J. Rubb. Res. Inst. Malaysia, **31(3)**, 188–203

A Forecasting Methodology as Applied to Rubber Prices

MOHD. NAPI BIN DAUD

The paper illustrates a statistical approach using Box and Jenkins' technique to forecast RSS 1 and RSS 2 prices. The technique developed begins with a generalised forecasting model followed by model specification namely identification, estimation and diagnostic checking.

A number of forecasting techniques have been developed depending upon the nature of the variable to be forecast and the purpose of the forecast. In this paper a technique developed by G.E.P. Box and G.M. Jenkins is used as a method to forecast the prices of two grades of rubber – RSS 1 and RSS 2.

There are two main advantages of Box and Jenkins method. First, using the traditional approach one selects arbitrarily a specific forecasting model. For example, in estimating monthly rubber price one may choose double exponential smoothing model when in fact mixed autoregressive moving average (ARMA) model would be more appropriate. Box and Jenkins' methodology starts with a broad, generalised model which takes all possible separate model combinations of moving average and autoregressive models. Using this broad model the forecaster rationally comes to the appropriate specific model. Second, the specific form of a given model has traditionally been the result of a trial-and-error procedure rather than a rational structured approach to the determination of a specific model. Box and Jenkins' structured approach eliminates various hit-and-miss tactics[1,2].

The Generalised Model

Time series data can be categorised as stationary and non-stationary data (*Figures 1* and *2*). Most are generated by a stochastic process. With a time dependent phenomenon and many unknown factors it is not possible to write a deterministic model[3].

Stationary models are based on the assumption that the process remains in equilibrium about a constant mean level. Suppose we have a stationary series having mean μ and the observations Z_t, Z_{t-1}, Z_{t-2}, Z_{t-4} ..... are taken at equal intervals. We define a_t, a_{t-1}, a_{t-2} ..... as 'white noise' or random shocks to the system. Then there are two ways to model the series as an autoregressive (AR) model; and, as a moving average (MA) model.

An autoregressive model can be expressed as:

$$\bar{Z}_t = \phi_1 \bar{Z}_{t-1} + \phi_2 \bar{Z}_{t-1} \ldots + \phi_p \bar{Z}_{t-p} + a_t \qquad \ldots 1$$

or using backshift operator B, we can write *Equation 1* as:

$$\phi(B)\bar{Z}_t = a_t \qquad \ldots 2$$

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

COMMUNICATION 727

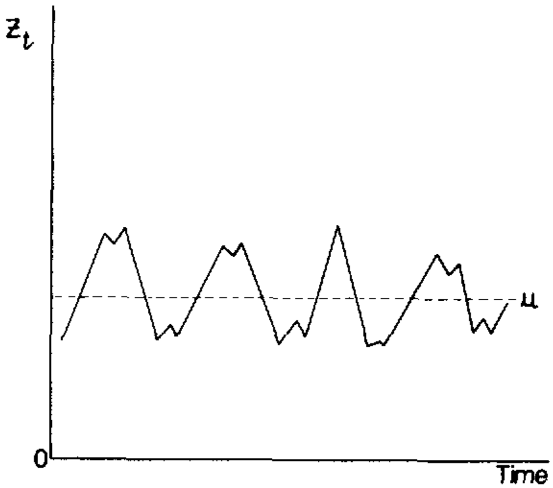

Figure 1. Stationary series.

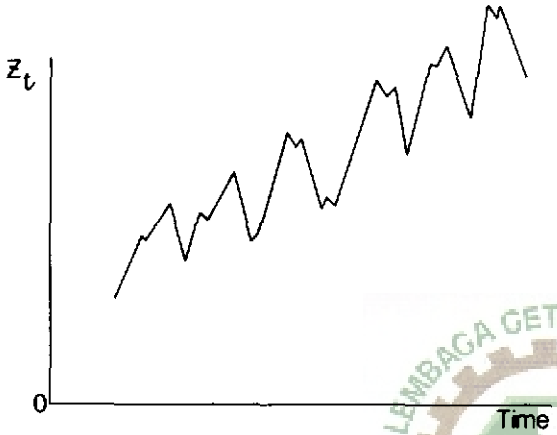

Figure 2. Non-stationary series.

$a_t \sim NI(O, \sigma a^2)$ and $\bar{Z}_t = \bar{Z}_t - \mu$

Sometimes a current deviation from the mean period t is made linearly dependent on all prior deviations back to period $(t-q)$. Therefore, the current deviation can be expressed as a linear function of the 'white noise' to the system. The model, called a moving average model can be expressed as:

$$\bar{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots \theta_q a_{t-q} \quad \dots 3$$

or $\bar{Z} = \theta_q(B)a_t$

$a_t \sim NI(O, \sigma a^2)$

Almost all of the stochastic or deterministic time series encountered in practice exhibit behaviour suggestive of non-stationarity. Thus, a forecaster should not restrict himself to either an autoregressive model or a moving average model. He should begin with a preliminary model which allows for stochastic and deterministic trend characteristics, non-seasonality and seasonality. The appropriate preliminary model to start with is an autoregressive integrated moving average model of the form:

$$\phi_p(B)(1-B)^d(1-B^s)^D Z_t = \theta_p + \theta_q(B)a_t \quad \dots 4$$

where $\theta_p = (1 - \phi_1 - \phi_2 - \dots \phi_p)\mu$

d = amount of regular differencing

s = length of a season

$\phi_p(B) = 1 - \phi_1 B^2 - \dots \phi_p B^p$

$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots \theta_q B^p$

D = amount of seasonal differencing

θ_p = represents a deterministic trend constant. Since most of the time series data are generated stochastically, the value of θ_p is normally set equal to zero.

Sometimes Z_t is in the form of logarithm or power transformation. This is to induce constant amplitude in the series over time so that the residuals from the fitted model will have a constant variance. By appropriately choosing certain levels of p, d, q and D, one can obtain an autoregressive model, moving average model or a combination of these.

For example, suppose that a particular series was generated stochastically, it was differenced and was found to arrive at stationarity at $d = 0$, $s = 4$, and $D = 1$. It was also identified that the model was an AR model of the second order, such that $p = 2$, $q = 0$, then our mathematical model, *Equation 4* will be:

$$\phi 2B\,(1-B)^{\circ}(1-B^4)^1 \bar{Z}_t = \theta_0(B)a_t \quad \ldots 5$$

From the previous definition, we see that $\phi 2(B) = 1 - \phi_1 B - \phi_2 B^2$ and $\theta_0(B) = 1$. Hence *Equation 5* becomes:

$$(1-\phi_1 B-\phi_2 B^2)(1-B^4)\bar{Z}_t = a_t \quad \ldots 6$$

Expanding

$$(1-\phi_1 B-\phi_2 B^2 -B^4 + \phi_1 B^5 + \phi_2 B^6)\bar{Z}_t = a_t \quad \ldots 7$$

then we can obtain:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + Z_{t-4} - \phi_1 Z_{t-5} - \phi_2 Z_{t-6} + a_t \quad \ldots 8$$

Methodology

The identification of a tentative model or set of models from the general class in *Equation 4* requires prior knowledge of the data pattern and the plots of sample autocorrelation and partial autocorrelation functions. The plots of autocorrelation and partial autocorrelation functions are the main indicators for identifying a model because each model has a different pattern of plot. A stationary series is required for identification because the theoretical autocorrelation and partial autocorrelation of stationary series have distinct patterns for various models. Thus, any non-stationary series must go through regular and/or seasonal differencing process until the series becomes stationary. The determination of seasonal length(s) can be made by looking at a plot of the series and where it is not easy to identify the seasonal length(s), one should try various values of s.

For each differencing pattern specified by d, D and s, Box and Jenkins calculate sample autocorrelation function, r_k of lag k as:

$$r_k = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t-k} - \bar{Z})}{\sum_{t=1}^{n} (Z_t - \bar{Z})^2} \quad \ldots 9$$

and the theoretical autocorrelation function, ρ_k ;

$$\rho_k = \frac{E(Z_t - \mu)(Z_{t-k} - \mu)}{\sigma Z^2} \quad \ldots 10$$

For the autoregressive model of order k for example, there exists an autocorrelation function ρ_j such that:

$$\rho_j = \phi_{k1}\ \rho_{j-1} + \ldots \phi_{k(k-1)}\ \rho_{j-1}\ 1 + \phi_{kk}\rho_{j-k},$$

$$j = 1, 2, 3 \ldots\ldots k \quad \ldots 11$$

where ϕ_{kk} is the last coefficient. This will lead to Yule-Walker Equation[4] which may be written in matrix form as:

$$P_k \phi_k = \rho_k \quad \ldots 12$$

The quantity ϕ_{kk} regarded as a function of lag k is the partial autocorrelation function. The estimates of partial autocorrelations as shown by Quenouille[5], of order $p + 1$ and higher are approximately independently distributed with variance:

$$\mathrm{Var}(\phi_{kk}) = \frac{1}{n}\ ; k \geqslant p + 1$$

where n = number of observation.

$$\mathrm{S.E.}(\phi_{kk}) = \frac{1}{\sqrt{n}} \quad \ldots 13$$

Generally, the theoretical autocorrelation function of the AR model will tend to damp off slowly as k increases. Meanwhile the theoretical autocorrelation function of the MA model is evidenced by Spikes in ρ_k for a particular k. In contrast, the partial autocorrelation function of the AR model drops quickly to zero after k lags, while the partial autocorrelation for the MA model tails off slowly as k increases[6].

Once a model has been identified, both ϕ_p and θ_q can be estimated by minimising the sum of squares residual term $[E(Z_t - \hat{Z}_t)^2 = \Sigma a_t^2]$, where $\hat{Z}_t$ is the estimated value of period t. This is done by using a non-linear least square estimation.

The final stage for model determination is the diagnostic checking. This involves the examination of residuals a_t. If the model is adequate these residuals will be normally and independently distributed with mean zero and constant variance, $\sigma^2 a$, (i.e. $a_t \sim NI(O, \sigma^2 a)$. Otherwise the residual autocorrelations should be re-examined, for potential improvement of the model, and the process of identification, estimation and diagnostic checking will continue until an adequate model has been found.

Case Study

This section presents rubber price data for RSS 1 and RSS 2 for the period 1970–81. The patterns of the original data are shown in *Figure 3*. Each series appears to follow a basic pattern of slow long-run growth with no seasonal fluctuation. Two irregular points appear in each series, in January 1974 and February 1980. A simple explanation would be that during that particular time the demand for natural rubber increased tremendously, but the supply of natural rubber was inelastic. Therefore, when the demand increased suddenly, the price would increase to a much higher level than the equilibrium price, but when the supply increased or the demand decreased with close substitutes, the price decreased.

Further, investigation of the data reveals that the non-constancy of error variance exists. As the yearly price average increases the error variance increases. Therefore, the original data need transformation to ensure homoskedasticity. The transformation function chosen is $Z_t = \sqrt{Y_t}$.

Identification. Each of the original series indicates that it has an increasing long-term trend. Therefore, each of the series needs at least one degree of regular differencing. *Figure 4* shows the results of the first degree of regular differencing for RSS 1 and RSS 2. The two plots suggest that the series have achieved their stationarity. Their autocorrelation and partial autocorrelation in *Figures 5, 6* and *7* suggest the more suitable application of the ARMA model. A general ARMA model is ARMA (p, d, q). Having $d = 1$ and varying p and q from 1 to 3 (*i.e.* $p = 1, 2, 3$, $q = 1, 2, 3$) there are nine combinations altogether. Therefore, in order to have a good model without going through the long process, prior knowledge of the data pattern and experience play an important part in building the specific model besides autocorrelation and partial autocorrelation. The suggested model for the two series is ARMA (2, 1, 2). The mathematical model is:

$$(1-B)(1-\phi_1 B-\phi_2 B^2)Z_t = (1-\theta_1 B-\theta_2 B^2)a_t \quad \ldots 14$$

This model, after estimating its parameters, is subject to diagnostic checking. If the diagnostic checking indicates that the residual autocorrelation is not random then one should look into the autocorrelation function for possible improvement of the model.

Estimation. Using monthly prices for RSS 1 and RSS 2 between January 1970

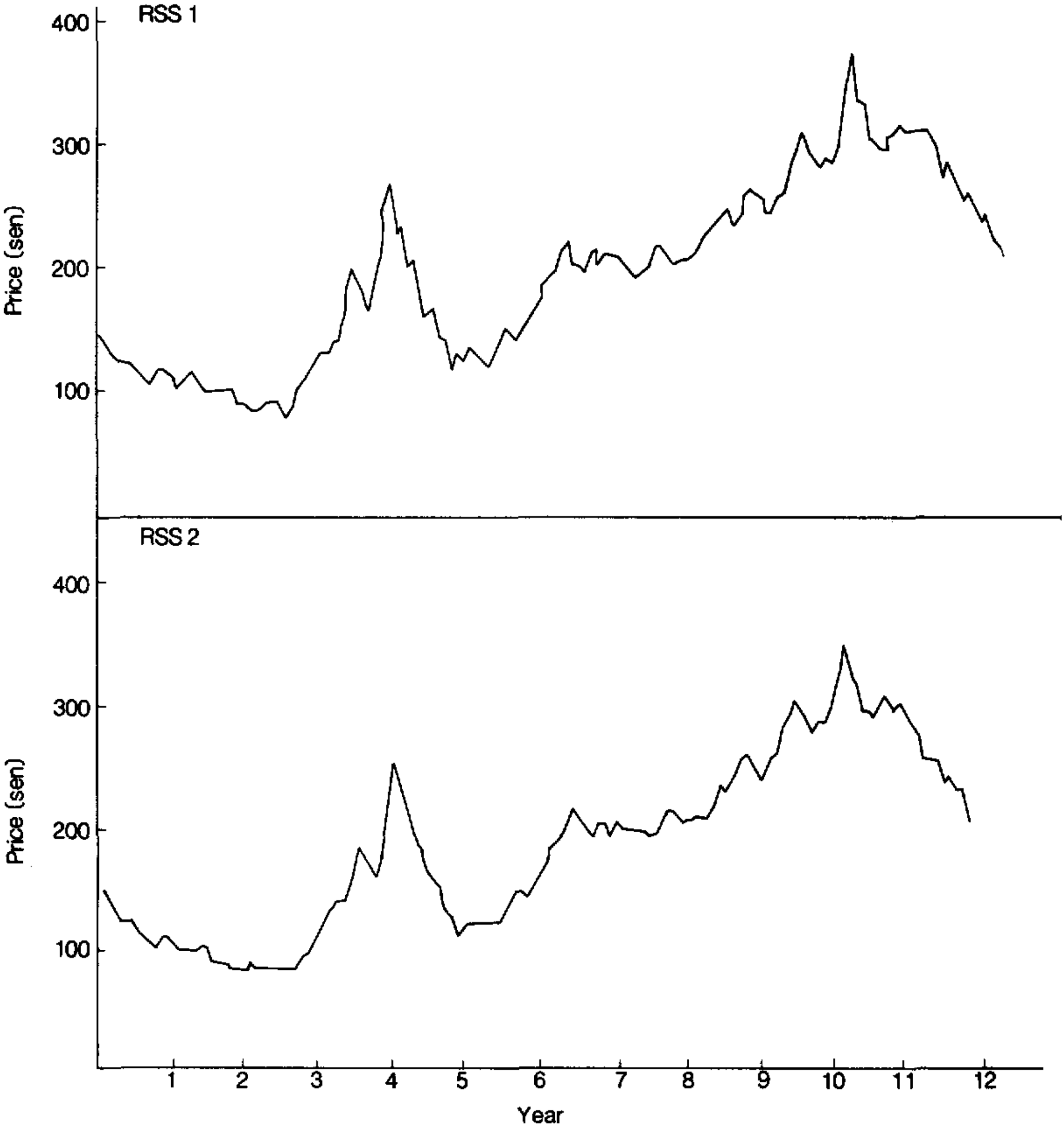

Figure 3. Plots of original data for RSS 1 and RSS 2 prices from 1970-81.

and December 1981 parametric values for *Equation 14* were estimated with an available computer routine. For RSS 1,

$$(1-B)\ (1 + 0.111B + 0.577B^2)\ Z_t = (1 + 0.398B + 0.738B^2)\ a_t \text{ with } \sigma a_t^2 = 0.18269 \qquad \ldots 15$$

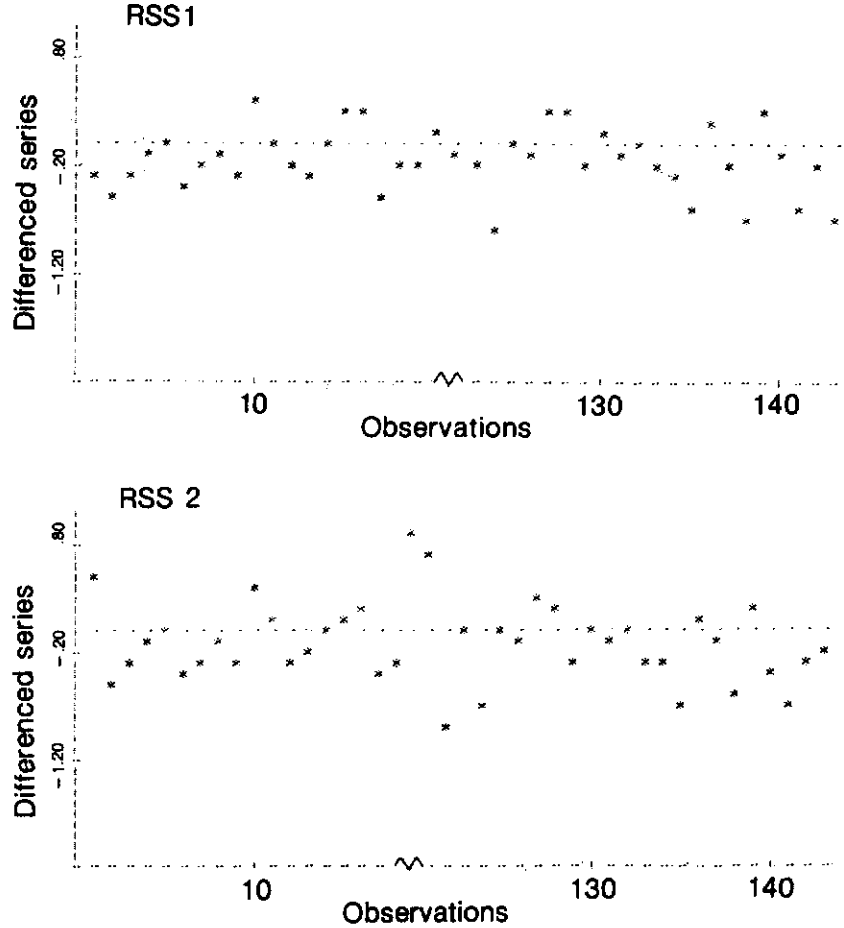

Figure 4. Plot of first degree of regular differencing for RSS 1 and RSS 2.

and for RSS 2;

$$(1-B)\ (1 + 0.025B + 0.395B^2)\ Z_t = (1 + 0.407B + 0.590B^2)\ a_t \text{ with } \sigma a_t^2 = 0.1298 \qquad \ldots 16$$

To illustrate how the models can be used for forecasting, prices for RSS 1 and RSS 2 were estimated for the next twelve months in 1982. *Tables 1* and *2* show the forecasted values, confidence intervals, actual values and percentage errors for the respective series. In these particular samples, the models appeared to predict reasonably well. The average percentage error forecasting for RSS 1 and RSS 2 for the first eight months of 1982 are 1.11% and 1.68% respectively. The graphic displays for forecasting at one time lead for RSS 1 and RSS 2 are shown in *Figures 8* and *9* respectively. The forecast points follow the original data very closely within 95% confidence limits. Whether the models are appropriate or not, however, can be determined in the diagnostic checking step.

Diagnostic checking. After the preliminary models have been identified

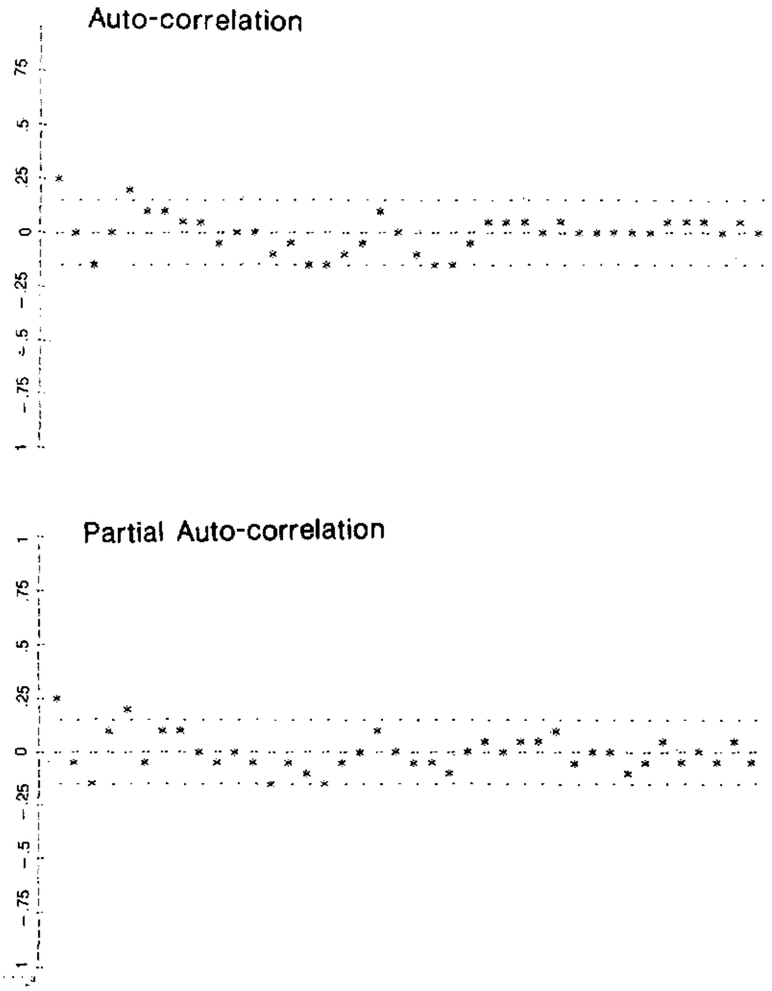

Figure 5. Autocorrelation and partial autocorrelation functions for RSS 1 with two standard error limits.

and estimated, one should examine the models to see whether they are appropriate or not, or whether any improvement can be made. The main objective of the models is to minimise the forecasting errors, a_t, into a set of independent and normally distributed errors with mean zero and constant variance. Therefore, the process of diagnostic checking involves testing whether the residual series is independent and normally distributed with mean zero and constant variance. *Figures 10* and *11* show the residual series of autocorrelation for RSS 1 and RSS 2 series respectively. The dotted lines are 95% confidence intervals. Although a number of residuals do approach the confidence intervals they do not exceed the limits. Therefore, the figures suggest that the residuals are randomly distributed. *Tables 3* and *4* are used to check the lack of fit of the models for RSS 1 and RSS 2

LAG	AUTO. CORR.	STAND ERR.
1	.346	.082
2	.065	.082
3	-.117	.082
4	-.010	.081
5	.161	.081
6	.140	.081
7	.140	.080
8	.019	.080
9	-.002	.080
10	-.047	.079
11	.012	.079
12	.018	.079
13	-.044	.079
14	-.092	.078
15	-.169	.078
16	-.165	.078
17	-.125	.077
18	-.055	.077
19	.050	.077
20	-.076	.076
21	-.205	.076
22	-.181	.076
23	-.114	.075
24	-.001	.075
25	.044	.075
26	.032	.074
27	.064	.074
28	.022	.074
29	.034	.073
30	-.049	.073
31	-.007	.073
32	.002	.073
33	.002	.072
34	.008	.072
35	.041	.071
36	.059	.071
37	.066	.071
38	.048	.070
39	.040	.070
40	-.013	.070

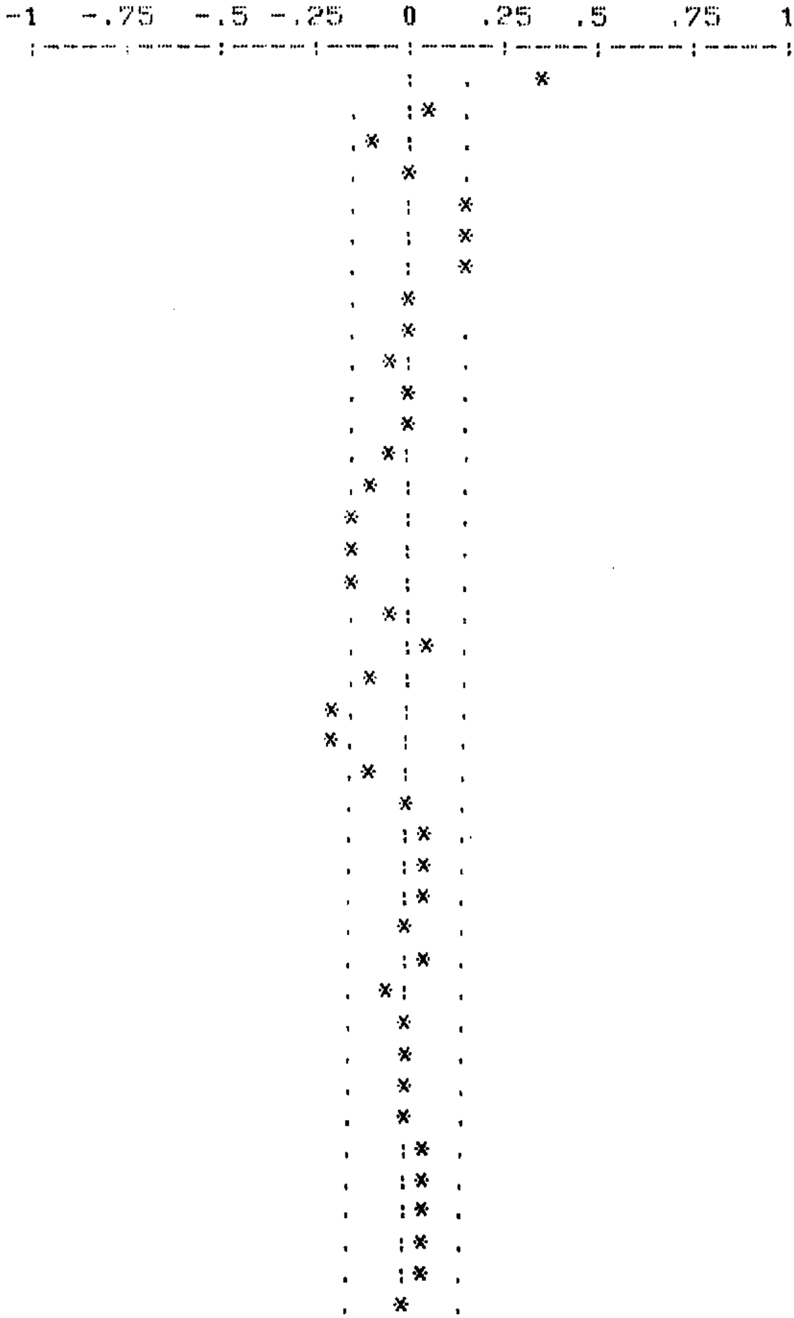

Figure 6. Autocorrelation function for RSS 2 with two standard error limits.

```
      PR-AUT  STAND.
LAG    CORR.   ERR. -1   -.75 -.5 -.25  0   .25  .5   75  1
                    |----|----|----|----|----|----|----|----|
  1     .346    .084                 .  :  .   *
  2    -.062    .084                 . *:  .
  3    -.137    .084                 *  :  .
  4     .093    .084                 .  : *.
  5     .167    .084                 .  :  *
  6     .003    .084                 .  *  .
  7     .093    .084                 .  : *.
  8    -.017    .084                 .  *  .
  9     .005    .084                 .  *  .
 10    -.054    .084                 . *:  .
 11     .027    .084                 .  :* .
 12    -.029    .084                 . *:  .
 13    -.082    .084                 .* :  .
 14    -.069    .084                 . *:  .
 15    -.115    .084                 .* :  .
 16    -.108    .084                 .* :  .
 17    -.055    .084                 . *:  .
 18    -.016    .084                 .  *  .
 19     .090    .084                 .  : *.
 20    -.105    .084                 .* :  .
 21    -.129    .084                 *  :  .
 22     .021    .084                 .  *  .
 23    -.037    .084                 . *:  .
 24    -.001    .084                 .  *  .
 25     .060    .084                 .  :* .
 26     .032    .084                 .  :* .
 27     .113    .084                 .  : *.
 28     .024    .084                 .  *  .
 29     .041    .084                 .  :* .
 30    -.105    .084                 .* :  .
 31    -.017    .084                 .  *  .
 32    -.034    .084                 . *:  .
 33    -.072    .084                 . *:  .
 34    -.044    .084                 . *:  .
 35     .035    .084                 .  :* .
 36    -.033    .084                 . *:  .
 37     .022    .084                 .  *  .
 38    -.029    .084                 . *:  .
 39     .047    .084                 .  :* .
 40    -.011    .084                 .  *  .
```

Figure 7. Partial autocorrelation function for RSS 2 with two standard error limits.

TABLE 1. FORECASTS FOR RSS 1, ORIGIN AT DECEMBER 1981 AND 95.00% CONFIDENCE LIMITS

Month, 1982	Lower confidence limit	Forecast	Upper confidence limit	Actual	Error (%)
Jan	182.03	205.35	230.08	204.50	0.42
Feb	168.43	205.76	246.83	203.75	0.99
Mar	156.14	204.65	259.71	204.00	0.32
Apr	149.23	204.53	268.54	207.50	1.43
May	144.37	205.19	276.67	205.25	0.03
Jun	138.62	205.18	248.75	204.50	0.33
Jul	133.03	204.80	292.00	200.00	2.43
Aug	128.73	204.85	298.57	199.00	2.94
Sep	124.90	205.06	304.99	–	–
Oct	120.92	205.01	311.18	–	–
Nov	117.16	204.89	317.00	–	–
Dec	113.85	204.94	322.61	–	–
Average	–	–	–	–	1.11

TABLE 2. FORECAST FOR RSS 2, ORIGIN AT DECEMBER 1981 AND 95.00% CONFIDENCE LIMITS

Month, 1982	Lower confidence limit	Forecast	Upper confidence limit	Actual	Error (%)
Jan	175.25	194.45	214.65	194.00	0.23
Feb	163.51	195.78	230.95	191.50	2.23
Mar	151.25	194.18	242.46	191.75	1.27
Apr	144.02	193.69	250.71	192.25	0.75
May	139.30	194.34	258.52	192.00	1.22
Jun	134.31	194.51	265.83	191.25	1.70
Jul	129.33	194.25	272.34	188.75	2.91
Aug	125.09	194.19	278.42	188.25	3.16
Sep	121.37	194.29	284.30	–	–
Oct	117.79	194.32	289.89	–	–
Nov	114.38	194.28	295.20	–	–
Dec	111.22	194.27	300.33	–	–
Average	–	–	–	–	1.68

respectively. The calculation for testing lack of fit is based on the assumption that:

$$Q = n \sum_{i=1}^{k} r_i^2 \; ; \; i = 1, 2, \ldots\ldots k \,, \qquad \ldots 17$$

where $n = N - d$, is approximately distributed as chi-square distribution with *(k-p-q)* degree of freedom[7], if the fitted model is appropriate. The fourth column in *Tables 3* and *4* indicate the levels of significance or probabilities. There is no lag that is significant at either

OBS	DATA	
96	198.750	
97	202.250	
98	205.250	
99	207.500	
100	207.750	
101	217.500	
102	235.500	
103	233.250	
104	241.000	
105	253.750	
106	259.000	
107	253.250	
108	239.000	
109	237.500	
110	250.500	
111	256.750	
112	279.250	
113	292.500	
114	304.250	
115	292.000	
116	284.750	
117	281.500	
118	285.750	
119	289.500	
120	295.500	
121	332.750	
122	369.250	
123	328.000	
124	322.500	
125	293.250	
126	294.750	
127	290.000	
128	300.250	
129	309.500	
130	303.000	
131	306.750	
132	305.000	
133	304.500	
134	299.000	
135	288.500	
136	267.250	
137	272.500	
138	267.000	
139	243.500	
140	254.000	
141	250.250	
142	230.500	
143	224.250	
144	203.500	
145	205.349	F

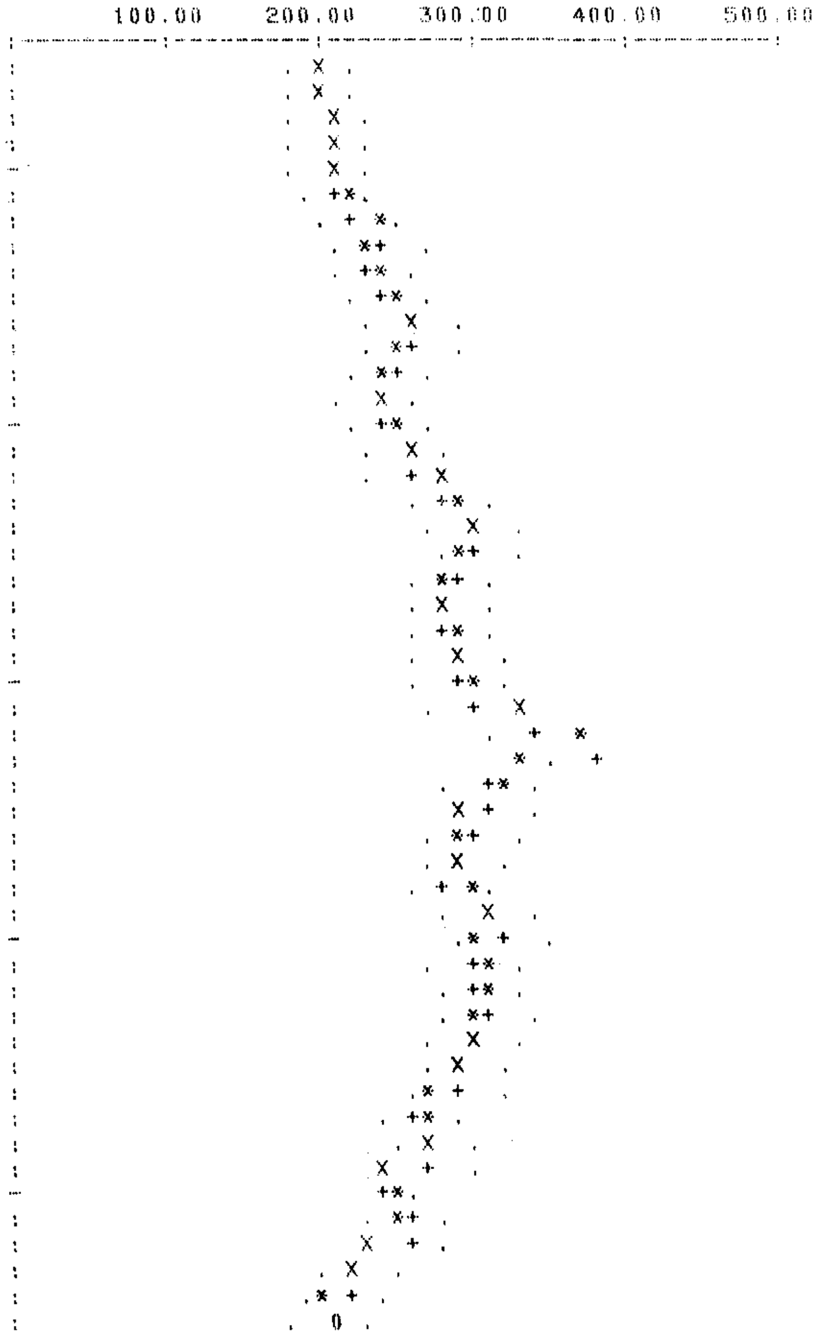

Figure 8. Graphic display for forecasting at one time lead for RSS 1.

OBS	DATA	
96	197.750	
97	200.250	
98	203.500	
99	203.750	
100	204.750	
101	212.750	
102	228.750	
103	226.500	
104	233.750	
105	248.000	
106	253.000	
107	249.000	
108	238.000	
109	236.500	
110	248.000	
111	254.750	
112	274.750	
113	286.000	
114	297.750	
115	285.500	
116	279.000	
117	275.000	
118	278.250	
119	281.500	
120	288.000	
121	320.500	
122	345.250	
123	312.750	
124	312.000	
125	288.000	
126	289.500	
127	285.500	
128	295.000	
129	302.000	
130	292.750	
131	294.000	
132	291.500	
133	290.750	
134	281.750	
135	271.750	
136	248.000	
137	252.250	
138	250.500	
139	231.500	
140	237.250	
141	226.500	
142	205.750	
143	196.500	
144	190.500	
145	194.454	F

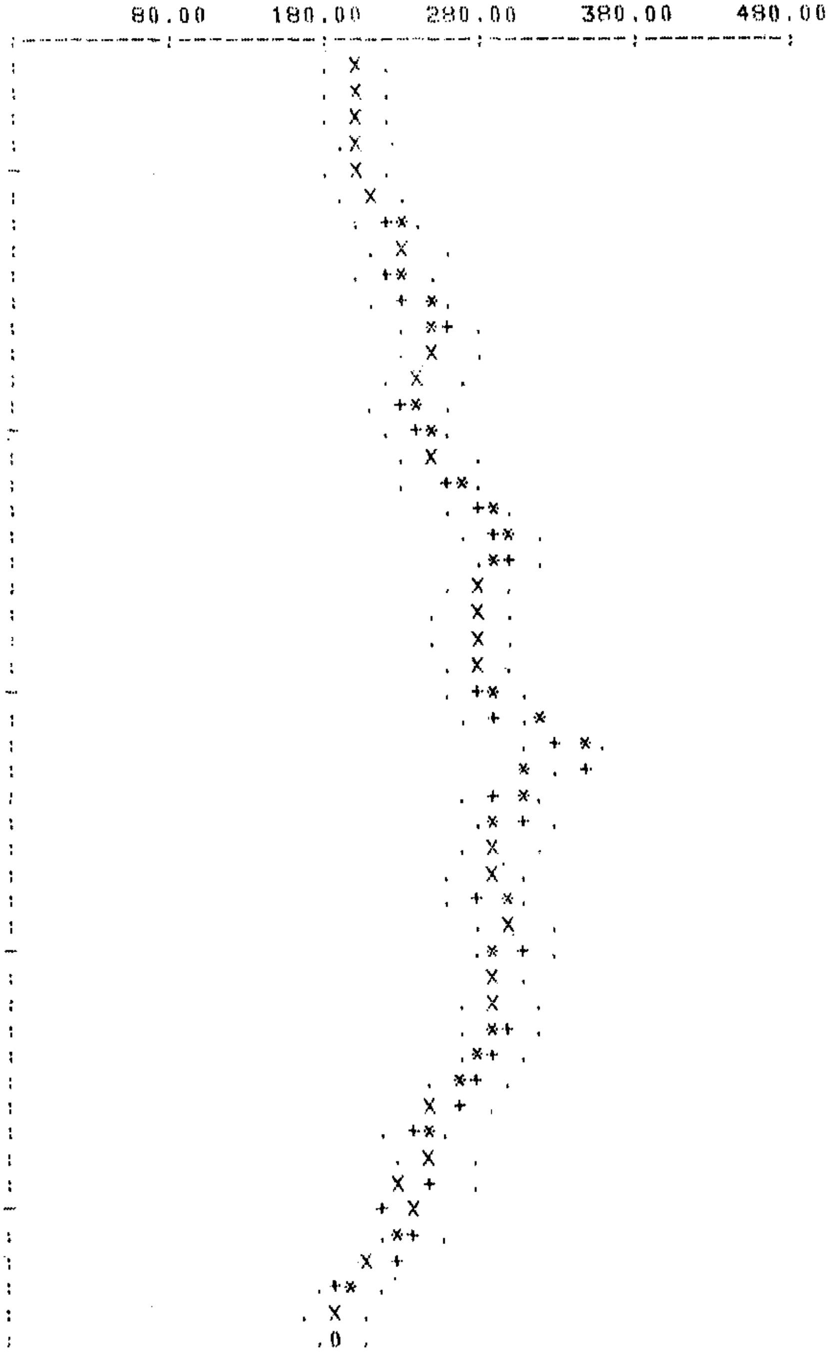

Figure 9. Graphic display for forecasting at one time lead for RSS 2.

```
      AUTO.  STAND.
LAG   CORR.    ERR.   -1  -.75  -.5 -.25  0   .25   .5  .75   1
                      |----|----|----|----|----|----|----|----|
  1   -.028    .083                    . *:  .
  2   -.021    .082                    .  *  .
  3    .010    .082                    .  *  .
  4   -.016    .082                    .  *  .
  5    .094    .082                    .  : *.
  6    .040    .081                    .  :* .
  7    .108    .081                    .  : *.
  8    .025    .081                    .  *  .
  9    .021    .080                    .  *  .
 10   -.037    .080                    . *:  .
 11    .006    .080                    .  *  .
 12   -.009    .079                    .  *  .
 13   -.090    .079                    .* :  .
 14   -.013    .079                    .  *  .
 15   -.094    .078                    .* :  .
 16   -.121    .078                    .* :  .
 17   -.069    .078                    . *:  .
 18   -.071    .077                    . *:  .
 19    .107    .077                    .  : *.
 20   -.034    .077                    . *:  .
 21   -.101    .077                    .* :  .
 22   -.082    .076                    .* :  .
 23   -.132    .076                    *  :  .
 24   -.036    .076                    . *:  .
 25    .034    .075                    .  :* .
 26    .047    .075                    .  :* .
 27    .014    .075                    .  *  .
 28    .010    .074                    .  *  .
 29    .084    .074                    .  : *.
 30   -.066    .074                    . *:  .
 31    .027    .073                    .  :* .
 32    .018    .073                    .  *  .
 33   -.001    .073                    .  *  .
 34   -.019    .072                    .  *  .
 35    .038    .072                    .  :* .
 36    .044    .071                    .  :* .
 37    .050    .071                    .  :* .
 38   -.050    .071                    . *:  .
 39    .064    .070                    .  :* .
 40    .019    .070                    .  *  .
```

Figure 10. Residual autocorrelation function for RSS 1 with two standard error limits.

```
     AUTO.  STAND.
LAG  CORR.    ERR.  -1   -.75   -.5  -.25    0    .25   .5    .75    1
                     :----:----:----:----:----:----:----:----:
  1    .004    .083                     .    *    .
  2   -.014    .082                     .    *    .
  3    .027    .082                     .    :*   .
  4   -.016    .082                     .    *    .
  5    .096    .082                     .    :   *.
  6    .048    .081                     .    :*   .
  7    .119    .081                     .    :   *.
  8    .019    .081                     .    *    .
  9   -.001    .080                     .    *    .
 10   -.062    .080                     .  *:     .
 11    .026    .080                     .    :*   .
 12    .016    .079                     .    *    .
 13   -.035    .079                     .  *:     .
 14   -.043    .079                     .  *:     .
 15   -.145    .078                     *    :    .
 16   -.064    .078                     .  *:     .
 17   -.079    .078                     .*   :    .
 18   -.088    .077                     .*   :    .
 19    .093    .077                     .    :   *.
 20   -.061    .077                     .  *:     .
 21   -.167    .077                     *    :    .
 22   -.079    .076                     .*   :    .
 23   -.081    .076                     .*   :    .
 24    .000    .076                     .    *    .
 25    .038    .075                     .    :*   .
 26    .008    .075                     .    *    .
 27    .048    .075                     .    :*   .
 28    .001    .074                     .    *    .
 29    .045    .074                     .    :*   .
 30   -.071    .074                     .  *:     .
 31    .030    .073                     .    :*   .
 32   -.000    .073                     .    *    .
 33    .004    .073                     .    *    .
 34   -.006    .072                     .    *    .
 35    .025    .072                     .    :*   .
 36    .065    .071                     .    :*   .
 37    .014    .071                     .    *    .
 38    .000    .071                     .    *    .
 39    .073    .070                          :*   .
 40    .004    .070                     .    *    .
```

Figure 11. Residual autocorrelation for RSS 2 with two standard error limits.

TABLE 3. DIAGNOSTIC CHI-SQUARE FOR RESIDUAL SERIES RSS 1

Lag	Chi-square	Df	Probability
6	1.75	2	0.4164
12	3.91	8	0.8655
18	10.57	14	0.7194
24	18.73	20	0.5395
30	21.46	26	0.7176
36	22.36	32	0.8973
40	24.21	36	0.9329

TABLE 4. DIAGNOSTIC CHI-SQUARE FOR RESIDUAL SERIES RSS 2

Lag	Chi-square	Df	Probability
6	1.87	2	0.3918
12	4.81	8	0.7775
18	11.57	14	0.6405
24	20.48	20	0.4281
30	22.44	26	0.6643
36	23.56	32	0.8598
40	24.66	36	0.9235

5% or 10% level (*i.e.* $\alpha = 0.05$ or $\alpha = 0.10$). Thus the residual series of autocorrelation and chi-square tests cast no doubt that the models are appropriate.

DISCUSSION AND VALIDITY OF THE MODELS

The models that are obtained and presented in this paper are univariate stochastic models. This simple approach cannot be expected to produce very accurate forecasts over long-term periods but they may be useful for short-term forecasting, and these models are important for the following reasons[8] :

- It may be impossible to obtain variables related to the variable being forecast, leaving univariate models as a choice.
- The development of univariate models provides a 'yardstick' with which more sophisticated models can be compared whenever the related variables may be used to improve the accuracy of the forecasts.
- Univariate models serve as a tool for screening data during the early stages for an analyst if the cause of large residuals can be identified.

The models developed are not ultimate models. The long-term forecasts for rubber prices may need to consider a few related substitutes, demand for rubber products, *etc.* Yet, the variables related to dependent variables can be incorporated into a multivariate transfer function which is equivalent to the simultaneous equation system.

We observe that the 95% confidence limits in *Tables 1* and *2* become wider as the time lead increases. This is a reflection of the increase in forecast error variance as the time lead increases. One may suggest that updating the parameter estimates as new data become available usually leads to better forecasts. Another way of obtaining better forecasts is by updating the forecast to a new origin $(t + 1)$ whenever a new deviation is available.

Rubber Research Institute of Malaysia
Kuala Lumpur
November 1982

REFERENCES

1. THOMSON, H.E. AND TIAO, G.C. (1971) Analysis of Telephone Data: A Case Study of Forecasting Seasonal Time Series. *Bell J. Econ. Mgmt. Sci.*, 515.
2. MABERT, V.A. AND RAOCLIFFE, R.C. (1974) A Forecasting Methodology as Applied to Financial Times Series. *The Accounting Review.*
3. BOX, G.E. AND JENKINS, G.M. (1971) *Time Series Analysis: Forecasting and Control.* San Francisco: Holden – Day.

4. WALKER, G. (1931) On Periodicity in Series of Related Term. *Proc. Royal Soc.*, A 131, 518.

5. QUENOUILLE, M.H. (1949) Approximate Test of Correlation in Time Series. *J. Royal Stat. Soc.*, B11, 68.

6. MAKRIDAKIS, S. AND WHEELWRIGHT, S. (1978) *Forecasting Methods and Applications*, pp 334 and 336. New York: John Wiley & Sons.

7. BOX, G.E.P. AND PIERCE, D.A. (1970) Distribution of Residual Auto-Correlations in Autoregressive-Integrated Moving Average Time Series Models. *J. Am. Statist. Ass.*, **64**, 1509.

8. JENKINS, G.M. (1979) Practical Experience with Modelling and Forecasting Time Series. *Forecasting (Anderson, O.D. ed.)*. Amsterdam, N. York: Nor-Holland Publishing Co.