What can we learn from univariate time series models? The case of sugar production in Mauritius 1879-1987

Lallmahomed, Naguib and Taubert, Peter

Universite de Paris I (Pentheon-Sorbonne)

23 February 1989

WHAT CAN WE LEARN FROM UNIVARIATE TIME SERIES MODELS ? THE CASE OF SUGAR PRODUCTION IN MAURITIUS 1879-1987 *

by

Naguib Lallmahomed and Peter Taubert

- University of Paris I (Panthéon-Sorbonne) -

In this paper, we attempt to show the validity and limits of univariate time series modelling applied to annual production of sugar in Mauritius from 1879 to 1987. We analyse the series through the main components of long-term growth and stationary dynamics of short-term coupled with the impact of exogenous shocks.

This paper was published in the 'Bulletin du GREED' of February 1989, a publication of the 'Groupe de Recherches et d'Etudes du Développement' of the University of Paris I (Panthéon-Sorbonne).

* The facilities made available to us at the Computer Centre of the University of Paris I are fully acknowledged.

GRAPHICAL ANALYSIS

The evolution of sugar production in Mauritius from 1879 to 1987 is shown in Figure 1. We can note three salient features:

(a) The series is clearly non-stationary in the mean and sugar production has grown seven-fold over one century.

(b) We can note the presence of several outliers. Whilst we know that some correspond to the violent cyclones which hit the island in 1892, 1945, 1960, 1975 and 1980 (Padya, 1984) we suspect the others to have economic origins and we leave it as a matter of discussion amongst economic historians. Outliers are known to cause heteroscedasticity and moreover, in our case, as they all cause depression, there will be a systematic bias in the estimated mean and residuals will not be gaussian. Hence, there will be a need for appropriate econometric modelling.

(c) We can finally notice two more or less distinct periods of growth: one of low level of production [1879-1945] not exceeding 400 thousand tonnes and another one of higher level [1946-1987] well above 400 thousand tonnes.

SPECIFICATION ISSUES

As our series describe a secular growth phenomenon, we are particularly interested in finding out whether such models can yield reliable predictions over a very long time horizon, say 30 years. Despite the fact that we have a series extending over 109 years, we should note that its frequency is however low. From a statistical point of view, 109 points less spaced in time than 109 annual points are equivalent but from an agronomic point of view, these two latter series have quite different aspects. Whilst many economic series such as price level, interest rates, etc... are fundamentally continuous, agricultural production data, on the other hand, are genuinely discrete.

The question then arises as to whether our original series should be transformed into natural logarithms or not. If we do not take logs, growth will be assumed to be linear whilst taking logs, it will follow an exponential path. In the linear case, the growth rate will tend to zero but in both cases, the production level will tend to infinity. We should bear in mind that there are certain obvious limits to growth, for example, surface area cultivated, choice of high-yielding disease/pest and cyclone/

drought resistant varieties of cane, the relative efficiency of fertilizers as well as improved irrigation and sugar extraction rate cannot all increase indefinitely. We can alternatively ask ourselves whether the same growth path is likely to hold for the next thirty years or so. It is more than obvious that such an outcome is most unprobable. One possible solution might be to use a logistic curve.

The three functional forms which we have tested for the long-term growth component and the exogenous shocks are:

(1) linear:

$$Y[t] = a0 + a1*t + ad'Z[t] + e[t]$$

(2) log-linear: let $y[t] = \ln(Y[t])$

$$y[t] = b0 + b1*t + bd'Z[t] + e[t]$$

(3) logistic:

$$y[t] = \ln(c2/(1+\exp(c0 - c1*t))) + cd'Z[t] + e[t]$$

the logistic function of which is written as:

$$f[t] = c2/(1 + \exp(c0 - c1*t))$$

implying that: $\lim_{t \to +\infty} f[t] = c2$ and $\lim_{t \to -\infty} f[t] = 0$

From a historical point of view, the above limits seem quite reasonable but presupposes the existence of a turning point.

Y[t] : endogenous variable in level
Z[t] : vector of exogenous shocks dummies

{ a0,a1,ad },{ b0,b1,bd },{ c0,c1,c2,cd } : parameters to be estimated

e[t] : stochastic disturbance term

If e[t] follows a stationary process, then we can conclude that ordinary least squares estimators will have the ideal properties of unbiasedness and rapid convergence to the true values of the parameters. This is different from the classical case where the matrix of uncentered cross-moments of the explanatory variables tends to a finite limit and residuals are assumed to be white-noise. If this is not the case, then we will have to difference the y[t] series in order to tackle the problem of the stochastic unit root.

In the first case we will get predictions of finite variance whereas in the second case, confidence intervals will increase indefinitely.

We should point out that an extraneous shock of a given intensity would normally cause a certain percentage loss of the sugar crop but for this to hold, the shocks dummies should be introduced in the multiplicative rather than the linear form.

TESTING FOR A UNIT ROOT

In order to establish whether the series contains a stochastic unit root or not, we compute a Dickey-Fuller (1979) test in the autoregressive form, including the dummy variables and a time trend:

$$[1-A(L)][1-L]\ y[t] = -aa*y[t-1] + a0 + a1*t + ad'z[t] + e[t]$$

where A(L) is a polynomial including lags up to the fifth order.

The null hypothesis of no unit root is rejected if aa>0 and its 't' ratio exceeds the critical value given in the Dickey-Fuller [1979] tables. For the linear model, we obtained a value of aa = -.180 [2.66] and, in the log-linear case, aa = -.507 [4.97] (note that the critical t-values are much higher than for the classical case). Only the second result allows us to reject the null. Accepting the absence of a unit root implies that the variance does not increase with the mean. Therefore, we can say with reliable confidence that sugar production tends to become more and more stable.

ANALYSIS OF THE STATIC MODELS

We started by estimating models (1) and (2) without including any dummy variables. For the five cyclonic years we obtained the expected outliers but three more were located (1931,1934,1944). Hence, the most general linear models that we estimated include a constant, a time trend and eight dummies. We will now refer to Table 2. There is no direct method of testing the linear against the log-linear model but it is possible, by a change of variable in the joint density function, to compare the log-likelihood values (Harvey, 1981, ch. 5.6). The log-linear model gave a log-likelihood of 60.3 while for the linear model the adjusted value was 5.4. Comparing these two values on the basis of information criteria, the linear model is seen to lag much behind.

The Jarque-Bera (1980) and Engle's (1982) ARCH tests indicate that neither model produces non-normal or auto-regressive conditional heteroscedastic residuals. The latter result is not surprising for the reason that the variance reflects exogenous random shocks rather than an endogenous persisting phenomenon. Godfrey's (1978) Lagrange Multiplier and the Box-Pierce (1975) tests show that residuals from both models are highly auto-correlated. However, those from the linear model are worse.

Hence, we retain the log-linear model. Note that the estimated coefficient of time is almost the same as the mean growth rate of sugar production in Table 1. We have further simplified this model by forming two weighted dummies, one representing the impact of cyclones scaled as 1,2,3 and the other to take into account the unidentified shocks.

The results from the non-linear logistic model were not conclusive. Estimation by the Gauss-Newton iterative algorithm (based on analytical first derivatives) produced the results given in the top-RHS of Table 2. The asymptotic production level is equal to c2 while the turning point satisfies:

$$f''[tp] = 0 \iff tp = c0/c1$$

The results for the unrestricted logistic model (3a) are very close to those for the log-linear one, their log-likelihood being almost identical. Hence, there is no strong evidence to support the logistic function. The asymptotic bound is rather poorly estimated and would reach seven times the contemporaneous production level, which obviously is too optimistic. The turning point falls in the future, in about one hundred years (i.e. year 2091) and not within the sample, say around 1945.

Note that we formulated the logistic model in an attempt to determine whether there is enough evidence in the sample for a limit to growth. On apriori grounds, such a limit can be imposed by fixing a value for c2. We tried c2 = 750 (equation 3b) and c2 = 1500 (equation 3c) in order to test for the acceptance of more reasonable bounds. Note that the usual t-test against c2 = 0 is inappropriate because of the fact that its distribution cannot be symmetrical around zero (c2 > 0). We reject clearly the lower variant (3b), but the second model (3c) could be regarded as a plausible alternative to the unrestricted maximum of likelihood and, moreover, to the log-linear model. As the two functional forms have completely different long-term implications, the crucial point is that of the quality of disposable apriori information. Though we believe that growth will cease in the long-run, we retain equation (2b) as our basic model for the remainder of this paper.

STABILITY ANALYSIS

Following our initial observations made from the graphical analysis of Fig. 1, we have divided our sample into two sub-periods: an ante-World War II period, [1879-1945] and a post-World War II period, [1946-1987] and have re-estimated the log-linear model (2b) for these two sub-samples.

The period following World War II has seen many major changes in the sugar industry of Mauritius. In 1946, the Cyclone and Drought Insurance Fund was established followed two years later by the Sugar Industry Labour Welfare Fund. In 1951, Britain signed with its colonies the first Commonwealth Sugar Agreement (CSA). Research which started at the Sugar Cane Research Station in 1930 was taken over in 1953 by the newly-created Mauritius Sugar Industry Research Institute (MSIRI). The University of Mauritius was set up in 1965 with the former College of Agriculture, which was founded in 1925, becoming the School of Agriculture of the University. Mauritius gained independance from Britain in 1968 and in the same year the Public Relations Office of the Sugar Industry (PROSI) was created and the CSA was replaced in 1975 by the Lome Convention. In 1980 the Bulk Sugar Terminal came into operation and in 1984, the so-called Mauritius Sugar Authority was set up to monitor the overall activities of the industry. (PROSI, 1987)

Hence, we have carried out a test of structural change as outlined by Chow [1960] : under the null hypothesis of no structural change the parameters estimated from the two sub-samples are not significantly different. The test gave $F(4,101) = 6.44$ clearly indicating the presence of a structural change after World War II.

Fig. 2 shows the growth path which we could have predicted in 1945 with, however, a knowledge of future cyclones. We can also observe that the level of production during the fifties and the sixties outpassed some historical trajectory but later rejoined it.

With slightly greater computational burden, we have produced a more detailed picture of the underlying stability problems, which, however, should be interpreted with care. Fig. 3 shows the evolution of the four coefficients of the log-linear model (equation 2b) estimated iteratively with increasing sample size from 1900 onwards and with only 22 initial observations.

The most interesting graph is the top-left one which shows the evolution of the coefficient of time. Each increment in the sample size adds new information which in turn will modify the estimated value of the coefficient. To an accelerating

Table 2 : ESTIMATION RESULTS FOR THE STATIC MODELS

equation	(2a)	(2b)	(1)
	log-linear		linear
endogenous:	sprod	sprod	SPROD
sample	1879-1987	1879-1987	1879-1987
nobs	109	109	109
ssr	2.099	2.111	424416.
see	.1456	.1416	65.47
adj. R2	.9437	.9456	.8534
dw	1.27	1.25	.769
q(15)	103.	95.4	355.
lm(6)	34.3	31.8	66.7
jb(2)	1.51	1.61	1.69
arch(6)	1.73	1.66	14.5
const	4.667 [163.]	4.669 [170.]	20.71 [1.61]
time	.0186 [40.6]	.0185 [42.5]	6.03 [29.3]
c1852	-.6500 [4.41]		-33.15 [.499]
c1945	-.9840 [6.72]		-256.80 [4.35]
c1950	-.7261 [4.94]		-279.27 [4.23]
c1975	-.3202 [2.17]		-137.73 [2.07]
c1980	-.3982 [2.69]		-160.85 [2.42]
cyclone		-.3377 [10.1]	
z1931	-.5513 [3.77]		-176.36 [2.68]
z1936	-.5197 [3.55]		-179.46 [2.73]
z1946	-.5994 [4.09]		-219.77 [3.34]
zz		-.5563 [6.65]	

ssr : sum of squared residuals
see : standard error of the estimate
dw : Durbin-Watson statistic
q(.) : Box-Pierce portmanteau test -> khi2(.)
lm(.) : Godfrey-Breusch Lagrange multiplier test -> khi2(.)
jb(.) : Jarque-Bera test against non-normality -> khi2(.)
arch(.) : Engle's Lagrange multiplier test agaist autoregressive conditional heteroscedasticity -> khi2(.)

Figures in square brackets are t-statistics for the null

sprod[t] = ln(c2/(1+exp(c0-c1*t))) + cc*cyclone[t] + cz*zz[t]

equation	(3a)		(3b)		(3c)	
			logistic			
endogenous:	sprod		sprod		sprod	
sample	1879-1987		1879-1987		1879-1987	
nobs	109		109		109	
ssr	2.098		3.188		2.213	
see	.142		.174		.145	
adj. R2	.9484		.921		.944	
dw	1.30		.90		1.26	
q(16)	98.		334.		125.	
jb(2)	1.66		5.36		2.56	
arch(6)	2.26		10.69		3.22	
c0	3.85	[3.06]	1.98	[41.5]	2.65	[80.7]
c1	.018	[11.1]	.031	[27.2]	.024	[38.8]
c2	5074.	[.79]	750	-	1500	-
cc	.338	[10.1]	.323	[7.91]	.337	[9.88]
cz	.568	[6.71]	.643	[6.24]	.555	[7.03]
tp	195.	[2.41]	64.	[43.3]	112.	[58.8]
(year)	2091		1960		2009	

Table 1 : SOME DESCRIPTIVE STATISTICS

sample: 1880-1987
number of observations = 108

series	mean	std.dev	minimum	maximum
SPROD	341.17593	200.06550	72.	718.
(1-L) SPROD	5.52778	79.37071	-344.	317.
sprod=ln(SPROD)	5.65500	.60704	4.27667	6.57647
(1-L) sprod	.01647	.24590	-0.89920	.85153

growth will correspond an upward movement in the estimated value. According to classical convergence theory, the overall effect should be more and more dampened as sample size increases but here we find out that the coefficient first accelerates, then slows down for some time to finally achieve a certain stability.

Does such a behaviour suggest the past-existence of a so-called production cycle in the sugar industry of Mauritius ? Needless to say that further appropriate investigation will have to be conducted before a sufficiently reliable answer can emerge.

DYNAMIC IMPLEMENTATION

Since the static model leaves much unexplained variance with auto-correlated residuals, it should be possible to improve and its fit and its forecasting power by introducing dynamic elements. We have tried two ways of tackling the problem: either to complete the model by allowing for auto-regressive terms, or by fitting an ARMA-structure for the deviations of the long-term path described by the static model, i.e. use of a two-step procedure. Strictly speaking, the coefficients of the shocks dummies are not separable from the short-run structure. If there are only auto-regressive terms, all the parameters can be estimated in one step, otherwise the likelihood function can become quite complicated.

None of these two methods yielded a much more precise fit: there is not enough information to extract from such a univariate series to reduce the error to, say 10 %. The introduction of autoregressive terms leads to the following model:

$$
\begin{aligned}
&[\ 1 + \underset{[3.49]}{.298}\ L + \underset{[1.97]}{.153}\ L2 + \underset{[1.59]}{.10}\ L3\ ]\ [\ 1 - L\ ]\ \mathrm{sprod}[t] = \\
&\quad -\underset{[6.47]}{.557}\ L\ \{\ \mathrm{sprod}[t] - \underset{[25.2]}{.0188} * t\ \} \\
&\quad + \underset{[6.61]}{2.634} - \underset{[6.56]}{.500} * \mathrm{zz}[t] - \underset{[10.9]}{.335} * \mathrm{cyclone}[t]
\end{aligned}
$$

sample = 1883-1987, nobs = 105,
ssr = 1.596, see = .128, adj. R2 = .732
dw = 1.83, jb(2) = 1.64, q(16) = 23.2, arch(6) = 1.55

This form of the model corresponds to an error-correction mecanism with respect to deviations from the trend. While the speed of adjustement is quite high (.557), the autoregressive structure seems to be unsatisfactory. The associated parameters are not well determined and there remains some autocorrelation of the residuals at lags 3-5. Nevertheless, the values of the diagnostic tests (jb, q, arch) are acceptable.

Figure 4 shows the dynamic simulation results in growth rates (dsprod = [1-L] sprod) over the whole sample. In the absence of new shocks, the simulated values tend geometrically to the unconditional mean.

The two-step procedure consists in saving the residuals (esprod) from the static model and identifying the generating ARMA process. We obtained the best results with the following restricted ARMA(3,4) model:

$$[\ 1 - \underset{[8.88]}{.721}\ L3\ ]\ esprod[t] =$$

$$[\ 1 - \underset{[4.21]}{.315}\ L2 + \underset{[7.76]}{.579}\ L3 - \underset{[5.56]}{.408}\ L4\ ]\ w[t]$$

sample = 1883-1987, nobs = 105,
ssr = 1.372, see = .116, adj. R2 = .313,
dw = 1.79, q(16 - 4) = 16.7

The significant lags which appear are somehow unusual but remember, our data are annual and the residuals of the above model can be considered as white-noise. In Figure 5 we give the graph of actual and fitted values, reflecting the inherent problems of the linear long-run specification.

CONCLUDING REMARKS

In this paper, using annual sugar production in Mauritius from 1879 to 1987, we have tried to show what can be gained by the appropriate modelling of univariate time series through the three main components of long-term growth, stationary short-run dynamics and impact of exogenous shocks. The results achieved so far have shown the inherent limitations of modelling with univariate time series. However, we follow Granger and Newbold (1986, pp. 151) to say that it allows us to assess how much of the variation in sugar production can be explained by its own past behaviour so as to form a clearer picture of what particular behaviour patterns require consideration of extraneous factors for their explanation.

Situated between latitudes 19 degrees 59' S and 20 degrees 32' S and longitudes 57 degrees 18' E and and 57 degrees 47' E, near the southern edge of the tropical belt, Mauritius is practically free from continental influences and has a mild maritime climate. The weather does not however remain the same for more than a few days at a stretch, being affected by disturbances of both temperate and tropical types (Padya, 1984). Sugar production in Mauritius hence depends heavily on such factors as rainfall, temperature, solar radiation and wind speed.

It makes no doubt that further research will have to take into account not only the climatic factors mentioned above but also economic variables likely to affect sugar production in Mauritius.

Fig. 1: EVOLUTION OF SUGAR PRODUCTION IN MAURITIUS
1879 - 1987

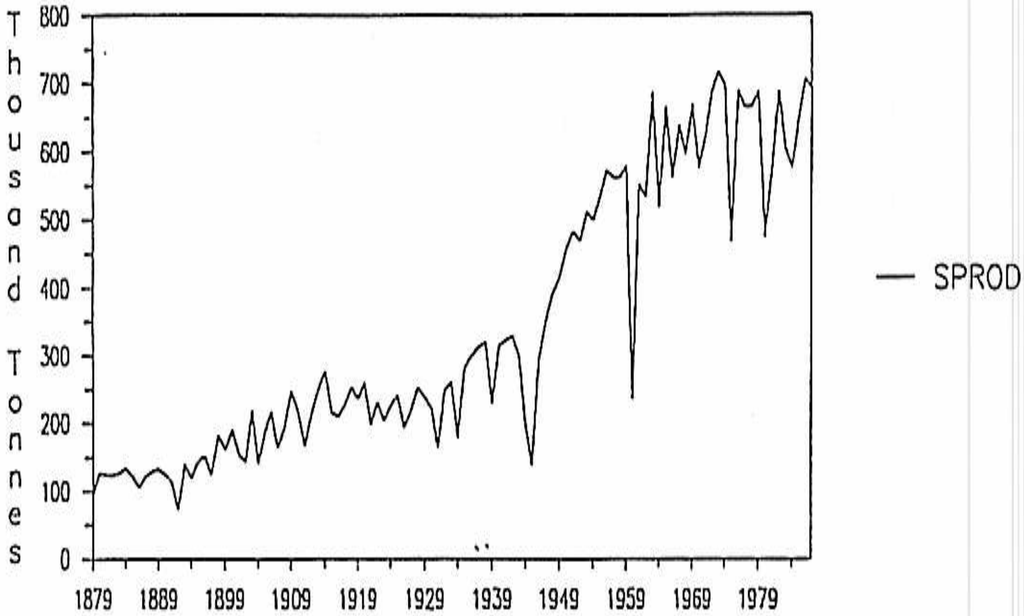


fig. 2: static log-linear model
estimated on 1879-1945 and forecasts for 1946-1987

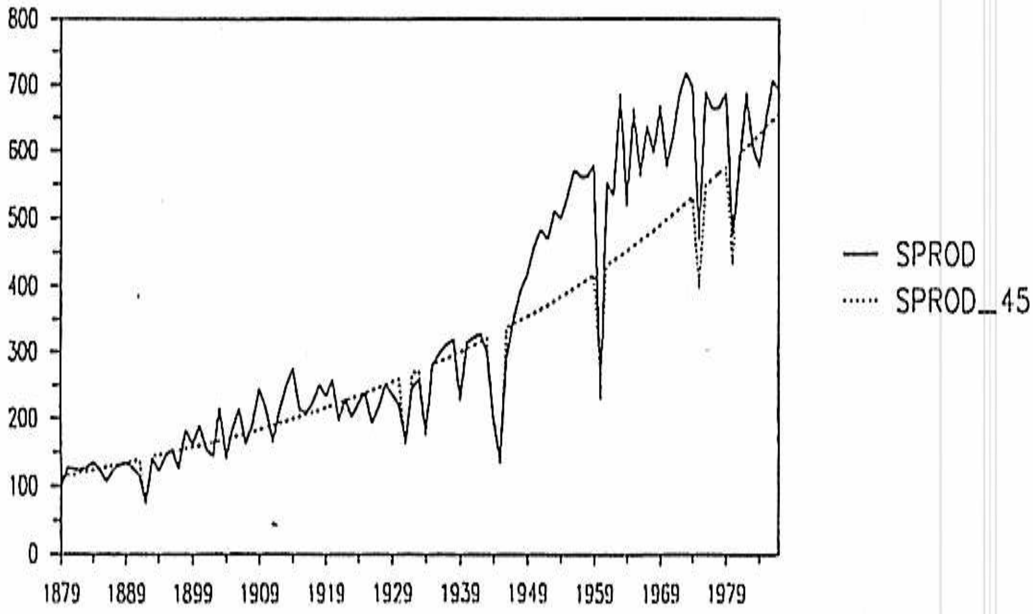
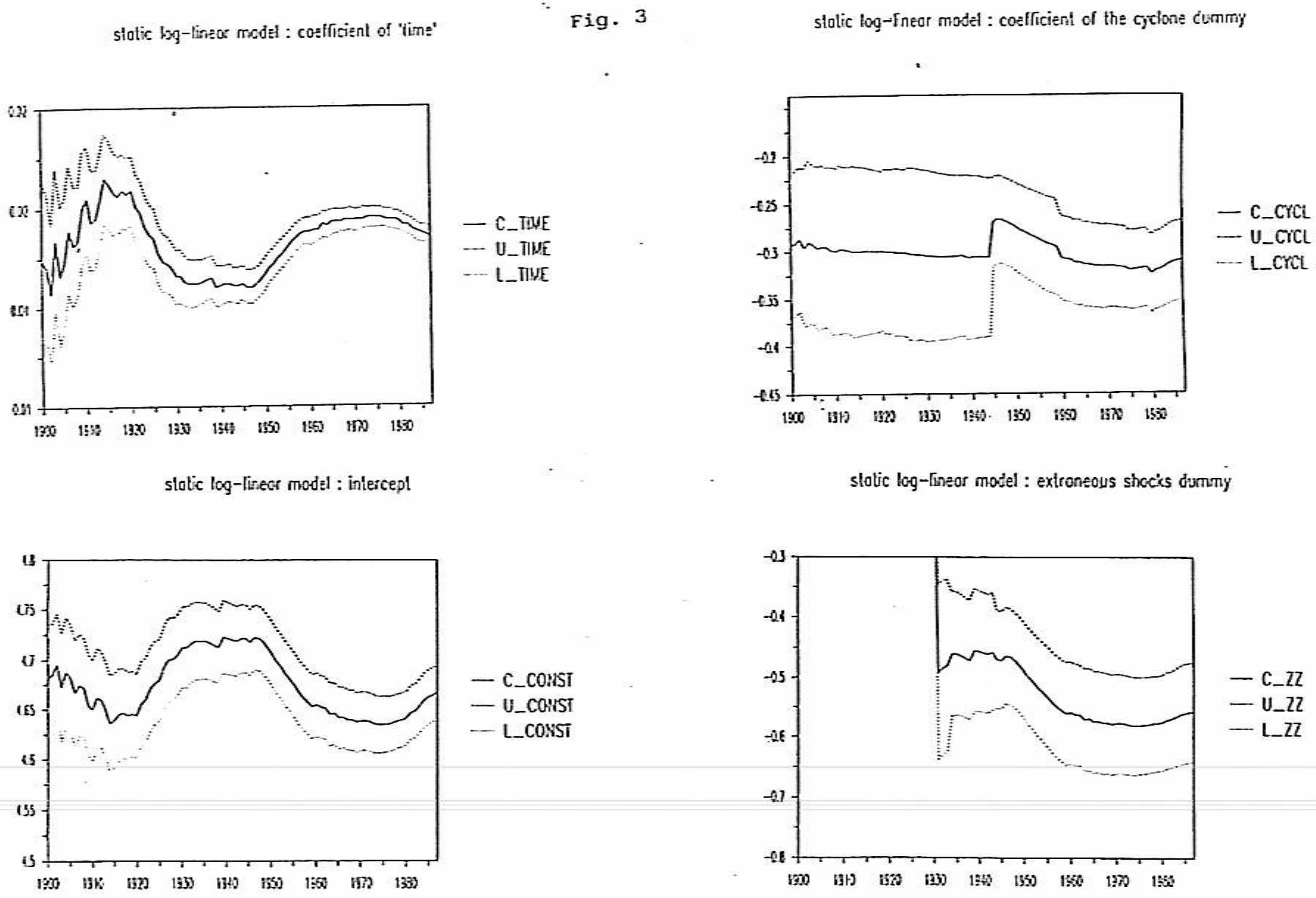


Fig. 3


static log-linear model : coefficient of 'time'
C_TIME
U_TIME
L_TIME
static log-linear model : coefficient of the cyclone dummy
C_CYCL
U_CYCL
L_CYCL
static log-linear model : intercept
C_CONST
U_CONST
L_CONST
static log-linear model : extraneous shocks dummy
C_ZZ
U_ZZ
L_ZZ

dynamic simulation of growth rate
fig. 4: auto-regressive log-linear model

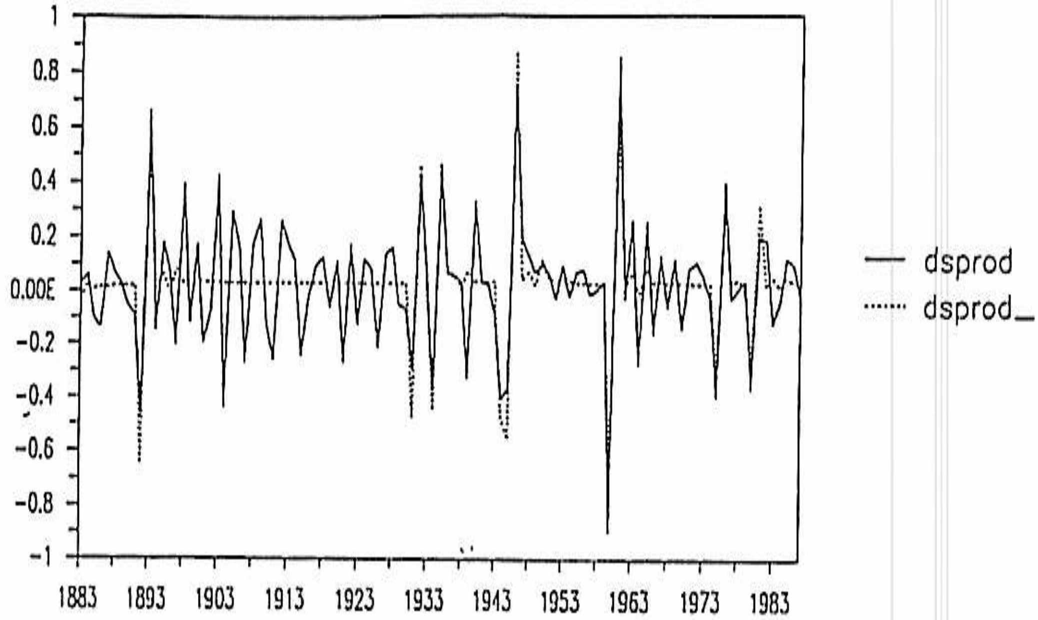
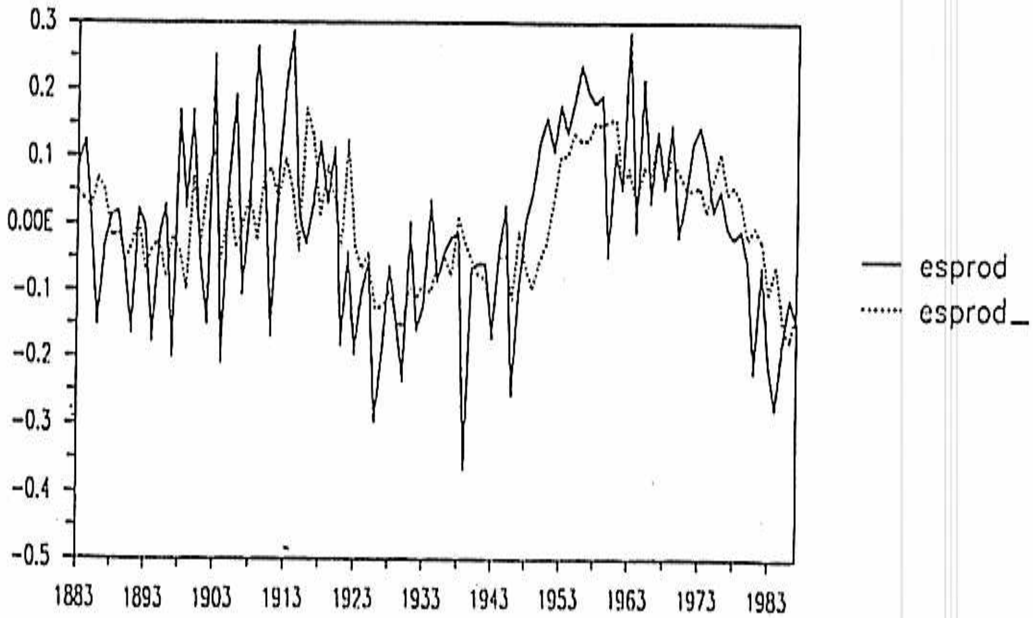


fig. 5: residuals from the static log-linear model
fitted by an arma[(3),(2-4)] model



REFERENCES

BOX, G.E.P. and D.A. PIERCE (1970), 'Distribution of residual autocorrelations in autoregressive-integrated moving average time series models', Journal of the American Statistical Association, 65, 1509-1526.

CHOW, G.C. (1960), 'Tests of equality between sets of coefficients in two linear regressions', Econometrica, 28, 591-605.

DICKEY, D.A. and W.A. FULLER (1979), 'Distribution of the estimator for autoregressive time series with a unit root', Journal of the American Statistical Association, 74, 427-431.

ENGLE, R.F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of the UK inflation', Econometrica, 50, 987-1007.

F.A.O. (undated), The World Sugar Economy in Figures 1880-1959, FAO, Rome.

GODFREY, L.G. (1978a), 'Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables', Econometrica, 46, 1293-1302.

GODFREY, L.G. (1978b), 'Testing for higher order serial correlation when the regressors include lagged dependent variables', Econometrica, 46, 1303-1310.

GRANGER, C.W.J and P. NEWBOLD (1986), Forecasting Economic Time Series, 2nd edition, Academic Press, Orlando.

HARVEY, A.C. (1981), The Econometric Analysis of Time Series, Philip Allan, Oxford.

JARQUE, C.M. and A.K. BERA (1980), 'Efficient tests for normality, homoscedasticity and serial independence of regression residuals', Economic Letters, 6, 255-259.

M.S.I.R.I. (1960-1986), Mauritius Sugar Industry Research Institute Annual Reports, Reduit, Mauritius.

PADYA, B.M. (1984), The Climate of Mauritius, 2nd edition, Meteorological Services, Mauritius.

P.R.O.S.I. (1987), Sugar in Mauritius, 4th revised edition, Port-Louis, Mauritius.

P.R.O.S.I. (1988), Bulletin Mensuel, 228, fevrier 1988, Port-Louis, Ile Maurice.

PARIS, 23/02/1989.