CHAPTER – III
METHODOLOGY

This section deals with the nature and sources of data and the methodology for examining market integration of selected agricultural commodities, the relationship between spot and futures prices and forecasting of futures and spot prices. For each objective of the study, the methodology, which was found to be the most appropriate, has been employed. This chapter is arranged into following broad headings:-

3.1 Nature and sources of data

3.2 Analytical framework

3.1 Nature and sources of data

The study was purely based on secondary data which was drawn from the official website of NCDEX. The data was comprise of daily closing spot prices and futures prices with near month (expiration month) maturity pertaining to the sample agricultural commodity derivatives traded on the exchange. The near month futures contracts were selected for the analysis as they are considered as highly liquid and the most active contracts. The data on trading volume and open interest for the sample agricultural commodity derivatives was collected on daily basis.

The period of study from 01st January 2010 to 31st December 2017 for coriander seed. The sample commodities and their respective data period as well as reference market for spot price is given in the following Table 1.

Table 3.1: Data Period for the Sample Agricultural Commodity Derivatives

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Commodity</th>
<th>Specification</th>
<th>Data Period</th>
<th>Futures Market</th>
<th>Spot Market</th>
</tr>
</thead>
</table>

Coriander was chosen for the purpose of study because India is the largest-producer and exporter of coriander in the global market. The exports have increased significantly in the past few years due to strong demand from the overseas markets. The changing pattern of food consumption or consumption of more spicy foods,
especially in developed countries and the large population of Indian origin in these
countries has resulted in good export orders for Indian spice exporters. The major
importers of coriander from India are Europe, the US, Singapore and the Gulf
countries. Exports touched an all-time high of 33,582 tonnes during 2004-05.
However, appreciation of the rupee in 2006-7 and 200708 has marginally affected
coriander exports. Coriander exports have once again increased in 2007-08, touching
26,000 tonnes, aggregating Rs110.25 crore. Due to higher export prices, coriander
exports have increased sharply in value terms by 47.75% Y/Y compared to a 26.83%
Y/Y rise in volumes in 2007-08 (Karvy Comtrade Limited, 2008).

3.2 Analytical framework

Keeping in view the objectives formulated for the study, the methods of
analysis are presented under the following heads:

3.2.1 To examine the extent of market integration between spot and future market
prices for coriander.

3.2.2 To understand the behaviour and pattern of causality between spot and future
market for coriander.

3.2.3 To assess the efficiency of coriander future market in its role of price
discovery and risk management function.

3.2.4 Forecasting of spot and future prices for coriander.

3.2.1 To examine the extent of market integration between spot and future market
prices for coriander.

3.2.1.1 Unit root test

In order to examine the cointegration of spot and futures prices for
establishing long-term relationship, it is customary to first examine the level of
integration in each of the given commodity price series. This can be done by
examining the unit root properties in each commodity price series. In this study,
augmented Dickey-Fuller (ADF) test will be employed to examine whether the
sample commodity price series are stationary or otherwise. Further, the necessary lag
length of the data series will be selected on the basis of Schwarz Information Criterion
(SC).
3.2.1.2 Johansen co-integration test

After testing the precondition of stationarity in the spot and futures market prices, Johansen and Juselius cointegration test will be employed to determine the existence of a long-run relationship between the selected agricultural commodities traded in NCDEX. Johansen’s Cointegration test is employed to examine long-run relationship among the variables after they are integrated in an identical order. The nature of long-run equilibrium relationship between the spot and futures market for the sample agricultural commodity derivatives is ascertained using Johansen’s Cointegration test. In practice many financial variables contain unit root and are thus said to be integrated at first difference. The cointegration test is based on maximum likelihood estimation and uses two test statistics, namely trace statistics ($\lambda_{\text{trace}}$) and maximum Eigen value statistics ($\lambda_{\text{max}}$) to determine the number of cointegrating vectors.

3.2.1.3 Lead-lag relationship

After identifying cointegration between spot and future prices of the selected agricultural commodities, the regression model will be employed to examine the lead-lag relationship between the two and it is presented below:

$$\Delta \text{SPOT}_t = C_1 + \sum_{k=1}^{x} \alpha_{1i} \Delta \text{SPOT}_{t-k} + \sum_{k=1}^{x} \beta_{2i} \Delta \text{FUT}_{t-k} + u_{1t}$$

$$\Delta \text{FUT}_t = C_2 + \sum_{k=1}^{x} \beta_{1i} \Delta \text{FUT}_{t-k} + \sum_{k=1}^{x} \alpha_{2i} \Delta \text{SPOT}_{t-k} + u_{2t}$$

Where, SPOT$_t$ and FUT$_t$ are spot and future market prices of individual agricultural prices at time $t$, $u_{1t}$ and $u_{2t}$ are white noise disturbance terms.

3.2.1.4 Vector Error Correction Model (VECM)

VECM model can only be used if variables have cointegration. Before evaluating efficiency and calculating VECM, lag length selection is important. To capture the deviations in the short run within the framework of cointegration analysis, which implies stable long run co-movement, adjustment to temporary deviation will be analysed using VECM. The VECM allows for the short run shocks and estimates the degree of convergence towards the long run relationship. Before estimating the VECM, the numbers of lags of the spot and futures price series that are included in the VECM will be identified on the basis of VECM lag order selection criteria. The equations for the test will be specified as,
Methodology

\[
\Delta F_t = \alpha + \delta u_{t-1} + \sum_{i=1}^{k} \beta_i \Delta F_{t-i} + \sum_{j=1}^{k} Y_i \Delta S_{t-i} + V_t
\]

\[
\Delta S_t = \alpha + \delta u_{t-1} + \sum_{i=1}^{k} \beta_i \Delta F_{t-i} + \sum_{j=1}^{k} Y_i \Delta S_{t-i} + V_t
\]

Where, F and S refer to futures and spot prices, respectively of the selected commodities and V is white noise.

3.2.2 To understand the behaviour and pattern of causality between spot and future market for coriander.

3.2.2.1 Granger Causality

Eagle Granger suggest that if cointegration exists between two variables in long run, then there must be either unidirectional or bidirectional Granger causality between these variables. The joint significance of all the lag of spot and futures prices included in the VECM will be examined using Granger causality which measures precedence and information contents. Null hypothesis is no causality between futures and spot prices. The rejection of the null hypothesis implies information transmission from spot prices to future prices and vice versa. If the lags of both series are significant, there is a bi-directional flow of information. The equations for the test are specified as

\[
F_t = a_0 + a_1 F_{t-1} + \ldots + a_p F_{t-p} + b_1 S_{t-1} + \ldots + b_p S_{t-p} + U_t
\]

\[
S_t = c_0 + c_1 S_{t-1} + \ldots + c_p S_{t-p} + d_1 F_{t-1} + \ldots + d_p F_{t-p} + V_t
\]

where, F and S refer to futures and spot prices, u, v are white noises.

3.2.3 To assess the efficiency of coriander future market in its role of price discovery and risk management function.

Time series data pertaining to the sample agricultural commodity derivatives will be analysed using various statistical and econometric tools. Firstly, the base data of spot and futures prices will be converted into continuous daily return series by taking natural logarithm. Natural logarithm of daily prices will be taken to minimize the heteroscedasticity in data. The rate of return for each day is defined as the
difference between the natural log of a particular day’s price and the natural log of the previous day’s price.

\[ Rt = [\ln(P_t) - \ln(P_{t-1})], \]

Where, \( Rt \) is the return for the day \( t \), \( \ln \) is the natural log, \( P_t \) and \( P_{t-1} \) are the closing prices for day ‘t’ and its previous trading day. However the base data of spot and futures prices will also be used at appropriate stages of the analysis. In order to assess the effectiveness of futures market for cotton the present study proposes two sets of tools comprising ratio tests and cointegration and related technique for examining the price discovery and market risk management. In the process, performance of coriander futures market will be assessed on the basis of extent of liquidity, price volatility and basis risk. On the other hand, the efficiency of the commodity is examined on the basis of the econometrics techniques of cointegration, VECM, Granger causality, impulse response and variance decomposition.

3.3.3.1 Descriptive statistics

The key statistics including mean, standard deviation, normality (using Jarque-Bera test), kurtosis and skewness will be investigated for both spot and futures prices.

3.2.3.2 Ratio Tests

To assess the performance of pepper futures market, ratio test employed are:

Extent of liquidity, price volatility and basis risk.

Total traded volume in transaction and total production of Coriander are taken in tonnes. Extent of liquidity is measured in terms of the proportion of the volume of transaction in the futures market to the total production in the country in a year in per cent. The volume of transaction is indication of performance. Increasing volumes indicates greater interest in futures trading and therefore generates higher liquidity in market. While volume could vary from year to year depending upon on the volatility in prices, a clear upward or downward trend is an indication of the stakeholder’s perceived utility of a futures exchange. Liquidity in a futures market has a direct bearing on the transaction cost and therefore affects participation in futures.

1. Extent of Liquidity = \( \frac{\text{Total traded volume of coriander seed}}{\text{Total production of coriander seed}} \)

Prices of spot and futures of Coriander are taken in Rs. /Quintals. Price volatility is calculated as ratio of S.D. of the futures prices to the S.D. of spot prices.
The ratio of S.D. of futures and spot prices are considered on a two monthly average basis to assess the futures market performance. If markets are efficient, day to day variations in cash and futures prices are purely a result of new information available. Therefore, for storable commodities, the extent of variation in cash and futures should be similar. Given the competitive structure of the physical markets for agricultural commodities, we assume that these markets are discovering prices efficiently. We also assume that the cost of carrying charges is negligible. The study has compared the extent of volatility in cash and futures prices to examine the extent to which information is incorporated in futures prices. However, with many restrictions on futures trading such as price and position limit, the price variations are artificially curtailed. Also, due to reasons such as the direction of causality, this measure is treated as only indicative. In such a formulation a ratio close to one indicates that a futures price is able to incorporate information efficiently. A ratio greater than one would indicate speculative activities. A ratio less than one indicate that information is not fully incorporated.

2. Price Volatility = \( \frac{\text{Standard deviation of future prices of coriander seed}}{\text{Standard deviation of spot prices of coriander seed}} \)

### 3.2.4 Forecasting of spot and future prices for coriander.

Forecasting of spot and futures prices was done using the ARIMA Model. The acronym ARIMA stands for “Auto-Regressive Integrated Moving Average.” Lags of the differenced series appearing in the forecasting equation are called “auto-regressive” terms, lags of the forecast errors are called “moving average” terms, and a time series which was differenced to be made stationary is said to be an “integrated” version of a stationary series. Random-walk and random-trend models, auto-regressive models, and exponential smoothing models (i.e., exponential weighted moving averages) are all special cases of ARIMA models. It was first given by Box and Jenkins in 1970. ARIMA Modelling consists of four operational steps:

(i) Identification : Using graphs, statistics, ACFs and PACFs, transformations, etc. to achieve stationary and tentatively identify patterns and model components

(ii) Estimation: Determine coefficient and estimate through software application of least squares and maximum likelihood methods.
(iii) Diagnostic: Using graphs, statistics, ACFs and PACFs of residuals to verify whether the model is valid. If valid then use the decided model, otherwise repeat the steps of Identification, Estimation and Diagnostics.

(iv) Forecast: Using graphs, simple statistics and confidence intervals to determine the validity of the forecast and track model performance to detect out of control situation.

3.2.4.1 ARIMA Model Assumption:

In ARIMA terms, a time series is a linear function of past actual values and random shocks, that is,

\[ Y_t = f(Y_{t-k}, e_{t-k}) + e_t \]

Where \( k > 0 \).

In ARIMA model, we do not have a forecasting model \textit{a priori} before Model Identification takes place. ARIMA helps us to choose “right model” to fit the time series.
ARIMA Notation:

- **AR(p)** Where \( p \) = order of autocorrelation
  (Indicates weighted moving average over past observations)
- **I(d)** Where \( d \) = order of integration (differencing)
  (Indicates linear trend or polynomial trend)
- **MA(q)** Where \( q \) = order of moving averaging
  (Indicates weighted moving average over past errors)