CHAPTER III

METHODOLOGY
CHAPTER III
METHODOLOGY

3.1 Nature of the Proposed Research

The proposed research is applied in nature in which the researcher shall be interested in establishing and studying the extent of regional imbalances in economic development of Goa. It shall also be the researcher's attempt to establish the historical, socio-economic and political causes for the regional disparities. At the end of the study, an attempt shall be made to evolve and recommend a development strategy to bridge the development gap in the identified regions.

This study is unique in that it is the only intra-state, sub-regional, micro-level (taluka-wise) study that is undertaken in Goa to study imbalances in economic development across regions (talukas). The only other such known sub-regional, intra-state study is the one undertaken by Rao (1984) for the state of Karnataka in "Regional Disparities and Development in India", however, her study is based on a comparative analysis of only land utilisation pattern, cropping pattern and industrialisation pattern pertaining to the years 1974-75 to 1979-80 and therefore, the present study is wider in scope and covers a longer period (1987-2006) enabling a more meaningful comparative analyses over time.

The broad aim of this study is to measure the regional (taluka-wise) inequalities in the economic development of Goa. The study starts with the hypothesis that there exists regional imbalances in the economic development of Goa and this imbalance exists
across agricultural sector, manufacturing sector and the service sector. The regional imbalance also exists in the various sub-sectors of the service sector.

3.2 Data and Variables

The study is based on the secondary data collected from the Directorate of Planning, Statistics and Evaluation, Government of Goa for the years 1987-2006 (20 years). The study takes into account the Agricultural Sector considering 15 parameters namely, total reporting area (Ha), net area sown (Ha), area sown more than once (Ha), gross cropped area (Ha), area under paddy (Ha), area under other cereals, millets/pulses & oil seeds (Ha), area under coconuts (Ha), area under arecanuts (Ha), area under cashewnuts (Ha), area under vegetables (Ha), area under garden crops (Ha) (a) average yield of rice per hectare (in kgs) i) kharif ii) rabi, and average yield of ragi per hectare (in kgs) kharif.

The Manufacturing Sector considers 8 parameters namely, number of factories registered under the Factories Act, 1948, number of Factories in operation registered under the Factories Act, 1948, estimated average number of daily workers employed in registered factories, number of small scale industries registered with the Industries Department, estimated average no. of daily workers employed in the small scale industries, number of large and medium scale industries, employment in large and medium scale industries and the number of sheds in industrial estates.

The Service Sector considers 4 sub-sectors of service activities namely, Education with 13 parameters that is literacy rate, number of schools, and number of students in primary,
middle, secondary and higher secondary level and colleges and university; Public Health with 6 parameters that is, number of specialised and general hospitals, number of community primary health centres, number of beds in government hospitals, number of private hospitals, number of beds in private hospitals, number of dispensaries (r.m.d.), number of urban health centres, and number of sub-health centres; Banking with 5 parameters that is, number of scheduled commercial banking offices, number of cooperative banks, Deposits mobilized, credits advanced; and Tourism with 3 parameters that is number of hotels and lodging houses and Number of domestic and foreign tourist arrivals. Thus, a total of 48 parameters are considered over the three sectors, for the 11 talukas over the 20 year period 1987 to 2006. In all, 10,560 observations are taken into consideration for the analyses. This analysis clearly indicates the regional (taluka-wise) inequalities in the development of each of the sectors and in the over all development of the State.

3.3 Index Number (Simple and Composite)

The study makes use of construction of the simple index for agriculture, manufacturing and each of the sub-sectors within the service sector. A combined index is then constructed for service sector by taking into consideration the sub-sector indices. Finally, a composite development index is constructed for the entire State by considering the sector-wise indices. The talukas are also then ranked according to the index values and categorised into high, medium and low development talukas.
Index number is one of the major statistical tools used. The study has used simple index and composite index to measure the differences in the growth pattern of agriculture, manufacturing and service sector activities, among 11 talukas. Each variable in each category (broad categories agriculture, industry and services and four sub-categories under services) is first converted into the index form by taking the average of that variable as the base across all talukas. Then each index value of the variables, in the index form, for a particular taluka, are combined to get the index for a particular category. For example, take the sub-category Education. The Number of Schools (NS) is a variable under the category education. This variable is nothing but the number of schools of each taluka. Therefore, this variable will have 11 observations. The number of schools in each taluka is converted into indices by taking average number of schools ($\Sigma NS/11$) as the base. These indices of each taluka will reflect the inequality among talukas, in relative terms, with respect to NS. Suppose there are 8 variables like NS in the category of education, then there will 8 such indices for each taluka. These 8 indices are combined to get one index value for each taluka for the category of education. Therefore, there will 11 such indices for 11 talukas in Goa. In the similar way, the indices for other categories of services are also prepared.

The Composite Indices are prepared by combining the indices of broad categories. There are two composite index numbers in this study. One, is the composite index of economic development of Goa, which is prepared by involving the index of agriculture, manufacturing and service sectors. The other composite index is the index for the service sector as there are 4 sub-categories in the service sector.
3.3.1 Categorisation of Talukas

The regions (talukas) are then ranked from 1 to 11 on the basis of the indices and categorised into low developed, medium and high developed regions (talukas) for agriculture, manufacturing and service sectors for each category of the service sector. The regions (talukas) are finally ranked and categorised on the basis of the Composite Development Index into high, medium and low development regions (talukas) to get an overall picture of economic development and the regional imbalances therein.

3.4 Spearman’s Rank Correlation Coefficient

When the data is in the form of ranks or are otherwise on an ordinal scale, we have an alternative measure of association between two variables. Most frequently used non-parametric measure of correlation between two variables is the Spearman Rank Correlation Coefficient denoted by ‘Rs’ (Aczel and Sounderpandian, 2006).

To compute the Spearman Correlation Coefficient, we first rank all the observations of one variable within themselves from smallest to largest. Then we independently rank the values of the second variable from smallest to largest. Spearman’s rank correlation coefficient is the usual correlation coefficient applied to the ranks.

The formula to compute the Spearman Correlation Coefficient is as follows.

\[ Rs = 1 - 6 \frac{\sum d_i^2}{n(n^2-1)} \]

\( d_i, i=1,2,\ldots,n \) are differences in the ranks of \( X_i \) and \( Y_i \), i.e., \( d_i = R(X_i) - R(Y_i) \)
The Spearman Correlation Coefficient satisfies the usual requirements of correlation measures. It is equal to 1 when the variables X and Y are perfectly positively correlated. It is equal to -1, if it is perfectly negatively correlated. It is equal to zero when there is no relation between X and Y. Values between these extremes give a relative indication of the degree of association between X and Y.

3.5 Co-efficient of Variation

The study also uses the coefficient of variation as a measure of dispersion to signify the extent of inequality.

The coefficient of variation is a percentage expression of variance, which has the same features as that of standard deviation in its relative measure. This measure of dispersion is a pure number, expressed as a percentage and is used in this study to compare between spatial and temporal inequalities occurring across the 11 talukas over the 20 year period of study.

It is calculated with the use of the formula:

\[ CV = \frac{\sigma}{\bar{x}} \times 100 \]

where,

- \( CV = \) coefficient of variation
- \( \sigma = \) standard deviation of index values
- \( \bar{x} = \) Average of index values
The coefficient of variation is then tabulated spatially and temporally for each category and observations are made regarding the extent of inequalities and the changes therein.

### 3.6 Gini Coefficient and Lorenz Curve

The Gini-Coefficient was developed by the Italian Statistician Corrado Gini (Gini, 1912) as a summary measure of income inequality in society. It is almost always associated with the plot of wealth concentration introduced by Max Lorenz (Lorenz, 1905). Since these measures were introduced they have being applied to topics other than Income and Wealth but mostly within Economics. (Li, Hongyi and others, 1998; Cowell, 1995, 2000; Jenkins, 1991; Sen, 1973).

The Gini-Coefficient is therefore based on the Lorenz Curve, a cumulative frequency curve that compares the distribution of a specific variable with the uniform distribution the represents equality (as shown in the figure)

![Lorenz and Equality Curves](image)

This equality distribution is represented by a diagonal line and the greater the deviation of the Lorenz Curve from this line, the greater the inequality.
While considering the variables, the cumulative proportion of the population is generally shown on the X axis and the cumulative proportion of the variable on the Y axis. The greater the distance from the diagonal line, the greater the inequality. The curve can be below or above the diagonal depending on the variable used when the variable is beneficial to the population the curve is found below the diagonal line. In contrast, when the variable is prejudicial, it is found above the line.

The Gini-Coefficient is a Lorenz Curve consistent tool for evaluating and comparing relative inequality in the distribution of populations access to each of the variables where:

\[ \text{Gini Coefficient} = \frac{\text{Area between Lorenz Curve and Diagonal}}{\text{Total area under the Diagonal}} \]

Thus, the Gini-Coefficient ranges from 0 to 1. 0 representing perfect equality and 1, total inequality. (Stuart and Ord, 1994). The Gini Coefficient has also been used widely in inequality studies conducted by the World Bank for the World Development Report and the United Nations for the UN Human Development Report.

In recent years, the best data on the extent of income inequality, expressed in terms of Gini-ratios, on a per-person income basis covering 49 countries, are available in recent issues of World Development Report. Lowest Gini values of less than 0.3, are found in the regions of the former communist bloc and the Welfare States of Western Europe. Ginis of a little over 0.4 are found for U.S. and China. Latin America, as always exhibits
very high Ginis and Malaysia, the only east asian economy in the data set is found in the high Gini (0.45-0.5) category (World Development Indicators, 1999). Categorization made by the World Bank with regard to inequalities as reflected by Gini Coefficients, considers Gini values of 0 to 0.3 as insignificant levels of inequality, Gini values for 0.3 to 0.5 as of medium significance and Gini values above 0.5 upto 1 as highly significant. However, this categorization made by the World Bank is in context of the world economy and therefore, very severe. The same measure need not be very relevant in a regional context, that too for a small developed state like Goa. Thus, in the present study, we redefine the categorization as Gini values of 0 to 0.2 as indicating insignificant inequality, Gini values of 0.2 to 0.4 as indicating inequalities of medium significance and Gini values above 0.4 as being indicative of highly significant inequalities.

Extensive academic literature thus, establishes the Gini-Coefficient as the most commonly used Lorenz consistent inequality measure. There are different methods to calculate the Gini-coefficient, this study uses the formula as developed by M. Brown (Brown, 1994), who presented a Gini-Style Index, seemingly calculated from two variables instead of one. The formula is as shown below E may,

\[
G = \left| 1 - \sum_{i=0}^{k-1} (Y_{i+1} + Y_i) (X_{i+1} - X_i) \right|
\]

The taluka-wise secondary data collected for 48 parameters across the 3 sectors for the period 1987 (Statehood of Goa) to 2006 (20 years) enables this study to calculate the Gini-Coefficient spatially and temporally and the steps followed are as given below:
1. Sort the two variables under consideration by the population (Xi), from the lowest to the highest.

2. Obtain the cumulative percentage of Xi variable, label it as Cum X%

3. Obtain the cumulative percentage of Yi variable (income, number schools or any other variable) and label it as Cum Y%

4. Calculate the sum of Cum X% and Cum Y% (Cum X% + Cum Y%), and call it as A

5. Calculate A* Cum X% for each observation and obtain the sum of the same.

6. Subtract 1 from the value obtained in the step (4) and that represents the Gini Coefficient.

7. Calculate the cumulative proportion of each of the two variables.

8. Calculate the Gini-coefficient using the formula.

The Gini-coefficient derived for each parameter is then added and the average Gini is calculated for each category and tabulated spatially and temporally to assess what is happening in the particular sectors / sub-sectors over the 20 year period of study.