to a sequence of consecutive wet (or dry) days. A sequence of wet and dry days Q, is obtained from the daily rainfall record as:

$$Q = \{X_1, X_2, X_3, \ldots, X_{n-1}, X_n\} \quad (1)$$

where X_1, X_2, X_3 or X_n is either 0 or 1 and the suffixes 1, 2,…,n denote the days when the records are taken. The sequence is said to fit a first order Markov chain model if the probability that it rains on day $t(X_1 = 1)$ depends only on the previous day's rainfall. The Markov chain is referred to as a two-state chain, since X_t is 0 or 1. The parameters of the model, $P_{al}(a = 0, or\ 1)$ are estimated as:

$$P_{a1}(t) = \mathrm{Prob}(X_t = 1/X_{t-1} = a) = \frac{\text{number of years } X_t = 1, \text{ and } X_{t-1} = a}{\text{number of years } X_{t-1} = a} \quad (2)$$

The present study presents two techniques for fitting mathematical functions to the model parameters that have previously been identified using

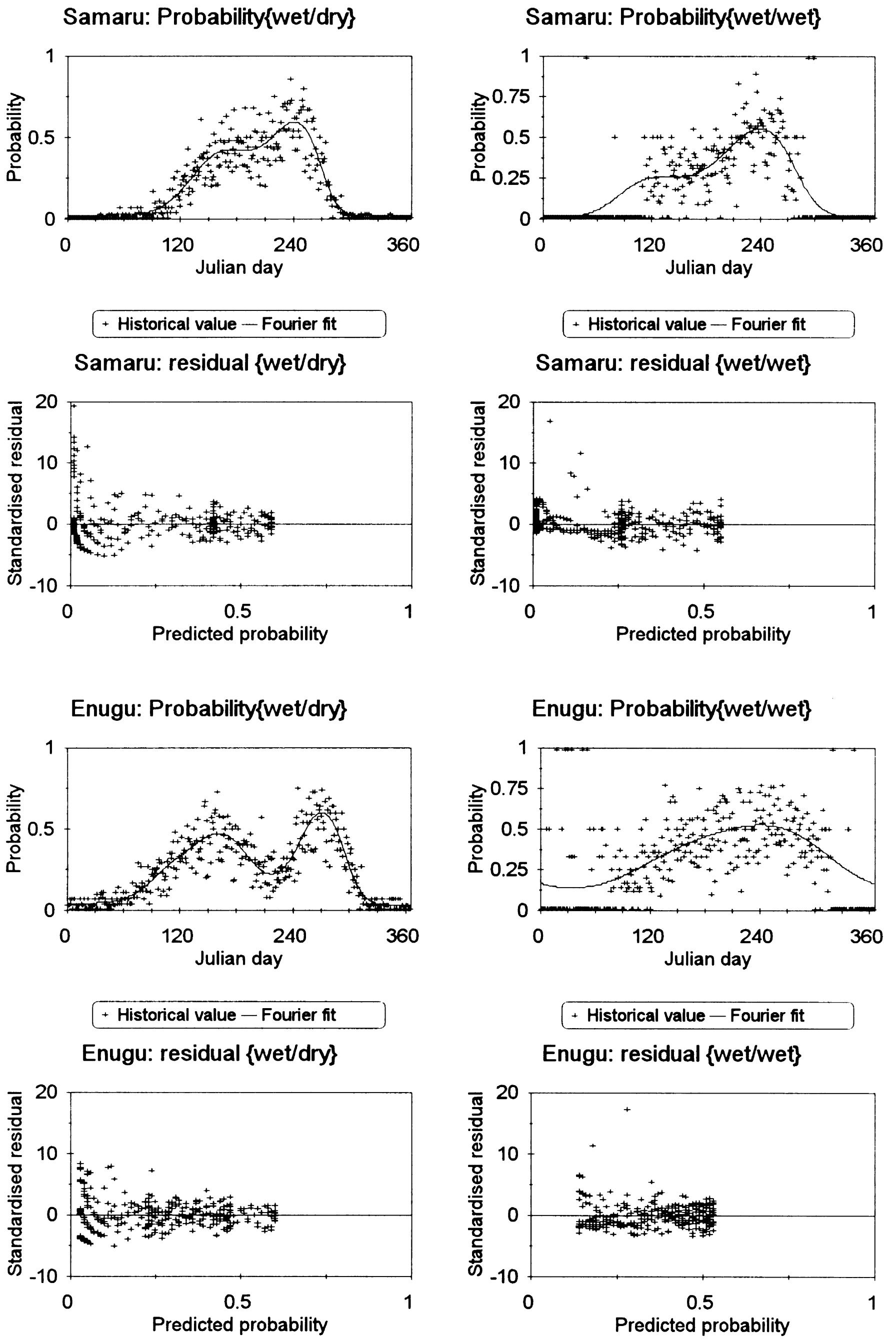

Fig. 3. Historical parameters (*H*1-0) and fitted Fourier function (*F*1-0).

Coefficients of Fourier function									
	A0	A1	B1	A2	B2	A3	B3	A4	B4
Samaru									
Pr{wet/dry	-2.29366	2.48820	0.60801	-0.17953	0.34400	-0.45301	-0.28336	0.12078	-0.13827
Pr{wet/wet	-1.98010	2.16147	0.47229	-0.92811	0.26635				
Enugu									
Pr{wet/dry	-1.47282	1.45193	0.38923	-0.68635	-0.26237	0.04871	-0.53190	0.25949	0.03750
Pr{wet/wet	-0.77221	0.71456	0.61766	-0.10845	-0.08214				

Fig. 3. (*continued*)

historical data. These techniques are Fourier series fitting, and averaging, coupled with Fourier fitting.

2.1. Fourier fitting at daily time step

The values of the model parameters on a daily time step estimated from the historical records (represented as H1-0 in this study) form the series of a binomial distribution (Collett, 1991). The H1-0 values may be expressed as a function of the Julian day using either a polynomial function or a Fourier series. A Fourier series is preferred to a polynomial function because the former is more flexible and thus may fit bi-modal functions more easily. The Fourier series to the parameters of a first order model at daily time step is expressed as

$$Y(t) = A_0 + \sum_{i=1}^{c_m} (A_i \cos(i\tau) + B_i \sin(i\tau)) \sum_{j=0}^{j=k} \beta_j x_{ji} \quad (3)$$

where i is the harmonic, c_m the maximum harmonic required for the series, A_i and B_i (or β_j) are the coefficients of the series, x_{ji} the cosine or sine of ($i\tau$) for a given harmonic I, $Y_p(t)$ is the transformed model parameter at day t, and $k = 2c_m + 1$.

The derivation of the coefficients is given in Appendix A of this article. The Fourier fitted values of the model parameters are denoted as the F1-0 set. The main problem encountered in adopting this technique to the seven stations in Nigeria is that values of model parameters estimated from the historical data (H1-0) are zero for some Julian days (see Fig. 1 for Bida station). Thus, it becomes impossible to obtain the logarithm of such values. The following assumptions were made during the fitting procedures so that transformed values of model parameters for all Julian days could be obtained.

- If the calculated $P_{a1}(t) = 0$, i.e. $N_{a1} = 0$; then set $N_{a1} = 0.01N_a$.
- If the calculated $P_{a1}(t) = 1$, i.e. $N_{a1} = N_a$; then set $N_{a1} = 0.99N_a$.
- If the calculated $P_{a1}(t) = \infty$,, i.e. $N_{a1} = 0$ and $N_a = 0$; then set $P_{a1} = 0.0066$, so that Logarithm P_{a1} equals 5.0.

The Fourier fitted (F1-0) and historical (H1-0) values for Samaru and Enugu stations are shown in Fig. 3. The figure shows that the variability of H1-0 around F1-0 is high, especially for P_{11}. The parameter P_{11} shows higher variability than P_{01} because there are fewer wet events lasting two or more days than lasting one day (Jimoh and Webster, 1996). The bi-modal shape of the P_{01} parameter set requires more harmonics than the uni-modal P_{11} set. The large number of harmonics cannot be defended on physical grounds, but does provide a systematic approach to smoothing of these complex and noisy functions. An alternative approach could be to model the beginning and end of the rainy season using an alternative procedure, whilst retaining a Fourier fit to the seasonal variation. Preliminary studies were carried out into fitting functions to different seasons, but these were not particularly successful.

The variability for Samaru is higher than for Enugu due to regional variation in the monthly number of wet days, with the number decreasing from the Coastal to the Sahel region (Jimoh, 1997). Samaru is broadly typical of stations in the Sahel and Midlands, whilst Enugu is typical of stations in the Coastal region (Fig. 2).

2.2. Fourier fitting at weekly time step

The value of the model parameters at a weekly time step is the arithmetic average of the daily values of the parameter within a seven days interval, and the

Table 2
Notation representing values of model parameters

	Description	Fourier fitted
*H*1-0	Values of parameters estimated at daily time step from the historical record.	*F*1-0
*H*7-0	Values of parameters at seven day time step, obtained by averaging *H*1-0 for the time interval.	*F*7-0
*H*1-*n*	Values of parameters at one day time step. They are obtained by averaging *H*1-0 at *n*-day time step and the intermediate values at one day time step obtained by linear interpolation.	*F*1-*n*

parameter sets are denoted as *H*7-0. In order to fit Fourier series to *H*7-0, Eq. A3 was modified to become:

$$\tau = \frac{\prod((\mathrm{wk}_{\mathrm{day}} - 3.5) - 182.5)}{182.5} \quad (4)$$

where $\mathrm{wk}_{\mathrm{day}}$ is the Julian day at the end of the week.

With this technique, the daily variability in the model parameters is smoothed, but the model parameters show the average value of the parameter for each time interval. The Fourier fitted values at a weekly time step (denoted as *F*7-0) have been evaluated, and plotted in the same style as Fig. 3. It is generally noted that the start and end of the wet season are better represented than with fitting over a daily time step. The use of linear interpolation to overcome the unnatural breaks at the end of a time period is discussed in the following Section.

2.3. Averaging technique and Fourier series at daily time step

The variability in the model parameters may also be described by a combination of arithmetic averaging of the parameter values and Fourier fitting. This procedure is described below, while Table 2 gives an explanation of the symbols used.

1. Use the historical records to estimate the model parameters at a daily time step (*H*1-0).
2. Average the *H*1-0 values over an *n*-day time step, noting the value at the mid-point of the time interval.
3. Interpolate the parameter values between the mid-points by linear interpolation, with the interpolated values denoted as *H*1-*n*.
4. Fit a Fourier series to the interpolated *H*1-*n* values, denoted *F*1-*n*.

This procedure was applied to the seven stations in Nigeria using *n* values of 7 and 15. The problem of having zero probability value, especially at the middle of the wet season (June–September) in *H*1-0 is eliminated in the *H*1-15 and *H*1-7 sets. An investigation of the variability of model parameters (*H*1-*n*) with *n* (*n* varying between 1 and 15) shows that the daily variability in *H*1-*n* decreases with the value of *n* increasing. However, the concept of linear interpolation assumed in obtaining *H*1-*n* values (step 2 above) is affected at high *n* values. Thus a balance between inter day variability and linear interpolation is required in selecting a value for *n*. It was found that values of *n* between 7 and 15 achieved a reasonable compromise.

The residual of the Fourier fitting to *H*1-*n* (*n* is 7 or 15) at Enugu, in the Coastal region is independent of the estimated probability of wet day. For the Sahel region using Samaru station for example, the residual is high at low probability, especially when the probability of wet day given a previous wet or dry day is less than 0.1. This suggests that a better Fourier fitting to *H*1-*n* values is achieved at the Coastal than at the Sahel regions. Generally, the *F*1-15 set is better than the *F*1-7 at all stations, and further discussion is based on the *F*1-15 set.

2.4. Performance of the fitting procedure

Synthetic sequences of wet and dry days were generated using the different sets of model parameters (*H*1-0, *F*1-0 and *F*1-15) in order to compare the performance of the different smoothing methods. The procedure for generating synthetic sequences of wet and dry days for an order 1 Markov model is described in Jimoh and Webster (1996). The synthetic sequences of wet and dry days were compared with those of the historical records, using the following criteria:

- the monthly number of wet days;
- the first and last wet day;
- average number of wet events and
- frequency duration curves for wet and dry events.

The results of this analysis using *H*1-0, *F*1-0 and

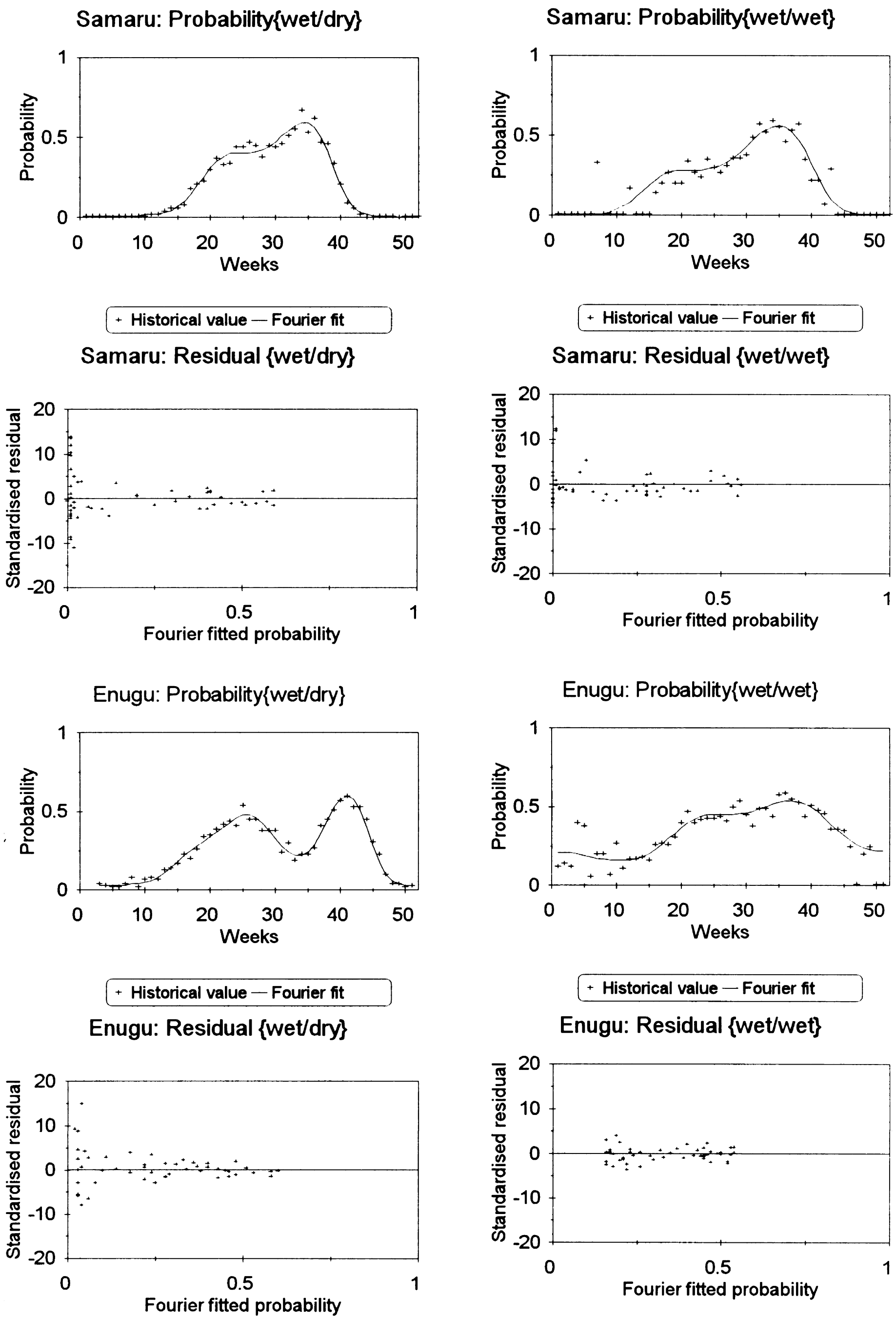

Fig. 4. Historical parameters (H7-0) and fitted Fourier function (F7-0).

	Coefficients of Fourier function								
	A0	A1	B1	A2	B2	A3	B3	A4	B4
	Samaru								
Pr{wet/dry}	-2.41839	2.65521	0.61961	-0.20781	0.41186	-0.42159	-0.42452		
Pr{wet/wet}	-2.39840	2.84090	0.63123	-1.18578	0.09056				
Pr{wet/dry}	Enugu								
Pr{wet/dry}	-1.54207	1.56843	0.37923	-0.74304	-0.27007	0.03409	-0.50742	0.29493	-0.00722
Pr{wet/wet}	-0.75281	0.62848	0.59882	0.00208	-0.02274	-0.06474	-0.18693		

Fig. 4. (*continued*)

*F*1-15 to estimate the average monthly number of wet days are presented in Fig. 4 for two stations. The figure. also shows the variation of the standardised residual with the predicted probability. The results show that there is no noticeable difference between the monthly number of wet days estimated by *F*1-0 and *F*1-15, and that of *H*1-0. The study also showed that the different fitting procedures present similar numbers of wet events and frequency duration curves. A comparison of the mean and standard deviation of the number of events with one and two days' duration is shown in Table 3. The results show good general agreement for both wet and dry events. There is a tendency for the standard deviation of one day wet events for generated sequences to be lower than those for observed sequences, though this is not universal. This trend has been noted by Smith and Schreiber (1973) and Gregory et al. (1992) and highlights a limitation of Markov chain models in representing seasonal variations in rainfall series. The use of a conditioned Markov model, as outlined in Jimoh and Webster (1998) provides a possible means of overcoming this problem.

This investigation showed that there is little difference between the *H*1-0, *F*1-0 and *F*1-15 in representing the characteristics of wet and dry days. For *F*1-0 and *F*1-15 however, the inherent variability in *H*1-0 has been smoothed. In addition, the smoothed curve represents the start of wet season better than the *H*1-0. The start of the planting season under rainfed agriculture depends on the start of wet season, and the proper representation of this criterion by smoothed curve is considered advantageous.

3. Regional variation in the model parameters

The principal benefit of the smoothing techniques is that it facilitates the comparison of parameter sets between different gauges. Model parameters for an order 1 Markov model have been evaluated for the seven gauges used in this study. The smoothed values for the P_{01} set are shown in Fig. 5 and those for the P_{11} parameter set in Fig. 6. Various features of these sets are summarised in Tables 4 and 5.

The following observations can be made from Fig. 5 about the P_{01} parameter set:

- The only set that is clearly uni-modal is that for Kano in the north. The set for Samaru is characterised by a turning point around Julian day 150, and all other sets are bi-modal.
- The timing of the peaks varies systematically in a north-easterly direction. The timing of the first peak varies from day 155 at Lagos and Calabar to day 175 at Samaru. The timing of the second peak varies over a greater range from day 285 at Lagos to day 230 at Kano.
- The modal values of probability for the first peak are reasonably constant in the north at around 0.4, with higher values in the south. There is less variability for the second peak with values in the range 0.5–0.6 for all stations, with the exception of Lagos. There is slight evidence of a decrease in a northerly direction.
- The set for Lagos is distinctive in that the first peak is higher than the second peak.
- The set for Calabar in the southeast also shows a distinctive behaviour, partly in terms of the area under the curve (i.e. average annual number of wet days), but also the consistently high value of probability between days 120 and 270.

In general terms, the timing of the peaks would appear to conform to general models of the movement of the rain bands. The observed timing suggests a generally north-eastward movement, with Lagos in

Table 3
Characteristics of events with durations of 1 and 2 days

Duration		Wet events				Dry events			
		1-day spell		2-day spell		1-day spell		2-day spell	
Station	Item[a]	Frequency (day)	Standard deviation (no).	Frequency (day)	Standard deviation (no)	Frequency (day)	Standard deviation (no)	Frequency (day)	Standard deviation (no).
Kano	Observed	24.5	4.8	5.6	1.6	11.6	3.5	7.3	2.4
	*H*1-0	23.9	4.2	6.1	2.2	11.5	3.4	6.8	2.4
	*F*1-15	24.2	4.1	6.6	2.4	11.5	3.4	6.8	2.4
Samaru	Observed	27.4	5.1	9.1	2.4	17.5	4.3	9.9	3.4
	*H*1-0	27.4	4.4	9.1	2.7	16.9	4.1	9.4	2.7
	*F*1-0	28.3	4.6	9.4	2.8	17.0	4.4	9.5	2.9
	*F*1-15	27.3	4.5	9.3	2.8	16.6	4.2	9.3	2.8
Minna	Observed	30.6	5.3	9.9	2.8	20.1	4.0	10.4	2.9
	*H*1-0	30.6	4.8	10.3	3.0	18.4	4.1	10.4	2.8
	*F*1-0	31.0	4.8	10.9	3.1	18.6	4.3	10.6	3.0
	*F*1-15	30.0	4.7	10.6	3.1	18.1	4.3	10.5	3.0
Bida	Observed	33.9	5.0	8.9	2.8	16.5	4.3	10.2	2.5
	*H*1-0	33.8	5.0	9.4	2.7	16.2	4.1	10.1	3.1
	*F*1-0	34.5	4.9	9.6	2.8	16.6	4.2	10.0	2.9
	*F*1-15	33.4	4.7	9.4	2.9	16.2	4.3	9.8	2.8
Enugu	Observed	33.9	6.7	11.5	3.3	19.7	5.8	12.0	4.5
	*H*1-0	33.6	5.1	12.7	3.3	19.4	4.3	11.8	3.2
	*F*1-0	34.1	5.3	12.9	3.3	20.1	4.6	11.4	3.2
	*F*1-15	33.8	5.2	12.6	3.3	19.3	4.6	11.3	3.1
Lagos	Observed	32.8	7.7	10.4	3.6	14.7	4.8	10.8	3.9
	*H*1-0	34.0	4.9	10.8	2.9	15.3	3.9	9.8	3.1
	*F*1-0	33.9	5.3	11.5	3.2	15.3	3.9	9.8	3.1
	*F*1-15	34.2	5.3	10.9	3.2	15.3	3.9	9.7	3.1
Calabar	Observed	35.5	4.5	15.8	3.2	31.7	5.0	15.5	4.7
	*H*1-0	36.4	5.4	14.9	3.5	30.6	5.1	15.5	3.6
	*F*1-0	36.4	5.7	15.5	3.7	30.4	5.2	15.8	3.7
	*F*1-15	36.6	5.8	15.3	3.6	30.3	5.1	15.7	3.7

[a] *H*1-0, *F*1-0 and *F*1-15, as defined in Table 2.

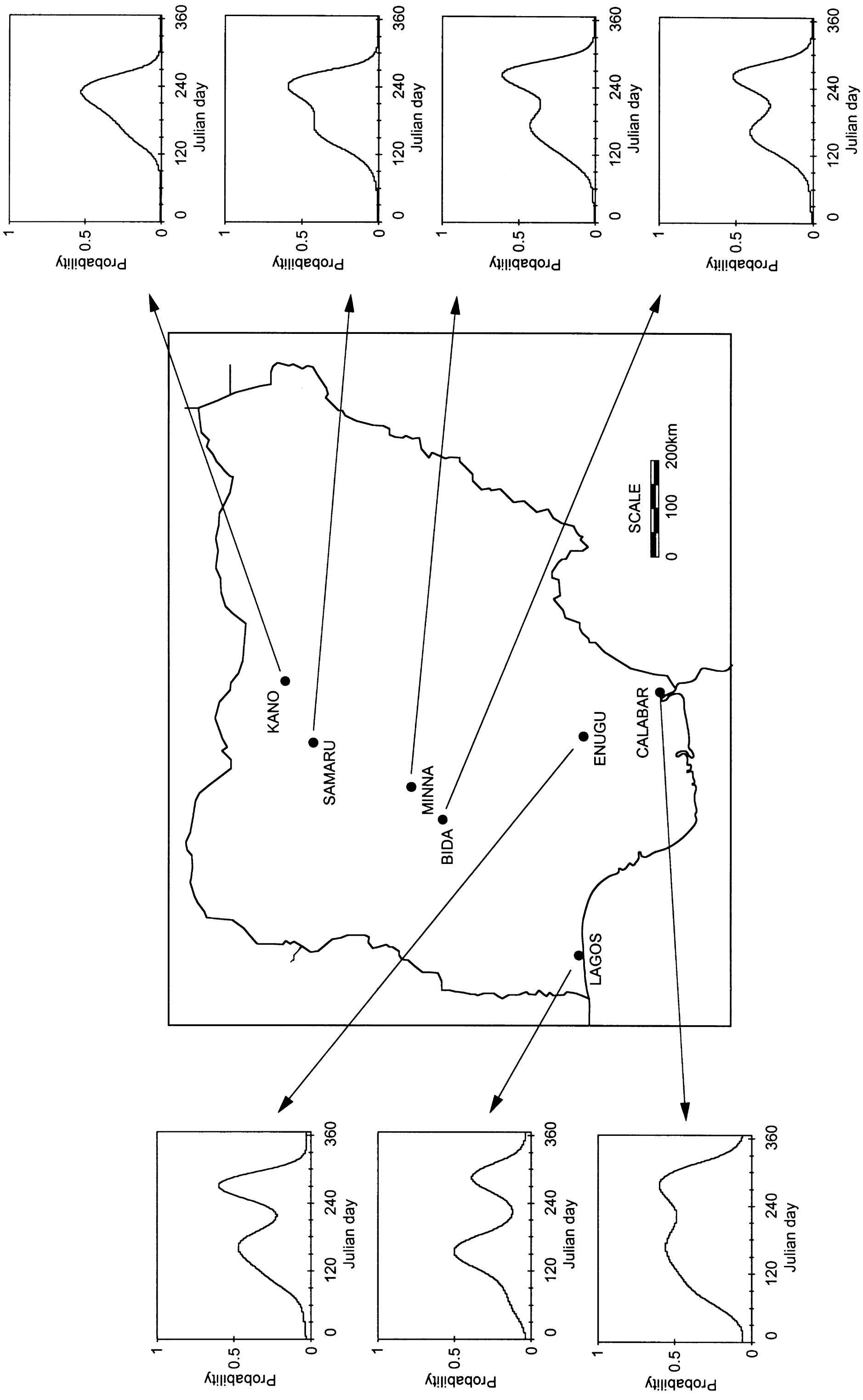

Fig. 5. Model parameters P_{01}.

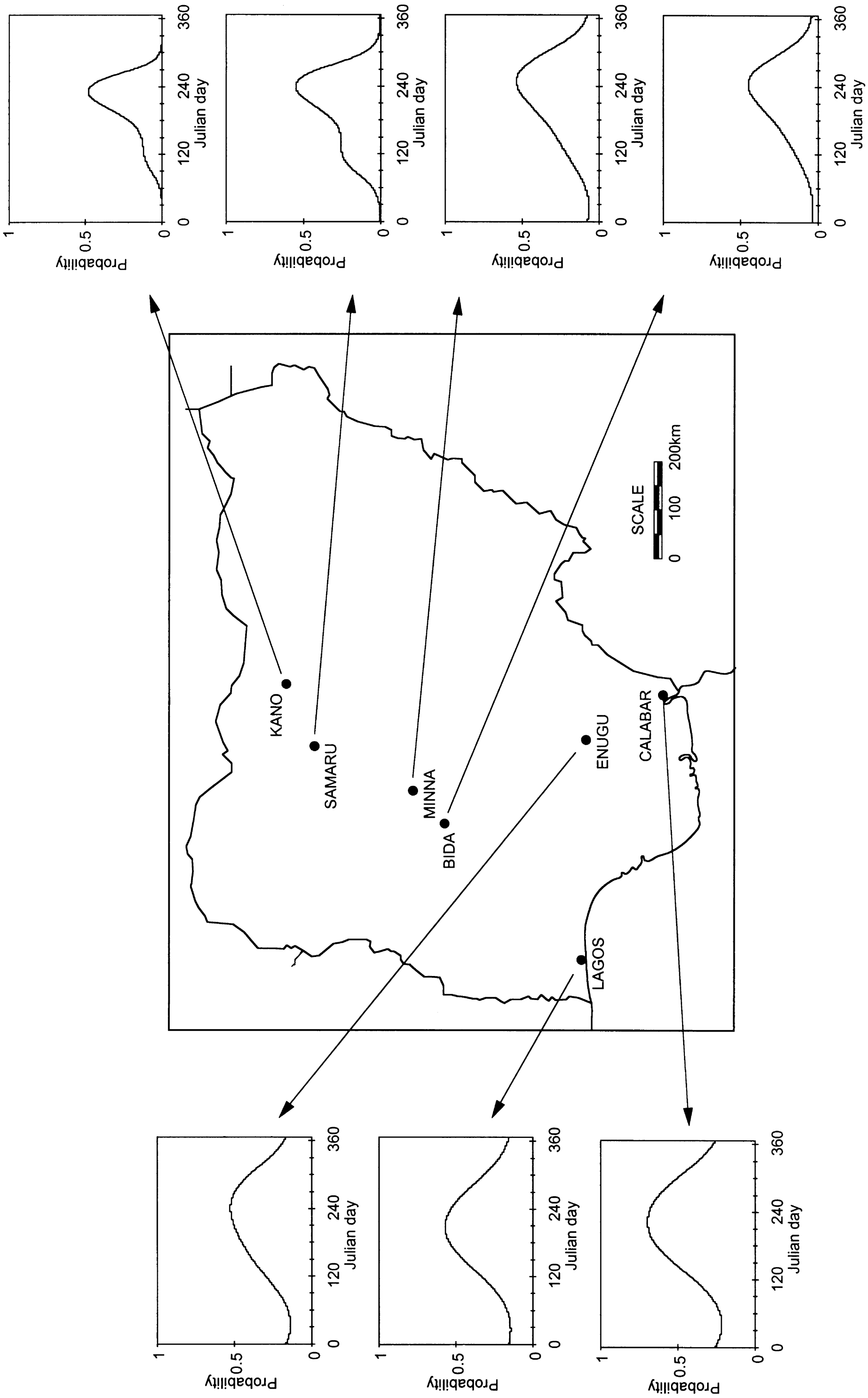

Fig. 6. Model parameters P_{11}.

Table 4
Timing in Julian days of peaks in P_{01} and P_{11} parameters sets

Station	P_{01} (Peak 1)	P_{01} (Peak 2)	Difference	P_{11} (uni-modal)
Kano		230		230
Samaru	160–190	240	50–80	240
Minna	170	260	90	250
Bida	165	265	100	240
Enugu	160	270	110	240
Calabar	165	280	115	225
Lagos	155	285	130	210

the southwest, experiencing the earliest appearance of the first peak. The greater range in timings of the second peak implies a slower retreat to the south of rain bands, and contributes to the relatively large range in timings between the peaks. The trend from bi-modal behaviour in the south to uni-modal behaviour in the north also confirms these trends in relative timing. The modal values of probability for each peak vary over a surprisingly small range 0.40–0.50 for the first peak (with the exception of Calabar) and 0.52–0.66 with the exception of Lagos for the second peak. Although some systematic trend may be identified, this is probably confounded by the relative timing of the peaks.

The observations that can be made from Fig. 6 and Tables 3 and 4 about the P_{11} parameter set are limited due to the decreased confidence in the parameter values. This follows from the smaller number of wet events lasting two days or longer relative to those of one day duration. Nevertheless, the following general observations can be made:

- The parameter sets are uni-modal throughout, although the set for Samaru is characterised by a turning point, as it was for the P_{01} set.
- The timing of the peak is reasonably uniform from day 210 (Lagos) to day 250 (Minna), although stations in the north show a slight reduction.
- The modal value of probability shows a general decline in a northerly direction.

Table 5
Magnitude of peaks in P_{01} and P_{11} parameters sets

Station	P_{01} (Peak 1)	P_{01} (Peak 2)	P_{11} (uni-modal)
Kano		0.52	0.47
Samaru	0.42	0.59	0.54
Minna	0.44	0.61	0.53
Bida	0.41	0.53	0.46
Enugu	0.48	0.60	0.53
Calabar	0.56	0.60	0.70
Lagos	0.49	0.39	0.57

The typical shape for the P_{01} set is negatively skewed, with a relatively steep rate of change from the second peak. In contrast, the typical shape of the P_{11} set is symmetrical. These typical shapes are also evident in Fig. 7 which explores the joint variation of the P_{01} and P_{11} parameter sets at three of the stations. The left-hand panels show the joint variation of the *H*-1 values, from which little can be gained. The right-hand panels show the variation of the Fourier fitted values, which reveal a hysteretic pattern throughout the season. The hysteresis is less pronounced for Kano and Samaru in the north than for Enugu in the south.

This behaviour of the parameter sets must relate to regional variation in the meteorological factors that are responsible for rainfall in the country. Rainfall in Nigeria occurs in response to the moist, warm, southwest trade wind. Further investigation on the position or boundary of this trade wind across Nigeria and its relationship with the parameters of a Markov model was subsequently studied.

4. Relationship with the ITCZ

The Inter-tropical Convergence Zone (ITCZ) is the boundary between the moist, warm southwest trade wind and the dry, cool, northeast trade wind. The position of the ITCZ also shows the position of the southwest wind. The position of the ITCZ may be determined using information on cloud, dew temperature, or wind at sea level. In this study, Highly

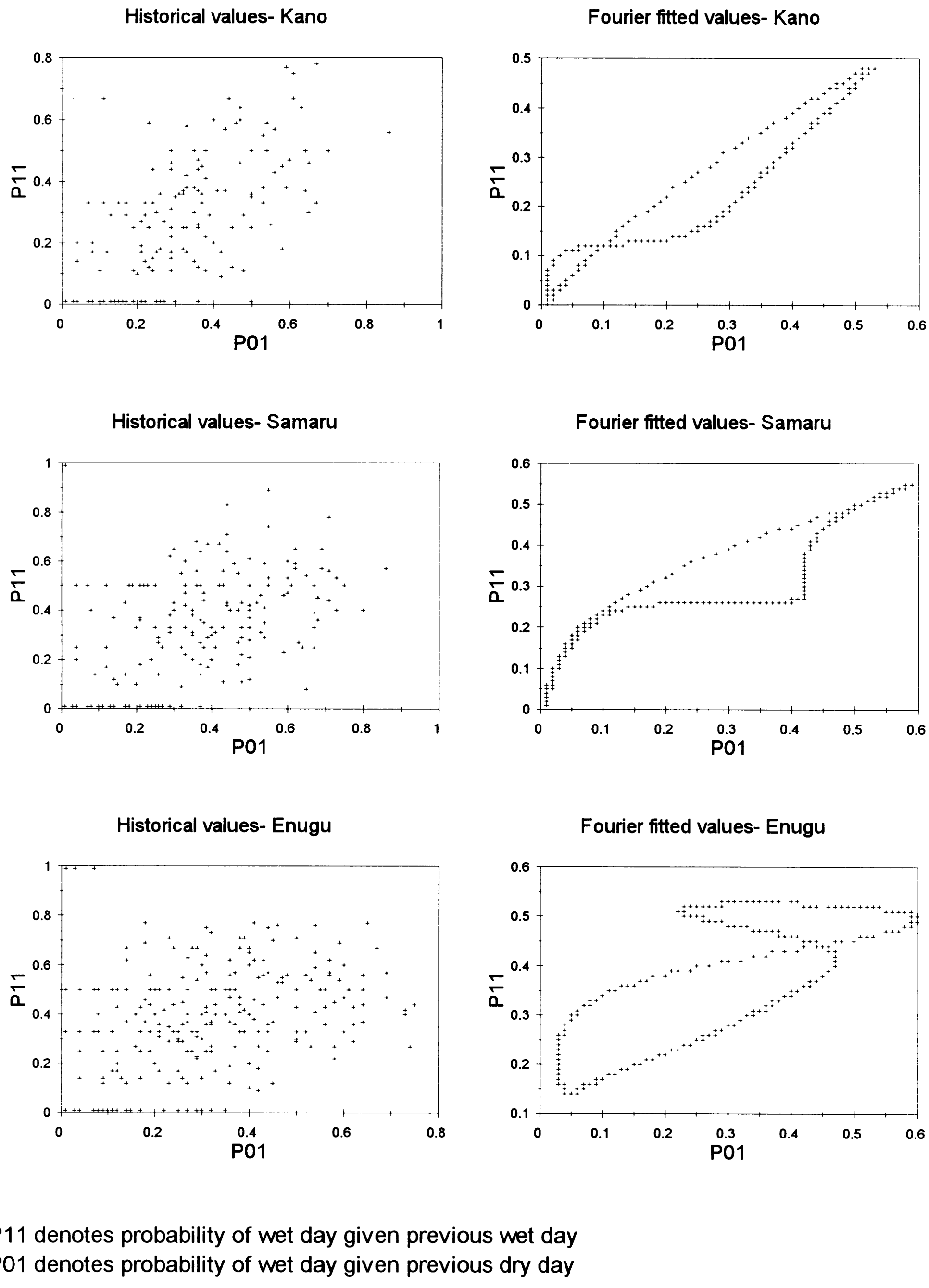

Fig. 7. Relationship between parameters of first order model.

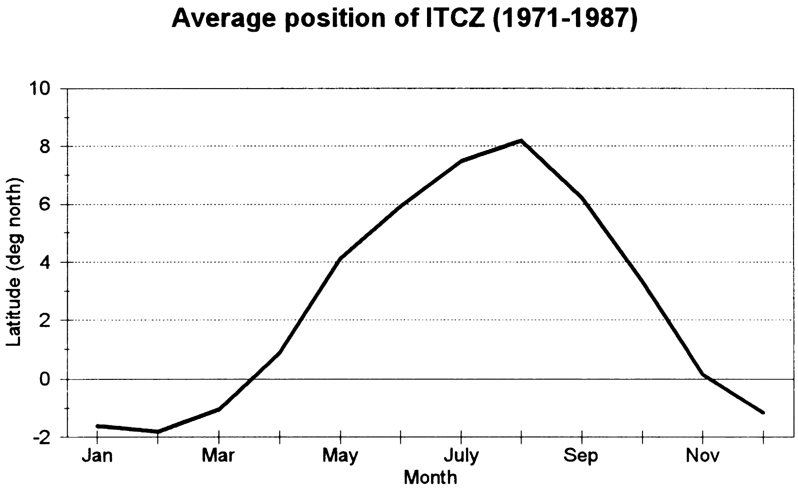

Fig. 8. Latitudinal position of ITCZ over Nigeria.

Reflective Cloud (HRC) data is used to define the position of ITCZ, and the merits of this technique relative to other methods are discussed in Jimoh (1997). The compilation of HRC data (Hastenrath, 1990) is based on National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites during January 1971–February 1978 and February 1979–December 1987, and Defence Meteorological Satellite Program (DMSP) during March 1978–January 1979. The HRC data set used for this study was supplied by Prof. Waliser, D.E of Institute for Terrestrial and Planetary Atmospheres, Sunny, Stony Brook, New York. For each month, the ITCZ latitude ($\bar{\theta}$) is calculated as the HRC-weighted mean latitude (Waliser and Gautier, 1993):

$$\overline{\theta(t)} = \frac{\int_{-25}^{+25} H(\theta, t)\theta \, \mathrm{d}\theta}{\int_{-25}^{+25} H(\theta, t) \, \mathrm{d}\theta} \qquad (5)$$

where t is the month index spanning the 204 months of HRC data; H the zonally averaged HRC over the given longitude domain; and θ represents latitude in degrees.

The area weighted mean technique is adopted for defining the position of ITCZ because of the following reasons:

1. It is felt that the ITCZ axis correlates with the axis of maximum convection.
2. The technique is less influenced by the occurrence of isolated convective cloud, and therefore less affected by outliers in the series.
3. There is perhaps, a consistent relationship between the axis of maximum convection and the northern boundary of the band of convective activity.
4. The technique is numerically straightforward.

Graphs showing the monthly latitudinal position of the ITCZ have been prepared for the period from 1971 to 1987 corresponding with the availability of the HRC imagery. It is evident from these graphs that there is considerable inter-annual variability in the position of the ITCZ, corresponding with the development of the Sahelian drought. The inter-decadal variation of the Markov model parameter sets, and their relationship with rainfall anomalies has been discussed in Jimoh and Webster (1998). Of interest in this article is the average behaviour of the ITCZ over the 17 year period of available record. Accordingly, the average latitudinal position was determined in Fig. 8, and compared with the parameter sets presented in Figs. 5 and 6, from which the following observations are made:

1. The annual behaviour of the ITCZ typically shows a period of advance and a period of retreat, giving rise to a predominantly uni-modal pattern. Occasional retreat is evident in individual years (e.g. 1984, 1987). However, the bi-modal pattern of rainfall and the P_{01} set may be construed from

the movement of the axis of maximum convection over a given gauge.

2. The P_{01} set is consistent with this general movement of the ITCZ, displaying typically a bi-modal pattern, with a trend towards a uni-modal pattern in the north. There is however, no evidence from the ITCZ position to explain the different magnitude of the first and second peaks.
3. It was observed that the P_{11} set was generally symmetrical at each station; a pattern which is similar to the advance and retreat of the ITCZ (Fig. 8). The time of peak for both patterns is also similar at between day 210 and 230.

It is therefore evident that there is general correspondence of the ITCZ position and the P_{01} parameter sets, although this does not appear to apply to the P_{11} set. However, the movement of the ITCZ cannot be determined with sufficient resolution to correlate its movement in detail with the specific shape of the parameter sets.

5. Concluding remarks

5.1. Methods of smoothing parameter sets

It is essential that parameter sets be smoothed in order to facilitate comparison between different gauges and different periods of record. The comparative work has shown that the use of different techniques of smoothing based upon Fourier fitting and averaging are equally good in terms of reproducing the characteristics of the observed record. The *F*1-15 curve appeared better able to reproduce the start of the wet season and has been used in all subsequent analysis.

5.2. Regional variation in parameter sets

The parameter sets have been evaluated for seven stations across Nigeria. The P_{01} set is predominantly bi-modal, though this tendency declines in a northerly direction, with the northernmost station being uni-modal. There is a systematic variation of the relative timing of the peaks, though the magnitude of the peak value is remarkably consistent for all stations studied. The P_{11} set is uni-modal at all stations. In contrast to the P_{01} set, the timing of the peak varies over a narrow range, whilst the magnitude shows a systematic decrease in a northerly direction.

5.3. Linking the parameter sets with movement of the ITCZ

The movement of the ITCZ has been deduced from Highly Reflective Cloud Imagery data. The movement of the axis of maximum convective activity over rainfall stations can give rise to the observed rainfall regimes and P_{01} parameter sets. It is therefore evident that there is general correspondence of the ITCZ position and the P_{01} and P_{11} parameter sets. However, the movement of the ITCZ cannot be determined with sufficient resolution to correlate its movement in detail with the specific shape of the parameter sets.

5.4. Non-stationary periods

A characteristic feature of the rainfall records is the non-stationarity during the period from about 1970 onwards. The relationship between rainfall anomalies and the smoothed parameter sets was investigated by Jimoh and Webster (1998). The work has identified a system for forecasting the anomaly based upon pre-season weather variables comprising sea surface temperatures and the position of the ITCZ. However, the predictive model is constrained by the limited duration for which information on the ITCZ position is available. It is hoped that the model can be refined using HRC imagery from 1987 onwards.

Acknowledgements

The rainfall data were obtained from the Federal Department of Meteorological Services, Lagos (Nigeria), the International Institute of Tropical Agriculture and National Horticultural Research Institute Project (Nigeria), and the Department of Applied Statistics, University of Reading (UK). The authors acknowledge these sources. This research was sponsored by the Commonwealth Scholarship Commission in the United Kingdom. Their financial support is appreciated. The assistance of Mary Andrews in producing Figs. 2, 5 and 6 is appreciated. We would also like to acknowledge the valuable contribution made by the anonymous reviewers of earlier drafts of the paper.

Appendix A

The Fourier series for the parameters of a first order model at a daily time step may be expressed as:

$$Y(t) = \log_e\left(\frac{P(t)}{(1 - P(t)}\right) \tag{A1}$$

and

$$P(t) = \frac{\exp(Y(t))}{1 + \exp(Y(t))} \tag{A2}$$

P denotes either the P_{01} or P_{11}

$$\tau = \frac{\prod(\text{Julian day} - 182.5)}{182.5} \tag{A3}$$

The maximum likelihood estimate of the coefficients of the Fourier series, β_j is expressed (Collett, 1991) as:

$$\bar{\beta}_{r+1} = \bar{\beta}_r + I^{-1}(\bar{\beta}_r)U(\bar{\beta}_r) \tag{A4}$$

where

$$U(\beta) = \sum_{i=1}^{n} n_i P(i)(1 - P(i))\frac{(y_i - n_i P(i))}{n_i}\frac{1}{P(i)(1 - P(i))}x_{ji} \tag{A5}$$

$$I(\beta) = \sum_{i=1}^{n} \frac{n_i[P(i)(1 - P(i))]^2}{P(i)(1 - P(i))}x_{ji}x_{qi} \tag{A6}$$

For $P = P_{01}$, n is the number of years with $X_t = 1$ and $X_{t\text{-}1} = 0$, and $E[y] =$ np.The maximum harmonic, c_{m} is determined using multiple regression techniques (Coe and Stern, 1982). A computer programme was written to perform these procedures, and the principal steps are listed below:

1. Determine the conditional probabilities and the log-transformed values.
2. Compute the sine and cosine values (x_{ji}) of the Fourier series for a given harmonic.
3. Determine the coefficients of the Fourier series using a least squares technique. The estimated coefficients are $\beta_0(\mathrm{I} = 0)$ values. Compute the deviance.

   Deviance = standardised residual

$$= 2\sum_{i=1}^{n}\left\{y_i \log\left(\frac{P(i)}{\bar{P}(i)}\right) + n_i y_i \log\left(\frac{1 - P(i)}{1 - \bar{P}(i)}\right)\right\} \tag{A7}$$

   where $\bar{P}(i) = P(i)/n_i$
4. Compute the efficient scores $U(\beta_r)$and information matrix $I(\beta_r)$ and solve for Eq. A4.
5. Compute the deviance and check whether β_{r+1} is better then β_r. The β values with the least deviance represent the best estimate.
6. Increase the harmonic and repeat steps (2)–(5).
7. Repeat step (6) until there is no justification for increasing the harmonic.

References

Coe, R., Stern, R.D., 1982. Fitting models to daily rainfall data. J. Appl. Meteor. 21, 1024–1031.

Collett, D., 1991. Modelling Binary Data, Macmillan, London pp. 1–165.

Garbutt, D.J., Stern, R.D., Dennett, M.D., Elston, J., 1981. A comparison of the rainfall climate of eleven places in West Africa using a two-part model for daily rainfall. Arch. Met. Geophy. Biokl. Ser. B. 29, 137–155.

Gregory, J.M., Wigley, T.M.L., Jones, P.D., 1992. Determining and interpreting the order of a two-state Markov chain: application to models of daily precipitation. Water Resour. Res. 28 (5), 1443–1446.

Haan, C.T., Allen, D.M., Street, J.O., 1976. A Markov chain model of daily rainfall. Water Resour. Res. 12 (3), 443–449.

Hastenrath, S., 1990. The relationship of highly reflective clouds to tropical climate anomalies. J. Climate 3, 353–365.

Jackson, I.J., 1981. Dependence of wet and dry days in the tropics. Arch. Met. Geophy. Biokl., Ser. B. 29, 167–179.

Jimoh, O.D., 1997. Modelling the occurrence of daily rainfall in Nigeria. Unpublished PhD thesis, The University of Birmingham, Edgbaston, England.

Jimoh, O.D., Webster, P., 1996. Optimum order of Markov chain for daily rainfall in Nigeria. J. Hydrol. 185, 45–69.

Jimoh, O.D., Webster, P., 1998. Weather conditioned model for the occurrence of daily rainfall in Nigeria. In: Wheater, H.S., Kirby, C. (Eds.), Hydrology in a Changing Environment, Proceedings of the First International Conference of the British Hydrological Society, vol. 1.

Smith, R.E., Schreiber, H.A., 1973. Point processes of seasonal thunderstorm rainfall. 1. Distribution of rainfall events. Water Resour. Res. 9 (4), 871–884.

Stern, R.D., 1980. Analysis of daily rainfall at Samaru Nigeria, using a simple two-part model. Arch. Met. Geophy. Biokl. Ser. B. 28, 123–135.

Stern, R.D., 1980. The calculation of probability distribution for models of daily precipitation. Arch. Met. Geophy. Biokl. Ser. B. 28, 137–147.

Waliser, D.E., Gautier, C., 1993. A satellite derived climatology of the ITCZ. J. Climate 6, 2162–2174.

Woolhiser, D.A., Pegram, G.G.S., 1979. Maximum likelihood estimation of Fourier coefficients to describe seasonal variations of parameters in stochastic daily precipitation models. J. Appl. Meteor. 18 (1), 34–44.

Zucchini, W., Adamson, P.T., 1984. The occurrence and severity of droughts in South Africa. Report No 91/1/84, Water Research Commission, Pretoria.