Chapter 3

METHODOLOGY
3. METHODOLOGY

Systematic designing of the study is sine-qua-non for any scientific inquiry. In this chapter, study area, selection of sample, sources for data and analytical tools and techniques employed in the study are presented.

3.1 Study Area

Study on “Cashew Plantation in India- Establishment, Management and Viability” is undertaken at Southern - India level; however, states with less importance in respect to production are exempted. The major cashew growing states in the country are in the southern part of India. Main cashew growing area in the Karnataka state is Dakshina Kannada, Quillon in Kerala, Dharmapuri in Tamil Nadu & Godhavari in Andhra Pradesh are selected for the study.

The cashew developments in other countries, which are the major suppliers of raw cashew to India like Guinea – Bissau, Ivory Costa, Tanzania, Senegal, Nigeria, Kenya, Mozambique and Benin are chosen at the global level to understand the present situation.

To study the different methods of raw cashew processing technologies available in India and to know the advantages and disadvantages of the each method, Quillon district of Kerala state is selected where drum roasting is followed as preliminary roasting; Dakshina Kannada district of Karnataka is selected to study the steam boiling method; and to study the outdated method of oil bath roasting, Palasa district of Andhra Pradesh is chosen. Goa is the only place in India where cashew fenny has been distilled from cashew apple which is selected as a study area to study the cashew fenny processing.

Study on “Demand and Supply Situation of CNSL in India” is undertaken at an all-India level; however, states with less importance both in respect of production and consumption are exempted. The major states covered under the study are Karnataka,
Kerala, Tamil Nadu, Andhra Pradesh, New Delhi, Maharashtra, West Bengal, Orissa and Goa. A detailed survey of CNSL manufacturers, users and traders is made at a national level to know the demand and supply of CNSL, problems in the CNSL industry and to profile the CNSL functionaries.

The growth in Indian cashew Industry and its trade performance is studied at all India level.

3.2 Selection of Samples
The survey covered all the key participants of the cashew industry such as the cashew growers, processors, traders, exporters, with different customized questionnaires.

They key reasons for this extensive primary survey were to understand and collect first hand information from the key players in fulfilling objectives of the study such as cost of cultivation, processing costs under different methods, demand and supply of CNSL, cashew industry problems etc. The questionnaires cover the vital aspects to achieve the objectives of the study.

Survey samples were made up of about 300 respondents including various stakeholders of the cashew industry.

Purely random sampling processes was adopted.

Face-to-face interviews, interviews through telephone, post and e-mail using structured questionnaires have been carried out; visits to the cashew growing, processing (raw cashew, cashew apple, CNSL) area and the markets were made; attended as a delegate for “World Cashew Congress” a international conference held at Cochin during February 23-25, 2001 to understand trends and practices that are fallowed by the industry and detailed discussions were conducted, besides getting answers through questionnaires.

3.3 Data Base
It is believed, the primary survey gives an inside view on the issues. But, the secondary source of information is vital for a complete study. In this context, the present study is based on both primary and secondary data.
The cultivation aspects were collected from five small, five marginal and five large farmers from Karnataka, Kerala, Tamil Nadu and Andhra Pradesh states are selected on random basis.

In order to assess the growth rates, the secondary data on year-wise area, production, productivity, exports of cashew kernel (quantity and value) and imports of raw cashew nuts (quantity and value) from 1960 to date are collected for India. The sources of the data are Cashew Export Promotion Council (CEPC), Cochin; Director General of Commercial Intelligence and Statistics (DGCIS), Calcutta; and USDA reports.

To study the cashew developments in other cashew producing countries which are supplying raw cashew nuts to India, the information is collected from cashew boards, traders, processors of respective countries, research done on cashew industry at those countries, etc.

To study the cost of raw cashew nut processing under different types of cashew processing technologies, six samples each from small, medium and large scale cashew processing units which are following different technologies of processing in Karnataka, Kerala and Andhra Pradesh are selected on a random basis.

Five cashew fenny extraction units in Goa were chosen to study the cashew fenny extraction process and cost involved for processing of fenny.

In addition to information collected through the questionnaire from grower, processors, importers, exporters, traders and other market functionaries the information is also gathered from officials of Cashew Export Promotion Council - Cochin, Directorate of Cashew and Cocoa Development Board - Cochin, Regional Research Centre for Cashew (Puttur), Professors from Agricultural Universities, Mangalore Cashew Manufacturers Association, Regional Research Laboratory – Trivandrum and Cashew Boards from other countries by direct interview, discussions and by observations.
A secondary survey of the various products of cashew apple is also undertaken.

Study on “Demand and Supply Situation of CNSL in India” is undertaken at an all-India level. As there are few persons are dealing with CNSL, the sample drawn from each state to collect information on CNSL supply were 11 from Karnataka, 3 from Kerala, 6 from Tamil Nadu, 2 from Goa, 2 from Orissa, 3 from Andhra Pradesh and one from Maharashtra. The sample drawn to collect the demand for CNSL were 4 from Andhra-Pradesh, 2 from Karnataka, 6 from Tamil Nadu, 20 from West Bengal, 6 from New Delhi and 3 from Gujarat.

To analyse the direction of cashew trade, the secondary data on country-wise exports of cashew kernels from India, country-wise raw nut imports to India for 9 years and year-wise exports of kernels and imports of raw nuts from 1990 are collected from Cashew Export Promotion Council, Cochin and DGCI & S, Calcutta.

To examine the market integration, weekly Indian FOB price, Rotterdam (Netherlands) and New York cashew kernel prices, exchange rates, etc., are collected from CEPC and other trade sources.

To analyse the strengths, weakness, opportunities and threats of the Indian cashew industry, the information is collected from the major grower, processor, trader, exporters and officials from the Cashew Export Promotion Council, Cochin in addition to the available secondary sources.

3.4 Analytical Tools Used in Research Work

3.4.1 Economics of cost and financial viability

In order to study the economics of cost and financial viability of commercial scale cashew plantation in India, Discounted Cashew Flow method is used.

There are three techniques for calculating the financial viability under this method.

The following steps are involved in the calculation of BC ratio and NPV

- Cash flow of the investment should be forecasted based on realistic assumptions
- Appropriate discount rate should be identified to discount the forecasted cash flow
- Present value of cash flow should be calculated using opportunity cost of capital as the discount rate
- NPV should be found out by subtracting present value of cash outflow from present value of cash inflow
- BC ratio should be found out by dividing present value of cash outflow from present value of cash inflow.

NPV

The Net Present Value (NPV) of a project indicates the expected impact of the project on the value of the firm. Projects with a positive NPV are expected to increase the value of the firm. Thus, the NPV decision rule specifies that the projects with a positive NPV should be accepted.

The NPV is calculated as the present value of the project's cash inflows minus the present value of the project's cash outflows. This relationship is expressed by the formula:

\[
NPV = \sum_{t=0}^{T} \frac{CF_t}{(1+r)^t} = CF_0 + \frac{CF_1}{(1+r)} + \frac{CF_2}{(1+r)^2} + \ldots + \frac{CF_T}{(1+r)^T}
\]

Where,

- \( CF_t \) = the cash flow at time \( t \) and
- \( r \) = the cost of capital.

BC ratio

The BC ratio is calculated as the present value of the project's cash inflows divided by the present value of the project's cash outflows. Thus, the BC ratio decision rule specifies that the projects with more than one BC ratio should be accepted.
**Internal Rate of Return (IRR)**

The Internal Rate of Return (IRR) of a project is the discount rate at which the Net Present Value (NPV) of a project equals zero. The IRR decision rule specifies that the projects with an IRR greater than the cost of capital should be accepted.

\[
NPV = 0 = \sum_{t=0}^{T} \frac{CF_t}{(1 + IRR)^t} = CF_0 + \frac{CF_1}{(1 + IRR)^1} + \frac{CF_2}{(1 + IRR)^2} + \ldots + \frac{CF_T}{(1 + IRR)^T}
\]

Where,
- \(CF_t\) = the cash flow at time \(t\) and
- \(r\) = the cost of capital

The sensitivity of the cashew plantation is calculated by using the decreasing cash inflow or increasing cash outflow.

The hypothesis to be tested is,
- \(H_0\): NPV <0, B: C ratio < 1, IRR < Interest Rate against (Project is economically not viable)
- \(H_1\): NPV >0, B: C ratio > 1, IRR > Interest Rate (Project is economically viable)

**3.4.2 Compound Annual Growth Rate (CAGR)**

Compound annual growth rates of area under cashew, raw cashew nut production, productivity, exports of cashew kernel and imports of raw cashew nut are analysed using the exponential growth function of the form,

\[
Y(t) = a \cdot b^u \cdot e^\text{t}
\]

Where,
- \(Y_t\) = area, production, productivity, export and import of cashew in India in year ‘t’
- \(t\) = year which takes value 1,2... \(n\)
- \(e\) = Exponent term (2.3018).
- \(u\) = Disturbance term.
- ‘a’ and ‘b’ are parameters to be estimated.
The logarithmic form of the equation \( Y(t) = a b^t e^{\theta} \) is, \( \ln Y = \ln a + t \ln b + e \).

The compound growth rate \((g)\) in percentage is computed from the relationship,

\[
(g) = (\text{Anti ln of ln } b) - 1 \times 100 \quad \text{or} \quad (g) = (b - 1) \times 100
\]

Where,

- \( g \) = estimated compound growth rate in per cent per annum.
- \( b \) = antilog of ln \( b \)

The standard error of the growth rate is estimated and tested for its significance with student’s \('t'\) statistic.

### 3.4.3 Co-efficient of Variation (Instability analysis)

In order to study the variability in export and import trade of cashew, an index of instability is developed as a measure of variability.

The coefficient of variation (CV) is calculated using the formula

\[
CV = \frac{\text{Standard Deviation}}{\text{mean}}
\]

The formula suggested by Cuddy and Della (1978) is used to compute the index of instability.

\[
\text{Index of Instability} = \frac{\text{Standard Deviation}}{\text{mean}} \times 100 \times \sqrt{1 - R^2}
\]

Coefficient of variation is multiplied by the square root of the difference between the unity and coefficient of multiple determinations \((R^2)\) in the cases where \(R^2\) is significant.

### 3.4.4 Markov Chain analysis

**Definition of Markov analysis**

The changes in the exports of cashew kernel to different countries are analyzed by developing a model with first order finite Markov chain property which captures the net effect in the change in exports of cashew kernel over a period of time.
We are often interested in characterizing or summarizing how economic parameters concerning cashew kernel exports have changed over time and predicting their behaviour in future. It is a common feature and meaningful too to associate occurrence of present to being related to occurrence in the recent past, which in fact is the essence of Markov chain analysis.

Markov chain analysis is a method of analyzing the current behavior of some variable in an effort to predict its future behavior. This procedure is developed by Russian mathematician Andrei A. Markov in 1912. He first used it to describe, and predict the behavior of particles of a gas in a closed container.

**Different orders of Markov analysis**

The Markov process can be of different order. The 'first order' Markov process is based on the assumption that the probability of the next event depends upon the outcome of the last event and not at all on any earlier buying behavior.

A 'second order' Markov process assumes that the probability of the next event depends upon the outcome of the past two events.

Similarly, a 'third order' process is based upon the assumption that future behavior is best predicted by observing its past three events.

In second and third order processes, however, the computation becomes more cumbersome and difficult. Studies suggest that using first order assumption for prediction purposes is not invalid, particularly if data appears to indicate that it follows a fairly stable pattern; that is, if the matrix of transition probabilities remain stable, they proceed to be reliable predictors of future behavior.

**Theoretical framework**

The 'first order' finite Markov model is a stochastic process \( (X_t) \) which describes the finite number of possible outcomes \( S \) \((i = 1, 2, 3... r)\), and is a discrete random variable
$X_i$ ($t = 1, 2, 3...t$). This is said to have the 'first order' Markovien property if the conditional probability distribution of $X_i$ and is dependent only on the state the system is in at step 'i', and not in steps 0, 1, 2, 3, i-1. Mathematically, the stochastic process ($X_i$) has Markovien property if,

$$P(X_{i+1} = S/X_i = r, X_{i} = T_1... X_{i-1} = t, X_i = r) = P(X_{i+1} = S/X_i = r)$$

Where, $P(X_{i+1} = S/X_i = r)$ is the one-step transitional probability of going from state 'r' at step 'i' to state 's' at step i+1. Thus, these represent the conditional probability of $X_{i+1}$ given $X_i$. If for each 'r', and 's', $P(X_{i+1} = S/X_i = r) = P(X_i = S/X_0 = r) = P_{rs}$ for all 'i', then the one-step transitional probability remains stationary.

One-step stationary transition probabilities takes into consideration one and only one state at each point in time, i.e., this will be useful for estimating the share of market for one step (year) from now only. As the present study uses annual export data for analysis for predicting the future exports of cashew kernel to different countries from n-step (year) from now, the n-step stationary transitional probability property will be appropriate.

The n-step transitional probabilities are defined as,

$$P_{rs}^{(n)} = P(X_{i+n} = S/X_i = r) = P(X_n = S/X_0 = r)$$

Where,

$$P_{rs}^{(n)} > 0 \text{ for all states } r, \text{ and } s; \quad n = 1, 2...$$

$$\sum_{s=0}^{n} P_{rs}^{(n)} = 1 \text{ for all states } r; \quad n = 1,2...$$

The above equation assumes there are $n+1$ possible states. Note that if the system is currently in state 'r', it must be in some state 'n' steps from now. In general, the 'n' step stationary transition probabilities can be calculated as follows.

$$P_{rs}^{(n)} = \sum_{j=0}^{n} P_{rj} P_{js}^{\frac{n}{n}}$$
Where the possible states are 0, 1, 2, 3…N. I.e., probability of going from state 'r' to state 's' in 'n' steps is the probability of going from state 'r' to state 's' in one step times the probability of going from state 'j' to state 's' in n-1 steps, summed over all j = 0, 1, 2, 3…n.

**Specification of the Markov Chain model**

The share of export of cashew kernel from India (Xᵢᵗ) to a particular country (j th) at time 't' is considered as a random variable, and this depends on only its past export with that country. Following the first order stationary Markovian property as discussed above, the model can be specified as follows:

\[ X_{jt} = \sum_{i=0}^{n} X_{i+1} P_{ij} + e_{jt} \]

Where,
- \( X_{jt} \) = the export of cashew kernel from India to j th country during the year ‘t’.
- \( X_{i+1} \) = the export to i th country during the year t-1.
- \( P_{ij} \) = the probability that exports will shift from i th country to j th country.
- \( e_{jt} \) = the error term independent of \( X_{i+1} \).
- \( n \) = the number of importing countries.

The transitional probability (\( P_{ij} \)) is the centre of the Markov chain model analysis and will have the following properties:

\[ 0 < P_{ij} < 1 \]

\[ P_{ij} = 1, \text{ for all 'i'} \]

The transitional probability \( P_{ij} \) indicates the possibility that exports will switch over from country ‘i’ to country ‘j’ with passage of time. The probabilities \( P_{ij} \) for i = j indicate the gains or losses in exports of each of the importing country. The probability \( P_{ij} \) for i = j (diagonal probabilities) indicate probability of retention of an importing country.

**Estimation of Markov Chain model**

There are several approaches to estimate the transitional probabilities of the Markov
chain model such as Unweighted Restricted Least Squares, Weighted Restricted Least Squares, Bayesian, Maximum Likelihood, Unrestricted Least Squares etc. In the present study, Minimum Absolute Deviations (MAD) estimation procedure was employed to estimate the transitional probabilities, which minimizes the sum of absolute deviations. The conventional Linear Programming (LP) technique was used as this satisfies the properties of transitional probabilities of non-negativity restrictions and row sum constraints in estimation.

The linear formulation is

\[ \text{Min } OP^* + I e \]
Subject to,
\[ XP^* + V = Y \]
\[ GP^* = 1 \]
\[ P^* > 0 \]

Where,
- \( O \) = the vector of zeroes.
- \( P^* \) = the vector of the probability \( P_{ij} \).
- \( I \) = an appropriately demonstrated identity matrix.
- \( e \) = a vector of absolute errors, \( I V I \).
- \( Y \) = the vector of exports of each country.
- \( X \) = the block diagonal matrix of lagged values of \( Y \).
- \( V \) = the vector of errors.
- \( G \) = the grouping matrix to add row elements of \( P \) arranged in \( P^* \) to unity.

**Goodness of fit**

To test whether the observed shares of exports to different countries and the predicted shares from the Markov chain model follow similar distributions, the chi square statistics of the following type is used.

\[ \sum_{(r-1)}^2 = \sum_{i=1}^T \sum_{i=1}^r N(t) \{ Y_{ij}(t) - y_{ij}(t) \}^2 / y_{ij}(t) \]
Where,

\( Y_i(t) \) = observed proportions of \( i^{th} \) country’s share at time \( t \).
\( \hat{y}_i(t) \) = predicted proportions of \( i^{th} \) country’s share at time \( t \).
\( N(t) \) = total number of countries at time \( t \).
\( r \) = countries.
\( T \) = year.

**Determination of equilibrium condition**

Regular stochastic matrices such as the one mentioned earlier have the property that when raised in power, all rows tend to converge to a unique vector, which may be called \( K \). Since \( K \) represents the final probabilities of being in each commodity category, it gives the equilibrium value of export of commodity at time \( t_1 \).

The same methodology is used to analyse the changes in the imports of raw cashew nuts from different countries.

**3.4.5 Co-integration (Market integration)**

The starting point of testing the Law of One Price (LOP) is the determination of the order of integration of the price series. If the series is stationary, it means that the series has basic statistical properties implying that it is invariant with respect to time. Such series will have a constant mean, a constant finite variance and covariance between observations that depend only upon their distance apart in time.

\[ P_t = P + e_t \]

Where, \( P_t \) is the observed value of the series at time ‘\( t \)’ and \( e_t \), a random disturbance term. The series \( P_t \) is said to be stationary, expressed as I (0). But, often price series tend to display an increase or a decrease, which violates the above condition. Successive differencing reduces the series to stationary, thus,

\[ P_t - P_{t-1} = e_t \]
\[ P_t = P_{t-1} + e_t \]
A series, which becomes stationary after first differencing is said to be integrated of order one and it is expressed as $I(1)$. Generally, a series may have to be differenced ‘d’ times to become stationary. In this case it is termed as $I(d)$. A major difference between $I(0)$ and $I(d)$ series is that the $I(0)$ series has a finite mean and variance, while in $I(d)$ series, these magnitudes do not exist.

Given the prices of cashew in India ($P_{1t}$) and the cashew prices in the overseas markets ($P_{2t}, P_{3t}, P_{4t}, \ldots$) in time ‘t’. The following three situations can be identified.

**i) Both the price series are having the order of stationarity equal to zero**

$P_{1t} \sim I(0)$ and $P_{2t} \sim I(0)$. In this case, the means and variances exist. This intern implies that the LOP holds on a long run relationship as both prices fluctuate around their mean. Differences in their mean are possible reflecting the fixed components such as middlemen’s profits and other factors. In such a situation, it is valid to regress $P_{1t}$ on $P_{2t}$ and test the restriction that the slope coefficient equals one and the intercept term equals zero.

**ii) Both the price series having the same order of integration, but greater than zero**

$P_{1t} \sim I(d)$, $P_{2t} \sim I(d)$, where, $d > 0$. In this case both the price series have the same order of integration which is greater than zero. Hence, additional information is needed to examine the validity of LOP. Such information is obtained from the theory of co-integration, which states that even though the same explosive pattern characterises both prices, yet there must exist a parameter which brings them together in the long run so that their linear combination is of a lower order of integration than the original series. In such a case $P_{1t}$ and $P_{2t}$ form a cointegrated system.

**iii) Both the price series having different order of integration**

$P_{1t} \sim I(d)$, $P_{2t} \sim I(b)$ where, $d \neq b$. In this case, prices have different order of integration and the LOP does not hold because at least one of either $P_{1t}$ or $P_{2t}$ will exhibit explosiveness. This can be understood if, $P_{1t}$ is $I(0)$ and $P_{2t}$ contains an explosive component, which cannot be explained by $P_{1t}$ alone.
In general, to determine whether $P_{1t}$ and $P_{2t}$ are cointegrated, the following regression is employed.

$$P_{1t} = a + b P_{2t} + e_t$$

Where, ‘$a$’ and ‘$b$’ are the parameters to be estimated. If $e_t$ is integrated by the order $b$ ($b < d$), $P_{1t}$ and $P_{2t}$ are said to be cointegrated. When $b = 0$, the LOP holds since both the prices move together in the long run. In particular, the LOP postulates that the cointegration parameter $b = 1$. Thus, cointegration test is transformed into a stationarity test of difference between the two price series. This has been illustrated hereunder.

Let $P_{1t}$ and $P_{2t}$ denote (non-stationary) prices of domestic and overseas market respectively. Central to LOP, $P_{1t}$ and $P_{2t}$ which form a cointegration system.

Let $P_{1t} = b P_{2t} + e_t$ be the cointegration regression where, $b$ represents the cointegration parameter. If ‘$b$’ is unity, it is interpreted as the long run counter part of the association between the own price series.

Expressing $P_{2t}$ in terms of $P_{1t}$ as, $P*_{2t} = b P_{2t}$

And substituting it in the cointegration regression yields,

$$P_{1t} = P*_{2t} + e_t$$

This has the cointegration parameter of one. Thus, a stationarity test of $e_t$ itself is a sufficient test of cointegration which can be expressed as

$$e_t = P_{1t} - b P_{2t}$$

The analysis of stationarity is carried out for the price series of domestic and overseas markets. Further, the order of stationarity within each market was examined by differencing the respective price series as follows,

$$\Delta P_{1t} = P_{1t} - P_{1t-1}$$

To determine the order of integration, the following procedure referred as the Dickey-Fuller test was employed which was based on the regression.
\[ P_t = a + b \, P_{t-1} \]

Subtracting \( P_{t-1} \) from both sides, we get

\[ P_t - P_{t-1} = a + b \, P_{t-1} - P_{t-1} \]

Thus,

\[ \Delta P_t = a + (b-1) \, P_{t-1} \]

\[ \Delta P_t = a + c \, P_{t-1} \]

Where, \( c = (b-1) \)

The regression coefficient should be negative and significantly different from zero for ‘b’ in equation \( P_t = a + b \, P_{t-1} \) to be zero. If not ‘b’ assumes a value, hence the series is not stationary.

The hypothesis to be tested is,

\[ H_0: P_t \text{ is not } I(0) \text{ against } H_1: P_t \text{ is } I(0) \]

\( H_0 \) is rejected if the estimate of ‘b’ is negative and significantly different from zero.

### 3.4.6 Tabular analysis

The primary data collected is subjected to tabular analytical technique with the help of percentages, averages and ratios in order to obtain meaningful results and the same are interpreted.