FORECASTING FOOD PRODUCTION USING VECTOR ERROR CORRECTION MODEL

K. Senthamarai Kannan[1], M. Velusamy[2] and S. Suresh[3]

[1,2]Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli – 627012
[3]Department of Statistics, Madras University, Chennai – 5
Email: senkannan2002gmail.com,velustat@gmail.com and sureshstat@gmail.com

ABSTRACT: Numerous past studies in Rice cultivation in India is criticized for using Johansson Co-integration test. Adjustment vector and also for analytical interpretation through Ordinary Least Squares creating spurious results for time series data the present paper estimates the modern techniques of co-integration with vector error correction model (VECM) restricts the long-run behaviour of the endogenous variables to converge their co-integrating relationships while allowing a wide range of short-run dynamics. In this article, attempts to predict the rice production and investigated the relationship among some other agriculture commodities like Wheat, cereals and pulse using Johansson co-integration test. In our result, the model fitted the data well and the stochastic seasonal variation was successfully model.Vector error correction model (VECM) was a proper method for modelling and forecasting the time series of yearly rice production data.

Key words: Multivariate time series, co-integration test, Akaike's Information Criteria, VECM and Durbin-Watson test statistic.

Introduction

Agriculture is the backbone of any country's economy. It is the main source of people's livelihood in most of the world mainly the developing and under-developing countries. Under-developing countries are suffering with food insecurity due to poor productivity of agricultural land. Soil degradation, lack of improved agricultural technologies, lack of irrigation water etc., are some of the important factors responsible for poor yields.

In pre-scientific agriculture, nine persons could produce enough food for fifteen persons in years of bad harvest; they could produce only enough for themselves. With the development of agricultural science and application of advanced technology, now five persons are able to produce enough food for more than

100 people. Early knowledge of agriculture was a collection of experiences transmitted from farmer to farmer verbally since there was no available agricultural technology

A time series is a sequence of observations collected at regular time intervals. Time series analysis comprises methods for analysing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values this multivariate time series model is used to discover the pattern and predict future values of real data. In the present work, we have implemented the VEC model to forecast the rice production in India.

Materials and Method

Scope of the study

The study consisting yearly commodity production data sourced from ministry of Statistics programme implementation (MOSBI) (1989-2015). Eviews econometric software package used to analyze the data.

Stationary Test

Time series stationary is the statistical characteristics of a series such as its means and variance over time. If both are constant over time, then the series is said to be stationary i.e. there is no random walk or unit root, otherwise the series has unit root.

Vector Error Correction model

This Multivariate time series model is used to discover the pattern and predict the future value of real data. The var (*p*) model can be expressed in the form,

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + e_t$$

The Vector Error Correction Model (VECM) involves expressing an nx1 vector of stationary time series (say Y_t) in terms of a constant, lagged values of itself and an error correction term. The standard VECM (p) model can be represented as follows,

$$\Delta Y_T = (C + Ø_{11+} \Delta X_{t-1} + Ø_{12} \Delta Yt\text{-}_1 + Ø_{13} \Delta Z_{t-1} + \alpha (Y_{t-1} - \beta X_{t-1} - \beta Z_{t-1}) + \mathcal{E}_t)$$

Where α (Y_{t-1} – βX_{t-1} - βZ_{t-1}) are known as Error Correction Term (ECT) $_{t-1}$ that is a product of an adjustment factor (α) and the co-integrating vector (β) The VECM specification restricts the long-run behaviour of the endogenous variables to converge to their co-integration relationships while allowing a wide range of short-run dynamics. The co-integration term is known as the error correction term since the deviation from long-run equilibrium is corrected gradually through a series of partial short-run adjustments.

A set of variables is cointegrated when there exists a stable long run relationship between them. Formally the set of variables which are interested as y_1, y_2... y_k Suppose that each variable is integrated of order one, viz.I (1) and therefore needs differencing in order to attain stationarity. If there exists linear combinations of the variables which are I (0), then the variables are said to be co-integrated, (i.e.) they have a stable long run relationship. Then the co-integrating vector can be estimated which quantifies the relationship between the concerned variables.

The Johansen Tests

The Johansen tests are called the maximum eigenvalue test and the trace test. Let r be the rank of Π. As the discussion above indicated, this is the same as the number of co-integration vectors. The Johansen tests are likelihood-ratio tests. There are two tests: 1. the maximum eigenvalue test, and 2. the trace test. For both test statistics, the initial Johansen test is a test of the null hypothesis of no co-integration against the alternative of co-integration. The tests differ in terms of the alternative hypothesis.

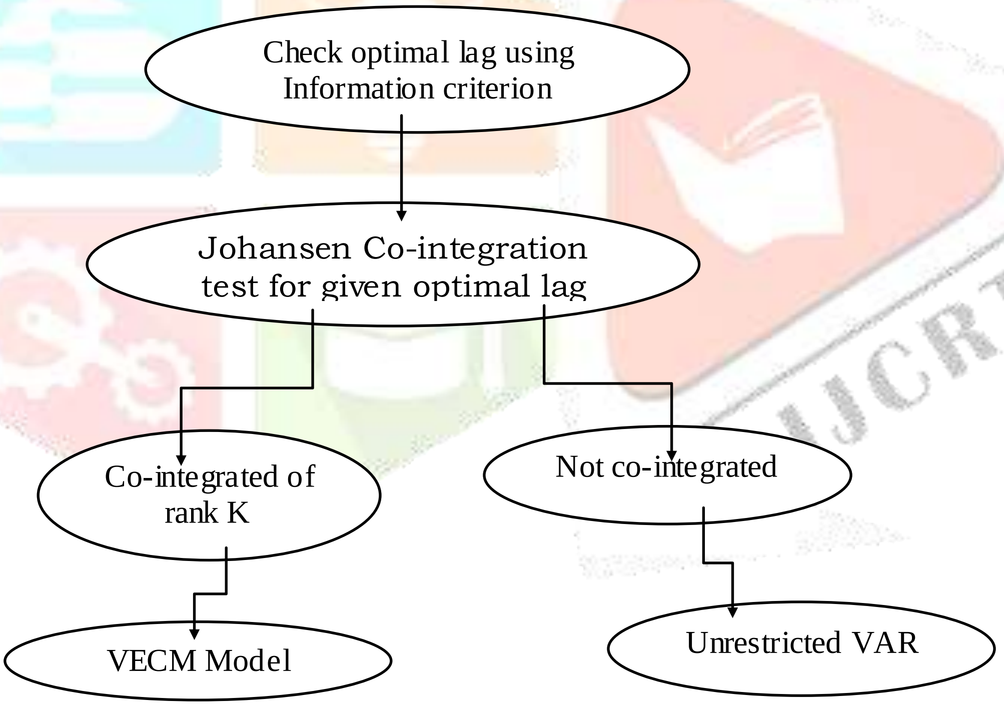

Figure1: Flow chart describing when to approach VAR and VECM

In this article, Time series analysis concepts are briefly discussed and reviewed. Multivariate Time series analysis is brief with their merits and demerits which includes Vector Autoregressive (VAR) model and Vector Error Correction model (VECM). Stationary and Co-integration is discoursed with concise explanations. VECM method adopted for rice production data to predict the future year rice production.

Results and Discussion

Table 1: lag length criteria

LAGS	LOG LIK	P(LR)	AIC	BIC	HQC
1	-210.31712	---	20.027575	**21.014962***	20.275900
2	-195.50042	0.02000	20.130471	21.907767	20.577456
3	-176.34920	0.00137	19.856452	22.423657	20.502097
4	-149.09998	0.00000	**18.878260***	22.235373	**19.722564***

The asterisks above indicate the best (that is, minimized) values of the respective information criteria, AIC = Akaike Information criterion, BIC = Bayesian Information criterion and HQC = Hannan - Quinn Information criterion. Optimal lag length is chosen by values of the criterion which is Minimum among the given lags. Here, 4 lags are chosen and lag 4 is chosen for the model based on AIC.

Johansen Co-Integration Test

Co-integration rank (rank of matrix) is estimated by using Johansen methodology. Johansen's approach derives two likelihood estimators for the CI rank: a trace test and a maximum Eigen value test. The CI rank (R) can be formally tested with the trace and the maximum Eigen value statistics. The trace statistic either rejects the null hypothesis of no co-integration among the variables or does not reject the null hypothesis that there is one co-integration relation between the variables. Start by testing $H_0: r = 0$. If it rejects, repeat for $H_0: r = 1$. When a test is not rejected by stop testing there and that value of r is the commonly-used estimate of the number of Cointegration relations.

Table 2: Unrestricted co-integration rank test (Trace)

Rank	Eigen value	Trace test [p-value]	Lmax test [p-value]
0	0.65639	48.945[0.0377]	25.638[0.0857]
1	0.48065	23.307[0.2389]	15.724[0.2513]
2	0.25354	7.5829[0.5180]	7.0180[0.4955]
3	0.023260	0.56484[0.4523]	0.56484[0.4523]

Johansen test: Number of equations = 4 Lag order = 3,

Estimation period: 1989 - 2015 (T = 26)

In this test, H0: r = 1 is not rejected at the 5% level. In other words, this trace test result does not reject the null hypothesis that variables are co-integrated. The final number of co-integrated vectors with three lags is equal to one, i.e. rank 1. Since, the rank is equal to 1 which is more than zero and less than the number of variables; the series are co-integrating among the variables.

Table 3: Co-integrating Vector (B) and Adjustment Vector (∝)

Commodities	**Co integrating Vector (Standard Error)**		**Adjustment vector**	
Rice	1 (0.00)	0 (0.00)	-1.444	2.402
Pulses	0 (0.00)	1 (0.00)	-0.0633	-0.3449
Wheat	-0.718 (0.031)	-0.055 (0.016)	0.7672	5.2363
Cereals	-0.962 (0.110)	-0.023 (0.056)	0.7674	-1.5996

The co-integrating vector (β) shows the long term equilibrium relationship between the concerned variables while the adjustment factors (α) shows the speed of adjustment towards equilibrium in case there is any deviation.

Table 4: OLS Estimates

	Coefficient	**std. error**	**t-ratio**	**p-value**
Const	-12.8920	49.6777	-0.2595	0.8027
d_Rice_1	0.120568	1.36493	0.08833	0.9321
d_Rice_2	0.176425	0.716897	0.2461	0.8127
d_Pulses_1	-1.56173	4.42356	-0.3530	0.7344
d_Pulses_2	-0.320092	2.91770	-0.1097	0.9157
d_Wheat_1	-0.917550	1.00918	-0.9092	0.3935
d_Wheat_2	-0.0812331	1.16730	-0.06959	0.9465
d_Cereals_1	-0.672171	1.90457	-0.3529	0.7345
d_Cereals_2	-0.528825	0.917714	-0.5762	0.5825
EC1	-1.44401	1.95546	-0.7385	0.4843
EC2	2.40210	5.20955	0.4611	0.6587

Mean of dependent variable = 0.778947, Standard deviation of dep. var. = 8.10227

Sum of squared residuals = 445.165, Standard error of the regression = 7.97465

Unadjusted R-squared = 0.6232, Durbin-Watson statistic = 1.91114

First-order autocorrelation coefficient = -0.123011

EC1 is the coefficient of error correction term at time t-1.It is the adjustment vector (α) for Rice, Pulses, Cereals, wheat. Co-integrating vector (β) for the equation is,

$$(\text{Rice}_{t-1} - \beta\,\text{pulses}_{t-1} - \beta\,\text{wheat}_{t-1} - \beta\,\text{cereals}_{t-1})$$

Error Correction Term (ECT) at t-1 is the product of adjustment vector (α) and co-integrating vector (β)

Table 5: Actual and Predicted Value

Year	Rice	Prediction	Std. Error	95% conf. interval
1989	70.49			
1990	73.57			
1991	74.29			
1992	74.68	75.64		
1993	72.86	75.64		
1994	80.30	80.28		
1995	81.81	73.82		
1996	76.98	81.93		
1997	81.73	83.36		
1998	82.54	81.05		
1999	86.08	87.04		
2000	89.68	86.89		
2001	84.98	87.06		
2002	93.34	90.26		
2003	71.82	81.08		
2004	88.53	87.60		
2005	83.13	82.72		
2006	91.79	91.57		
2007	93.36	88.20		
2008	96.69	87.57		
2009	99.18	96.86		
2010	89.09	99.99		
2011	95.98	104.04	4.840	94.55 - 113.53
2012	105.30	99.36	4.997	89.57 - 109.16
2013	105.24	106.47	6.084	94.55 - 118.40
2014	106.65	106.21	6.490	93.49 - 118.93
2015	104.80	105.24	6.899	91.71 - 118.76
2016		103.54	7.229	89.37 - 117.71
2017		109.39	7.614	94.47 - 124.31
2018		112.75	8.088	96.89 – 128.60

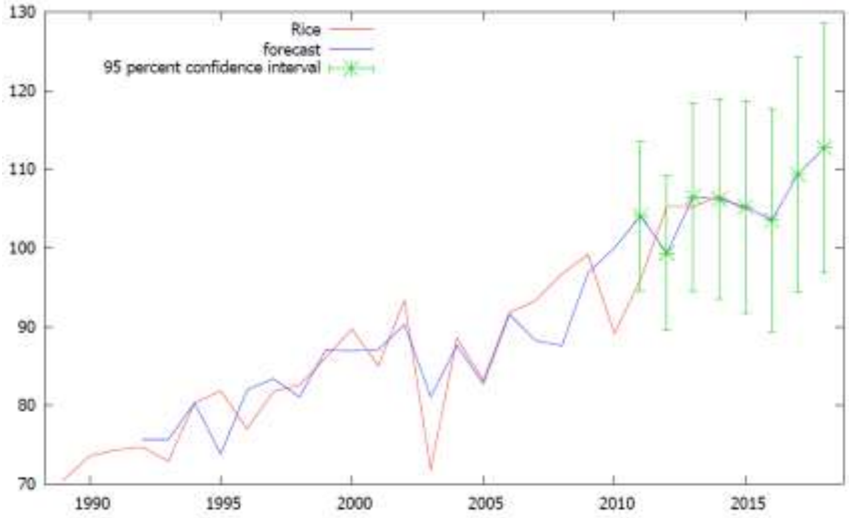

Figure 2: Actual and Forecast Value Graph for Rice Prediction

Conclusion

This paper identifies the relationship among the agriculture commodities using co-integration test. It reveals that there is a long run relationship among the endogenous variable. Vector Error Correction Model is used to estimate the parameters and using these parameter values to predict the rice production values from the year 2017 to 2018. The actual and forecasted values are presented in table-4 and graphically displayed in figure 2.

Acknowledgement:

The Second author thanks the University Grants Commission, New Delhi for awarding fellowship under the scheme of UGC – Basic Science Research fellowship to carry out this work.

References

1. Asche, F., H. Bremmes, and C.R. Wessels. 1999. "Product Aggregation, Market Integration, and Relationships between Prices: An Application to World Salmon Markets". *American Journal of Agricultural Economics*, 81:568-581.
2. Banerjee, Anindya, Juan Dolado, JohnW. Galbraith, and David F. Hendry (1993)
3. Chris, Brooks (2008). Introductory Econometrics for Finance Second Edition, Cambridge University Press, Cambridge, UK.
4. Dickey, D.A. and W. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root". *Journal of American Statistical Association*, 74(366a): 427-431.

5. Dickey, D.A. and W.A. Fuller (1979), Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-431.
6. Engle R.F. and C.W.J. Granger. 1987. "Co-Integration and Error Correction: Representation Estimation and Testing". *Econometrica*. 55:251-276.
7. Geweke J. and S. Porter-Hudak (1983), the estimation and application of long memory time series models, *Journal of Time Series Analysis*, 4, 221-238.
8. Hendy, D. F. & Juselius K. (2000). Explaining Co-integration Analysis: Part I, *Energy Journal* 21, 44.
9. Johansen, S. (1988). Statistical Analysis of Co-integrating Vectors, *Journal of Economic Dynamics and Control* 12,231—54.
10. Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.
11. Johansen, S. (2002). A small sample correction for the test of co-integration rank in the vector autoregressive model. Econometric, 70, 1929-1961.
12. MacKinnon, J. (1996). Numerical Distribution Functions for Unit Root and Co-Integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data, Oxford University Press, Oxford.
13. Quarterly Review of the Federal Reserve Bank of Minneapolis, winter.
14. Ranjit Kumar Paul (2013).Co-integration, *Indian Agricultural Statistics Research Institute*, New Delhi.
15. Reimers, H-E (1992). Comparsions of tests for multivariate cointegration, Statistical Papers 33, 335-359.
16. Sims, Christopher A. (1988). Are Forecasting Models Usable for Policy Analysis, Cointegration Tests, *Journal of Applied Econometrics*, 11, 601-618.