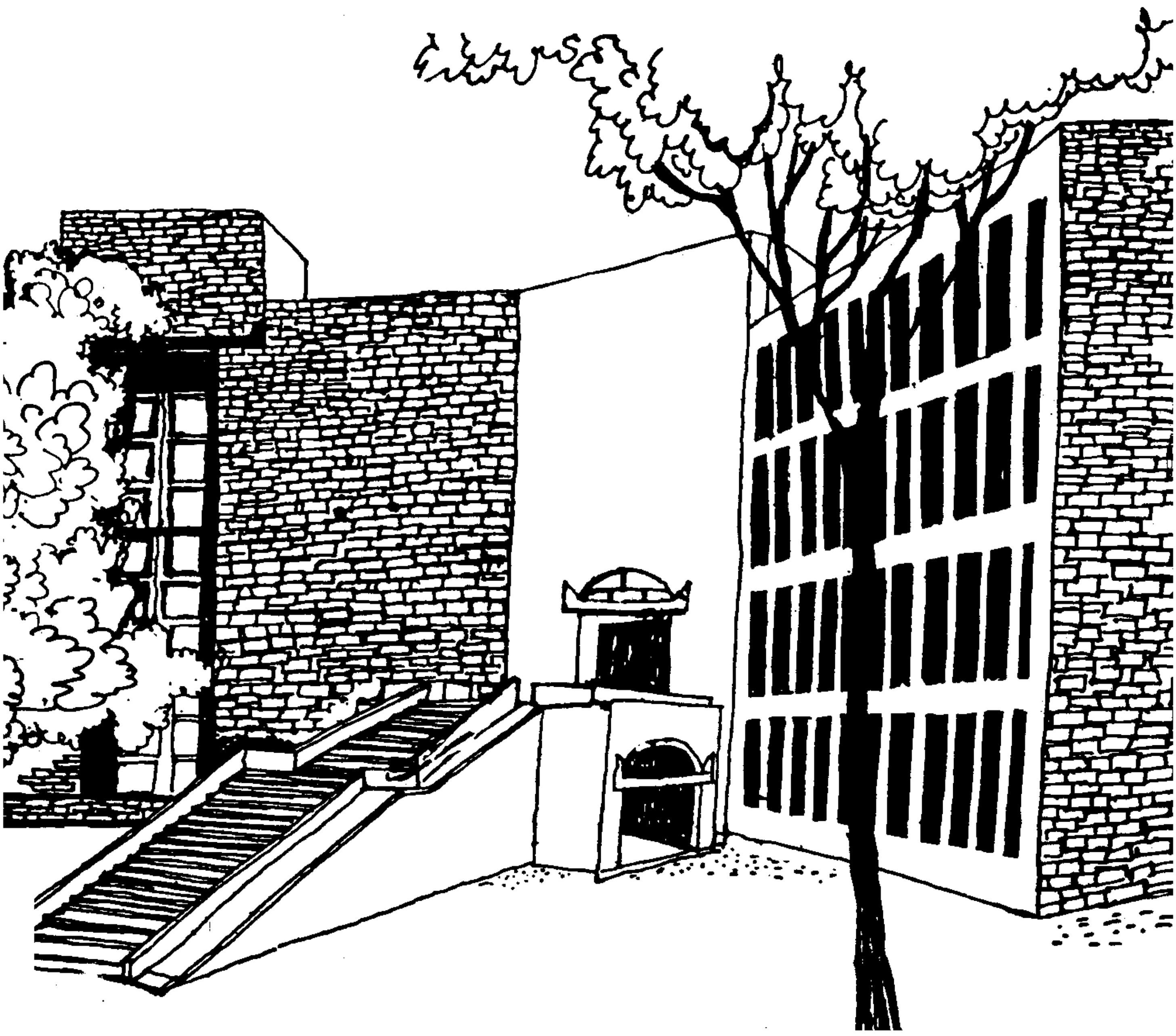
विद्याविनियोगाद्विकासः

IIM AHMEDABAD

Working Paper



ARIMA MODEL FOR AND FORECASTS ON TEA PRODUCTION IN INDIA

By

G. S. Gupta

W P No. 1097
April 1993

The main objective of the working paper series of the IIMA is to help faculty members to test out their research findings at the pre-publication stage.

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 015
INDIA

ARIMA Model For And Forecasts on Tea Production in India

G.S. GUPTA *

Autoregressive integrated moving average (ARIMA) model was advanced by Box and Jenkins (hence also known as Box-Jenkins' model) in 1960s for forecasting a variable. Though it has become quite popular in the West, its application to Indian data is still rare. This is basically because it is quite complicated and its appropriate use requires long time series data. With the requisite softwares, as well as the availability of reliable and long time series data now even for India, resort to this method is highly desirable. An effort is made in this paper to develop an ARIMA model for tea production in India and to apply the same in forecasting the variable under question.

ARIMA method is an extrapolation method for forecasting and, like any other such method, it requires only the historical time series data on the variable under forecasting. Among the extrapolation methods, this is the most sophisticated method, for it incorporates the features of all such methods, does not require the investigator to choose the initial values of any variable and values of various parameters a priori or through iteration, and it is robust to handle any data pattern. As one would expect, this is quite a difficult model to develop and apply as it involves transformation of the variable, identification of the model, estimation through the non-linear method, verification of the model and derivation of forecasts. In what follows, we first explain the ARIMA model, then develop the same for tea production using monthly data for India during January 1979 through July 1991, and finally apply the same to forecast the values of the variable during the future 12 periods.

1. ARIMA Model:

In its general form, the ARIMA model is expressed as follows:

ARIMA (p, d, q) (P, D, Q) ^{s}

Where
- p = order of non-seasonal autoregression (AR)
- d = order of non-seasonal difference
- q = order of non-seasonal moving average (MA)
- P = order of seasonal AR
- D = order of seasonal difference
- Q = order of seasonal MA
- s = length of season (=4 in quarterly data, 12 in monthly data, and so on).

---

* The author acknowledges the assistance from Mrs. Hemlatha Ramani and Mr. H. Keshava in terms of some data collection and computer runs for this paper.

If X denotes the variable, the model could be expressed in the form of an equation as below:

$$(1-a_1 B - a_2 B^2 - \ldots - a_p B^p)(1-a^*_1 B^s - a^*_2 B^{2s} - \ldots - a^*_P B^{Ps})$$
$$(1-B)^d (1-B^s)^D X_t = (1-b_1 B - b_2 B^2 - \ldots - b_q B^q)(1-b^*_1 B^s - b^*_2 B^{2s} - \ldots - b^*_Q B^{QS}) e_t \qquad \ldots\ldots\ldots(1)$$

which can be condensed as

$$a_p(B)\, a^*_P(B^S)\,(1-B)^d (1-B^s)^D X_t = b_q(B)\, b^*_Q(B^S)\, e_t \qquad \ldots\ldots\ldots(2)$$

where
- X = Variable under forecasting
- B = lag operator
- e = error term ($=X-\hat{X}$, where $\hat{X}$ is the estimated value of X)
- t = time subscript
- $a_p(B)$ = non-seasonal AR
- $a^*_P(B^S)$ = seasonal AR
- $(1-B)^d$ = non-seasonal differences
- $(1-B^s)^D$ = seasonal differences
- $b_q(B)$ = non-seasonal MA
- $b^*_Q(B^S)$ = seasonal MA

as, a^*s, bs and b^*s are parameters.

The model as expressed in equation 1 or 2 contains p+q+P+D parameters, which need to be estimated. The model is non-linear in parameters so long as either the data on X have both seasonal as well as non-seasonal elements or/and the model contains the moving average component (ie. bs and b^*s are non-zeros). Since the model is quite complicated, its easy and most popular form, viz. ARIMA $(1,1,1)(1,1,1)^4$ is elaborated below:

$$(1-a_1 B)(1-a^*_1 B^4)(1-B)(1-B^4)X_t = (1-b_1 B)(1-b^*_1 B^4)\, e_t \qquad \ldots\ldots(3)$$

Expansion of equation (3) yields

$$[1-(1+a_1)B-a_1 B^2+(1+a^*_1)B^4 - (1+a_1+ a^*_1+a_1 a^*_1)B^5$$

$$+(a_1 + a_1 a^*_1)B^6 - a^*_1 B^8 +(a^*_1 + a_1 a^*_1)B^9- a_1 a^*_1)B^{10}] X_t$$

$$= [1 - b_1 B - b^*_1 B^4 + b_1 b^*_1 B^5] e_t$$

which on further expansion yields

$$X_t = (1+a_1) X_{t-1} - a_1 X_{t-2} + (1+a^*_1)X_{t-4} - (1+a_1+a^*_1+a_1 a^*_1)X_{t-5}$$

$$+ (a_1+a_1 a^*_1)X_{t-6} - a^*_1 X_{t-8} +(a^*_1+ a_1 a^*_1)X_{t-9} - a_1 a^*_1 X_{t-10}$$

$$+ e_t - b_1 e_{t-1} - b^*_1 e_{t-4} + b_1 b^*_1 e_{t-5} \qquad \ldots\ldots(4)$$

Thus, in the simple model of equation (3), though there are only four parameters (a_1, a^*_1, b_1, b^*_1), the corresponding forecast equation has many more terms. The coefficients of various terms are made up of either single parameters or of some combinations of parameters. The presence of lagged values of X variable indicates the autoregressive component and of error terms the moving average component.

Given the parameters' values and historical data on X variable, forecasts for X_t can be obtained through equation (4). Forecasts for periods beyond 't' are obtained through the procedure of boot strapping, wherein the estimated value for X_t is used in generating forecast for X_{t+1}, for X_t and X_{t+1} are used for forecasting X_{t+2}, and so on.

2. ARIMA Model For Tea Production:

Development of ARIMA model for any variable involves three steps:

identification
estimation
verification

Each of these steps is now explained for tea production.

2.1 Model Identification:

Identification is concerned with deciding the appropriate values for p, d, q, P, D and Q. This is done through two stages. In the first stage, the values for d and D are decided and in the second stage those for p, q, P and Q. Details on these follows:

(a) ARIMA model is estimated only after transforming the variable under forecasting into a stationary series. The stationary series is the one whose values vary over time only around a constant mean and constant variance. There are several ways to ascertain this, an application of which to tea production data (Table-1) follows.

One way to check stationarity is to just compute mean, and compare it with the minimum and maximum values of the variable during the sample period. Mean of tea production stands at 524 and its range between 61 and 970 (standard deviation=281). Since these reflect a wide fluctuation, the variable is not stationary. The variable is also seasonal, for the means of various months data range between 102 (Feb.) and 839 (Sept.).

The second method of checking stationarity is through examining the graph of the data. Figure-1 reveals wide fluctuations seasonally (as there are peaks and troughs each year) and somewhat rising trend non-seasonally.

There is a statistical method also to ascertain stationarity. This is through the computation of autocorrelation coefficients of various orders. The values of these autocorrelations for tea production are provided in Table-2. As many of these are significantly different from zero, tea production is far from stationary. Further, a careful examination of the autocorrelations would reveal that these are higher for orders 12, 24 and 36 as compared to other orders and neighbouring values. This indicates that the series is non-stationary seasonally (monthly) as well.

Non-stationarity in mean is corrected through appropriate differencing of the data. Since the variable under study is non-seasonally as well as seasonally non-stationary, one needs to take both these differences once or more until the stationarity is achieved.

Thus, if X denotes the original value, non-seasonal difference is given by

$$Y_t = X_t - X_{t-1}$$

and then seasonal (monthly) difference is given by

$$Z_t = Y_t - Y_{t-12}$$
$$= (X_t - X_{t-1}) - (X_{t-12} - X_{t-13})$$

Or, $Z_t = X_t - X_{t-1} - X_{t-12} + X_{t-13}$ .... (4)

The newly constructed variable Z_t could now be examined for stationarity and if it is still non-stationary, one can go on taking successive non-seasonal and seasonal differences until stationarity is achieved. The autocorrelation coefficients of various orders for Z_t are contained in Table-3. The results reveal that while the first autocorrelation is significant, they drop significantly thereafter until seasonal factor gets repeated. Thus, variable Z_t appears to be quite stationary.

Since each of the non-seasonal as well as seasonal difference was carried out only once to arrive at a stationary series, the value for each of d and D in the ARIMA model is unity.

In case the original or the transformed variable has a non-stationary variance, the transformation required for its correction would be to either work with the logarithm of the variable or some power function (e.g. square or cube of the variable) of the variable. The graph of tea production data in Figure-1 does not reveal any non-stationarity in variance and so there was no need to perform this transformation for tea production.

(b) There is need to specify the orders of autoregression (AR) and moving averages (MA) before proceeding with the estimation of ARIMA model. This can be done in two ways. One, choose separately these orders for non-seasonal (p, q) and seasonal components (P, Q). Two, choose these orders for the whole series, ignoring non-seasonal and seasonal components. The TSP (Time series Processor) package which we have used in estimating the ARIMA model follows the second procedure and accordingly that is the one explained below.

In order to choose the appropriate values for the orders of AR and MA, we need the plot of the autocorrelations and partial autocorrelations of the transformed (stationary) variable (Z). The same is available in Table-3. The autocorrelation graph helps choosing the appropriate values for MA and partial autocorrelation graph those for AR. The rule in this regard is simple. Whichever of these values are significantly different from zero, the corresponding are the orders for MA and AR. A careful examination of autocorrelation graph would reveal that these are highly significant for order 1 and somewhat significant for orders 12, 13 and 24. Thus, the model could be MA(1, 12, 13, 24) or some other combinations of these. Similarly, partial autocorrelation graph reveals that while it is highly significant for order 1, it is also reasonably significant for orders 2, 4 and 23. Thus, the appropriate AR model could be AR(1), or AR(1,2,4, 23) or some other combination of these.

This completes the identification of the ARIMA model. To summarize, d=1, D=1, and there are alternatives for AR and MA. The alternatives are:

AR(1) and MA(1)
AR(1,2,4,23) and MA(1,12,13,24)
any other combinations of above orders.

.2 Model Estimation:

The most popular estimation method for a single equation is the ordinary least-squares (OLS) method. However, OLS is not an appropriate method for an ARIMA model. This is for two reasons.

(a) OLS method is available for models with linear parameters only and the ARIMA model contains non-linear parameters if it has non-zero MA component. The latter can be seen even in its simplest version, viz. ARIMA (1,0,1):

$$\hat{X}_t = a_0 + a_1 X_{t-1} + b_1 e_{t-1} \qquad \ldots\ldots\ldots(6)$$

Where a_0 is the constant term, which can even be taken as zero. substituting for e_{t-1} in equation (6), we have

$$\hat{X}_t = a_0 + a_1 X_{t-1} + b_1 (X_{t-1} - \hat{X}_{t-1})$$

Again substituting for $\hat{X}_{t-1}$, from one period lagged version of equation (6), we get

$$\hat{X}_t = a_0 + a_1 X_{t-1} + b_1 X_{t-1} - b_1[a_0 + a_1 X_{t-2} + b_1 e_{t-2}]$$

Or,

$$\hat{X}_t = a_0 (1-b_1)+(a_1 + b_1)X_{t-1} - a_1 b_1 X_{t-2} + b_1^2 e_{t-2} \quad \ldots(7)$$

In equation (7), the intercept term as well as the coefficients of X_{t-2} and e_{t-2} are non-linear.

(b) OLS method does not yield unbiased and consistent estimates when the explanatory variables include lagged endogenous variables. Since the AR component in ARIMA model specifies lagged endogenous variables as explanatory variables, the OLS method is inappropriate even in the absence of MA component in the ARIMA model.

In view of the above difficulties, the OLS method is inappropriate for estimating an ARIMA model. Marquardt (1963) has designed a powerful algorithm for estimating ARIMA models through iterative improvement. where some preliminary estimates are first chosen and then the computer programme refine them iteratively so as to minimize the sum of squared residuals. The TSP package contains this procedure and the same has been used for developing the ARIMA model for tea production in India.

The alternative models identified above under Section 2.1 have been estimated through the Marquardt procedure using TSP package and the results are provided in Table-4. Comparing the three alternative models on the basis of statistics such as $\bar{R}^2$, t-values and Durbin Watson values, one finds that the most elaborate form is the most representative ARIMA model for tea production in India. Thus, the estimated ARIMA model for tea production in India is

$$\hat{Z}_t = -5.0629 - 0.6384\,Z_{t-1} - 0.2823\,Z_{t-2} - 0.0692\,Z_{t-4} + 0.1221\,Z_{t-23} - 0.1166\,e_{t-1} - 0.5350\,e_{t-12} + 0.2135\,e_{t-13} - 0.5132\,e_{t-24} \qquad \ldots\ldots(8)$$

2.3 Model Vertification:

The model verification is concerned with checking the residuals of the model to see if they contain any systematic pattern which can still be removed to improve on the chosen ARIMA. This is done through examining the autocorrelations and partial autocorrelations of the residuals of various orders, both individually and collectively. For this purpose, the various correlations upto 25 lags were computed and the same along with their significances are provided in Table-5. As the results indicate, none of these correlations is significantly different from zero at a reasonable level. This rules out any systematic pattern in the residuals.

There is a Box-Pierce Q test to see if a number of autocorrelations together are significantly different from zero. Their Q statistic is given by

$$Q=n\sum_{k=1}^{m}\gamma_k^2 \quad ....(9)$$

where n = sample size
m = length of the lag considered
γ = autocorrelation coefficient of order k.

The Q statistic has a Chi-square distribution with m degrees of freedom. The computed value of Q for m=25 equals 25.857 (vide Table-5) and the theoretical Chi-square value for 25 degrees of freedom at 5% significance level equals 37.65. Since the computed value is less than the theoretical value, the joint test indicates that the group of autocorrelations is insignificant. This proves that the selected ARIMA model is an appropriate model.

Thus, the ARIMA model for tea production in India is ARIMA, d=1, D=1, p=23, q=24 with coefficients of AR terms 3,5 - 22 and MA terms 2 - 11, 14-23 as zeros. The estimation results of this chosen model are available in the bottom third part of Table-4.

3. Forecasting with ARIMA Model:

ARIMA models are developed basically to forecast the corresponding variable. There are two kinds of forecasts: sample period and post-sample period forecasts. The former are used to develop confidence in the model and the latter to generate genuine forecasts for use in planning and other purposes. The ARIMA model can be used to yield both these kinds of forecasts.

3.1 Sample Period Forecasts:

The sample period forecasts are obtained simply by plugging the actual values of the explanatory variables in the estimated equation (8). The explanatory variables here are the lagged values of Z_t and the estimated lagged errors, and the dependent variable is Z_t. The so obtained values for

$\hat{Z}_t$ together with the actual values of Z_t are included in Table-6.

Since Z_t happens to be the stationary component of the true variable X_t (tea production), we must use their definitional equation (5) to derive the series on X_t, which would give the sample period forecasts for tea production. Thus,

$$\hat{X}_t = \hat{Z}_t + X_{t-1} + X_{t-12} - X_{t-13} \qquad \ldots.(10)$$

Using this equation, we have derived the sample period forecasts for tea production (X), which, together with the corresponding actual values are included in Table-6.

To judge the forecasting ability of the fitted ARIMA model, important measures of the sample period forecasts' accuracy, both for the tea production (X) as well as its stationary component (Z), were computed and the same are reported in the bottom line of Table 6. The mean absolute percentage error (MAPE) for tea production turns out to be 14.18% while that for the transformed variable at 23.91%. The Theil U_2 measure for these two versions turns out to be 0.46 and 0.49, respectively. These measures indicate that the forecasting inaccuracy is low and that the ARIMA forecasts are far better than the naive forecasts.

3.2 Post-Sample Period Forecasts:

The principal objective of developing an ARIMA model for a variable is to generate post sample period forecasts for that variable. This is done through using equation (8) for the transformed variable Z and equation (10) for tea production. These equations need data on a few lagged values of the variable under forecasting and thus forecasts can be derived only upto the time gap between the current period and the lowest lagged value appearing as an explanatory variable. In our select model, one period lagged value appears as an explanatory variable and thus forecasts are possible only upto one future period. To generate longer period future forecasts, which are needed for planning and other purposes, the boot-strapping method is recommended. Under this method, the forecast for period t+1 is used as the true value of the variable in period t+1 while generating forecasts for t+2 period, and so on. Using this method, we have derived forecasts for the transformed variable (Z) as well as tea production for the next 12 periods (months), and the same are reported in Table-7. Thus, the forecasts for tea production during August 1991 through July 1992 are available in column 3 of Table 8. As the results would reveal, tea production is forecasted to be 89,600 tonnes in August 1991, rise to 89,700 tonnes in October 1991, fall to 14,300 tonnes in February 1992, and then rise upto 90,000 tonnes in July 1992.

Since the data on actual tea production in India beyond July 1991 are yet not available atleast in published form, there is no way to check our forecasts with their actual counterparts. However, one can examine these forecasts through a look at their seasonal variations. A careful evaluation of the data over the last over 12 years (Vide Table-1) would reveal that tea production was on the highest level during August and September, and on the lowest level during January and February every year. It is heartening to note that the forecasts in Table-8 do confirm rather exactly such a seasonal variation.

4. An Alternative ARIMA Model:

The model presented in section 2 and used for forecasting in section 3 above, is our chosen model. However, an alternative model is presented in this section to aid readers to weigh the comparative benefits and costs of using a simple model.

It is rather apparent from Figure-1 that tea production in India is highly seasonal but contains a poor trend. Under such a situation, while seasonal difference is a must, non-seasonal difference may not be very useful in transforming the variable into a stationary series. To examine this, we have also developed an ARIMA model for tea production on the basis of seasonal difference only. The procedure followed is as follows:

$$Y_t = X_t - X_{t-12} \qquad .....(11)$$

and $$\hat{X}_t = \hat{Y}_t + X_{t-12} \qquad .....(12)$$

The results of ARIMA model for Y variable are given in Tables 8-12. The post-sample forecasts for tea production obtained through this model (Vide Table-12, Col.3) also conforms to the seasonal movements, where tea production is the highest during August and September, and the lowest during January and February. Of the two sets of forecasts (Table-7 Vs. Table-12) which one is more accurate will be known only when the actual data on tea production for recent periods become available. However, our chosen model is the earlier one and thus the values of Table-7 constitute our forecasts.

5. Conclusion:

ARIMA model offers a good technique for predicting the magnitude of any variable. Its strength lies in the fact that the method is suitable for any time series with any pattern of change and it does not require the forecaster to choose a priori the value of any parameter. Its limitations include its requirement for a long time series (large sample size), and a rather sophisticated technique. Like any other method, this technique also does not guarantee perfect forecasts. Nevertheless, with the easy accessibility of computers, appropriate softwares, and the availability of long time series data, the ARIMA method is gaining popularity and its use is only going to increase over time.

References

1. Makridakis, S., S.C. Wheelwright and V.E. McGee (1983): Forecasting: Methods and Applications, John Wiley and Sons, New York.

2. Marquardt, D.W. (1963): An Algorithm for Least Squares Estimation of Non-Linear Parameters, Journal of the Society for Industrial and Applied Mathematics, Vol.II, pp.431-41.

3. Pindyck, R.S. and D.L. Rubinfeld (1991): Econometric Methods, McGraw-Hill, New York.

Table-1: Tea Production in India

('00 Tonnes)

							TEA PRODUCTION ('00 Tonnes)						
	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Jan	100	99	117	76	100	107	126	126	109	132	103	143	143
Feb	112	71	67	61	121	113	120	110	83	100	79	143	143
Mar	212	153	72	66	183	237	265	199	216	323	169	291	291
Apr	243	360	429	326	342	479	509	434	521	536	542	615	595
May	524	669	573	434	332	561	630	473	457	620	523	603	605
Jun	571	676	651	661	747	662	718	954	769	753	838	792	900
Jul	722	823	821	765	752	805	864	791	886	905	807	876	906
Aug	687	740	718	754	859	866	834	845	839	906	942	890	
Sep	806	675	763	785	803	814	848	874	939	899	970	897	
Oct	706	711	725	730	734	734	780	768	827	811	844	844	
Nov	549	560	541	593	580	574	544	543	569	657	673	582	
Dec	230	221	256	309	355	338	292	329	569	569	346	417	

Source: C.S.O.: Monthly Abstracts of Statistics, various issues.

Figure-1: Tea Production in India

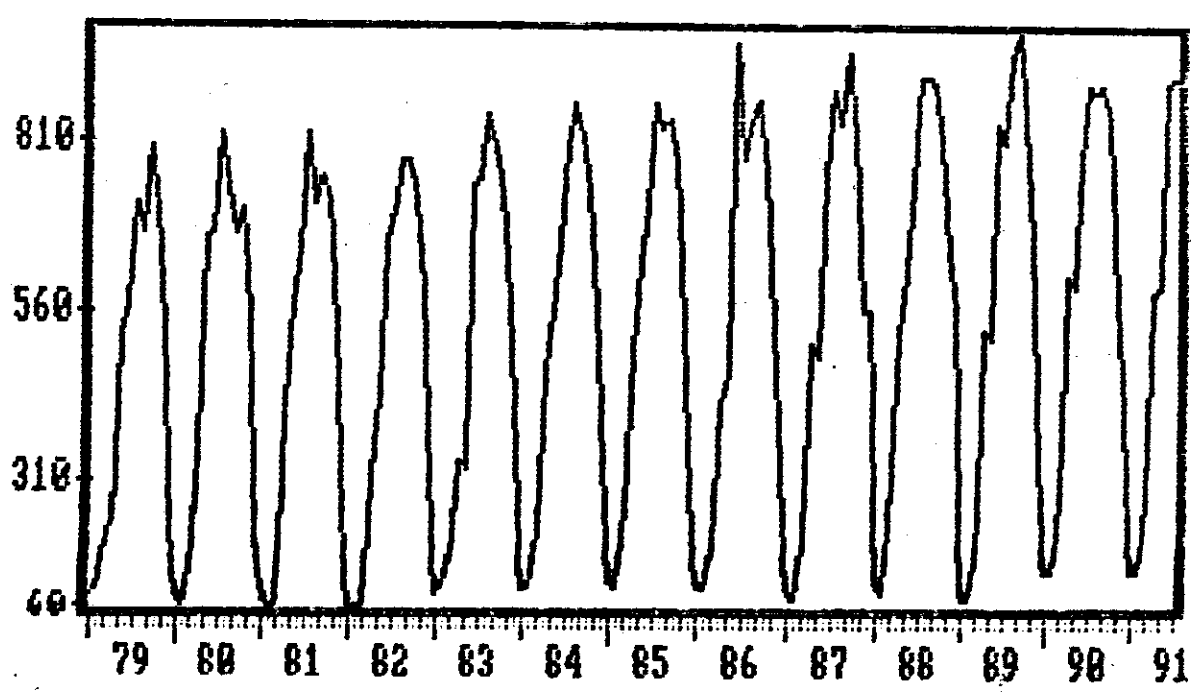


Table 2: Autocorrelations for Tea Production

Lag Period	Autocorrelation coefficient	Lag Period	Autocorrelation coefficient
1	0.798	19	-0.631
2	0.445	20	-0.352
3	-0.010	21	0.018
4	-0.432	22	0.404
5	-0.712	23	0.692
6	-0.792	24	0.794
7	-0.684	25	0.659
8	-0.384	26	0.361
9	0.024	27	-0.027
10	0.444	28	-0.366
11	0.762	29	-0.598
12	0.892	30	-0.665
13	0.730	31	-0.568
14	0.399	32	-0.332
15	-0.025	33	0.003
16	-0.408	34	0.357
17	-0.664	35	0.615
18	-0.731	36	0.720

Table 3: **Autocorrelation and Partial Autocorrelation Plots for Stationary Series (Z)**

Autocorrelations	Partial Autocorrelations		ac	pac
*******¦ .	*******¦ .	1	-0.556	-0.558
. ¦*.	****¦ .	2	0.091	-0.320
. ¦ .	**¦ .	3	0.020	-0.149
**¦ .	***¦ .	4	-0.128	-0.259
. ¦**	.*¦ .	5	0.154	-0.094
.*¦ .	.*¦ .	6	-0.070	-0.048
.*¦ .	**¦ .	7	-0.077	-0.191
. ¦*.	**¦ .	8	0.086	-0.162
. ¦ .	. ¦ .	9	0.015	-0.023
.*¦ .	**¦ .	10	-0.076	-0.120
. ¦**	. ¦*.	11	0.172	0.099
****¦ .	**¦ .	12	-0.290	-0.192
. ¦***	**¦ .	13	0.211	-0.118
.*¦ .	**¦ .	14	-0.064	-0.123
. ¦ .	.*¦ .	15	0.001	-0.069
. ¦*.	. ¦ .	16	0.063	-0.033
**¦ .	**¦ .	17	-0.129	-0.121
. ¦**	. ¦*.	18	-0.192	0.102
**¦ .	.*¦ .	19	-0.159	-0.048
. ¦ .	.*¦ .	20	0.030	-0.090
. ¦*.	. ¦*.	21	0.073	0.073
**¦ .	.*¦ .	22	-0.122	-0.068
. ¦***	. ¦****	23	0.240	0.330
****¦ .	***¦ .	24	-0.345	-0.208
. ¦**	***¦ .	25	0.117	-0.233
. ¦**	. ¦ .	26	0.132	-0.016
**¦ .	.*¦ .	27	-0.133	-0.086
. ¦*.	.*¦ .	28	0.041	-0.098
. ¦ .	**¦ .	29	0.010	-0.183
. ¦ .	. ¦ .	30	-0.026	0.030
. ¦*.	**¦ .	31	0.072	-0.116
. ¦ .	. ¦*.	32	0.033	0.060
**¦ .	. ¦ .	33	-0.183	0.014
. ¦**	.*¦ .	34	0.188	-0.068
**¦ .	. ¦*.	35	-0.182	0.068
. ¦***	. ¦ .	36	0.235	0.029

Q-Statistic (36 lags) 143.794

S.E. of Correlations 0.085

Table 4: Estimates of Alternative ARIMA Models (Z)

Model 4.1

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-0.3976263	7.6687544	-0.0518502	0.959
MA(1)	-0.9194988	0.1992330	-4.6151933	0.000
AR(1)	0.0368575	0.1725283	0.2136304	0.831
AR(2)	0.1086706	0.1107948	0.9808268	0.327

R-squared	0.488679	Mean of dependent var	0.507353
Adjusted R-squared	0.477056	S.D. of dependent var	105.5654
S.E. of regression	76.33932	Sum of squared resid	769255.3
Durbin-Watson stat	2.138940	F-statistic	42.05157
Log likelihood	-780.5312		

Model 4.2

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-0.1602216	4.4310041	-0.0406729	0.968
MA(1)	-0.6643151	0.1706710	-3.8923719	0.000
MA(12)	-0.4870484	0.0873199	-5.5777505	0.000
MA(13)	0.4923767	0.1157711	4.2530179	0.000
AR(1)	-0.1667604	0.1459496	-1.1425890	0.253
AR(2)	-0.0395569	0.0982629	-0.4025622	0.687
AR(4)	-0.1438328	0.0586346	-2.4530351	0.014

R-squared	0.599178	Mean of dependent var	-1.007463
Adjusted R-squared	0.580242	S.D. of dependent var	105.2239
S.E. of regression	68.17324	Sum of squared resid	590244.0
Durbin-Watson stat	1.996636	F-statistic	31.64152
Log likelihood	-752.2980		

Model 4.3

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	-5.0629035	3.3596155	-1.5069890	0.135
MA(1)	-0.1186026	0.1766201	-0.6715122	0.504
MA(12)	-0.5350475	0.0954354	-5.6063833	0.000
MA(13)	0.2135119	0.1168400	1.8273862	0.071
MA(24)	-0.5132327	0.1070570	-4.7940151	0.000
AR(1)	-0.6384420	0.1376964	-4.6365916	0.000
AR(2)	-0.2822516	0.1005014	-2.8084350	0.006
AR(4)	-0.0692411	0.0644334	-1.0746146	0.285
AR(23)	0.1221058	0.0643036	1.8988940	0.061

R-squared	0.665285	Mean of dependent var	-0.217391
Adjusted R-squared	0.640023	S.D. of dependent var	107.9985
S.E. of regression	64.79703	Sum of squared resid	445057.4
Durbin-Watson stat	1.766190	F-statistic	26.33588
Log likelihood	-638.1869		

Table 5: Autocorrelation and Partial Autocorrelation Plots for Residual (Z)

Autocorrelations	Partial Autocorrelations		ac	pac
. ¦*.	. ¦*.	1	0.113	0.113
.*¦ .	.*¦ .	2	-0.082	-0.096
**¦ .	**¦ .	3	-0.163	-0.145
. ¦*.	. ¦**	4	0.097	0.130
. ¦*.	. ¦*.	5	0.097	0.048
. ¦ .	.*¦ .	6	-0.021	-0.051
.*¦ .	. ¦ .	7	-0.061	-0.007
. ¦ .	. ¦*.	8	0.020	0.039
. ¦**	. ¦*.	9	0.116	0.082
. ¦*.	. ¦*.	10	0.078	0.052
. ¦**	. ¦***	11	0.166	0.198
**¦ .	**¦ .	12	-0.127	-0.146
. ¦ .	. ¦*.	13	-0.023	0.077
. ¦*.	. ¦**	14	0.099	0.122
. ¦**	. ¦*.	15	0.168	0.079
. ¦*.	. ¦*.	16	0.075	0.093
. ¦ .	. ¦ .	17	-0.020	0.037
. ¦*.	. ¦*.	18	0.080	0.113
. ¦ .	.*¦ .	19	-0.003	-0.075
. ¦*.	. ¦*.	20	0.077	0.078
. ¦**	. ¦**	21	0.117	0.170
. ¦ .	.*¦ .	22	-0.010	-0.110
.*¦ .	. ¦ .	23	-0.063	0.015
.*¦ .	.*¦ .	24	-0.048	-0.080
. ¦**	. ¦*.	25	0.123	0.058

Q-Statistic (25 lags) 25.857

S.E. of Correlations 0.093

Table 6: Actual and Estimated Values of Stationary Component (Z) and of Tea Production (X)

TEA PRODUCTION : MODEL ARIMA:d =1. D = 1: ar(1,2,4,23) ma(1,12,13,24)

MONTH	Z(t) ACT.	Z(t) EST.	X(t) ACT.	X(t) EST.	MONTH	Z(t) ACT.	Z(t) EST.	X(t) ACT.	X(t) EST.	MONTH	Z(t) ACT.	Z(t) EST.	X(t) ACT.
1982.01	-76	-41.477	76	111	1985.03	21	-27.585	265	216	1988.05	114	114.938	620
1982.02	35	53.579	61	80	1985.04	2	40.164	509	547	1988.06	-138	-120.739	753
1982.03	0	82.165	66	148	1985.05	39	50.580	630	642	1988.07	34	-10.260	905
1982.04	-97	22.189	326	445	1985.06	-13	59.747	718	791	1988.08	48	15.870	906
1982.05	-36	28.383	434	498	1985.07	3	-35.791	864	825	1988.09	-107	-86.325	899
1982.06	149	146.582	661	659	1985.08	-91	-65.369	834	860	1988.10	24	63.278	811
1982.07	-123	-111.941	708	719	1985.09	66	88.031	848	870	1988.11	104	19.835	457
1982.08	179	87.959	784	693	1985.10	12	28.423	780	796	1988.12	-88	-162.596	569
1982.09	-44	-56.705	785	772	1985.11	-76	-57.615	544	562	1989.01	-29	35.563	103
1982.10	-17	-28.871	730	718	1985.12	-16	77.176	292	385	1989.02	8	93.154	79
1982.11	47	4.309	593	550	1986.01	46	13.112	126	93	1989.03	-133	-49.625	169
1982.12	1	35.319	309	343	1986.02	-10	-23.911	110	96	1989.04	166	80.438	542
1983.01	-29	-32.837	100	96	1986.03	-56	-29.146	199	226	1989.05	-109	-120.400	523
1983.02	36	20.230	121	105	1986.04	-9	76.170	434	519	1989.06	182	127.803	838
1983.03	57	68.655	183	195	1986.05	-82	-55.998	473	499	1989.07	-183	-179.690	807
1983.04	-101	48.386	342	491	1986.06	393	185.887	954	747	1989.08	134	30.541	942
1983.05	-118	35.434	332	485	1986.07	-309	-274.869	791	825	1989.09	35	-67.055	970
1983.06	188	180.658	747	740	1986.08	84	85.442	845	846	1989.10	-38	-91.779	844
1983.07	-42	-88.098	752	706	1986.09	15	36.546	874	896	1989.11	-17	-17.231	673
1983.08	22	-14.808	850	813	1986.10	-38	-37.960	768	768	1989.12	-239	-113.337	346
1983.09	-48	15.633	803	867	1986.11	11	19.788	543	552	1990.01	263	219.952	143
1983.10	-14	37.363	734	785	1986.12	38	57.568	329	349	1990.02	24	-28.480	143
1983.11	-17	-21.016	580	576	1987.01	-54	-88.177	109	75	1990.03	58	-107.011	291
1983.12	59	82.594	355	379	1987.02	-10	7.697	83	401	1990.04	-49	-90.437	615
1984.01	-39	-23.593	107	122	1987.03	44	6.154	216	178	1990.05	7	-21.869	603
1984.02	-15	3.503	113	132	1987.04	70	22.340	521	473	1990.06	-126	-18.735	792
1984.03	62	54.428	237	229	1987.05	-63	-68.397	497	492	1990.07	135	74.206	896
1984.04	83	84.427	479	480	1987.06	-210	-67.605	768	910	1990.08	-141	-149.518	890
1984.05	92	24.358	561	493	1987.07	281	187.702	886	793	1990.09	-21	21.992	897
1984.06	-314	-151.552	662	824	1987.08	-101	-130.801	839	809	1990.10	73	83.445	844
1984.07	138	184.457	805	851	1987.09	71	0.533	939	869	1990.11	-91	-100.827	582
1984.08	-37	-71.121	866	832	1987.10	-6	-25.525	827	807	1990.12	162	62.094	417
1984.09	-5	-1.893	814	817	1987.11	-33	-35.263	569	567	1991.01	-71	-113.345	143
1984.10	-11	39.831	734	785	1987.12	214	82.062	569	437	1991.02	0	-11.353	143
1984.11	-6	-39.453	574	541	1988.01	-217	-197.898	132	151	1991.03	0	1.514	291
1984.12	-11	23.627	338	373	1988.02	-6	74.153	100	180	1991.04	-22	-64.656	593
1985.01	36	9.345	126	99	1988.03	90	56.070	323	289	1991.05	22	14.220	603
1985.02	-12	-26.258	120	106	1988.04	-98	-67.785	530	560	1991.06	108	-4.821	900
										1991.07	-98	-135.368	906

MAPE(Xt) 14.18272 THEIL U2(Xt) 0.464352 MAPE(Zt) 23.91312 THEIL U2(Zt) 0.486373

Table 7: Post Sample Period Forecasts for Stationary Component (Z) and Tea Production (X)

	Forecast Z(t)	X(t-1) 906.0000	Forecast X(t)	X(t-13)
1991.08	-20.683		879.3171	896
1991.09	-2.3112		884.0059	890
1991.10	-42.889		788.1169	897
1991.11	-11.427		514.6900	844
1991.12	56.628		406.3180	582
1992.01	-62.713		69.60484	417
1992.02	0.68250		70.28735	143
1992.03	-76.838		139.4492	143
1992.04	-6.7948		434.6544	291
1992.05	-3.7812		440.8732	593
1992.06	7.6611		745.5343	603
1992.07	-52.145		699.3889	900
				906

X(t) = Z(t)+X(t-1)+X(t-12)-X(t-13)

Table-8: Autocorrelation and Partial Autocorrelation Plots for Stationary Series (Y).

Autocorrelations	Partial Autocorrelations		ac	pac
. ¦ .	. ¦ .	1	-0.017	-0.017
. ¦*.	. ¦*.	2	0.100	0.100
. ¦ .	. ¦ .	3	0.032	0.036
.*¦ .	.*¦ .	4	-0.072	-0.082
. ¦*.	. ¦*.	5	0.084	0.076
.*¦ .	.*¦ .	6	-0.073	-0.058
.*¦ .	.*¦ .	7	-0.089	-0.105
. ¦*.	. ¦*.	8	0.050	0.054
. ¦ .	. ¦*.	9	0.010	0.049
.*¦ .	.*¦ .	10	-0.057	-0.084
. ¦ .	. ¦ .	11	0.031	0.020
***¦ .	***¦ .	12	-0.232	-0.207
. ¦*.	. ¦*.	13	0.097	0.081
. ¦ .	. ¦ .	14	-0.005	0.028
. ¦ .	. ¦*.	15	0.021	0.043
. ¦*.	. ¦ .	16	0.047	0.001
.*¦ .	. ¦ .	17	-0.058	-0.033
. ¦*.	. ¦*.	18	0.099	0.061
**¦ .	**¦ .	19	-0.134	-0.165
.*¦ .	. ¦ .	20	-0.046	-0.026
. ¦ .	. ¦ .	21	-0.015	0.020
**¦ .	**¦ .	22	-0.134	-0.146
. ¦ .	. ¦ .	23	-0.005	-0.012
*****¦ .	******¦ .	24	-0.364	-0.436
. ¦ .	. ¦*.	25	-0.023	0.057
. ¦*.	. ¦**	26	0.082	0.122
.*¦ .	.*¦ .	27	-0.082	-0.064
. ¦ .	. ¦ .	28	0.023	-0.010
. ¦*.	. ¦ .	29	0.043	0.008
. ¦*.	. ¦*.	30	0.046	0.111
. ¦*.	.*¦ .	31	0.099	-0.100
. ¦ .	. ¦*.	32	0.009	0.049
**¦ .	**¦ .	33	-0.149	-0.129
. ¦*.	.*¦ .	34	0.065	-0.094
.*¦ .	. ¦ .	35	-0.103	-0.011
. ¦*.	**¦ .	36	0.100	-0.155

Q-Statistic (36 lags) 51.289

S.E. of Correlations 0.085

Table-9: Estimates of ARIMA Model (Y)

Y

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	11.592664	2.9127968	3.9799083	0.000
MA(12)	-0.2786417	0.1320450	-2.1102015	0.038
MA(24)	-0.6652080	0.1413423	-4.7063610	0.000
AR(12)	-0.4212854	0.1110798	-3.7926370	0.000
AR(24)	-0.2617521	0.0787888	-3.3221987	0.001

R-squared	0.523851	Mean of dependent var	12.83478
Adjusted R-squared	0.506537	S.D. of dependent var	74.37646
S.E. of regression	52.24718	Sum of squared resid	300274.5
Durbin-Watson stat	1.815408	F-statistic	30.25508
Log likelihood	-615.5603		

Table-10: Autocorrelation and Partial Autocorrelation Plots for Residual (Y).

Autocorrelations	Partial Autocorrelations		ac	pac
. ¦*.	. ¦*.	1	0.091	0.091
. ¦*.	. ¦*.	2	0.086	0.078
. ¦*.	. ¦*.	3	0.068	0.055
. ¦ .	.*¦ .	4	-0.027	-0.045
. ¦ .	. ¦ .	5	0.031	0.028
. ¦ .	. ¦ .	6	-0.025	-0.029
**¦ .	**¦ .	7	-0.146	-0.145
. ¦ .	. ¦ .	8	-0.007	0.016
. ¦ .	. ¦ .	9	-0.009	0.020
.*¦ .	.*¦ .	10	-0.100	-0.090
. ¦*.	. ¦*.	11	0.053	0.062
. ¦ .	. ¦ .	12	0.007	0.024
. ¦**	. ¦**	13	0.124	0.121
. ¦ .	.*¦ .	14	-0.023	-0.084
. ¦*.	. ¦*.	15	0.095	0.104
. ¦ .	. ¦ .	16	0.025	-0.004
.*¦ .	**¦ .	17	-0.092	-0.134
. ¦*.	. ¦*.	18	0.089	0.115
.*¦ .	.*¦ .	19	-0.078	-0.072
.*¦ .	.*¦ .	20	-0.064	-0.043

Q-Statistic (20 lags) 12.551 S.E. of Correlations 0.093

Table-11: Actual and Estimated Values of Tea Production (X) Through Stationary Variable Y

(Y

TEA PRODUCTION: MODEL ARIMA: D = 1. ar(12,24), ma(12,24)

MONTH	X(t)	X(t) Est.	MONTH	X(t)	X(t) Est.	MONTH	X(t)	X(t) Est
1982.01	76	80	1985.03	265	186	1988.06	753	719
1982.02	61	60	1985.04	509	511	1988.07	905	905
1982.03	66	145	1985.05	630	602	1988.08	906	858
1982.04	326	423	1985.06	718	729	1988.09	899	894
1982.05	434	530	1985.07	864	819	1988.10	811	834
1982.06	661	699	1985.08	834	838	1988.11	657	613
1982.07	708	773	1985.09	848	843	1988.12	569	442
1982.08	784	782	1985.10	780	777	1989.01	103	159
1982.09	785	806	1985.11	544	582	1989.02	79	139
1982.10	730	738	1985.12	292	349	1989.03	169	256
1982.11	593	557	1986.01	126	132	1989.04	542	507
1982.12	309	307	1986.02	110	115	1989.05	523	605
1983.01	100	68	1986.03	199	215	1989.06	838	856
1983.02	121	55	1986.04	434	479	1989.07	807	843
1983.03	183	171	1986.05	473	536	1989.08	942	904
1983.04	342	454	1986.06	954	802	1989.09	970	869
1983.05	332	480	1986.07	791	852	1989.10	844	801
1983.06	747	759	1986.08	845	832	1989.11	673	645
1983.07	752	782	1986.09	874	859	1989.12	346	371
1983.08	850	834	1986.10	768	798	1990.01	143	152
1983.09	803	825	1986.11	543	580	1990.02	143	136
1983.10	734	758	1986.12	329	351	1990.03	291	200
1983.11	580	559	1987.01	109	124	1990.04	615	555
1983.12	355	349	1987.02	83	102	1990.05	603	562
1984.01	107	114	1987.03	216	191	1990.06	792	808
1984.02	113	98	1987.04	521	491	1990.07	896	873
1984.03	237	204	1987.05	497	540	1990.08	890	886
1984.04	479	477	1987.06	768	824	1990.09	897	939
1984.05	561	536	1987.07	886	813	1990.10	844	857
1984.06	662	756	1987.08	839	867	1990.11	582	625
1984.07	805	834	1987.09	939	866	1990.12	417	382
1984.08	866	819	1987.10	827	787	1991.01	143	193
1984.09	814	830	1987.11	569	606	1991.02	143	179
1984.10	734	763	1987.12	569	389	1991.03	291	332
1984.11	574	561	1988.01	132	144	1991.04	593	561
1984.12	338	338	1988.02	100	125	1991.05	603	657
1985.01	126	98	1988.03	323	249	1991.06	900	825
1985.02	120	72	1988.04	530	545	1991.07	906	921
			1988.05	620	602			

MAPE 11.68840 THEIL U2 0.409340

Table-12: Post Sample Period Forecasts for Stationary Component (Y and Tea Production (X)

	Forecast Y(t)	Forecast X(t)	X(t-12)
1991.08	5.6959	895.6958	890
1991.09	-24.076	872.9235	897
1991.10	-13.849	830.1511	844
1991.11	47.179	629.1791	582
1991.12	54.557	471.5568	417
1992.01	29.366	172.3661	143
1992.02	8.2232	151.2231	143
1992.03	-61.290	229.7095	291
1992.04	-39.480	553.5204	593
1992.05	-13.750	589.2495	603
1992.06	-24.512	875.4883	900
1992.07	-19.372	886.6280	906

X(t) = Y(t)+X(t-12)