On Some Aspects of Climatology Associated with a Dam

T G J DYER
[CLIMATOLOGY RESEARCH GROUP, UNIVERSITY OF THE WITWATERSRAND, 1 JAN SMUTS AVENUE, JOHANNESBURG 2001]

Abstract

Rainfall series for Gaborone (Botswana) and inflow relating to a local dam have been analysed for the period 1909/10 to 1975/6. Variance spectra for both series showed peaks at wave lengths found in series elsewhere over southern Africa. Neither series possess notable oscillations with periods greater than about 10 years. Trigonometric regressions fitted to these data accounted for 71 per cent of the rainfall variance and 87 per cent of the inflow variance. These regression equations were subsequently used to obtain some estimated rainfall and inflow values for the future. This was carried out for the raw data, and for lightly smoothed data. The latter case therefore indicates the possible pattern of intervals of positive and negative anomalies. It is suggested that these estimates may be useful in reaching a decision regarding the completion date for a new dam in the same locality as the data analysed in this paper.

Introduction

The work described in this paper arose from the project that involves the proposed construction of a new dam in the Gaborone (Botswana) area. A question of interest was that of answering, when would be the most suitable year or period of years by which the dam should be completed, in order to make full use of a time of good rains and its associated variables. That there might be a "best" time stems from the knowledge that rainfall, and hence inflow to the dam, varies over time in at least a pseudo-periodic manner (Tyson *et al.*, 1975; Dyer, 1976a; Dyer, 1977a).

The behaviour of spatial mean rainfall series has been described elsewhere for a number of rainfall regions over South Africa (Dyer, 1975; Dyer, 1976b; Dyer and Tyson, 1977). These mean series are too general to describe rainfall patterns at a particular locality and it is difficult to overstress this point. This is particularly so when the locality under consideration is close to the boundary of a rainfall region for which a spatial series has been defined. In the present instance, not only is the site close to a regional boundary, but also there is no adjacent region defined to the west of it. Therefore, it was decided to analyse some local climatological factors relating to the proposed dam.

The rainfall station at Gaborone is situated at 24°40'S and 26°E, and this is also the location of Gaborone Dam. The new dam is Bokaa which is situated on the Metsemothlaba River at 24°24'S and 26°E. From available inflow data for Bokaa it was found to have 0,4 of the corresponding Gaborone inflow. The inflow data analysed as time series is therefore applicable to both dams.

Some Theoretical Aspects

Time series may be analysed over the time or frequency domain and, therefore, a combination of both was chosen. For the case of the frequency domain, the total variance of a series is broken down into components that can be represented by sinusoidal functions. When carrying out this type of work it is commonly felt that cycles are being hunted for in the data. From experience, it is appreciated that cycles just do not exist, certainly in this type of data where a cycle may be taken as a particular case of an oscillation in which the distances between successive peaks, or troughs, remains constant over time. What may be expected are pseudo-cycles, that is, sinusoidal type motions with varying distances between the peaks. To be able to describe such patterns still remains useful as long as the range of the wavelengths, distances between peaks, is not too large. A good example of a cycle which is not one, is the so-called solar or sun-spot cycle. This is taken as eleven years, but can be between 9 and 15 years, so it is an oscillation. An inspection of Figure 1 which shows annual rainfall totals, and dam inflow suggest pseudo-periodic behaviour of these series (hydrological years).

The conventional approach to a frequency domain analysis of time series is that of obtaining the spectra. This involves a Fourier transform of the autovariance function. The method depends on choosing a suitable maximum lag in the auto-covariance function and this is by no means a trivial matter. Too large a lag may bring about spurious peaks in the spectrum and too small a value tends to have too dampening an effect on it.

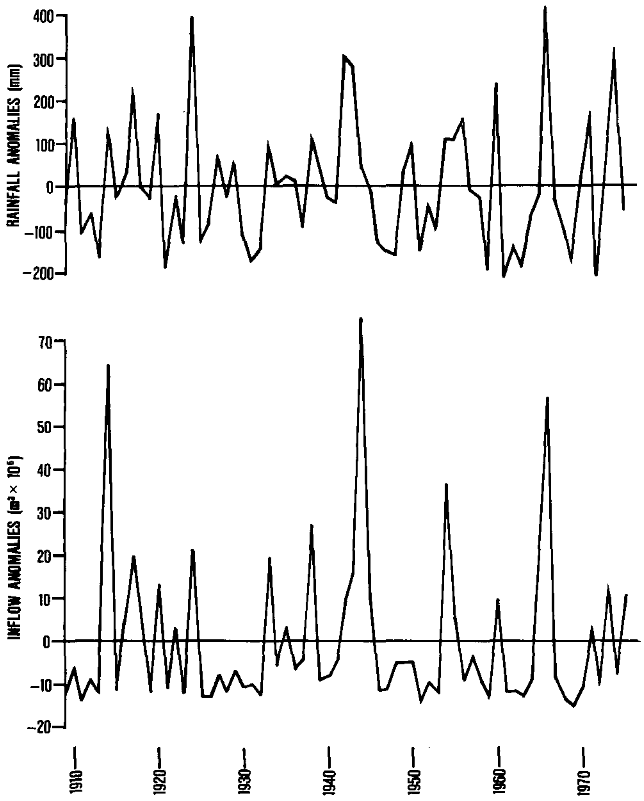

Figure 1
The temporal behaviour of rainfall anomalies (upper) and inflow anomalies (lower) over the period 1909/10 to 1975/6 (Bokaa Dam)

A possible aid to the correct choice of this lag has been described by Dyer (1976c). A further restriction with this approach is that the series of observations needs to be long.

For short series there is the relatively new approach known as the maximum entropy technique (Ulrych and Bishop, 1975; Dyer, 1976c). However, this again requires that an arbitrary choice be made for the filter length used. A method for determining this length has been described by Akaike (1970), but has been shown to be unsuitable for sets of data drawn from geophysical data (Jenkinson, 1977; Dyer, 1977b). Finally, the statistical distributional properties of this technique have not yet been well developed.

Yet another approach to time series analysis is to use harmonic analysis. The resulting periodogram displays the spectral ordinates (variances associated with different frequencies) directly, and is not dependent upon arbitrary choices of parameters. However, it too does have problems relating to the number of spectral ordinates, because these are a fixed fraction of the data length. This problem has been recently solved by Jenkinson (1977). He has derived a method which provides a systematic and quasi-continuous set of spectral ordinates over the frequency range (0—0,5). This enables the correct proportions of the total variance and degrees of freedom of the series to be allocated to any element Δf of the frequency range. Because details of this approach are probably not readily available, we will present a brief outline of this technique.

An arbitrary number (250), which can easily be varied, of wave lengths L_m are given by

$$L_m = 2\exp(\lambda m) \quad \ldots\ldots (1)$$

for m = 0, 1, 2, ,., 249. The parameter λ can be called the tuning, and it determines the fineness of the spacing of the frequencies at which spectral ordinates are calculated. Jenkinson suggests values of $\lambda = 0{,}005$ which gives 250 wave lengths in the range L = 2 to L = 6,95, and $\lambda = 0{,}02$ giving the same number of wave lengths but in the range 2 to 291 units of time.

The periodogram estimates are $A_m \cos(2\pi f_m t - \phi_m)$ where

$$A_m = \left(a_m^2 + b_m^2\right)^{1/2} \quad (2)$$

$$a_m = \frac{2}{n}\sum_{t=1}^{n} y_t C_{mt} \quad (3)$$

$$b_m = \frac{2}{n}\sum_{t=1} y_t S_{mt} \quad (4)$$

$$\phi_m = \tan^{-1}(b_m/a_m) \quad (5)$$

and y_t is the time series elements taken as deviations from the mean. The standardised variance is

$$H_m = A_m^2/V \quad (6)$$

where

$$V = 4\delta^2/n \quad (7)$$

and δ^2 is the variance, or its estimate, of the time series. The degrees of freedom associated with the mth wave are given by

$$\nu_m = 2n\lambda/L_m \quad (8)$$

except at m = 0, when $\nu_0 = n\lambda/2$. The χ^2 contribution for each wave length, the power, is given by

$$\chi_m^2 = \nu_m H_m \quad (9)$$

The percentage of the total variance accounted for by each wave length is $100\,\chi_m^2/n$. The test of significance for the wave band m = j, j + 1, ..., k can be made as a χ^2 test,

$$\sum_j^k \chi_m^2 \text{ with } \sum_j^k \nu_m \quad (10)$$

degrees of freedom.

This approach to time series analysis is very powerful because the width of a wave band can be tested for significance, and the preciseness, or sharpness, of a wave length can also be determined. The wave lengths of oscillations found in such analyses can be used as a means of auto-projecting the time series into the future using such methods as those described by Jenkinson (1977) and Dyer (1977a).

Some results

Rainfall series

Since each spectrum consists of 250 ordinates, it is not practical to portray complete spectra. For the annual rainfall at Gaborone (October to September), only those portions of main interest are presented. In Figure 2, that part of the spectrum is given for wave lengths from 2 to 10 years. Unlike the spatial mean series for the summer rainfall region as a whole, there is no evidence of oscillations with long wave lengths (Dyer, 1976a).

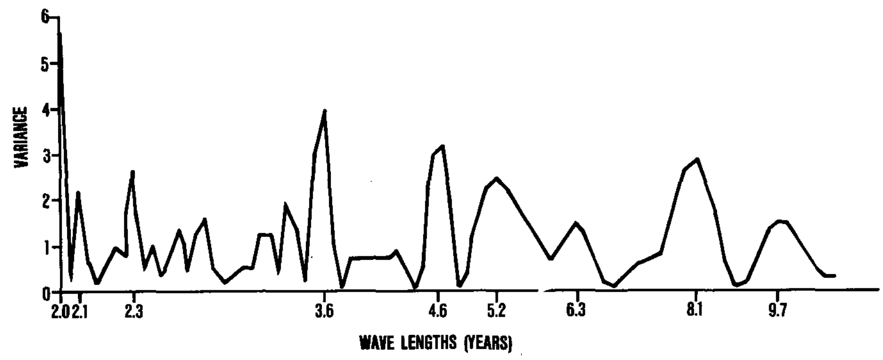

Figure 2
The variance spectrum for rainfall over the wavelength interval 2,0 to 10,0 years

The large peak at 2 years is difficult to interpret because it can be either a peak in its own right, or a result of aliasing. This latter phenomenon is a folding back of variance present over those wave lengths for which the spectrum is undefined, namely 0 $\leqslant$ wave lengths < 2 units of time. The problem can be removed by filtering the data before the spectral analysis is carried out. However, even the lightest smoothing would attenuate some spectral peaks of interest. The 2,35 year or biennial oscillation, and the 3,6 year one, both of which have been found elsewhere over South Africa and further afield, would be lost. A different solution is provided by increasing the sampling rate, that is, decreasing the sampling interval. Fortunately monthly rainfall totals are available and these were submitted to spectral analysis. It is not necessary to burden the reader with a detailed account of this spectrum. Suffice it to say that although the 2 year wave was in evidence, there was the annual

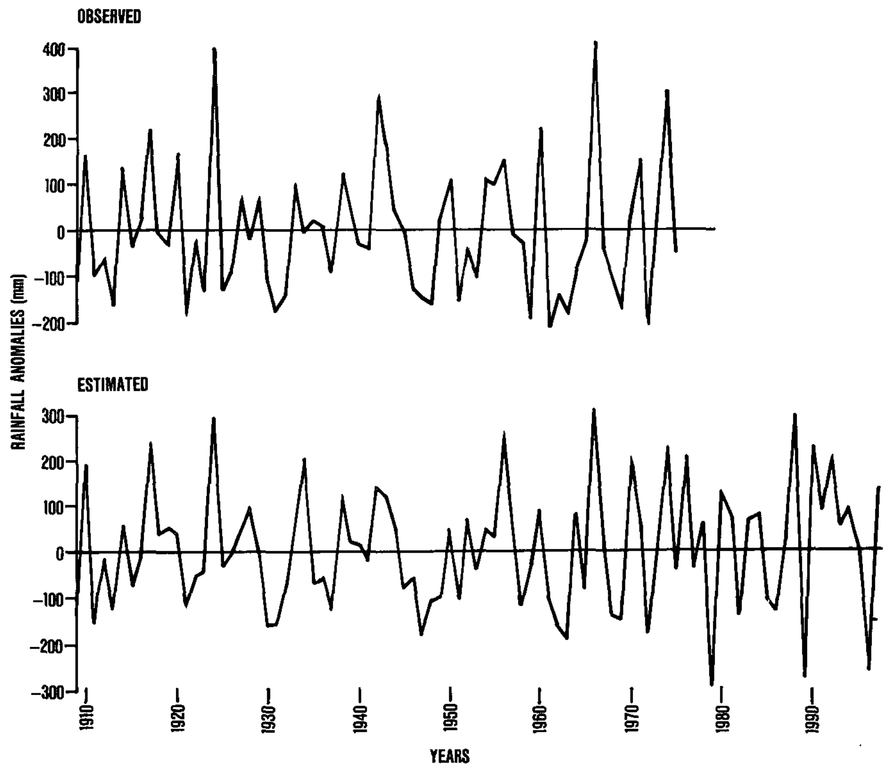

Figure 3
A trigonometric regression fitted to rainfall anomalies (upper) to give the estimated values (lower) together with some forecasts for future years

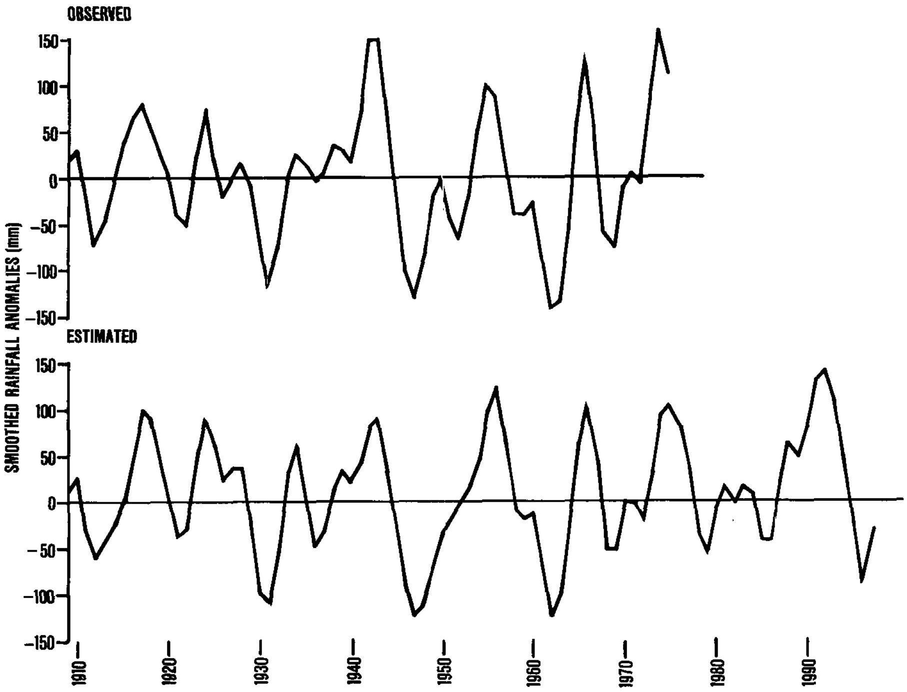

Figure 4
The effect of smoothing rainfall (upper) and estimated rainfall (lower) with a five term binomial filter. These curves provide an indication of likely periods of positive and negative anomalies

wave and harmonics thereof. The annual variation was represented at a wave length of 0,94 years and this is due to the asymmetry of the rainfall year over time.

The wave lengths of the oscillations obtained from the spectra were used as first estimates in an iterative procedure for fitting a trigonometric regression model to the rainfall annual totals. This model may be defined as

$$R_t = \sum_{m=1}^{p} \left(a_m \cos \frac{2\pi t}{\lambda m} + b_m \sin \frac{2\pi t}{\lambda m} \right) + \epsilon_t \qquad (11)$$

where $0 \leqslant t \leqslant n - 1$, p is the number of waves fitted, λ_m the wave length of the m^{th} wave, and a_m and b_m the k^{th} regression coefficients, ϵ_t is the error variable. No constant term is employed in equation (11) because rainfall deviations from either the mean or, if appropriate, a linear trend straight line are used.

The observed and estimated rainfall totals together with future estimates up to the late 1990's are shown in Figure 3. All the component waves included in the regression, and there are 11 of them, make a significant contribution to the sum of squares accounted for by the equation at the 95% level. The overall sum of squares accounted for by the regression equation is high, having a value of 71 percent of the total. It will be seen that there is good agreement between the observed and estimated plots.

It may, however, be of more practical value to possess information relating to periods of positive and negative rainfall anomalies. Therefore, both the observed and estimated rainfall values were lightly filtered, using a 5-term binomial filter. This operation was carried out after the estimated values had been obtained from the regression equation. That is, unsmoothed data were used for the actual fitting process. The results of this

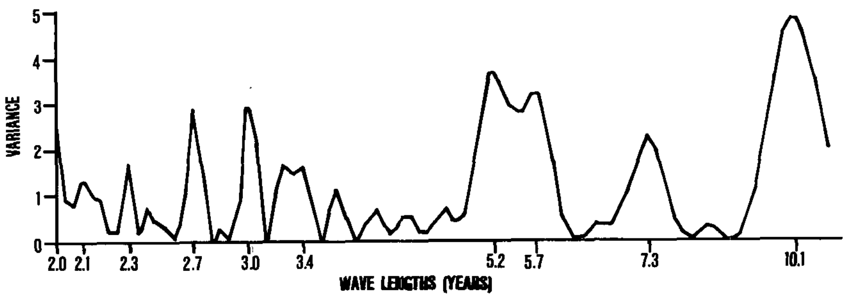

Figure 5
The variance spectrum for inflow over the wavelength interval 2,0 to 10,0 years

exercise are presented in Figure 4, where it will be seen that the two curves are in good agreement. But how is a smoothed curve interpreted? In the present case, consider the peak at 1991. This maximum is the result of combining the five rainfall totals of which 1991 is the middle one. That is, the weighted linear combination of rainfall totals for the years 1989 to 1993 inclusive result in a large positive anomaly centred at 1991. In other instances one or more of the individual five years could have negative anomalies but still result in an overall positive one for the complete interval, this will depend upon the magnitude of the anomalies. Notice that the deep trough or negative anomaly at 1988 in Figure 3 does not show up in the smoothed curve given in Figure 4.

Inflow series to dam

In Figure 5 part of the spectrum of the time series of annual inflow totals to the dam is portrayed. A notable feature of this plot is the strength of the 10 year wave. It can be observed quite easily in Figure 1. Without further study, it would not be prudent to assume that this, and possibly the 9,7 year wave in rainfall, has any relationship with the sun-spot cycle. Again, wave lengths of long period are not present in this series. Both the rainfall and inflow spectra have much in common, and one can see this from an inspection of Figures 2 and 5.

Equation (11) was fitted to the raw inflow data and the result is shown in Figure 6 together with estimated future values. The observed and estimated series are in good agreement and the fit accounts for 87 percent of the variation in the inflow data.

When these two series are smoothed using the binomial filter, the graphs in Figure 7 result. From this figure it is possible to view intervals of overall positive and negative anomalies in the pattern of inflow to the dam.

Concluding remarks

From an inspection of graphs of rainfall totals (hydrological year) and inflow to a dam it seemed that their patterns over time were, at least, pseudo-periodic. This provided justification for submitting the data to spectral analysis in order to determine the wave lengths of any oscillations that might be present. The two spectra exhibited peaks in variance many of which have been found not only in other parts of South Africa, but

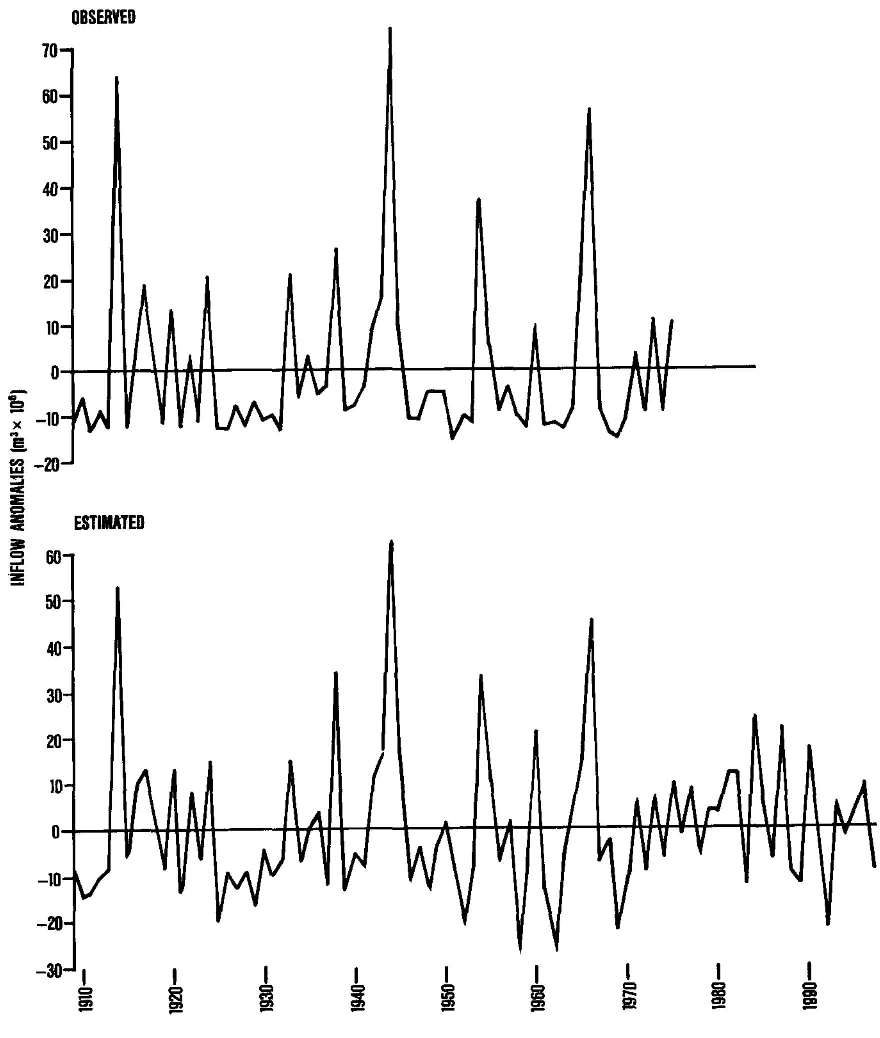

Figure 6
A trigonometric regression fitted to inflow anomalies (upper) to give the estimated values (lower) together with some forecasts for future years

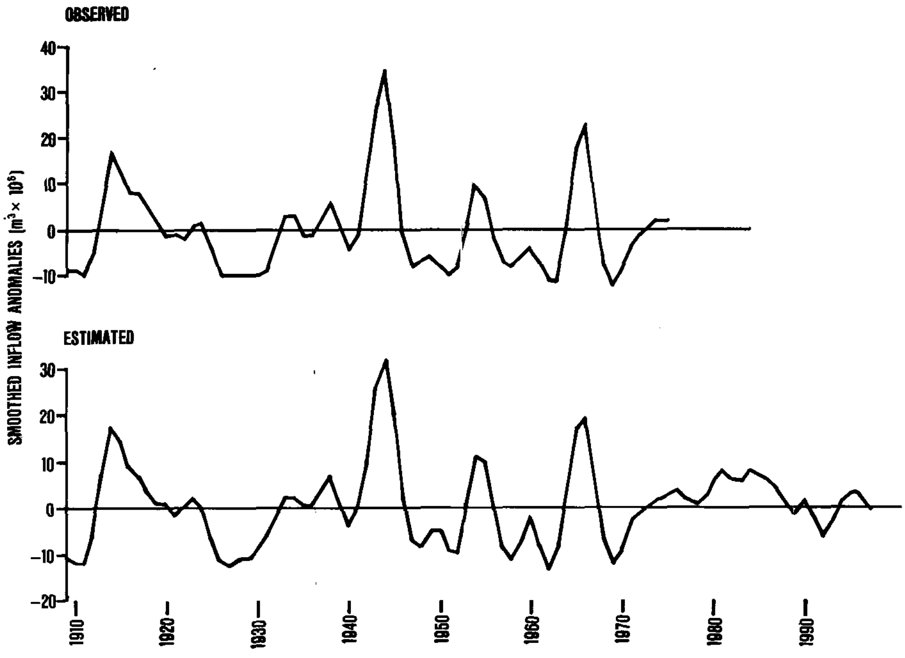

Figure 7
The effect of smoothing inflow (upper) and estimated inflow (lower) with a five term binomial filter. These provide an indication of likely periods of positive and negative anomalies

wider afield as well. However, these spectra have some features not in common in other spectra for rainfall over much of South Africa (Dyer, 1977a). It was noted that oscillations with long wave lengths were conspicuous by their absence. This may be because the site is positioned where the influence of the 16—20 year oscillation on rainfall has faded in a spatial sense. Alternatively, it could be because in previously published work, calendar rather than hydrological years have been used. This is a subject for future investigation which has already been commenced.

The wave lengths obtained from the spectra were then used as initial estimates in an iterative procedure to fit a trigonometric regression to the two series. The fits were in both cases very good as measured by the percentage variance accounted for by the equations. Most of the unaccounted for variance was associated with extreme values in the data, i.e. the residuals tended to be larger for these data points. Some data transformations were used in an attempt to resolve this but they provided little improvement. Ideally, one would have liked to have saved some data to investigate how well the models simulated future estimates. However, in the present study, there were too few data points available to make this a very meaningful exercise. Finally, the estimated series were smoothed using a five term binomial filter. These graphs enable one to gain an idea of runs of above and below average rainfall and inflow. These latter graphs probably are of greater practical value than the unfiltered series. It is suggested that the Bokaa dam should be completed in time for the expected period of positive rainfall anomalies over the interval 1980/90.

Acknowledgements

Mr B.G.A. Lund for useful discussion and data provision. From 1959 credit to the Principal Hydrological Engineer, Department of Water Affairs, Republic of Botswana for data. Prior to 1959 data from Dr W. Pitman, Hydrological Research Unit, at the University of the Witwatersrand. To Mr A.F. Jenkinson, U.K. Meteorological Office for provision of computer programme for his spectral technique.

References

AKAIKE, H. (1970) Statistical predictor identification. *Ann. Inst. Statist. Math.* **22** 203.

DYER, T.G.J. (1975) The assignment of rainfall stations into homogeneous groups: an application of principal component analysis. *Quart. J. Roy. Met. Soc.* **101** 1005.

DYER, T.G.J. (1976a) Meridional interactions between rainfall and surface pressure. *Nature* **264** 48.

DYER, T.G.J. (1976b) Expected future rainfall over selected parts of South Africa. *S. Afr. J. Sci.* **72** 237.

DYER, T.G.J. (1976c) An analysis of Manley's central England temperature data: I. *Quart. J. Roy. Met. Soc.* **102** 871.

DYER, T.G.J. (1977a) On the application of some stochastic models to precipitation forecasting. *Quart. J. Roy. Met. Soc.* 103.

DYER, T.G.J. (1977b) An analysis of Manley's data: I; a comment. *Quart. J. Roy. Met. Soc.* **103** 817.

DYER, T.G.J. and TYSON, P.D. (1977) Extended forecasts of above and below normal rainfall over South Africa. *Jour. Appl. Meteorol.* **16** 145.

JENKINSON, A.F. (1977) A powerful elementary method of spectral analysis for use with monthly, seasonal or annual meteorological time series. *Met. 0 13 Branch Memorandum* No. 57, U.K. Met. Office, Bracknell.

TYSON, P.D., DYER, T.G.J. and MAMETSE, M. (1975) Secular changes in South African rainfall: 1880—1972. *Quart. J. Roy. Met. Soc.* **101** 817.

ULRYCH, T.J. and BISHOP, T.N. (1975) Maximum entropy spectral analysis and autoregressive decomposition. *Rev. Geophys. and Sp. Phys.* **13** 183.