TIME SERIES ANALYSIS OF PHILIPPINE FISH AND FISH PREPARATION USING BOX-JENKINS METHOD[1]

by Estela B. Tormon[2]

1. Introduction

Fitting models to the export and import time series of a commodity may be used by economists and policymakers in gaining a better understanding of the nature of our external trade and in making short-term forecasts. The model may help in identifying factors or variables which may have generated the series. These changes may be attributed to devaluation of the currency, to inflation, to the imposition of control or decontrol measures. The fitted model may be used also to predict future import values of the commodity. These values in turn may guide planners in determining the appropriate control measures necessary to regulate the importation of that commodity leading to a favorable balance of trade.

In this particular study, the Box-Jenkins method of model-building is applied to the Philippine fish and fish preparation import series. First, a tentative model is identified for the series. Then, the unknown parameters of the model are estimated. Finally, diagnostic checks are performed to determined the adequacy of the fitted model.

The data used in this study are monthly observations for 25 years (1953-1977) of f.o.b. export and import values, in thousand US dollars, of fish and fish preparation as published by the National Census Statistics Office. A computer program written in FORTRAN IV was used for identification purposes while the TSERIES package was utilized for preliminary and final estimation and diagnostic checking.

2. The Box-Jenkins Non-Seasonal Models

In the Box-Jenkins method of model-building, the first step is to postulate a general class of models from which a particular model may be identified for the observed time series. The different kinds of non-seasonal models -- autoregressive, moving average, autoregressive-moving average and integrated models -- comprising this general class of models will be discussed in this section.

---

[1]This paper is taken from the author's masteral thesis completed at the U.P. Statistical Center.

[2]The Author is a Senior Census Statistical Coordinator and Officer-in-Charge, Population and Housing Census Branch of the National Census and Statistics Office.

2.1 The Autoregressive Model

Let a_t be a purely random process or white noise, that is, it consists of a sequence of mutually independent and identically distributed random variables with mean zero and constant variance σ_a^2. The process z_t is said to be an autoregressive process of order p, abbreviated to AR(p), if

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots \phi_p z_{t-p} + a_t .$$

This expression can be rewritten as

$$(1 - \Phi(B))z_t = a_t,$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$$

and B is called the backward shift operator defined as $B^j z_t = z_{t-j}$.

Due to its similarity to a multiple regression model the term "autoregressive" is coined to describe this model. But, rather than regressing on independent variables z_t is regressed on its own past values.

Initially, the order of autoregressive process to fit to an observed time series is not known. The problem is analogous to deciding on the number of independent variables to be included in a multiple regression. An analysis of the correlogram of the partial autocorrelation function:

$$\phi_{hh} = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(1) & \cdots & \rho(1) \\ \rho(1) & 1 & & \cdots & \rho(2) \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ \rho(h-1) & . & . & \cdots & \rho(h) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(h-1) \\ \rho(1) & 1 & & \cdots & \rho(h-2) \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ \rho(h-1) & . & . & \cdots & 1 \end{vmatrix}}$$

may help solve this problem.

In general, for an autoregressive process of order p, the partial autocorrelation function ϕ_{hh} will be nonzero for $h \leq p$ and zero for $h > p$ which implies that the partial autocorrelation of a pth order AR process has a cut-off after lag p.

2.2 Moving Average Model

A moving average process of order q, abbreviated to MA(q), is defined by

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \tag{2.2.1}$$

where the symbols $-\theta_1, -\theta_2, \ldots, -\theta_q$ are the finite set of weight parameters. The expression (2.2.1) may be written in the equivalent form $z_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)a_t$ or $z^t = \theta(B)a_t$.

The process is called a moving average process of order q because the observations are a moving average in the disturbances reaching back to q periods.

The autocorrelation function of a MA(q) process is zero, beyond the order q of the process. In other words, the autocorrelation function of a moving average process has a cut-off at lag q.

2.3 Mixed Autoregressive-Moving Average Model

To achieve parsimony it may be necessary to include both autoregressive and moving average terms. The mixed autoregressive-moving average of degree p and, q or abbreviated to ARMA (p,q) is defined by

$$z_t = \phi_1 z_{t-1} + \ldots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \tag{2.3.1}$$

Using the back-shift operator, this can be expressed as

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q) a_t,$$

or simply

$$\phi(B) z_t = \theta(B) a_t \tag{2.3.2}$$

where $\phi(B)$ and $\theta(B)$ are polynomials of degree p and q in B.

The correlogram of the autocorrelation function of an ARMA (p,q) process consists of a mixture of damped exponentials and/or damped sine waves if $q-p < 0$. This general pattern, however, is not followed if $q-p \geq 0$.

Although the partial autocorrelation function of a mixed process is infinite in extent, eventually, it behaves like a pure moving average process in that its correlogram is also dominated by a mixture of damped exponentials and/or damped sine waves.

2.4 Autoregressive Integrated Moving Average Model

The models described in the preceding sections are applicable to stationary processes only. Since most of the observed time series are non-stationary, these are reduced to stationary ones by suitable differencing. These models are called autoregressive integrated moving average or ARIMA processes. The term "integrated" is used because the stationary model fitted to the differenced data has to be summed or integrated to provide a mode for non-stationary data.

The ARIMA (p,d,q) has the form $\phi(B)\nabla^d z_t = \theta(B)a_t$ or simply $\phi(B)\Delta w_t = \theta(B)a_t$ where $\nabla w_t = \nabla^d z_t$. The model transforms the non-stationary process z_t into mutually independent and identically distributed random variables, a_t.

3. Preliminary Identification

The objective of preliminary identification is to obtain an idea of the kind of representational model which later on will be fitted to the data and checked. At this stage of model building a great deal of judgment must be exercised because of the inexact nature of the identification procedure.

The first step in the preliminary identification procedure is to difference the observed time series as many times as needed to produce stationarity. Then, the resulting ARMA model is identified.

3.1 Identifying the Degree of Differencing

The autocorrelation function of a stationary mixed auto-regressive moving average process of order (p,o,q), $\phi(B)z_t = \theta(B)a_t$, satisfies the difference equation

$$\phi(B)\rho(h) = 0 \quad h > q$$

Also, if $\phi(B) = \sum_{i=1}^{p} (1 - G_i B)$, assuming distinct roots, the solution of this difference equation for the hth autocorrelation is of the form

$$\rho(h) = A_1 G_1^h + A_2 G_2^h + \ldots A_p G_p^h, \quad h > q\text{-}p.$$

The stationarity requirement that the zeroes of $\phi(B)$ lie outside the unit circle implies that the roots $G_1, G_2, \ldots, G_p$ lie outside the unit circle.

If none of the roots lie close to the boundary of the unit circle the autocorrelation function of a stationary model will quickly fade for moderate and large h; however, if a root close to unity exists the function will fall off slowly and very nearly linearly. The estimated autocorrelation function tends to follow the behavior of the theoretical autocorrelation function; therefore, nonstationarity is indicated if the estimated autocorrelation function fails to die out rapidly.

It is possible that the process may become stationary if differencing is performed. The "dying-out" fairly quickly of the autocorrelation function indicates that stationarity has been achieved at that degree d of differencing. In practice it is usually sufficient to examine the first 20 estimated autocorrelations and its first or second differences.

3.2 Identifying the Resultant ARMA Model

The order p and q for ARMA processes may be determined through the characteristic behavior of the correlograms of its autocorrelation and partial autocorrelation functions. The auto-correlations of an autoregressive process of order p tails off while its partial autocorrelation function has a cut-off after lag p. Conversely, the autocorrelation function of a MA(q) process has a cut-off after lag q, while its partial autocorrelation tails off. A mixed process is indicated if both autocorrelations and partial autocorrelations exhibit exponential decay. Furthermore, its autocorrelation function consists also of a mixture of exponentials and damped sine waves after the first q - p lags while its partial autocorrelation function is dominated by a mixture of exponentials and damped sine waves after the first p - q lags.

3.3 Tentative Identification With Initial Estimates

As mentioned earlier the autocorrelation and the partial auto- correlation functions are useful not only in helping guess the form of the model but also in obtaining preliminary estimates of the parameters. The PEST command in the TSERIES computer program computes the preliminary estimates of the parameters of the specified model. In addition to the preliminary estimates for a specified ARMA model, this command also produces the following outputs: the sample autocorrelation function of the working series, the sample partial auto-correlation of the working series and the chi-square statistic associated with the "portmanteau test" for the sample autocorrelations.

4. Maximum Likelihood Estimation of Model Parameters

After having identified a tentative model for a given time series, the next step in the Box-Jenkins model building is to obtain estimates of the model parameters which make efficient use of the data. If the estimation procedure is not efficient, then it could be possible that inadequacy of fit is due to inefficient fitting, and not because the assumed model is incorrect. It may be shown that estimates which maximize the likelihood function are generally inefficient when the number of observations is large.

4.1 The Conditional Maximum Likelihood Function

Let $z_1,\ldots,z_N$ be the observed series and $w_1,\ldots,w_n$ be the differenced series, where $n = N - d$ and $w_t = \nabla^d z_t$. Suppose that the identified model for the observed series is ARIMA (p,d,q). Then the differenced series can be fitted to the ARMA (p,q) model which may be written as

$$a_t = w_t - \phi_1 w_{t-1} - \phi_2 w_{t-2} - \cdots - \phi_p w_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} \qquad (4.1.1)$$

Unfortunately, the a_t's cannot be obtained directly from (4.1.1) because of the difficulty of starting up the difference equation. However, the values of a_t can be calculated successively if preliminary estimates of the parameters $\underline{\phi}' = [\phi_1, \phi_2,\ldots,\phi_p]$ and $\underline{\theta}' = [\theta_1, \theta_2,\ldots,\theta_q]$ are given together with the starting values $\underline{w}^* = [w_0, w_{-1},\ldots,w_{1-p}]$ and $\underline{a}^* = [a_0, a_{-1},\ldots,a_{1-q}]$.

Now, on the assumption that the a_t's are independently and normally distributed their joint density is given by

$$p(a_1,a_2,\ldots,a_n) = (2\Pi\sigma_a^2)^{-n/2} \exp\left| -\frac{1}{2}\sigma_a^2 \sum_1^n a_t^2 \right| \qquad (4.1.2)$$

Since the Jacobian of the transformation is unity, direct substitution of (4.1.1) in (4.1.2) gives the following joint density of the w_t's:

$$p(\underline{w}|\underline{\phi}, \underline{\theta}, \sigma_a^2) = (2\Pi\sigma_a^2)^{-n/2} \exp -\frac{1}{2}\sigma_a^2 \sum_{t=1}^{n} (w_t - \phi_1 w_{t-1} - \ldots)^2$$

Given the data $\underline{w}$, the likelihood function associated with the parameter values $(\underline{\phi}, \underline{\theta}, \sigma^2)$, conditional on the choice of $\underline{w}^*$ and $\underline{a}^*$ is therefore

$$L(\underline{\phi}, \underline{\theta}, \sigma_a^2 | \underline{w}^*, \underline{a}^*, w) = (2\Pi\sigma_a^2)^{-n/2} \exp\left| - S_*(\underline{\phi}, \underline{\theta})/2\sigma_a^2 \right|$$

where

$$S*(\underline{\phi}, \underline{\theta}) = \sum_{t=1}^{n} a_t^2 \; (\underline{\phi}, \underline{\theta}, \underline{w}_*^*\; \underline{a}, \underline{w})$$

Since the interest is only in the relative magnitude of the likelihood, it suffices to consider only the log likelihood function given by

$$L^*(\underline{\phi}, \underline{\theta}, \sigma_a^2 \; \underline{w}_*^*, \underline{a}^*, \underline{w}) = -n\ell n\sigma_a^2 - S*(\underline{\phi}, \underline{\theta})/2\,\sigma_a^2 \qquad (4.1.3)$$

In the conditional log likelihood function given by (4.1.3) the parameters $\underline{\phi}$ and $\underline{\theta}$, which are estimated with the use of the data $\underline{w}$, enter only through the sum of squares $S(\underline{\phi}, \underline{\theta})$. Therefore, to maximize the likelihood function it is necessary only to minimize the sum of squares function and the values of the parameters. In other words, the conditional maximum likelihood estimates are equivalent to least squares estimates.

4.2 Nonlinear Estimation

In general, model (4.1.1) is nonlinear in the parameters. To linearize the model, a_t is first expanded in Taylor series about its value $a_{t,0}$ corresponding to some guessed set of parameter value.

The equation may be expressed in matrix form as

$$[\underline{a}_0] = \underline{X}\,(\underline{\beta} - \beta_0) + [\underline{a}]$$

where $\underline{a}_0$ and $\underline{a}$ are column vectors and

$$\underline{X} = \begin{bmatrix} X_{1,1} & Z_{1,2} & \cdots & X_{1,p+q} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,p+q} \\ \cdot & & & \\ \cdot & & & \\ X_{n,1} & X_{n,2} & \cdots & X_{n,p+q} \end{bmatrix}$$

The coefficients $\beta_i - \beta_{i,0}$ can then be estimated as

$$B - B_0 = (X'X)^{-1}X\;[\underline{a}_0]$$

by the method of ordinary least squares. Given the estimate $B - B_0$, the initial estimate β_0 is adjusted by $\beta'_0 = \beta_0 + (B - B_0)$ and the process is repeated until satisfactory convergence is obtained.

The nonlinear estimation procedure described above may converge very slowly or not at all. A more efficient procedure used in this study is the Marquadt algorithm (1963) which is essentially a compromise procedure between the linearization technique discussed above and the method of "steepest descent" described in Draper and Smith (1966).

5. Model Validation and Modification

At the start of model building one is dealing with uncertainties -- possible models are only "guessed" at in the identification stage and the true coefficients of the model are only estimated in the final stage. Unless some kind of test is performed on the fitted model one cannot use it with confidence for forecasting. It is for this reason that diagnostic checking is necessary as the last step in the Box-Jenkins method of model building. Not only does it provide a way of assessing the adequacy or inadequacy of a fitted model for a given data set but it may also lead to alternative models. Box and Jenkins describe several methods of diagnostic checking, but for this study residual analysis and the portmanteau lack-of-fit test was employed.

The rationale behind the use of the residuals as a means of checking for model adequacy is that the residuals will constitute a white noise process if the fitted model is the correct one. Anderson has shown that the sample autocorrelations "white noise" residuals are uncorrelated and normally distributed with mean zero and standard deviation $1/\sqrt{n}$ for moderately large samples. One can then make use of the sample autocorrelations to assess any deviations from white noise behavior. It should be reiterated that this holds true only if the form of the model is correct and the real parameter values are known.

Box and Pierce (1970) devised a chi-square test of model adequacy, better known as the portmanteau lack-of-fit test wherein the autocorrelations of the estimated residuals are considered together rather than individually. The null hypothesis that the true model residuals are white noise is tested using the statistic

$$Q = \sum_{h=1}^{H} r_a(h), \qquad h = 1,2, \ldots, H, \quad = N - d$$

which has a chi-square distribution with degrees of freedom equal to (H-p-q). The null hypothesis is rejected for large values of Q.

6. Box-Jenkins Model for Import Values of Fish and Fish Preparation

As an illustration, a model will be fitted to the import values of fish and fish preparation from 1953 to 1977. To determine the general behavior of the series, the 300 observations were first plotted (Figure 1). The computer program on identification was run and the computed autocorrelations and partial autocorrelations were plotted in Figure 2. As shown in Figure 2, d = 0, the behavior of the autocorrelation function of fish import series strongly suggests an AR(1) process. It is characterized by a damped exponential. The correlogram

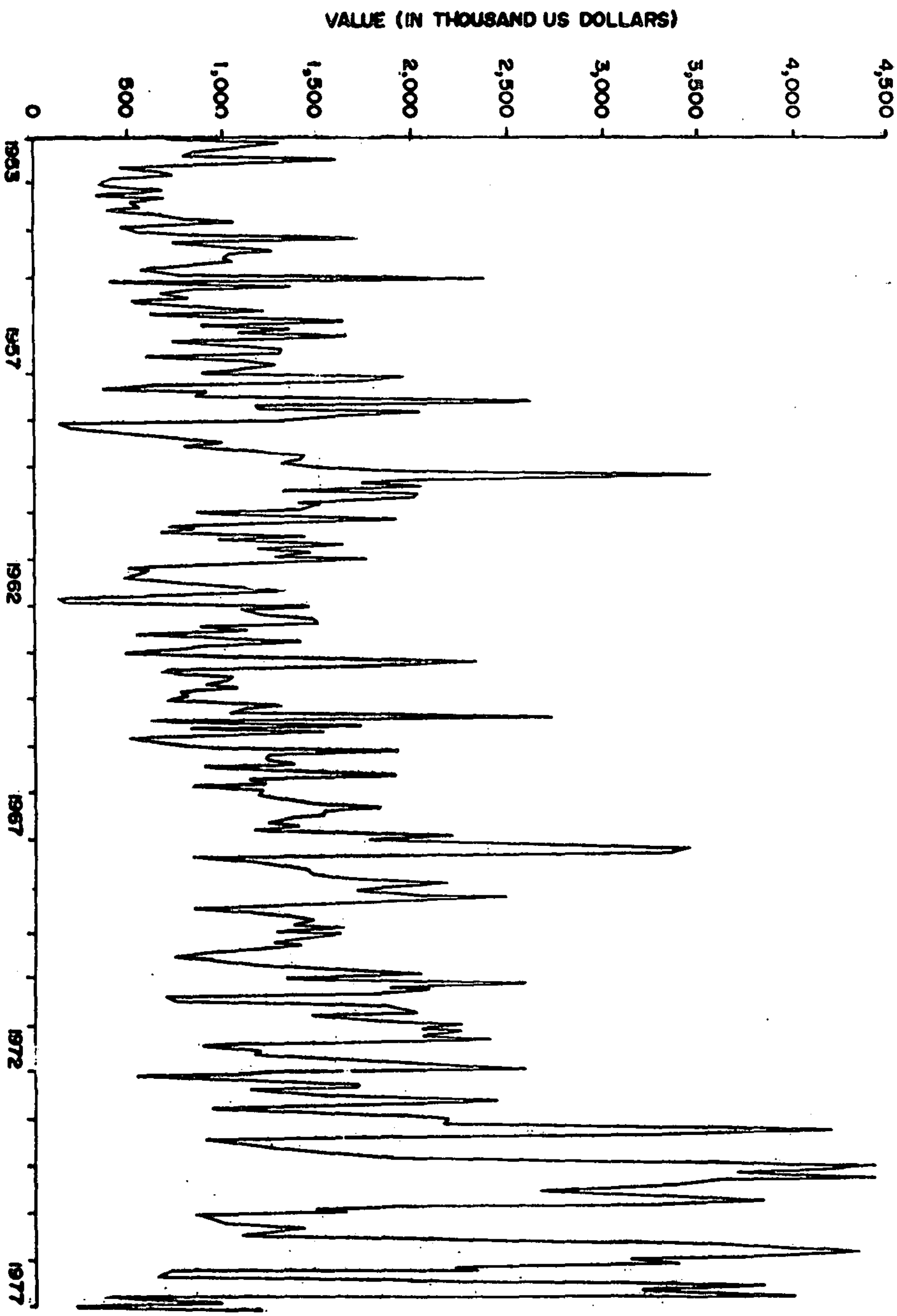

Fig.1 - PLOT OF MONTHLY IMPORT VALUES OF FISH AND FISH PREPARATION FROM 1953 TO 1977.

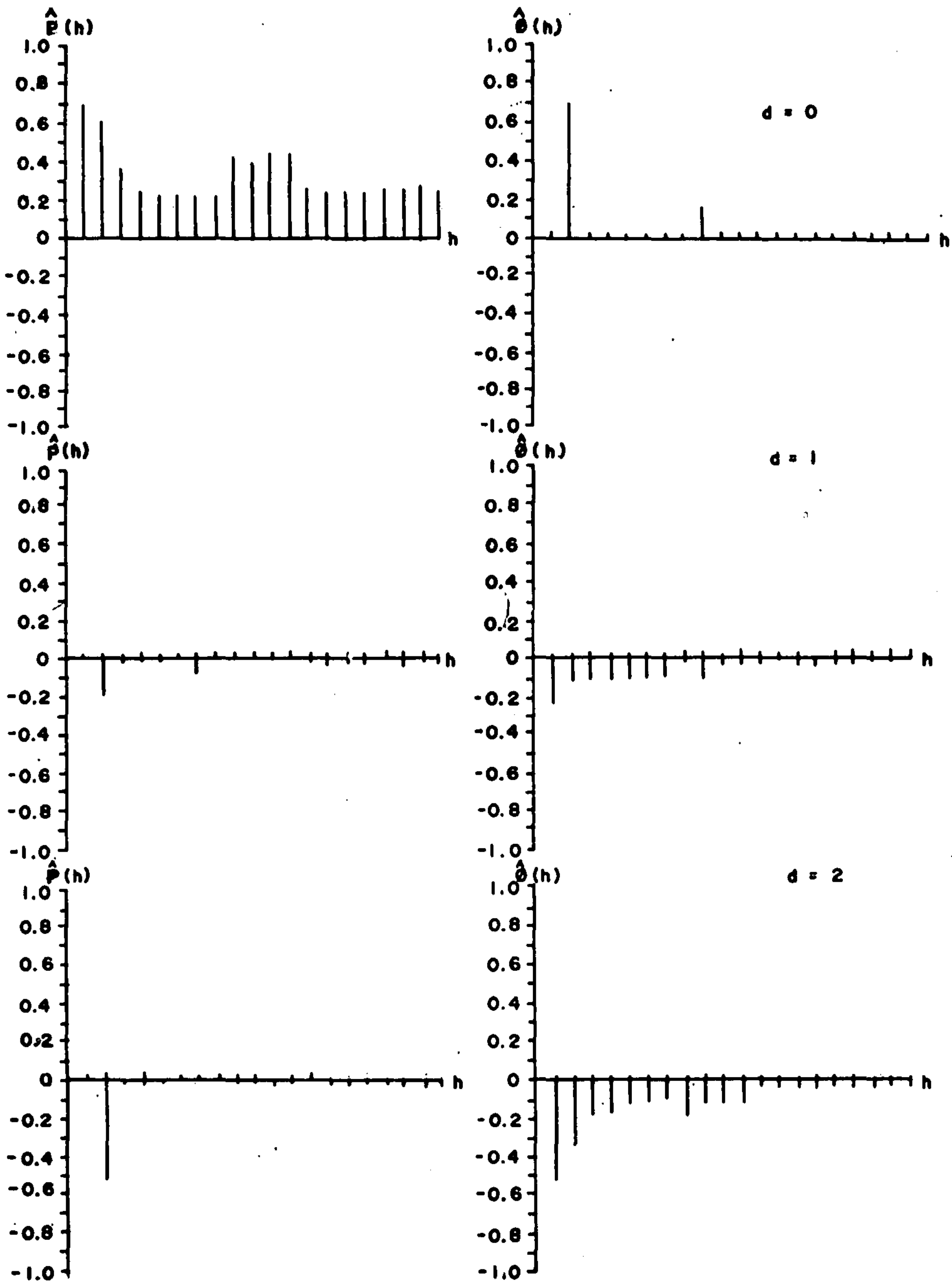

Fig. 2 - ESTIMATED AUTOCORRELATIONS AND PARTIAL AUTOCORRELATIONS OF FISH AND FISH PREPARATION - IMPORT TIME SERIES.

of the partial autocorrelations supports this guess as the values are zero after the first lag. Not much information can be observed from the first differenced series, but in the same figure for d = 2, an alternative model may be suspected. The autocorrelations have damped exponential behavior which suggests a tentative IMA(0,2,1) model.

Using the PEST command in the TSERIES package program, preliminary estimates of the parameters of the specified model for fish import were computed. There were two proposed models for this series - the AR(1) and IMA(o,2,1) models. The preliminary estimates of the first model had the form

$$Z_t = 0.6988\ Z_{t-1} + a_t$$

and the second,

$$\Delta^2 a_t = a_t - 1.3368\ a_{t-1}.$$

At this point, we can eliminate the possibility of an IMA(o,2,1) model for fish import because the estimate of θ_1, of θ, does not satisfy the invertibility condition of a MA process. To be invertible, θ_1, must lie in the range $-1 < \theta_1 < 1$. Since θ_1 which is equal to 1.3368 is greater than 1, then the model IMA (0,2,1) is not invertible.

The parameters of the identified model were estimated by the method of maximum likelihood employing the command ESTI in the same package program. The correlograms of the series indicated an AR(1) model with the following estimates,

$$Z_t = 434.5715 + 0.7007\ Z_{t-1} + a_t$$

and a residual variance of 3.86×10^5. A constant term was included in the model because the mean of the series is non-zero and d = 0. The final estimate of the model differs from the preliminary estimate by the constant term. The value of the constant term in the preliminary estimate was truncated because the number of digits computed was larger than the number specified in the program. However, the preliminary and final estimates of θ_1 are nearly the same.

The ESTI command also provided for various outputs for residual analysis and diagnostic checking of the model and it turned out that the AR(1) model fitted to fish import time series was adequate. The observed Q-value of 20.353 was not significant at both 5% and 10% levels of significance. The approximate upper bound for the standard error of a single autocorrelation is $1/\sqrt{n}$ or $1/\sqrt{300}$ or 0.06. Referring to Figure 3, $\rho(9) = 0.12$, $\rho(11) = 0.10$, $\rho(12) = 0.10$ and $\rho(19) = 0.08$ are large compared with 0.06. However, taking the results as a whole the Q-value obtained turned out to be not significant. Hence, it may be safely concluded that the AR(1) model is adequate for fish and fish preparation import series.

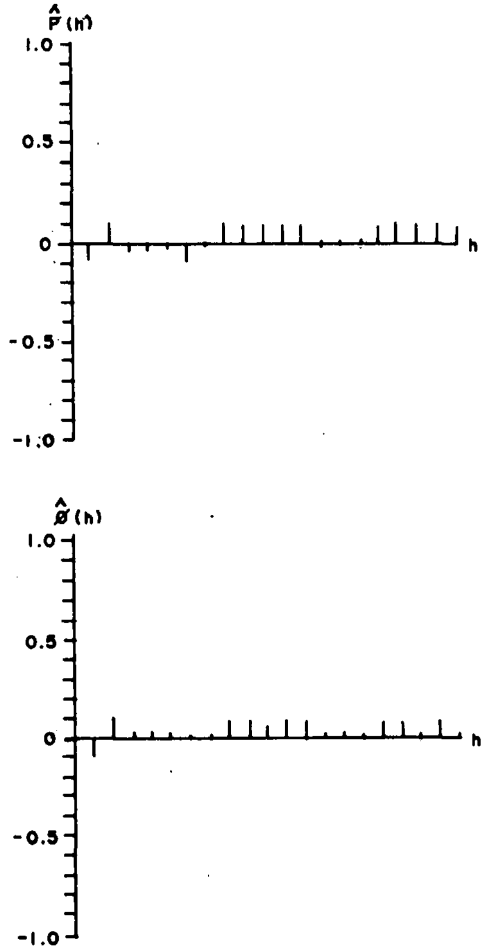

Fig. 3 - PLOT OF ESTIMATED AUTOCORRELATIONS AND PARTIAL AUTOCORRELATIONS OF RESIDUALS FROM THE A R (1, 0, 0) MODEL FITTED TO THE FISH AND FISH PREPARATION IMPORT TIME SERIES.

7. Summary

In this study, the Box-Jenkins method of model building was employed to construct a model for Philippine export and import time series of fish and fish preparation. The preliminary identification and the initial estimation of the model parameters were made using the correlograms of the autocorrelation and partial autocorrelation function as primary tools. The final estimates of the model parameters were obtained by the method of maximum likelihood using a constrained optimization method commonly referred to as the Marquadt algorithm. Diagnostic checking of the fitted model was performed using the Box-Pierce portmanteau lack-of-fit test.

8. References

Anderson, O.D. (1975) Time Series and Forecasting -- The Box-Jenkins Approach, Butterworth and Co., Redding, Massachusettes

Anderson, T.W. (1971) The Statistical Analysis of Time Series, John Wiley and Sons, New York.

Box, G.E.P. and Jenkins, G.M. (1976) Time Series Analysis, Forecasting and Control, Revised Edition, HoldenDay, San Francisco.

Chatfield, C. (1975) Analysis of Time Series, Theory and Practice, Chapman and Hall, London.

Chatfield, C. and Prothero, D.1. (1973) "Box-Jenkins Seasonal Forecasting: Problems in a Case Study", (With Discussion), Journal of Royal Statistical Society, 136; pp. 295-352.

Draper, N.Er. and Smith, H. (1966) Applied Regression Analysis, John Wiley and Sons Inc., New York.

Granger, C.W.J. and Newbold, P. (1977) Forecasting Economic Time Series, Academic Press.

Marquadt, D.W. (1963) "An Algorithm for Least Squares Estimation of Nonlinear Parameters", Journal Soc. Ind. Appl. Math. 11; pp. 431-441.

Meeker, William Q. Jr. (1977) TSERIES User's Manual, Iowa State University.

National Census and Statistics Office, (1969) "The Central Problems and Issues of Philippine Foreign Trade", Journal of Philippine Statistics, Volume 20, Number 1, First Quarter 1969, Manila, Philippines.

__________, (1970) "A Comparative Study of the Components of Import and Exports of the Philippines" 1968 and 1967", Journal of Philippine Statistics, Volume 21, Number 1, First Quarter 1970, Manila, Philippines.

__________, (1972) "Philippine Imports and Exports" 1968-1970", Journal of Philippine Statistics, Volume 23, Number 4, Fourth Quarter 1972, Manila, Philippines.

__________, (1974) "A Study of the Export and Import of the Philippines, 1971 to 1973", Journal of Philippine Statistics, Volume 25, Number 4, Fourth Quarter 1974, Manila, Philippines.

Nelson, C.R. (1973) Applied Time Series Analysis for Managerial Forecasting, Holden-Day, San Francisco.

Twenty-Five Years Economic and Financial Statistics in the Philippines (1974), Volume III, Department of Economic Research, Central Bank of the Philippines, Manila, Philippines.