INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **20**: 211–228 (2000)

CYCLICAL COMPONENTS OF LOCAL RAINFALL DATA

R.P. MENTZ*, M.A. D'URSO, N.M. JARMA and G.B. MENTZ
Universidad Nacional de Tucuman and CONICET, Facultad de Ciencias Economicas, Casilla de Correo 209, 4000 Tucuman, Argentina

Received 9 September 1998
Revised 24 April 1999
Accepted 9 May 1999

ABSTRACT

This paper reports on the use of a comparatively simple statistical methodology to study local short time series rainfall data. The objective is to help in agricultural planning, by diminishing the risks associated with some uncertainties affecting this business activity.

The analysis starts by assuming a model of unobservable components, trend, cycle, seasonal and irregular, that is well known in many areas of application. When series are in the realm of business and economics, the statistical methods popularized by the US Census Bureau and US National Bureau of Economic Research are used for seasonal and cyclical estimation, respectively. The flexibility of these methods makes them good candidates to be applied in the meteorological context, and this is done in this paper for a selection of monthly rainfall time series.

Use of the results to help in analysing and forecasting cyclical components is emphasized. The results are interesting. An agricultural entrepreneur, or a group of them located in a single geographical region, will profit by systematically collecting information (monthly in our work) about rainfall, and adopting the scheme of analysis described in this paper.

KEY WORDS: rainfall in Argentina; unobservable components; X-12-ARIMA seasonal procedure; US National Bureau of Economic Research; dating cyclical turning points

1. INTRODUCTION

1.1. General

The purpose of this paper is to report on the use of a simple statistical methodology to describe local records of rainfall data coming in short series. The aim is to facilitate agricultural planning by helping to reduce some of the risks associated with the activity.

The records are local since it is desirable to differentiate among geographical locations which are comparatively close to each other in space, in so far as they differ in their climatic experiences. The series to be studied are short, in the order of 15 or 20 years of monthly data, which are shorter than those used in other types of meteorological studies.

The statistical methods that we use arose historically in the treatment of empirical time series. For them, it was found useful to propose models of components associated with distinctive systems of causes: trend, cycles, seasonal and irregular. It is postulated that these unobservable components are combined to form an observable time series, in a multiplicative way, $T \cdot C \cdot S \cdot I$, or in an additive way, $T' + C' + S' + I'$. Other types of combinations have also been considered in the literature. See, for example, Nerlove *et al.* (1979) for an interesting review, that includes early references to applications in astronomy, meteorology, the study of diseases, and others.

Some authors have studied the presence of long-term trends in rainfall data, and tried to relate this to more general issues of global climate change; see, for example, the analysis of annual data in Karl *et al.*

* Correspondence to: Universidad Nacional de Tucuman and CONICET, Facultad de Ciencias Economicas, Casilla de Correo 209, 4000 Tucuman, Argentina.

(1996). However, for rainfall series as short as those considered here, we can safely disregard the effect of trend components. Our rainfall series have large seasonal and random components, and cyclical components. The seasonal component is large, and the variability of the data is different in the different seasons, in the extreme case the dry season contains zeros.

Simple statistical methods to study these components are, for example, moving averages to estimate a trend component, and the so-called 'ratio (or difference) to moving average' to estimate a seasonal component. For series arising in the economics and business areas, these ideas led to the development in the USA of computer programs identified as the X-11 family, which have been used for many years, in various countries and by numerous researchers.

In a manner that parallels these developments, in the US National Bureau of Economic Research (NBER) a procedure was developed and made computationally available, to assign dates to business cycles. It is also true that these methods have been in use for many years, and a considerable amount of experience has been accumulated.

In this paper we analyse monthly rainfall series by using modern versions of the X-11 seasonal procedure and of the NBER business cycle analysis program. One main purpose of our study is to explore how useful these methods are to deal with time series other than those of the economic area, for which they were originally developed. Our basic argument is that for rainfall data (as well as for other meteorological variables), the indicated decomposition has full meaning, and that flexible methods can work reasonably well, in spite of the different nature of the component sources.

We use the seasonal and cyclical procedures in a descriptive way, in the sense that no attribution of physical causes is attempted.

1.2. Some alternative statistical approaches

In the analyses we present in the following sections we show that our choice of statistical procedures is adequate for our purposes. Several other statistical methods have been used in the literature, to deal with various aspects of the description and analysis of rainfall and similar types of series. It is useful to keep in mind two dichotomies: (a) deterministic versus stochastic components or models; (b) time versus frequency domain analyses. Note that those we discuss are not causal procedures; as such, they will differ from causal models developed by meteorologists.

To start our presentation, let us consider a very simple deterministic regression model for monthly rainfall data, formed by the addition of the following components: (i) trend, modelled by a low degree polynomial; (ii) cycle, modelled by a sinusoidal term with a given period; (iii) a stable seasonal pattern represented by 12 constants; and (iv) random independent error terms. This seasonal component can also be written as the sum of sine and cosine functions. In general this approach will prove to be too rigid. Variants to allow for more flexibility include using other smoothing functions of time to model trend, a sum of sinusoidal terms to model the presence of more than one cycle, and time trending seasonal constants.

Non-deterministic trends can be dealt with by smoothing or filtering procedures, one of the most widely used being moving averages, which can also be used in cyclical analysis. A stochastic formulation that is frequently used is that of seasonal autoregressive integrated moving average (ARIMA) models, as described in Box *et al.* (1994); they will take into account trends and seasonal effects. ARMA models can be used in the regression context to give more flexibility to the error terms.

The structure of an observed time series can often be seen more clearly in the frequency domain, by studying the estimated spectral density of a series, or of some of its transformations. There are several statistical procedures to obtain these estimates. The analysis of seasonality and the effect of seasonal adjustments can be evaluated in the frequency domain: spectral density estimates show peaks clearly at seasonal frequencies, even in series with only moderately regular seasonal components (Nerlove, 1964).

Combinations of some of these elements have been used in the literature to study monthly series of rainfall or similar questions. Kashyap and Ramachandra Rao (1976) considered 'the class of autoregressive models with deterministic sinusoidal terms added to them' (p. 249) to model monthly river flows.

They compare their forecasting performances with ARMA and seasonal ARIMA models. These authors present other case studies, some related to rainfall, and compare results of the analysis of daily, monthly and yearly data. However, they do not emphasize cyclical analysis. Percival and Walden (1993) also study river flows among other series, and emphasize the use of spectral methods, including tests of the significance of periodicities. Burroughs (1992) considers studies of cycles based on monthly data, by 'using both spectral analysis and filtering techniques' (p. 39); the results he studies are directly relevant to our purposes, as will be commented on in Section 6.3.

There are many studies of trends and cycles based on annual data, that therefore do not have to deal with the intra-annual seasonal variation. They use several of the types of models and methods we discussed in the section. In general they search for cycles longer than those treated in our work.

Two useful sources of information are the book by Murphy and Katz (1985) and also the article by Gani (1985).

2. STATISTICAL METHODS

2.1. Seasonal analysis

To perform a seasonal analysis of a given series, we used the X-12-ARIMA (Beta Version 1.0) program of the US Bureau of the Census, made available in 1996 on the Internet (see also Findley *et al.*, 1998). This program uses an iterative semi-automatic procedure that works in stages. In the first stage, an ARIMA model can be attempted for the series, so that if a reasonably good fit is achieved, 1 year of predicted observations are added at one or both ends of the series. Stages two to four consist of the following: a 'trend–cycle' component (trend and cycle are not separated in these programs) is estimated by means of moving averages which are centred at a given observation for central data; this estimate is used to define a detrended series, and this in turn to estimate a seasonal component responding to an additive or multiplicative model, which must be specified by the user of the program. The estimated seasonal component is used to define a seasonally adjusted series. In the second and third stages, the seasonally adjusted series is the input of the following stage, and leads to new estimates of trend–cycle and seasonal with procedures similar to those of the preceding stage. In the fourth and last stage, the program uses the output of the third stage to automatically obtain the final estimates of trend–cycle and seasonal components. In the third and fourth stages, trend–cycle moving averages can be of lengths 9, 13 or 23, a choice is made automatically by the program, but the user can alter this choice.

Program X-12-ARIMA incorporates the so-called 'sliding spans' methodology, which is a diagnostic procedure that compares seasonal estimates obtained from partially overlapping subseries of the given series, and aims at evaluating the reliability of the final estimates by computing and analysing a set of statistical measures.

The complete output of the X-12-ARIMA program for monthly series, contains more than 50 tables, each with $12n$ figures, where n is the number of years that are processed, plus a set of graphs and a set of additional statistical measures. Some of the main results appear in the following tables: B1, original series; D10, final seasonal factors; D11, final seasonally adjusted series; D12, final trend–cycle; D13, final irregular series; F, summary measures; S, sliding spans. It is very convenient for the user that the program allows that the output tables be transferred to independent archives, so that other tables or graphs can be developed using any kind of software.

Processing a series with the X-12-ARIMA program leads to a set of summary measures (table F) to be used to evaluate the quality of the seasonal fit. There are 11 quality control measures, the M-statistics, with values between 0 and 3, and acceptance regions between 0 and 1. Each statistic evaluates a certain characteristic of the seasonal analysis, and gives the user a measure of success along this direction. All measures are combined to define a Q-statistic, a global measure of the quality achieved by the program with the given series.

The X-11 method was described in detail in Shiskin *et al.* (1967). The ARIMA variant of X-11 is described, for example, in Dagum (1983), and the sliding spans methodology in Findley *et al.* (1989). Useful information about seasonal (and cyclical) adjustment procedures can be found in Mentz *et al.* (1989), and in the recent paper with discussion (Ghysels *et al.*, 1996). For a recent detailed description of the final version of the new X-12-ARIMA package, see Findley *et al.* (1998).

The monthly rainfall series used in our study have some values equal to zero, which correspond to months without rain. (However, they do not have long strings of zero values, as they might have in a very dry region.) Under these circumstances, the computer program automatically chooses an additive model. Since the multiplicative option tends to be more popular, we tried it with data transformed to avoid the zero values, by adding a constant. However, the results were poor and are not included in this report.

An analysis of other common transformations was also performed. Since logarithms are not feasible due to the zero values, power transformations were explored. Powers 1/4, 1/2 and 3/4 of the observations were considered, and some results are given in Table I. It is apparent that not much is gained in general by using these transformations. Also, graphs of the estimated seasonal components were studied. The conclusion is that the original observations could be retained without much loss.

One reason for our preoccupation with the quality of the seasonal fit, is that it may affect the cyclical analysis that is carried out with deseasonalized observations. Our conclusion is that if an X-11 type of seasonal adjustment is used, not much is expected to be gained by standard data transformations.

2.2. Cyclical analysis

Once the seasonal analysis is completed, series D11 and D12 are used to perform cyclical analyses, for which we used the program 'Turning Points Determination' made available by the Center for International Business Cycle Research (CIBCR) of Columbia University. It is based on the NBER methodology. This program identifies the maxima and minima of a seasonally adjusted series, and assigns dates to them. In the standard or 'classical' business cycle analysis of economic time series (the option being the growth cycle analysis), this operation is performed with series D11, final seasonally adjusted series; with our rainfall data we tend to rely more on the dates coming from the D12 series, which is 'irregular free'.

Cycles identified by these procedures can be qualified as short cycles, as compared with other lengths considered in the literature for meteorological data. On the other hand, oscillations shorter than 15 months are deemed irrelevant, and are not included into the dated cycles. For a description of these procedures, see, for example, Bry and Boschan (1971) and for applications Mentz *et al.* (1989). Considerations about the duration of the rainfall cycles will be made in Section 6.

After completing the cyclical analysis, the main results appear in tables, some of which facilitate comparisons among locations.

3. DATA USED IN THE STUDY

Our initial motivation for this study was to generate information useful for some local agricultural firms, and this explains the nature of the available data. We have rainfall series for four locations in the State of Tucuman, Argentina, three in the central-northern part of the state, chosen with reference to the Agricultural Experimental Station providing the information, one (Las Cejas) is usually regarded as drier than the other (La Ramada). The fourth location in Tucuman (La Cocha) is in the southern-most part of the state. For purposes of comparison, data from two other states are considered, Santa Fe and La Pampa, which are part of the traditional Argentine prairies.

Table I. Seasonal analysis, rainfall data, additive model, X-12-ARIMA: figures of merit

Locations	Period	Transformation	*F*-statistics analysis			Month of cyclical dominance	M-statistics analysis				Sliding spans analysis		
			Stable seasonality	Moving seasonality	Seasonal fit is:		Number of M > 1	Detail	*Q*-statistic	Seasonal fit is:	*S* (%)	MM (%)	Seasonal fit is:
Agricultural Experimental Station Obispo Colombres (Tucuman)	1975–1995	Original	69.36	1.28	Stable 0.1%	12	1	M5	0.71	Accepted	91.7	98.2	Unlikely
	1975–1995	1/2 power	71.90	0.99	Stable 0.1%	12	2	M3, M5	0.91	Cond. accep.	52.8	77.8	Unlikely
	1975–1995	1/4 power	71.90	0.99	Stable 0.1%	12	2	M3, M5	0.91	Cond. accep.	55.6	77.8	Unlikely
	1975–1995	3/4 power	63.83	1.06	Stable 0.1%	12	2	M3, M5	0.94	Cond. accep.	92.6	87	Unlikely
La Ramada (Tucuman)	1975–1989	Original	39.02	1.25	Stable 0.1%	12	4	M1, M2, M3, M5	1.01	Cond. reject.	91.7	84.7	Unlikely
	1975–1989	1/2 power	57.44	0 88	Stable 0.1%	12	4	M1, M2, M3, M5	1.05	Cond. reject.	77.1	95.1	Unlikely
	1975–1989	1/4 power	57.44	0.88	Stable 0.1%	12	4	M1, M2, M3, M5	1.05	Cond. reject.	77.1	95.1	Unlikely
	1975–1989	3/4 power	49.78	1.05	Stable 0.1%	12	4	M1, M2, M3, M5	1.01	Cond. reject.	96.5	88.9	Unlikely
Las Cejas (Tucuman)	1975–1989	Original	36.54	2.24	Stable 0.1%	12	2	M3, M5	0.83	Cond. accep.	90.7	83.3	Unlikely
	1975–1989	1/2 power	52.36	1.89	Stable 0.1%	12	5	M1, M2, M3, M5, M6	0.99	Cond. reject.	71.1	86.8	Unlikely
	1975–1989	1/4 power	57.44	0.88	Stable 0.1%	12	4	M1, M2, M3, M5	1.05	Cond. reject.	75.7	95.1	Unlikely
	1975–1989	3/4 power	42.48	2.31	Stable 0.1%	12	5	M1, M2, M3, M5, M6	0.99	Cond. reject.	96.5	82.6	Unlikely
Agricultural Experimental Station Obispo Colombres (Tucuman)	1982–1995	Original	39.67	0.78	Stable 0.1%	12	1	M5	0.71	Accepted	92.6	89.8	Unlikely
	1982–1995	1/2 power	51.62	0.32	Stable 0.1%	12	2	M1, M3	0.93	Cond. accep.	77.1	77.8	Unlikely
	1982–1995	1/4 power	51.47	0.33	Stable 0.1%	12	2	M3, M5	0.99	Cond. accep.	62.0	65.7	Unlikely
	1982–1995	3/4 power	43.55	0.54	Stable 0.1%	12	2	M3, M5	0.96	Cond. accep.	92.6	87.0	Unlikely
La Cocha (Tucuman)	1982–1995	Original	13.18	0.98	Stable 0.1%	12	4	M1, M2, M3, M5	1.23	Rejected	95.8	95.1	Unlikely
	1982–1995	1/2 power	23.22	1.06	Stable 0.1%	12	6	M1, M2, M3, M5, M10, M11	1.25	Rejected	90.7	94.4	Unlikely
	1982–1995	1/4 power	20.11	1.89	Stable 0.1%	12	6	M1, M2, M3, M5, M10, M11	1.23	Rejected	62.0	91.7	Unlikely
	1982–1995	3/4 power	19.82	9.6	Stable 0.1%	12	4	M1, M2, M3, M5	1.31	Rejected	97.9	93.8	Unlikely
Santa Fe	1975–1989	Original	7.18	1.01	Stable 0.1%	12	7	M1, M2, M3, M5, M8, M10, M11	1.50	Rejected	94.4	95.4	Unlikely
	1975–1989	1/2 power	9.70	1.01	Stable 0.1%	12	7	M1, M2, M3, M5, M8, M10, M11	1.52	Rejected	86.1	94.4	Unlikely
	1975–1989	1/4 power	9.01	1.16	Stable 0.1%	12	7	M1, M2, M3, M5, M8, M10, M11	1.35	Rejected	21.3	88.9	Unlikely
	1975–1989	3/4 power	8.53	0.95	Stable 0.1%	12	7	M1, M2, M3, M5, M8, M10, M11	1.49	Rejected	94.4	95.4	Unlikely
La Pampa	1975–1989	Original	13.47	0.71	Stable 0.1%	12	4	M1, M2, M3, M5	1.28	Rejected	95.1	92.4	Unlikely
	1975–1989	1/2 power	16.34	0.65	Stable 0.1%	12	4	M1, M2, M3, M5	1.30	Rejected	83.3	93.8	Unlikely
	1975–1989	1/4 power	12.64	0.78	Stable 0.1%	12	4	M1, M2, M3, M5	1.37	Rejected	10.4	84.0	Unlikely
	1975–1989	3/4 power	15.42	0.66	Stable 0.1%	12	4	M1, M2, M3, M5	1.30	Rejected	98.6	94.4	Unlikely

A summary of the available information is as follows:

State	Location	Period	Average monthly rainfall (mm)
Tucuman	Agricultural Experimental Station (EEAOC)	1975–1995	93.2
	Agricultural Experimental Station (EEAOC)	1982–1995	88.8
	La Ramada	1975–1989	96.0
	Las Cejas	1975–1989	69.5
	La Cocha	1982–1995	81.1
Santa Fe		1975–1989	74.5
La Pampa		1975–1989	70.5

The first series has 21 years of monthly data, while the remaining series are 14- or 15-years long. The second series is just a segment of the first one.

4. MAIN FINDINGS ON SEASONAL AND CYCLICAL ANALYSES

1. Results of the exploration of certain power transformations under the additive version of the X-12-ARIMA seasonal program, are given in Table I. For each series, the results of processing the original and the power transformations for 1/2, 1/4 and 3/4 are compared. This analysis leads us to retain the original data for analysis, since differences are small.

2. As indicated in Section 1.1, rainfall data are known to be highly variable, so that in our terms, controlling or eliminating the irregular component is expected to be a source of difficulty. Our results, in terms of the goodness of fit statistics, confirm this general idea.

3. About the goodness of fit statistics mentioned in Section 2, high values for the F-statistic for the presence of stable seasonality, and low values for the F-statistic for the presence of moving seasonality, are an indication of a strong and comparatively regular seasonal component. Hence, the part of Table I containing these F-statistics shows in general that our series are satisfactory in this respect.

An often-used indicator of the quality of a seasonal adjustment of a series, is the so-called 'month of cyclical dominance' (MCD), which comes from relating the size of the irregular to that of the trend–cycle component: a large MCD means that the movements of the irregular component are large enough to make difficult a reliable estimation of trend–cycle. In our work we often find values higher than 6 for MCD, when small values are 1 or 2, with 12 being the largest possible. In Table I, all reported MCDs are equal to 12, which is the least favourable value.

We note that the number of M-statistics that exceed the threshold value of 1 varies from 1 to 7. The summary measure Q, a weighted average of the 11 M-statistics, has values either smaller than 1 or slightly larger. Further, the M7 statistic is never found to be larger than 1: 'Values of M7 larger than 1 are usually interpreted as evidence that the seasonal component is evolving too rapidly for the X-11 procedure' (Findley, 1996).

Finally, Table I shows the measures associated with the sliding spans, which indicate uniformly that the seasonal fits are considered poor. This points to difficulties in trying to predict or extrapolate seasonal effects.

4. In view of the preceding comments, we processed with the turning points program, not only the seasonally adjusted series (D11), but also the final trend–cycle series (D12) in which the irregular has been smoothed out.

5. A graphical analysis associated with the first (and longest) series mentioned in Section 3 is incorporated as an example in Figures 1–4. Graphs are, in terms of the definitions introduced in Section 2, respectively, those of serie*s* B1, D10, D11, and D12 with its dated turning points and superimposed average (mean) line.

6. For the same series, Table II summarizes the distances among dates presented in the graph of series D12. The mean, median and standard deviation of the cyclical durations are presented. Cycle durations are measured peak–trough–peak and trough–peak–trough, which in cases of asymmetries need not coincide.

7. Table III facilitates the comparison among the available series, of the average durations of cycles estimated in the indicated way.

8. For the series reported in Table II, Table IV summarizes the information in the graphs of series D12 in a different way: cycles are now defined as comprising consecutive periods of permanence of the series over and under the average line, which is shown in all graphs. It is thought that this is a manner reasonably easy to understand, to exhibit the results of the analysis.

9. Besides the preceding analysis, there is a good deal of information in the graph of series D12. For example, comparing the graphs for La Cocha (Tucuman) in Figure 5, where the risks associated with the behaviour of rainfall are well known, with La Pampa in Figure 6, part of the traditional Argentine

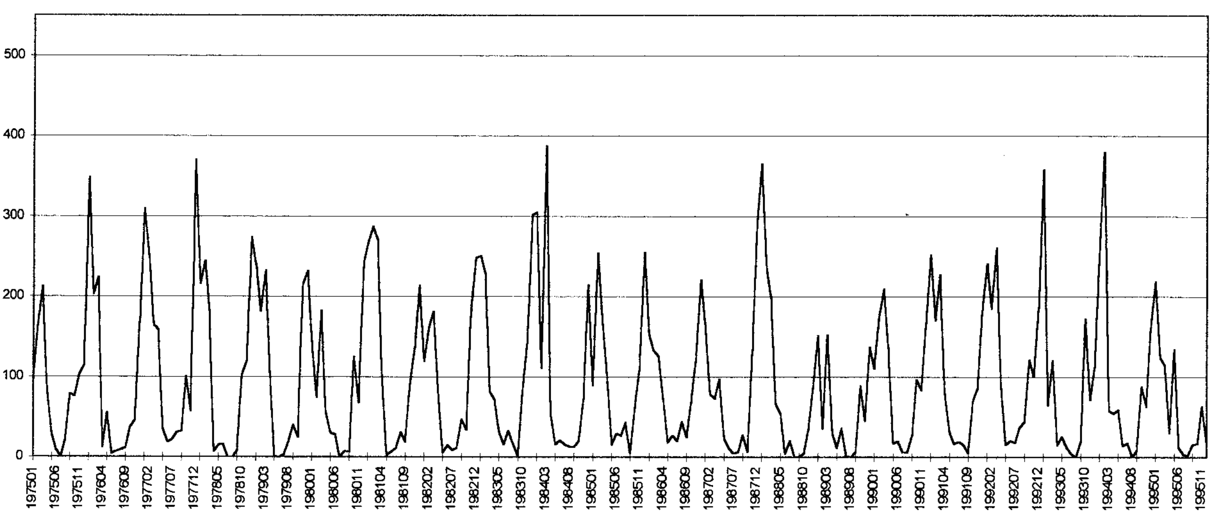

Figure 1. Monthly rainfall in Agricultural Experimental Station (Tucuman). January 1975–December 1995. Original series

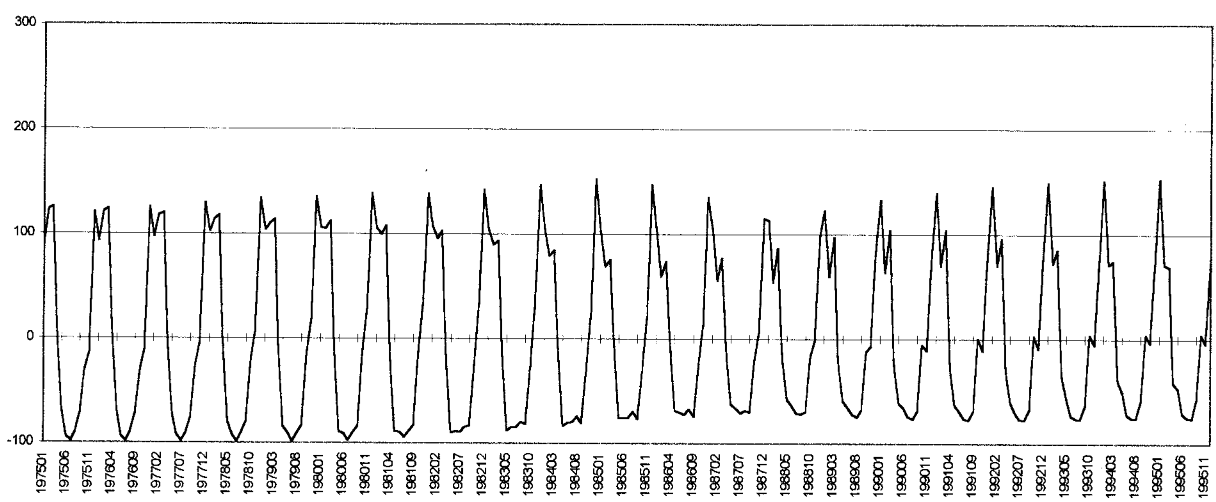

Figure 2. Monthly rainfall in Agricultural Experimental Station (Tucuman). January 1975–December 1995. Seasonal factor, additive model, X-12-ARIMA

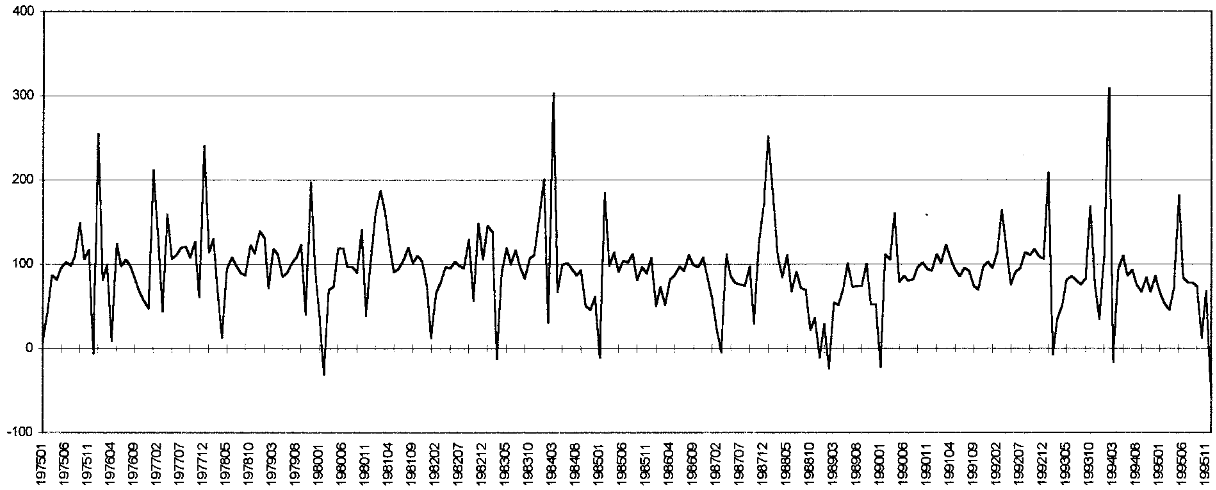

Figure 3. Monthly rainfall in Agricultural Experimental Station (Tucuman). January 1975–December 1995. Seasonal adjusted series, additive model, X-12-ARIMA

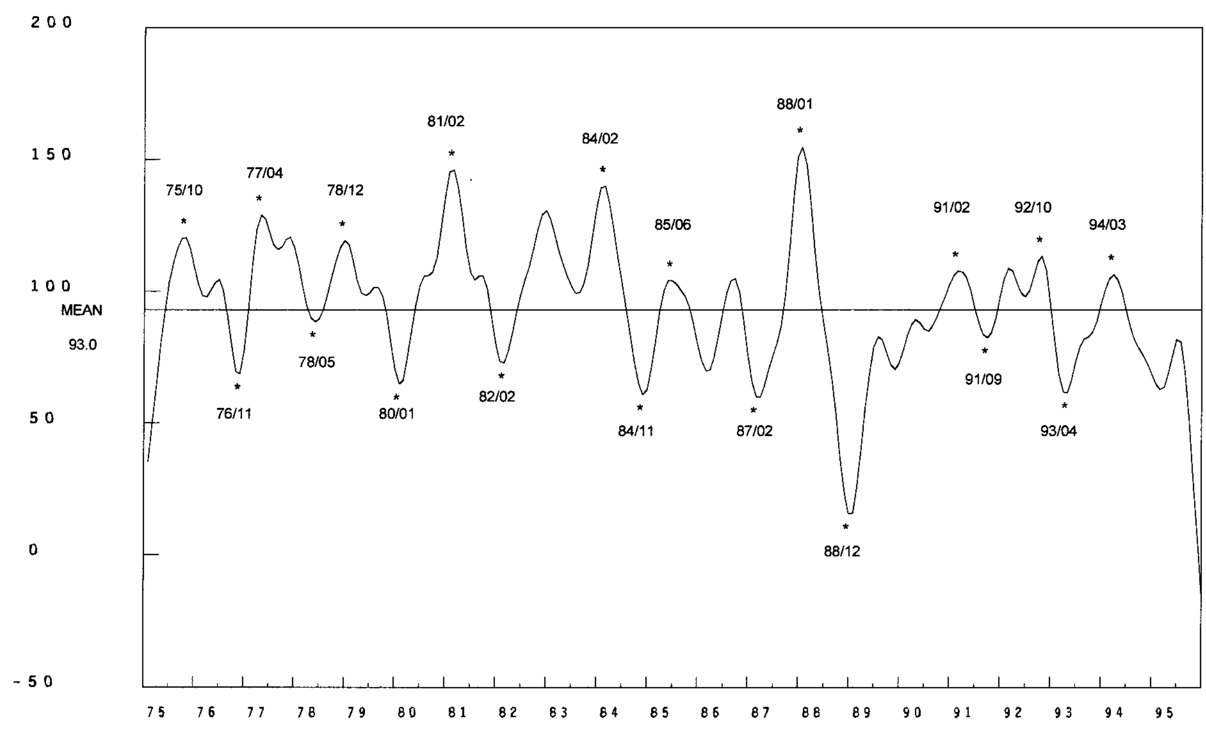

Figure 4. Monthly rainfall in Agricultural Experimental Station (Tucuman). January 1975–December 1995 trend–cycle component, additive model, X-12-ARIMA, dated by CIBCR procedure, and superimposed mean line

prairies, we observe that the difference is not so much in the duration of the observed cycles as in the amplitude of the movements of the series D12. Since the two graphs are drawn in the same scale, we conclude that the risks associated with the cyclical behaviour of rainfall are much higher in La Cocha than in La Pampa, in agreement with experience. In fact, La Pampa is, in a large proportion, a part of an important area known precisely as 'la pampa húmeda' (humid prairies).

5. FORECASTING

In trying to forecast one of our rainfall series, attention must be paid to its seasonal, trend and cyclical components. The estimated seasonal components comprise table D10 of the output in the X-12-ARIMA program, and appears in Figure 2 for one of the series. The corresponding analysis was presented in Section 4. The X-12-ARIMA program also provides a 'year ahead' set of 12 estimates of the seasonal component, that could be used in forecasting. In our case we recall a warning coming from the sliding spans measures of fit (see Section 3).

More interesting is trying to forecast the other components. Trends and cycles are treated jointly in the X-12-ARIMA program as the 'trend–cycle' component. The estimates appear in table D12, and are shown in Figures 4–6 and 11 for various series. We recall that we disregard the contribution of trend, in view of the limited lengths of our series. Hence, we take the indicated estimates as representing cyclical components only. We aim at forecasting these by extending or extrapolating the corresponding series.

Table II. Monthly rainfall in Agricultural Experimental Station Obispo Colombres (Tucuman), January 1975–December 1995. Final turning points of trend–cycle component, additive model, X-12 ARIMA dated by CIBCR procedure

(a) Peak–trough–peak analysis

Peaks		Troughs		Peaks		Duration in months		
Year	Months	Years	Months	Years	Months	Trough–peak	Peak–trough	Total cycle
1975	10	1976	11	1977	4	13	5	18
1977	4	1978	5	1978	12	13	7	20
1978	12	1980	1	1981	2	13	13	26
1981	2	1982	2	1984	2	12	24	36
1984	2	1984	11	1985	6	9	7	16
1985	6	1987	2	1988	1	20	11	31
1988	1	1988	12	1991	2	11	26	37
1991	2	1991	9	1992	10	7	13	20
1992	10	1993	4	1994	3	6	11	17
1994	3							
Means						11.6	13.0	24.6
Medians						12.0	11.0	20.0
Standard deviations						4.1	7.4	8.2

(b) Trough–peak–trough analysis

Troughs		Peaks		Troughs		Duration in months		
Years	Months	Year	Months	Years	Months	Peak–trough	Trough–peak	Total cycle
1976	11	1977	4	1978	5	5	13	20
1978	5	1978	12	1980	1	7	13	26
1980	1	1981	2	1982	2	13	12	36
1982	2	1984	2	1984	11	24	9	16
1984	11	1985	6	1987	2	7	20	31
1987	2	1988	1	1988	12	11	11	37
1988	12	1991	2	1991	9	26	7	20
1991	9	1992	10	1993	4	13	6	17
1993	4	1994	3			11		
Means						13.0	11.4	25.4
Medians						11.0	11.5	23.0
Standard deviations						7.4	4.4	8.4

Table III. Final turning points of trend–cycle components, rainfall data, additive model, X-12-ARIMA. Dated by CIBCR procedure

Locations	Period	Peak–trough–peak			Trough–peak–trough		
		Trough–peak	Peak–trough	Total cycle	Peak–trough	Trough–peak	Total cycle
Agricultural Experimental Station Obispo Colombres	1975–1995	11.4	15.0	24.6	13.0	11.4	25.4
La Ramada (Tucuman)	1975–1989	15.6	15.8	32.8	15.8	16.5	32.3
Las Cejas (Tucuman)	1975–1989	10.7	13.1	23.9	13.1	11.2	25.5
Agricultural Experimental Station Obispo Colombres	1982-1995	11.3	16.7	24.4	13.4	11.3	26.5
La Cocha (Tucuman)	1982–1995	14.2	14.2	28.4	14.2	15.3	30.3
Santa Fe	1975–1989	12.8	11.0	24.8	12.8	12.8	25.7
La Pampa	1975–1989	16.4	16.5	28.5	14.8	16.4	31.2

Table IV. Monthly rainfall in Agricultural Experimental Station Obispo Colombres (Tucuman), January 1975–December 1995. Final turning points of trend–cycle component, additive model, X-12-ARIMA dated by CIBCR procedure. Peak–trough–peak analysis referred to the mean

Over the mean				Under the mean				Months of rainfall		
Start		End		Start		End		Over the mean	Under the mean	Total cycle
Year	Month	Year	Month	Year	Month	Year	Month			
1975	5	1976	7	1976	8	1976	12	14	4	18
1977	1	1978	3	1978	4	1978	6	14	2	16
1978	7	1979	9	1979	10	1980	4	14	6	20
1980	5	1981	10	1981	11	1982	5	17	6	23
1982	6	1984	7	1984	8	1985	3	25	7	32
1985	4	1985	10	1985	11	1986	6	6	7	13
1986	7	1986	10	1986	11	1987	8	3	9	12
1987	9	1988	5	1988	6	1990	9	8	27	35
1990	10	1991	6	1991	7	1991	11	8	4	12
1991	12	1992	12	1993	1	1993	11	12	10	22
1993	12	1994	5	1994	6	1995	10	5	16	21
Mean								11.5	8.9	20.4
Medians								12.0	7.0	20.0
Standard deviations								6.3	7.1	7.6

To do this, we use the information given, for example, in Tables II and IV for one of the series. We take as our main result the estimated average duration of the cycles. To agree with our purposes, we resort to the duration reported in Table IV, which is based on periods of rainfall over or under the mean. This average duration of cycles may differ slightly from those in Table II. At the end of a period of observation (in our examples it is always the month of December), if the last dated turning point is a

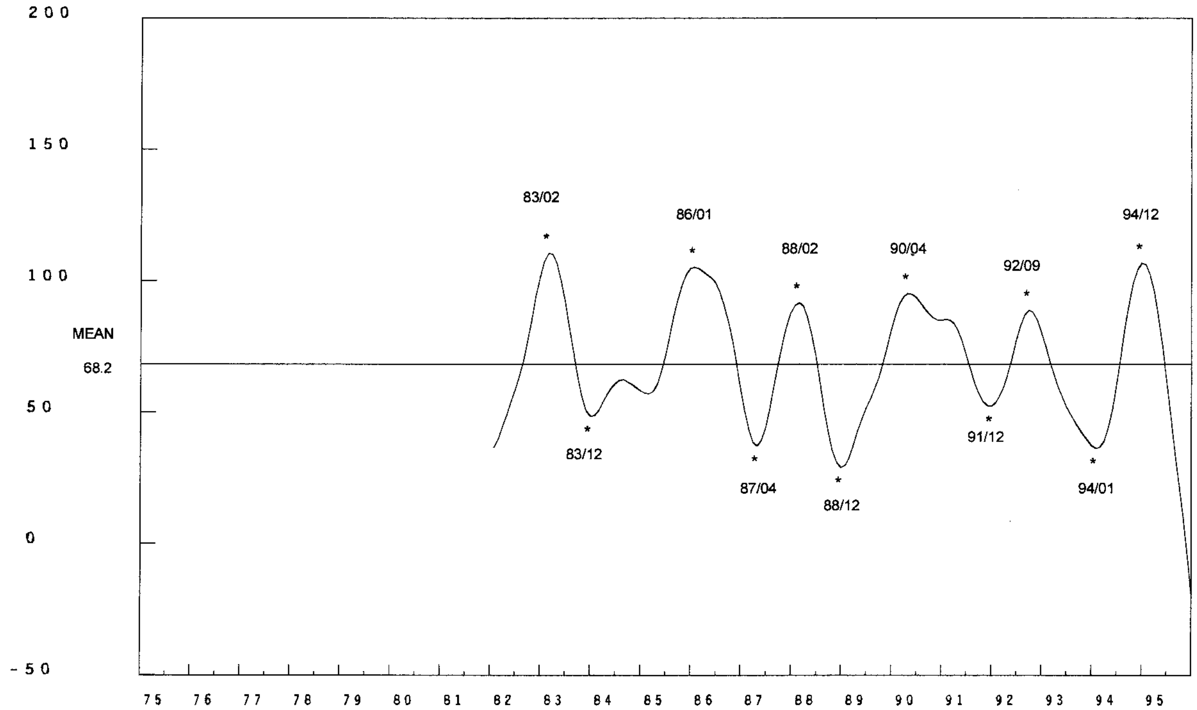

Figure 5. Monthly rainfall in La Cocha (Tucuman). January 1982–December 1995. Trend–cycle component, additive model, X-12-ARIMA, dated by CIBCR procedure, and superimposed mean line

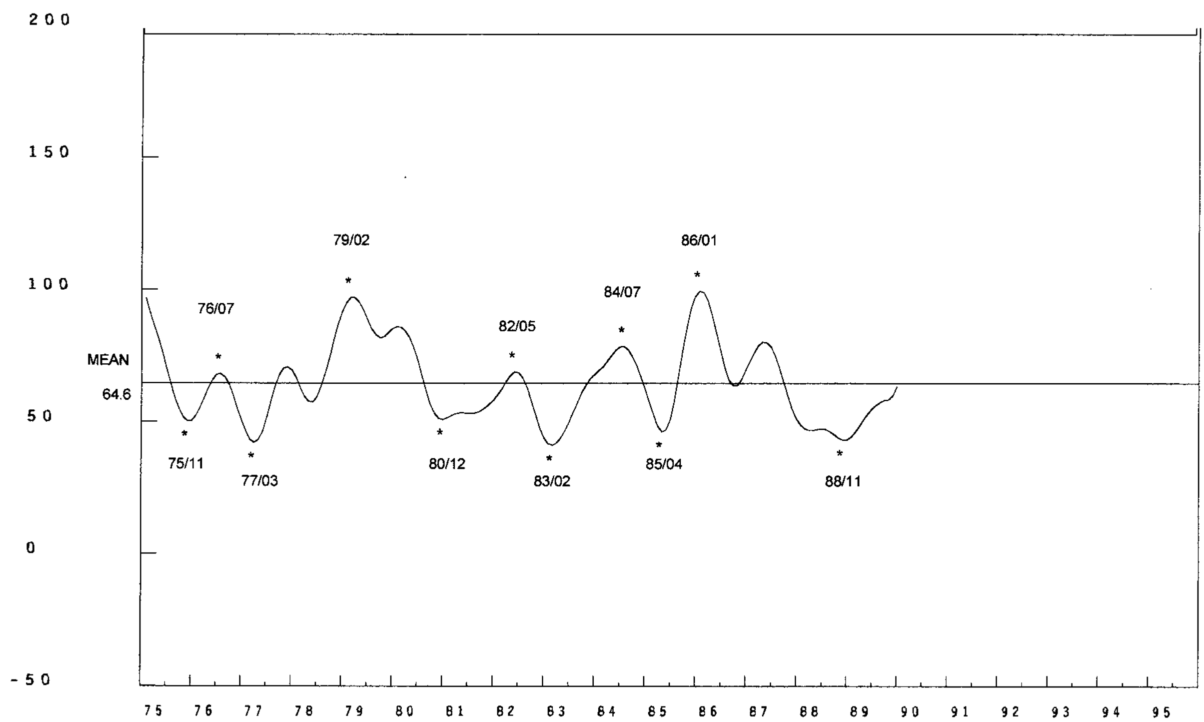

Figure 6. Monthly rainfall in La Pampa. January 1975–December 1989. Trend–cycle component, additive model, X-12-ARIMA, dated by CIBCR procedure, and superimposed mean line

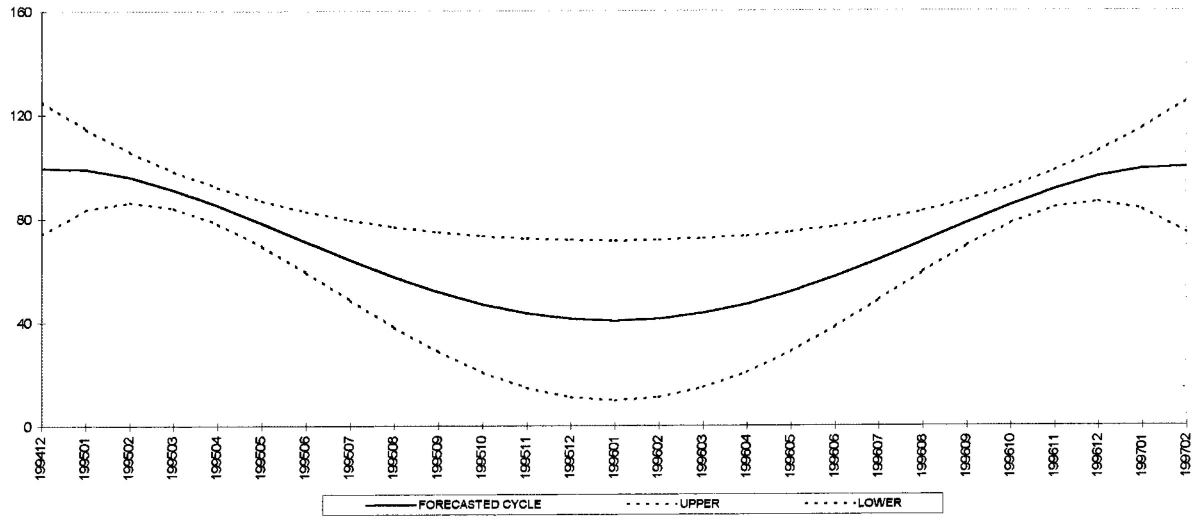

Figure 7. Monthly rainfall in La Cocha (Tucuman), January 1982–December 1995, forecasted cycle December 1994–February 1997 and upper and lower 90% confidence limits (fourth degree polynomial)

peak, we extend from there for the estimated average duration of a cycle, and consider that in this period a full cycle occurs ending in another dated peak. A parallel argument is used when the last dated turning point is a trough. In another example, for the series appearing in Figure 5 we add a new cycle: this is shown in Figure 8, and is calculated by the procedure we describe below. This will then indicate in which months to expect extra high or low levels of rain, in addition to the seasonal pattern. This is what we claim is of use in agricultural planning.

A 'forecasted cycle' is defined by the following steps: (i) two points are computed by averaging amplitudes at peak dates and trough dates; (ii) one of these is considered twice, as end of a cycle and beginning of the next one; (iii) two points are computed by averaging dates at which periods of permanence over the mean are started and ended, respectively; (iv) to these five points polynomials of degrees four, three and two are fitted and analysed. The forecasted cycle is that showing best fit on an inspection basis.

A further question is how to measure the reliability of this forecasted cycle. A confidence band is computed by the following steps: (i) confidence intervals are set at each of the five points mentioned above, considering that three of them are means of the amplitudes of certain quantities of peaks and troughs, and the other two are means of durations as those calculated in Tables II and III; (ii) for a confidence coefficient of 90%, the Bonferroni procedure is used to set the confidence band, with the t-table used to extract the needed values.

For La Cocha, Figure 7 shows the forecasted cycle which extends for 27 months (26.6), beginning at the observed peak of December 1994, and extending until February 1997, with a fourth degree fit. We also show upper and lower 90% confidence limits. Figure 8 shows this same forecasted cycle drawn immediately after the trend–cycle component values reach December 1995. We note that we took a very simple approach to the important issue of timing the beginning of a cycle.

6. ANALYSIS AND CONCLUSIONS

6.1. *Comparing locations*

An interesting question is how to compare results obtained for different locations. Two attempts in this direction appear in Figures 9 and 10. Figure 9 is based upon the following idea: to stress the oscillatory nature of the cycle component, graphs associated with it (e.g. Figure 4) are often presented with superimposed shaded strips covering the periods from one dated peak to the following dated trough. In these periods we have decreasing parts of the graph. This is usually performed in business cycle analysis, and will be used here for purposes of comparison.

In Figure 9 the dates assigned by the CIBCR cycle procedure to the turning points are compared for six locations. The shaded areas for a given location, correspond to periods of decreasing rainfall, from peak to trough. For example, for the series EEAOC (1975–1995), its diagram is an abridgement of Figure 4: most details are omitted, and only the dates appearing in Table II are retained.

Two series are said to coincide in a given month, if both are going from (dated) peaks to troughs, or both from troughs to peaks. Figure 9 shows a good level of coincidence among the four locations in Tucuman in the period when we have data for all of them (1982–1989), and among three locations for another period (1975–1982). Santa Fe tends to coincide to some extent, but La Pampa shows larger differences.

Figure 10 is based on the number of months of coincidence for the series in Figure 9, except La Pampa. The measure reported is the proportion of the absolute difference between the number of series that coincide and those that differ in a given month. For five series these differences can be 1 (three of a kind and two of the other), 3 (four of a kind and one of the other) or 5 (all of the same kind), and hence the proportions can be 0.20, 0.60 or 1.

6.2. *Reference to El Niño*

In studying meteorological variables, it pays to consider what is known as the El Niño phenomenon. We have information about the Southern Oscillation Index (SOI), which is defined as the normalized difference (i.e. difference divided into its standard deviation) between the pressure anomalies (monthly means minus long-term means) in two locations, Tahiti and Darwin, as defined by Troup (1965).

Information about El Niño is frequently used in meteorological studies in the place of a possible causal factor. In our case, and in view of the descriptive approach used throughout, we will only apply to the SOI series the same kind of methods we have discussed so far.

Figure 11 presents the final result of applying to the SOI monthly series, from January 1975 to September 1997, the seasonal and cyclical methods described in the previous sections, and therefore it is the estimated trend–cycle component dated by the CIBCR procedure. The mean duration of the cycle is

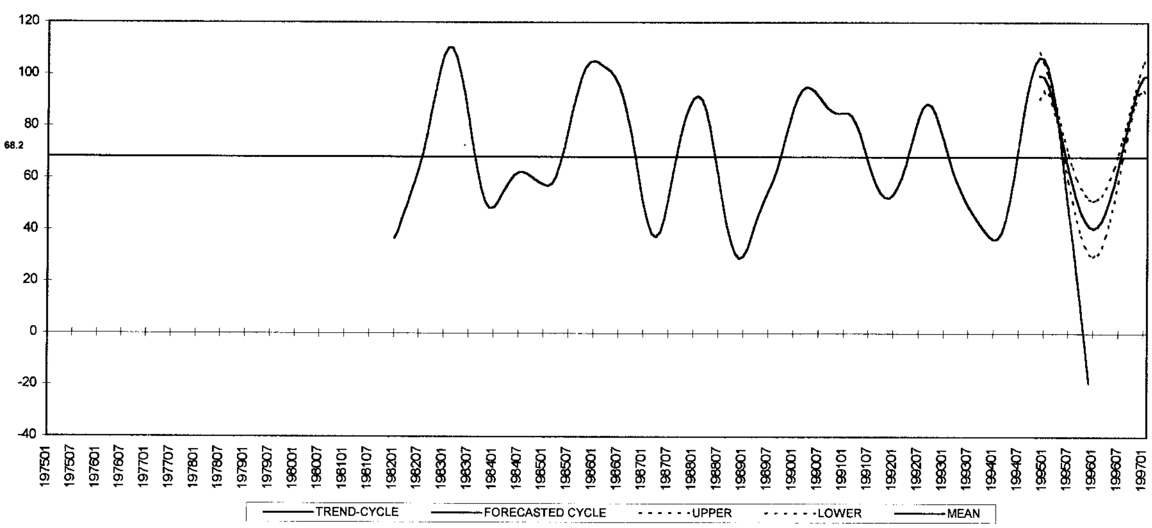

Figure 8. Monthly rainfall in La Cocha (Tucuman). January 1982–December 1995. Trend–cycle component, additive model, X-12-ARIMA, dated by CIBCR procedure. Superimposed mean line, forecasted cycle (fourth degree polynomial). December 1994–February 1997

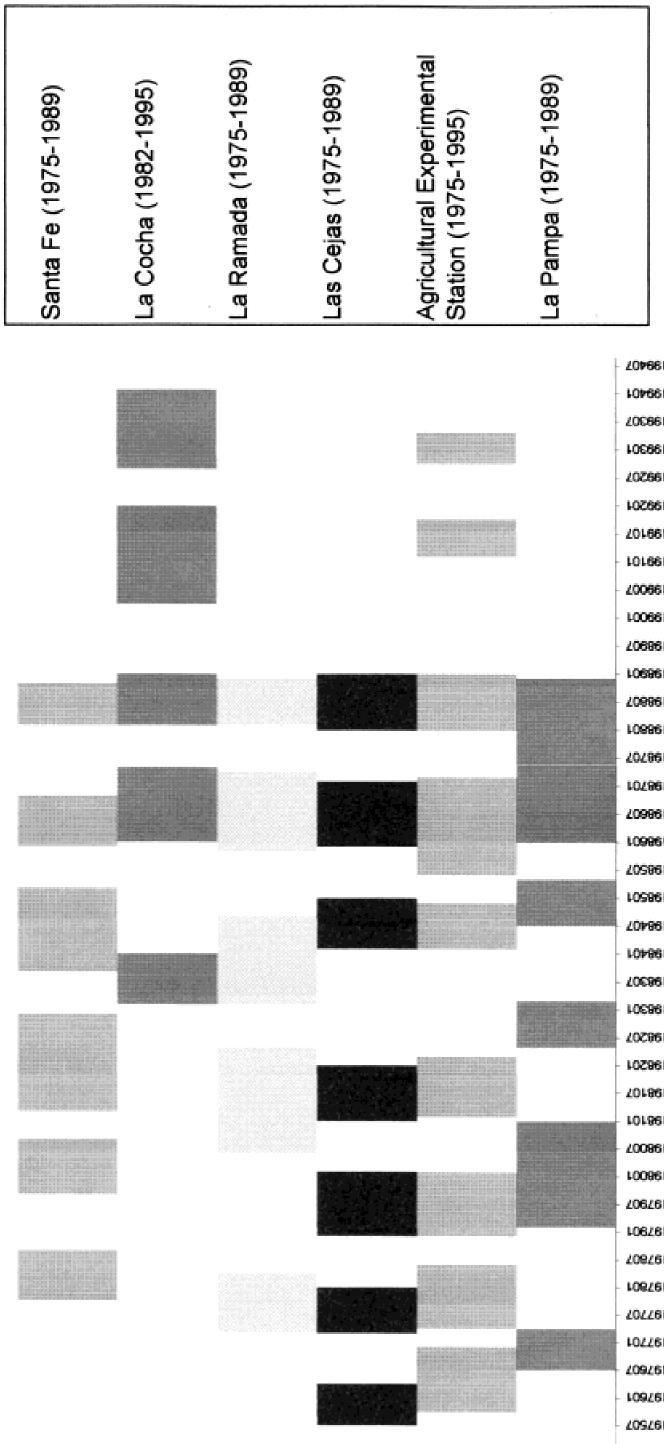

Figure 9. Final turning points of trend–cycle component of six locations, January 1975–December 1995: comparison of dates assigned by CIBCR procedure. Shaded areas correspond to periods from peak to trough

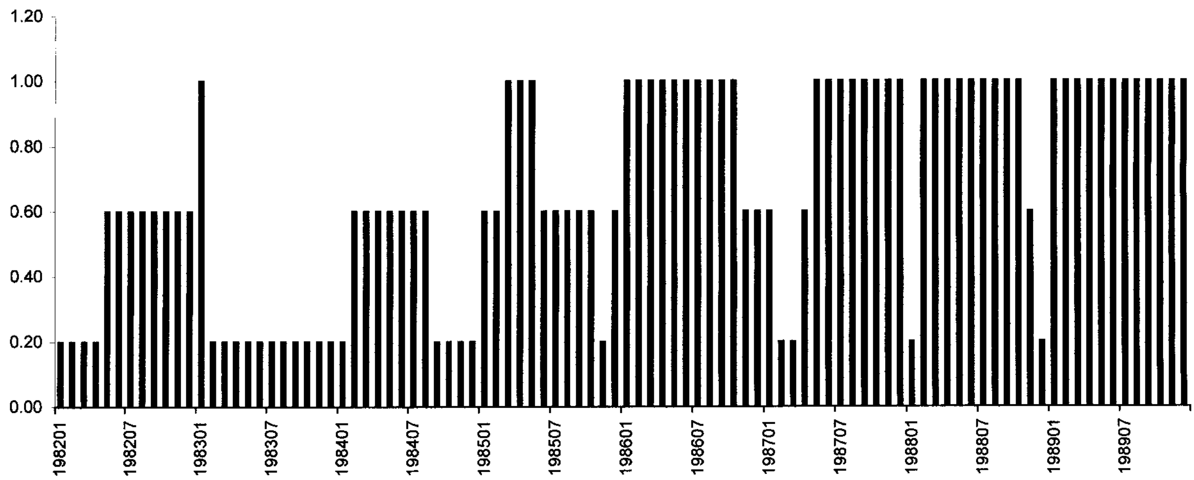

Figure 10. Proportion of coincidences among five series in Figure 9

35.7 months for peak–trough–peak and 34.2 months for trough–peak–trough, and the corresponding median durations are 32 months in both cases. The means are comparable with those in Table III, being very slightly larger.

In general, and in relation to our cyclical analysis of rainfall in the locations that we considered, we expect some degree of relation between the SOI and rainfall in La Pampa. The direction of this relation is that the lower value of the SOI should coincide with periods of high rainfall in La Pampa.

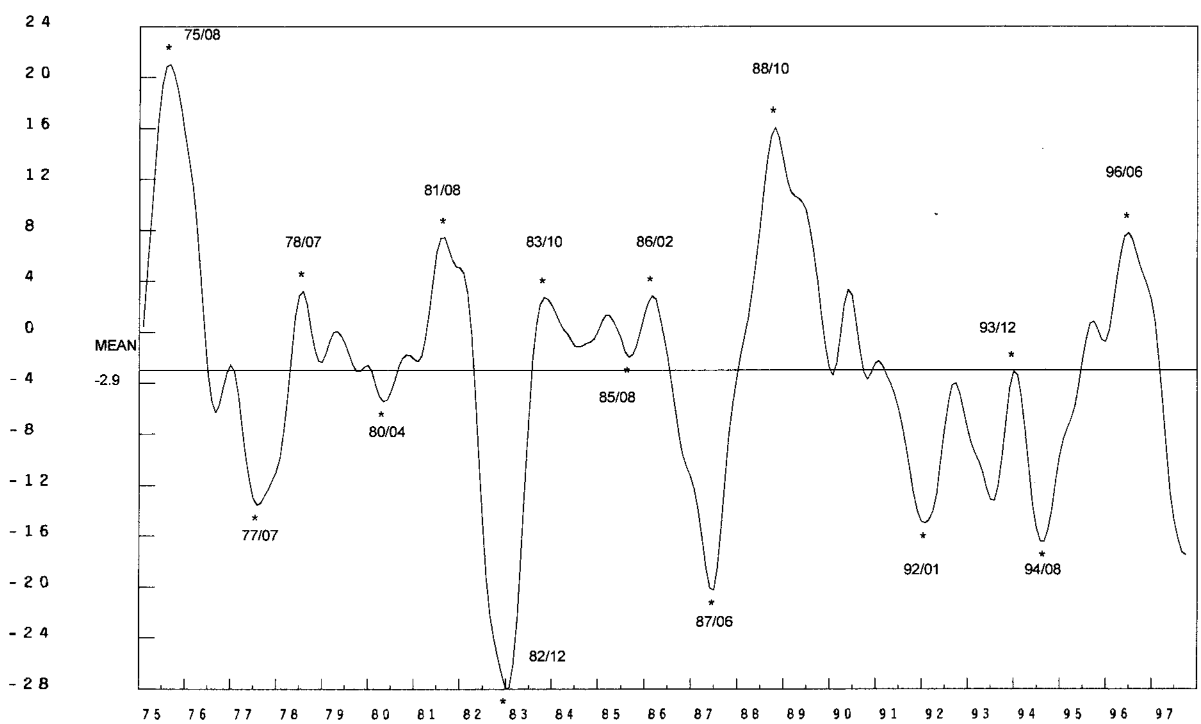

Figure 11. Monthly values of SOI January 1975–September 1997. Trend–cycle component, additive model, X-12-ARIMA, dated by CIBCR procedure, and superimposed mean line

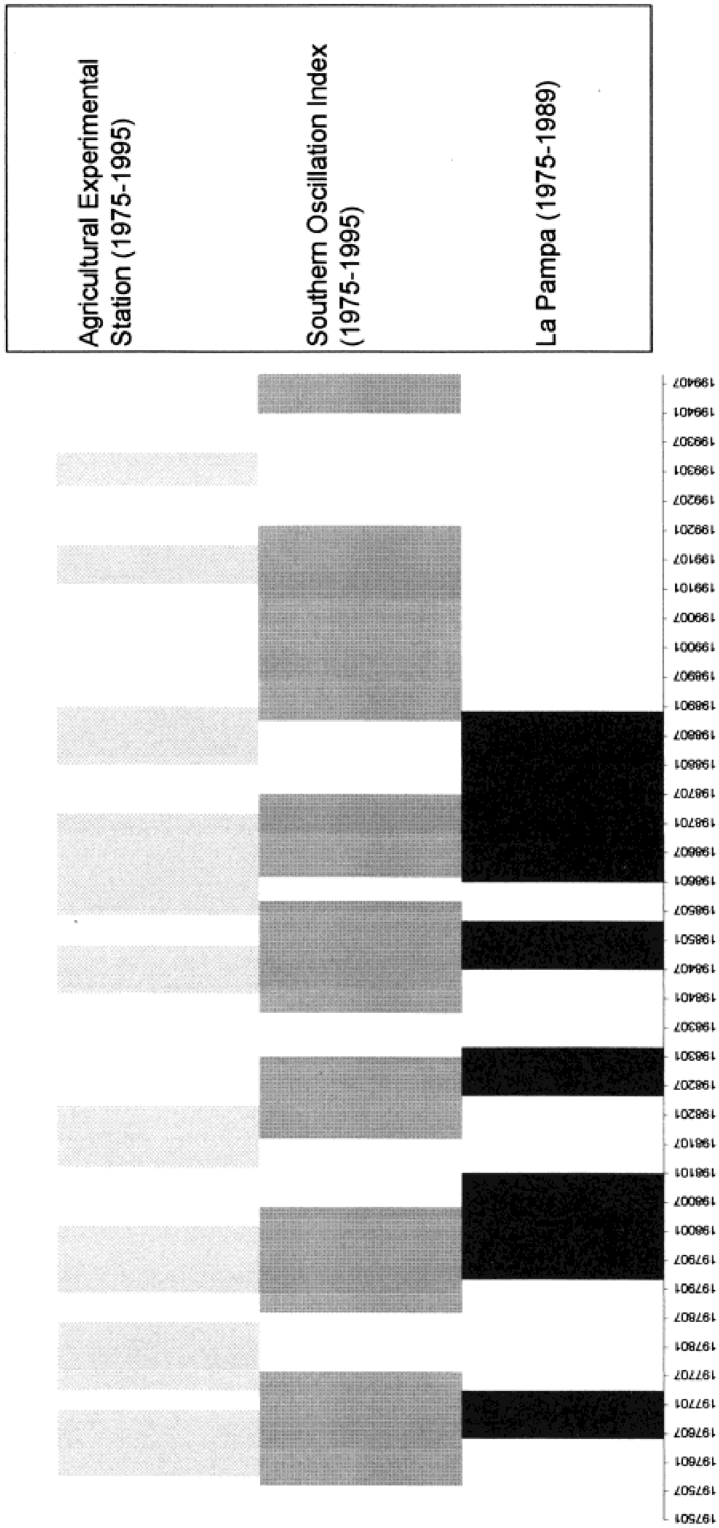

Figure 12. Final turning points of trend–cycle component of SOI and two series in Figure 9, January 1975–December 1994: comparison of dates assigned by CIBCR procedure. Shaded areas correspond to periods from peak to trough

The dates assigned to the SOI are compared with those of La Pampa in Figure 12, where Tucuman's EEAOC is also included for comparison. In 1975 and 1985, the dated troughs for the SOI occur when La Pampa is in periods of increasing rainfall (i.e. going from trough to peak). In 1980 and 1982 the dated troughs for the SOI occur shortly before the beginning of the period of increasing rainfall for La Pampa. We find reasonable agreement of these findings with what was expected. The last period of decreasing rainfall reported for La Pampa covers from 1986 to 1988, and its relation to the SOI is less clear.

The coincidences between the SOI and Tucuman's EEAOC is much less regular and interesting than for La Pampa, in agreement with what is expected.

6.3. Final comments

Information about weather variables and conditions are of great importance for agricultural management. Good forecasts of weather conditions are of real value since they help in reducing some of the risks associated with the activity. Good forecasts may be used to program yearly agricultural activities in such a way, that over a period of several years this anticipatory behaviour of the programmers leads to important benefits, in comparison with a traditional program of some sort of average behaviour over the years.

One important component of weather variables is the amount of rainfall, which in our case has been observed monthly. Other (shorter) periods could also be considered, as has often been done in practice. In our case we consider the description and analysis of monthly rainfall data for local areas, understanding that, farms or groups of them may differ in weather conditions, even when they are not too far apart. This has also been considered by other investigators, leading to the definition of areas of certain radius, so that homogeneity of weather conditions holds within areas, and heterogeneity between them. We have not explored this here.

The methods we have used deal, separately, with the seasonal and cyclical components, which are assumed to enter as unobservables in a given time series. The new X-12-ARIMA package is a flexible procedure for seasonal analysis, that can be very useful in the analysis of rainfall. Given the regularity and importance of the seasonal components, even in the presence of a large random or irregular component, we expect reasonable results of the proposed seasonal analysis.

The output of the seasonal adjustment program is an estimate of the trend–cycle component of a series, that for the case of rainfall can be safely associated with the cyclical component, since trend is not expected to be of practical significance.

The cycles that are considered through the use of the CIBCR program, are short in duration. This is in agreement with the intuitive ideas behind our work, and is in agreement with standard meteorological analysis. In effect, Burroughs (1992), section 3.12, writes: 'The QBO (quasi-biennial oscillation) is the most widely observed feature in the records, and must clearly be regarded as a real feature of almost all meteorological records'. The identification of the QBO is used for any periodicity in the range 2.2–2.8 years (26.4–33.6 months), and is then consistent with our findings in Table IV and in the duration of the cycles found in the SOI.

The analysis that we propose is then useful in interpreting the cyclical component of monthly rainfall data. We have come to the point where we believe that for one of the local areas, and having available short series (say, of up to 20 years of monthly data), a systematic statistical analysis of the data will prove useful. The average duration of the estimated cycles, and the inspection of the amplitude of the cyclical oscillations, will be useful indications for those engaged in the planning and management aspects of agriculture. One drawback lies in the variability of the observations: our indicators of quality tend to detect in the present cases even more variability than that for economic series processed for Argentina. Hence, a word of caution in the interpretation of the results should be expressed, as we pointed out in various parts of this work.

ACKNOWLEDGEMENTS

This study was started with data for La Cocha collected by Lic. Antonio Benito, whose interest in interpreting the results of our calculations we greatly appreciate. The rainfall series for the other locations were provided by Eng. Cesar Lamelas, meteorologist from EEAOC, whose interest in our work we sincerely appreciate. We thank Dr Martin Grondona for helpful comments and information about El Niño, and Professor Pedro Morettin for help with the references. We also thank two referees and the editors for many comments that we used to improve our presentation.

REFERENCES

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. 1994. *Time Series Analysis: Forecasting and Control*, 3rd Edition. Prentice Hall, New Jersey.
Bry, G. and Boschan, C. 1971. *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. US National Bureau of Economic Research, New York.
Burroughs, W.J. 1992. *Weather Cycles Real or Imaginary?* Cambridge University Press, Cambridge.
Dagum, E.B. 1983. *The X-11-ARIMA Seasonal Adjustment Method*. Statistics Canada, Ottawa.
Findley, D.F. 1996. 'Comment', *J. Business Econ. Stat.*, **14**(3), 389–393.
Findley, D.F., Monsell, B.C. and Shulman, H.B. 1989. 'Sliding spans diagnostics for seasonal and related adjustments', in Mentz, R.P. *et al.* (eds), *Statistical Methods for Cyclical and Seasonal Analyses*. IASI, Panama.
Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. and Chen, B.-C. 1998. 'New capabilities and methods of the X-12-ARIMA seasonal-adjustment program', *J. Business Econ. Stat.*, **16**(2), 127–177 (with discussion).
Gani, J. 1985. 'Statistics in meteorology', in Kotz, S. and Johnson, N.L. (eds-in-chief), *Encyclopedia of Statistical Sciences*, volume 5. Wiley, New York, pp. 453–460.
Ghysels, E., Granger, C.W.J. and Siklos, P.L. 1996. 'Is seasonal adjustment a linear or nonlinear data-filtering process?', *J. Business Econ. Stat.*, **14**(3), 374–397 (with discussion).
Karl, T.R., Knight, R.W., Easterling, D.R. and Quayle, R.G. 1996. 'Indices of climate change for the United States', *Bull. Meteorol. Soc.*, **77**(2), 279–292.
Kashyap, R.L and Ramachandra Rao, A. 1976. *Dynamic Stochastic Models from Empirical Data*. New York, Academic Press.
Mentz, R.P., de Alba, E., Espasa, A. and Morettin, P.A. (eds) 1989. *Statistical Methods for Cyclical and Seasonal Analyses*. Inter-American Statistical Institute (IASI), Panama.
Murphy, A.H. and Katz, R.W. 1985. *Probability, Statistics and Decision Making in the Atmospheric Sciences*. Westview Press, Boulder, CO.
Nerlove, M. 1964. 'Spectral analysis of seasonal adjustment procedures', *Econometrica*, **32**, 241–286.
Nerlove, M., Grether, D.M. and Carvalho, J.L. 1979. *Analysis of Economic Time Series*. Academic Press, New York.
Percival, D.B. and Walden, A.T. 1993. *Spectral Analysis for Physical Applications: Multitaper and Coventional Univariate Techniques*. Cambridge University Press, Cambridge.
Shiskin, J., Young, A.H. and Musgrave, J.C. 1967. *The X-11 Variant of the Census Method. II. Seasonal Adjustment Program*. Bureau of the Census, Washington, DC.
Troup, A.J. 1965. 'The southern oscillation', *Q. J. R. Meteorol. Soc.*, **91**, 490–506.

