Chapter

Design of the study
DESIGN OF THE STUDY

The present study is an attempt to evaluate the performance of almost all the time-series models in forecasting the prices of the essential commodities viz. Paddy, Rice, Coir, Arecanut, Palm-jaggery, Tapioca and Coconut in Kanyakumari District, Tamilnadu. We also present here the details of the study area, method of data collection and the methodology of different time-series and econometric models used in this study.

Study Area

Kanyakumari district was purposely selected because of the familiarity of the researcher with the district and also on the prior information on the availability of the data and also the sources.

This district was carved out from the erstwhile Travancore Cochin state and merged with Tamilnadu during the state reorganization in the year 1956. This district comprises of four revenue taluks namely Agastheeswaram, Thovalai, Kalkulam and Vilvancode, which also represent the four Agricultural divisions namely Nagercoil, Boothapandy, Thuckalay and Kuzhithurai.
respectively. There are nine development blocks, four municipalities and one
township.

Kanyakumari is the smallest district in the State with an area of
1672 Sq.km. The present population is 16.69763 lakhs where 8.29542 lakhs
are males and 8.40221 lakhs are females (as per 2001 Census) with a density of
957 per sq.km. The present population in literacy is 13.20564 lakhs where
males are 6.74939 lakhs and females are 6.45625 lakhs. The present
population having age between 1 to 6 are 1.70693 lakhs, where males are
86,912 thousand and females are 84,051 thousand. The district lies between
77°.05' and 77°.36' of the eastern longitude and 8°.03' and 8°.35' of the
northern latitude.

It is bound by Tirunelveli district in the north and northeast, by Kerala
state in the northwest and has sea in the southeast (Bay of Bengal), south
(Indian Ocean) and west (Arabian Sea).

Based on the agro climatic and topographic conditions, the district can
be divided into three regions, namely,

1. The up lands - Comprising of hills and hill bases suitable for
growing crops like Rubber, Cloves, Nutmeg, Pepper, Pineapple etc.
2. The middle lands - Comprising of plains and valleys fit for growing crops like paddy, tapioca, banana, coconut etc.

3. The low lands - Comprising the coastal belt ideal for growing coconut, cashew etc.

In Kanyakumari district there are two cropping seasons one is Kumbapoo (long duration) and another is Kannipoo (short duration). They uniformly cultivate brown rice called Samba throughout the district. This rice is prefered by this district people for meals as well as for preparing tiffen like idly, kuzhai puttu etc., Another short term crop is Ponmani is the reason behind choosing these two varieties for forecasting the prices.

Here the prices quoted for 1 unit = 100 kgs.

Tapioca is the staple food for this district people. There is no gradation in this. However they also consume modified forms of this crop one is chips and the other is flour. Flour is used for animal feeding, also for the preparation of pasting materials. Hence in the case of tapioca, we have considered prices of tapioca and also its major bi-products. Here the prices quoted are

i) Chips - 1 unit = 100 kgs

ii) Flour - 1 unit = 100 kgs
In the case of coconut there is only gradation the biggest is taken as large (called grad I), normal size as medium (called grad II) and the smallest size as small (called grad III). Thus study of these three grades separately became essential. This grading is mainly due to the proportion of oil content. 

Here the prices quoted for 1 unit = 100 nuts.

A bi-product of coconut is the coir. In the coir industry the gradation is based on the fineness of coir fibre. Based on this, the coir is divided into fine and medium.

Here the prices quoted for 1 unit = a bundle consisting of 50 pieces, each of length 40 feet.

Even today in Kanyakumari district in town as well as villages, people use Palm product called as Palmgur or Palm-jaggery as main ingredient for coffee making. Since coffee made out of these palm–jaggery does the minimum health hazard to individual. The people also have the habit of eating certain special form of jaggery as raw in between free time at work. Hence two types of these products are manufactured the medium for coffee and fine for direct consumption.

Here the prices for 1 unit = 100 kgs.
There is no gradation in Arecanut, since only local varieties are used for regular consumption alone with betel wine leaves.

Here the prices for 1 unit = 1000 nuts.

**Vital Statistics**

Land Classification (FASLI 1408) 1998 – '99 in hectares

1. Forests : 54155
2. Barren & uncultivated land : 3338
3. Land put to non-agricultural use : 25091
4. Cultivable waste : 145
5. Permanent pastures and other grazing lands : 75
6. Land under miscellaneous tree crops and groves not included is net area sown : 417
7. Current fallows : 1002
8. Other fallow lands : 1134
9. Net area sown : 82667

Total : 16,7184
1. FORESTS
2. BARREN & UNCULTIVABLE LANDS
3. LAND PUT TO NON-AGRICULTURAL USES.
4. CULTIVABLE WASTE
5. PERMANENT PASTURES AND OTHER GRAZING LANDS
6. LAND UNDER MISCELLANEOUS CROPS
7. CURRENT FALLOWS.
8. OTHER FALLOW LANDS
9. NET AREA SOWN.
Climate and Rainfall

The district is having a favourable agro climatic condition as to grow a number of crops. Even though the district is small, it can accommodate various foods and non-food crops growing under different agro climatic conditions. Because of it’s being nearer to the equator, its topography and other climatic factors favour the growth of varied crops.

The general climate of the district is pleasant. Both the southwest monsoon and the northeast monsoon influence the climate of the region, besides the proximity of the sea and the dwindling heights of the western-ghats. The normal rainfall for the district is 1443 mm. The month wise and season wise normal rainfall is given below:

<table>
<thead>
<tr>
<th>Month</th>
<th>Rainfall (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>29.70</td>
</tr>
<tr>
<td>February</td>
<td>20.60</td>
</tr>
<tr>
<td>March</td>
<td>47.80</td>
</tr>
<tr>
<td>April</td>
<td>111.70</td>
</tr>
<tr>
<td>May</td>
<td>158.20</td>
</tr>
<tr>
<td>June</td>
<td>211.40</td>
</tr>
<tr>
<td>July</td>
<td>150.00</td>
</tr>
<tr>
<td>August</td>
<td>87.60</td>
</tr>
<tr>
<td>September</td>
<td>102.70</td>
</tr>
<tr>
<td>October</td>
<td>246.80</td>
</tr>
<tr>
<td>November</td>
<td>206.20</td>
</tr>
<tr>
<td>December</td>
<td>70.30</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1443.00</strong></td>
</tr>
</tbody>
</table>
Season wise rainfall details (in mm)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter (January – February)</td>
<td>50.3</td>
<td>17.8</td>
<td>77.1</td>
<td>41.2</td>
<td>19.5</td>
<td>10.6</td>
<td>2.8</td>
<td>52.7</td>
</tr>
<tr>
<td>Summer (March – May)</td>
<td>317.7</td>
<td>205.2</td>
<td>267.6</td>
<td>509.6</td>
<td>213.5</td>
<td>269.2</td>
<td>240.9</td>
<td>458.6</td>
</tr>
<tr>
<td>South west Monsoon (June – September)</td>
<td>551.7</td>
<td>477.3</td>
<td>589.5</td>
<td>502.8</td>
<td>571.7</td>
<td>760.3</td>
<td>667.8</td>
<td>590.4</td>
</tr>
<tr>
<td>North east Monsoon (October-December)</td>
<td>523.3</td>
<td>1039.5</td>
<td>630.2</td>
<td>607.9</td>
<td>646.5</td>
<td>653.4</td>
<td>1069.3</td>
<td>602.0</td>
</tr>
<tr>
<td>Total</td>
<td>1443</td>
<td>1739.5</td>
<td>1564.4</td>
<td>1661.5</td>
<td>1451.2</td>
<td>1693.5</td>
<td>1980.8</td>
<td>1703.7</td>
</tr>
</tbody>
</table>

Soils

In this District vast area comes under Utisol in between Nagercoil and Kaliakavilai. These Utisols are heavily leached soils, reddish and yellowish in colour, low organic matter content, low nitrogen, phosphorous and potash. These soils are acidic (pH 6.0) in soil reaction. Cation exchange capacity is low and phosphate fixation is high.

Altisols are observed in cultivable tracts of Agasteeswaram and Thovalai areas. Weathered granite is the main parent material. These soils are
low in organic matter, almost neutral pH (6.5 to 7.5) low to medium status of nitrogen content.

Entisols and Inceptisols are particularly found in coastal areas. A layer of sand alternated with a layer of silt on clay is observed.

Entisols are observed in Aralvoimozhi area. These soils are mainly associated with Granite rock and are richly associated with limenite and manganesite clays. These soils are low in organic matter and contain much insoluble material. pH range of these soils is around 5.5 to 7.5. These soils are low in N, P, and K content, with low cation exchange capacity.

**Major Crops**

<table>
<thead>
<tr>
<th>Crop</th>
<th>I Season</th>
<th>II Season</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paddy</td>
<td>-</td>
<td>20,000 Ha</td>
<td>40,000 Ha</td>
</tr>
<tr>
<td>Pulses</td>
<td>-</td>
<td>4,000 Ha</td>
<td></td>
</tr>
<tr>
<td>Tapioca</td>
<td>-</td>
<td>9,000 Ha</td>
<td></td>
</tr>
<tr>
<td>Banana</td>
<td>-</td>
<td>4,000 Ha</td>
<td></td>
</tr>
<tr>
<td>Coconut</td>
<td>-</td>
<td>21,000 Ha</td>
<td></td>
</tr>
<tr>
<td>Rubber</td>
<td>-</td>
<td>19,475 Ha</td>
<td></td>
</tr>
<tr>
<td>Fruits &amp; Vegetables</td>
<td>- 7,500 Ha</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Normal area and productivity of different crops

<table>
<thead>
<tr>
<th>Sl.No</th>
<th>Crop</th>
<th>Area in hectares</th>
<th>Normal yield in Kgs/Ha.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Paddy</td>
<td>40,000 (Both crops)</td>
<td>5824 Kgs. of paddy</td>
</tr>
<tr>
<td>2</td>
<td>Tapioca</td>
<td>9,554</td>
<td>15000 Kgs.</td>
</tr>
<tr>
<td>3</td>
<td>Coconut</td>
<td>21,670</td>
<td>10000 Nuts</td>
</tr>
<tr>
<td>4</td>
<td>Banana</td>
<td>4,000</td>
<td>250 Kgs.</td>
</tr>
<tr>
<td>5</td>
<td>Pulses</td>
<td>4,000</td>
<td>1200 Kgs</td>
</tr>
<tr>
<td>6</td>
<td>Rubber</td>
<td>19,450</td>
<td>2000 Kgs</td>
</tr>
<tr>
<td>7</td>
<td>Cashew</td>
<td>1,800</td>
<td>20000 Kgs</td>
</tr>
<tr>
<td>8</td>
<td>Mango</td>
<td>1,750</td>
<td>500000 Nuts</td>
</tr>
<tr>
<td>9</td>
<td>Arecanut</td>
<td>750</td>
<td></td>
</tr>
</tbody>
</table>
Cropping pattern

This district produces paddy, tapioca and oilseeds such as groundnut and coconut besides commercial crops like cashew, rubber, fruits and spices. The important feature of this district is the production of off-season mangoes.

Paddy is the main food crop of this district. It is grown in two seasons. First crop is sown in the month of April – June (Kannipoo) and second crop is raised in the month of September – October (Kumba poo).

Tapioca is raised as a subsidiary food crop in this district. The main planting season is April – May. In some pockets, September – October planting is also done as second season crop. It is purely raised as rain fed crop in Kalkulam and Vilavancode taluks and as irrigated crop in some area of Agastheeswaram taluk of this district.

Coconut is an important cash crop of this district. The main planting season is May to July.

Pulses are raised in rice – fallow and as inter crop in tapioca. The important pulses are blackgram, greengram, horsegram, cowpea and redgram. In
the months of April – May and September – October pulses are grown as intercrop with Tapioca. In the months of February – March pulses are raised as pure crop in rice-fallows.

**Vegetables** are cultivated during January – February and July – August. Vegetable is also grown as 3rd crop after the harvest of 2nd crop paddy in some pockets of this district.

**Banana** is cultivated mainly during March to May and September – October.

**Groundnut** is raised during March – April and October – November as rain fed crop.

In addition, pepper, cloves, arecanut, betelvine, cocoa, gingelly, ginger, turmeric, arrowroot etc. are also grown in this district.

**Crop sequences followed in Kanyakumari district**

1. Paddy – Paddy – Fallow (One year rotation)
2. Paddy – Paddy – Pulses (One year rotation)
3. Paddy – Banana – Paddy (Two year rotation)
4. Paddy – Banana – Tapioca (Two year rotation)
5. Paddy – Banana – Banana (Two year rotation)
Source of data

Data pertaining to the 14 items namely arecanut, coconut-big, coconut-medium, coconut-small, coir-fine, coir-medium, paddy-ponmani, paddy-samba, palm-fine, palm-medium, rice-ponni, rice-samba, tapioca-chips, tapioca-flour, were collected from the District Statistical office from 1978 to 2000.

Two types of models are used namely

1. Regression models

2. Time Series models

In the Regression models Ordinary Least Square was used to estimate the coefficients on the prediction equation. In the time-series models moving averages were mostly used to smoothing the data before forecast is being done.

1. Regression models

This includes models of the form,

1 \( y_t = b_0 + b_1 t \), called the Linear model
2 \( y_t = b_0 + b_1 t + b_2 t^2 \), called the Quadratic model
3 \( y_t = e^{b_0 + b_1 t} \), called the Growth model
4 \( y_t = b_0 + b_1 \log t \), called the logarithm model
5. \[ y_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3 \] called the cubic model

6. \[ y_t = e^{b_0 + b_1/t} \] called the S-Curve model.

7. \[ y_t = b_0 e^{b_1 t} \] called the Exponential Growth model

8. \[ y_t = b_0 + b_1/t \] called the Inverse model

9. \[ y_t = b_0 t^{b_1} \] called the Power model

10. \[ y_t = \frac{1}{1/u + b_0 b_1^t} \] called the Logistic model

where \( b_0, b_1, b_2, b_3 \) are constants, \( u \)-upper bound and \( t \)-time.

2. Time series models

Single moving average

This is simply a numerical average of the last \( N \)-data points that are used for purposes of making a forecast. As each new data point is acquired it is included in calculating the average, and the data point for the \( N^{th} \) period preceding the new data point is discarded. In general, the moving average at time \( t \), taken over \( N \) periods, is

\[ M_t^{[1]} = \frac{X_t + X_{t-1} + \ldots + X_{t-N+1}}{N} \]

Where \( X_t \) is the observed response at time \( t \).
Another way of stating the above equation to minimize required calculations is

\[ M_t^{[1]} = M_{t-1}^{[1]} + \frac{[X_t - X_{t-N}]}{N} \]

Double moving average

The single moving average will be misleading if the trend is either linear or quadratic, to get over this the double moving average is prepared. To get the double moving average we treat the first moving average as the data point and compute the moving average of this moving average.

The double moving average, \( M_t^{[2]} \) is given by

\[ M_t^{[2]} = \frac{M_t^{[1]} + M_{t-1}^{[1]} + \ldots + M_{t-N+1}^{[1]}}{N} \]

The forecast equation will be of the form\( Y_{t+T} = a_t + b_t \cdot T \), where \( T \) is the number of time periods from the present time \( t \), to the period we are forecasting, and

\[ a_t = 2 \cdot M_t^{[1]} - M_t^{[2]} \]

\[ b_t = \left( \frac{2}{N-1} \right) (M_t^{[1]} - M_t^{[2]}) \]
3 Exponential smoothing models

Single exponential smoothing

The moving average overcomes some of the difficulties encountered with regression, because with moving average it is relatively easy to take each new data point into consideration in developing forecast. The moving average requires large amount of data storage, but the exponential smoothing has most of the advantages of moving averages, but it needs only minimum data storage.

From the equations for moving averages, we have

\[ M_t^{[1]} = \frac{1}{N} X_t + (1 - 1/N) M_{t-1}^{[1]} \]

This equation can be used to calculate \( M_t^{[1]} \) if no data were stored. The exponential smoothing model is obtained by putting \( \alpha = 1 / N \) and \( S_t^{[1]} = M_t^{[1]} \).

We have the single exponential smoothing model

\[ S_t^{[1]} = \alpha X_t + (1-\alpha) S_{t-1}^{[1]} \]

(or) New Estimate = \( \alpha \) (New Data) + \( (1 - \alpha) \) (Previous Estimate)
\( \alpha \) is called the smoothing constant, which should lie between 0.01 and 0.3. Here the parameter \( \alpha \) is oscillated from 0.01 to 0.3 with step 0.01.

**Double exponential smoothing**

The double exponentially smoothed statistics is

\[
S_t^{[2]} = \alpha S_t^{[1]} + (1-\alpha) S_{t-1}^{[2]}
\]

Here we have to assume values for \( S_0^{[1]} \) and \( S_0^{[2]} \).

These smoothed statistics are further used to forecast. The forecast equation is of the form

\[
Y_{t+T} = a_t + b_t \times T
\]

where

\[
a_t = 2 S_t^{[1]} - S_t^{[2]}
\]

\[
b_t = (\alpha / 1-\alpha ) (S_t^{[1]} - S_t^{[2]})
\]

**Triple exponential smoothing**

If the data exhibits the curvature, then the better model for forecasting is triple exponential smoothing. The triple exponential smoothing statistics are calculated by
\[ S_t^{[3]} = \alpha S_t^{[2]} + (1-\alpha) S_{t-1}^{[3]} \]

The forecast equation is of the form

\[ Y_{t+T} = a_t + b_t T + c_t T^2 \]

Where

\[ a_t = 3 \ S_t^{[1]} - 3 \ S_t^{[2]} + S_t^{[3]} \]

\[ b_t = \frac{\alpha}{2(1-\alpha^2)} \left[ (6 - 5\alpha) S_t^{[1]} - 2(5 - 4\alpha)S_t^{[2]} + (4 - 3\alpha)S_t^{[3]} \right] \]

\[ c_t = \alpha^2 / 2(1-\alpha)^2 \left[ S_t^{[1]} - 2S_t^{[2]} + S_t^{[3]} \right] \]

Winter's method

Winter's is a procedure for revising the forecasting equations as new information becomes available. This method begins with the initial information and then incorporates exponential smoothing to update estimates of the intercept, the slope of the trend line, and the seasonal factors.

The basic equations for doing this are presented first, at time \( t \),

The estimate of the current intercept is

\[ a_t = \alpha (X_t / F_{t,N'}) + (1-\alpha) (a_{t-1} + b_{t-1}) \]

The estimate of the slope of the trend line is

\[ b_t = \beta(a_t - a_{t-1}) + (1-\beta) b_{t-1} \]
The updated seasonal factor is

\[ F_t = \sigma \left( \frac{X_t}{a_t} \right) + (1-\alpha) F_{t-N'} \]

The forecast equation is

\[ Y_{t+T} = (a_t + b_t^T) F^* \]

where

- \( X_t \) - the actual observation of time t.
- \( a_t \) - the estimated intercept of the trend line at time t.
- \( b_t \) - the estimated slope of the trend line at time t.
- \( N' \) - the number of observations comprising the periodicity of the data
- \( F_t \) - the estimate of the multiplicative seasonal factor for period t
- \( F_{t-N'} \) - the estimate of the seasonal factor \( N' \) periods in the past.
- \( F^* \) - used to denote our best estimate of the seasonal factor in period \( t+T \).

\( \alpha, \beta, \sigma \) - exponential smoothing constants where \( 0 < \alpha, \beta, \sigma < 1 \)

The parameter \( \alpha, \beta, \sigma \) were oscillated from 0 to 1 with step 0.1
Consider any time series, denote it by $y_t$, and suppose it depends on $p$ of its previous values according to a linear regression, called an autoregression (AR) of order $p$;

$$\text{AR}(p): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \nu_t$$

Where $\nu_t$ are serially uncorrelated perturbations with an unchanging distribution over time and $\nu_t \sim N(0, \sigma^2)$

Similarly a moving average (MA) of order $q$ is defined as

$$\text{MA}(q): y_t = \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2} - \ldots - \theta_q \nu_{t-q}$$

A time series that combines the above two equations are called a mixture of an auto regression and moving average.

$$\text{ARMA}(p,q) = \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2} - \ldots - \theta_q \nu_{t-q}$$
ARIMA model (Box – Jenkins model)

Box – Jenkins is a more general model called an Auto Regressive Integrated Moving Average (ARIMA) process. Box and Jenkins (1970) reverted to a method based on an autoregressive series, in which no prior assumption are made about discounting factors. They deal with the problem of trend by working on the differences of the series, and allow the residuals to be correlated by regarding them as a finite moving average of random $\varepsilon$’s. Thus, writing $\nabla$ for the backward difference,

$$\nabla U_t = U_t - U_{t-1}$$

**Box and Jenkins** Consider the (non-seasonal) model

$$\nabla^d U_t - \alpha_1 \nabla^d U_{t-1} - \alpha_2 \nabla^d U_{t-2} - \ldots - \alpha_p \nabla^d U_{t-p}$$

$$= \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q}, \quad \text{where} \quad \nabla U_t = U_t - U_{t-1}$$

If $B$ is the backward shift operator defined by $BU_t = U_{t-1}$, so that $\nabla = 1 - B$, this can be written

$$\alpha(B)(1-B)^d U_t = \beta(B) \varepsilon_t, \quad \text{where} \quad \alpha(B) \text{ and } \beta(B) \text{ are polynomials in } B \text{ of order } p \text{ and } q \text{ with the finite moving average of random } \varepsilon\text{'s.}$$
Input Supply

The fertilizer requirement were met by the agricultural co-operative credit societies and private dealers.

Transport and Communications

Transport

There exists a well-developed transport system, which helps the farmers in the transporting the goods and the communication system, helps them to know the current technology. The entire district is well covered with tare road.

Communications

The District has 2 head post offices, 77 departmental sub-post offices, 29 extra departmental sub-post offices, 157 extra departmental branch post offices and all the post offices having public call office / telegraph facilities available. It has 34 telephone exchanges, 90,500 telephone connections and 2100 public call offices. This enable to them to know the demand and supply position of various commodities quickly.
Finance

This district has number of financial institution extending credit facility to the farmers. Most of them avail their credit through the commercial and Co-operative banks. The Government has at present introduce Insurance scheme for these crops.

Marketing

The Kanyakumari district Market Committee was established on 30.12.1965. The following regulated markets are functioning in this district from the dates mentioned against each market.

1. Eathamozhi Regulated Market : 19.4.1967
2. Vadasery Regulated Market : 18.4.1968
There are three agricultural check posts to ensure whether cess due for the agricultural commodity notified have been remitted.

1. Aralvoimozhi check post along Kanyakumari - Tirunelveli Road
2. Anjugrammam check post along Kanyakumari - Tirunelveli Road
3. Chunkankadai check post along Nagercoil - Trivandrum Road

Extension

The state department of agriculture is providing necessary training programme frequently to solve the field problems and also adopting improved varieties and modern technologies. The state Agricultural University is the key source for all these information.

Limitation

It is to be noted that the time series approach does not entertain other influencing variables. Due to this reason even though the regression analysis permits the addition of other informative variable the researcher restricted to the same set of variables in both the types of analysis, since it helps in comparison also.