CHAPTER -3

RESEARCH METHODOLOGY

CHAPTER - 3
RESEARCH METHODOLOGY

Research methodology is an important part of the research and it is used to know what the methods are used for the analysis the study. The following research methodology is adopted to fulfil the objectives of the study.

3.1) Collection of Data

This study is based on secondary data, which were collected from the published reports and Department of Economics and Statistical Analysis, Government of Haryana. The study has been conducted district -wise and zone -wise analysis for the period 2001-02 to 2014-15, pertaining to areas, production and productivity of selected major food grain and cash crops.

3.2) Research Methods

For measure growth rate different methods are used such as exponential growth rate, kinked exponential growth, CAGR (Compound Annual Growth Rate). This study used CAGR by zone -wise and district-wise disparities of area, production and productivity of major selected crops in Haryana. Haryana divided four zones, northern, central, and western zone. Northern zone included Ambala, Panchkula, Yamunanagr, Kurukshetra, Karnal and Panipat. This zone is endowed with more assured irrigation and fertile soil. Central zone included 5 districts, namely Kaithal, Sonipat, Rohtak, Jhajjar, and Jind. This zone is more or less like zone as far as range of temperature during summer as well as during winter is concerned, but there is a slight difference in soil texture and annual average rainfall. Western zone included 5 districts namely Bhiwani, Fathebad, Sirsa and Hisar. This zone is arid and dry. Southern zone included 6 districts namely Faridabad, Mahendergarh, Rewari, Gurgaon, Mewat and Palwal. Although the general agro-climatic outlook except rainfall in zone -IV are more or less similar to that of zone-III.

σ-convergence model has been used for significance of standard deviations of area, production and productivity over time and β convergence model has been used to find the significance of these growth rates by regression on their initial values. These models deal the total production as a sum of major selected crops. Productivity is calculated by the formula (production/ area)*100 in this study.

3.3) CAGR (Compound Annual Growth Rate)

To fulfil the objectives the Compound Average Growth Rate for areas, production and productivity during the period 2001-02 to 2014-15 (CAGR) has been used which states change in variable year over year.

Compound Growth Rate: $\mathbf{Yt=Yo(l} + r)_t$

Here Yt and Yo are the areas, production and productivity for terminal and initial period of a particular state. r is the compound growth rate.

3.4) Co-efficient of Variations

Co-efficient of Variations is the ratio of the standard deviation to the mean. The higher the coefficient of variation, the greater level of dispersion around the mean.

C.V = Standard Deviation/Mean

Standard Deviation $= [\Sigma (xi-x)^2/N]^{1/2}$

Where, N=Total Number of Observation

Xi=Area/Production or Productivity

X=Mean of the Distribution

3.5) Convergence Model

On the basis of neoclassical growth models, countries with lower GDP per capita will indicate to grow faster than richer ones. However, convergence is not always confirmed. This

means that economies are converging but the steady-state level is not always common, so countries may converge to own level of steady-states. At the same time, the word 'convergence' can be explained by different ways.

According to Barrow Sala-i-Martin there are three concepts of convergence, σ-convergence, β-convergence and conditional convergence. But this study used two convergences; σ- convergence and β-convergence. Absolute convergence hypothesis following considers: group of all members of countries which have access to the same technology the same population growth and the same propensity to save and on different in term of their initial capita labour ratio. Than we should except all countries to converge to the same stead capital- labour ratio, output per capita and consumption per capita and consumption per capita and of course, the same growth.

3.5.1) σ-Convergence

While β--convergence emphasis on determine possible catching-up processes, σ-convergence simply refers to a diminishing of disparities among regions in time. The two concepts are of course closely related. Formally, β-convergence is necessary but not adequate for σ - convergence. Instinctually, this is either because economies can converge towards one another but random shocks push them apart or because, in the case of conditional β-convergence, economies can converge towards different steady-states.

This study deals with σ-convergence of dispersion of area, production and productivity across the districts over time. This model is based on calculating the standard deviation (S.D) across four zone and nineteen districts of Haryana for each year starting from 2001-02 to 2014-15. As the σ- convergence measures the inter-regional inequality, it may very well infer that the inter-regional inequality among the districts have increased during given time period. When the dispersion of area, production and productivity across districts falls over time, there is σ - convergence. The model is

$$SD_t = \eta + \tau t + \omega_t$$

There, SD = Standard Deviation (Area, Production and Productivity)

η = Intercept, τ = Slope coefficient, ω = Error term, t= time (2001-02to 2014-2015)

3.5.2) β- convergence

β -convergence concern to a process in which poor regions grow faster than rich ones and therefore catch up on them. The concept β -convergence is directly related to neo-classical growth theory (Solow, 1956) where one key assumption is that factors of production, in particular capital, are subject to diminishing return. Accordingly, the growth process should conduct economies to a long-run steady-state characterised by a rate of growth which depends only on the (exogenous) rates of technological progress and labour force growth. Diminishing return also implies that the growth rate of poor economies should be higher and their income and per head levels should catch up with those of rich economies. When all economies are assumed to converge towards the same steady-state (in terms of GDP per head and growth rate), β-convergence is also called absolute convergence. However, the steady-state may depend on features specific to each economy, in which case convergence will still take place, but not necessarily at the same long-run levels. This will be the case when GDP per head is supposed to depend on a series of determinants such as factor endowment or institutions, which can vary from one economy to the other even in the long-run. Beta-convergence is then said to be conditional.

β -convergence is empirically examined by estimating regression of the annual growth rate of production and productivity on the initial level of production and productivity. Hence, the test for the absolute β-convergence hypothesis is performed by estimating the following equation by the ordinary least squares (OLS) method.

$$G.W_i = \alpha + \beta X_{0+} \mu_i$$

Here, G.W. = Growth rate (2001 to 2015), Production (000 tonnes) and Productivity (kg/ha)

α = Intercept, β = Slope Coefficient, μ_i = Disturbance term,

X_0= Initial area, production and productivity, i = Districts

Partial correlation between growth of area, production and productivity over time and its initial level is negative, there is β -convergence.

3.6) Regression Analysis

To identify the inputs contributing significantly in agricultural production based on selected crops, for all the districts under study, over the period 2001-02 to 2014-15 at first stage simple regression analysis was carried out for each district by taking total production as dependent variable and inputs as independent variable. Using the following model:

3.6.1) Multiple Regression Model

It shows relationship the between a dependent variable and two or more independent variables. To assess the influence of agriculture input on agriculture production, the multiple linear regression analysis by Ordinary Least Square (OLS) estimation has been applied by Eviews Software. The functional form of multiple linear regression models is given below:

$$Y_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \mu_i$$

Where,

Y_i = Dependent Variable- Agriculture production of selected food grain crops

α = is the constant or intercept

(X_1..... X_5) are Independent Variables

X_1 = Irrigation

The irrigation is the most important factor for agriculture in the Haryana district, for increase productivity .Irrigation facility can help in solving the problems created insufficient, uncertain and irregular rains

X_2 = Fertilizer

The proper use of chemical fertilizers can considerably increase the productivity of soil. Increase agricultural production is related to increased consumption of fertilizers.

X_3 = Pesticides

Pesticide is a chemical preparation for destroying plant, fungal, or animal pests.

X_4 = Rainfall

The food grain production of India has a proportional relationship to the monsoon, with its critical dependence on the onset, duration and distribution of rainfall**.** In this study we measured rainfall in centimetre.

X_5 = Tractors

Agricultural machines increase productivity of land and labour by meeting timeliness of farm operations and increase work out-put per unit time

($\beta_{1.....}$ $\beta_{5)}$ = the slope (beta coefficient) for ($\mathbf{X}_1$.....$\mathbf{X}_5$)

μ_i = Random Error or Stochastic Disturbance Term

After regression this study checks multicollinearity.

3.7) Multicollinearity

The term Multicollinearity is introduced by 'Ranger Frish'. Multicollinearity is a high degree of correlation among all explanatory variables.

Problem Multicollinearity: If independent variables contain essentially same information to a large extent, one gains little by using both the regression model. Multicollinearity leads to unstable estimates as it tends to increase the variances of regression coefficients.

Symptoms of multicollinearity may observe in situations:

1. R- Square for regression is very high

2. Coefficient may have very high standard error

High degree of correlation coefficient do not necessary implies Multicollinearity. Than can make a judgement by checking Tolerance and VIF (Variance Inflation Factor). The ith tolerance value is defined as $(1-R^2_K)$, R^2_K is the coefficient of determination for the ith independent variable on all the other independent variables. VIF is the reciprocal of a tolerance value. VIF is a measure of the effect of multicollinearity on the variance parameter estimates. High VIF suggest the presence of multicollinearity.

- VIF > 10 indicate presence of multicollinearity

If VIF value is lowest 10 then there is not multicollinearity