Journal of Advanced Zoology

ISSN: 0253-7214

Volume 44 Issue 02 Year 2023 Page 280:300

Analysis of State Wise Yield of Major Food Crops in India for the Year 2014-15

Bhaskarjyoti Talukdar[1,*], Pranjal Kumar Chakravartty[2], Bandana Sharma[3]

[1]Ph.D Research Scholar, Department of Statistics, Cotton University, Guwahati, Pincode-781001, District-Kamrup Metro, Assam, India. ORCID ID: 0009-0004-2768-7479, E-mail: bhaskarphd13@gmail.com
[2]Data Analyst, Assam Rural Infrastructure and Agricultural Services (ARIAS) Society, Guwahati, Pincode-781022, District-Kamrup Metro, Assam, India. E-mail ID: khukustar@gmail.com
[3]Associate Professor, Department of Statistics, Cotton University, Guwahati, Pincode-781001, District-Kamrup Metro, Assam, India. E-mail ID: bandana0666@gmail.com
***Corresponding Author's Email**: bhaskarphd13@gmail.com*

Article History

Received: 01 June 2023
Revised: 07 Aug 2023
Accepted: 27 Aug 2023

Abstract

Agriculture plays a pivotal role in India's economy and food security. Understanding the state-wise yield variations of major food crops is essential for informed policymaking and sustainable agricultural practices. This study conducts a comprehensive analysis of state-wise crop yield data for major food crops, including rice, wheat, pulses, and sugarcane. Our findings reveal significant disparities in crop yields across different states in India. Factors such as climate variability, soil quality, irrigation infrastructure, and government interventions have profound effects on crop production. While states like Punjab and Haryana consistently exhibit high yields of wheat and rice, others face challenges in achieving similar productivity. The state-wise variability in crop yields underscores the need for region-specific agricultural strategies. In this study, covariance matrix of nine major food crops have been generated along with graphical representation of state wise food grains. Scatter plot represented nine variables considering major crops in India. Variation has been examined using PCA technique and justification has been done using Scree Plot. Biplots are plotted for all the five components of PCA technique. Kruskal-Wallis Rank Sum Test is conducted for Staple and Non-staple food crops and found heterogeneity among six zones in India.

Keywords: *Biplots, Covariance matrix, Kruskal-Wallis test, PCA, Scatter plot, Sree Plot*

1. Introduction

Agriculture is a major part of our lifestyle today. Without it, we probably would not have many of the features available to us today. Agriculture is the basis of all civilization, the basis of writing, of cities and towns, of our governing systems, of our technology, of almost everything we have. The term "AGRICULTURE" is late Middle English adaptation of Latin "AGRICULTURA", from 'ager', "field" and from 'cultura', "cultivation" or "growing", which means cultivation of land. i.e., In relation to crop farming and livestock farming, the term "agriculture" may be defined as: the art and science of growing plants and other crops and the raising of animals for food, other human needs, or economic gain. There are many online queries on what is agriculture although its history started more than 10,000 years ago. Agriculture has a large coverage as science, business, and for other purposes including legal matters, and with new technologies and specialized fields. No definition can be exacting for everybody and for all purpose, although there are some definitions where its coverage is limited to crop production and livestock production.

(a) Agriculture is the growing of both plants and animals for human needs [1]; (b) Agriculture is the deliberate effort to modify a portion of Earth's surface through the cultivation of crops and the raising of livestock for sustenance or economic gain [2]; (c) Agriculture includes farming in all branches and, among other things, includes the cultivation and tillage of soil, dairying, the production, cultivation, growing and harvesting of any agricultural and horticultural commodities, the raising of livestock or poultry, and any practices performed by a farmer on a farm as an incident to or in conjunction with such farming operations, but does not include the manufacturing or processing of sugar, coconuts, abaca, tobacco, pineapple or other farm products [3]; (d) Agriculture is the science of cultivating the soil, harvesting crops, and raising livestock and also as the science or art of the production of plants and animals useful to man and in varying degrees the preparation of such products for man's use and their disposal [4].

Due to its large population Indian agricultural is largely dominated by the predominance of the food crops which occupy 60.0% (in 2021) of the total cropped area of the country [5]. These crops are bound to have priority in years to come despite the emerging trends of commercialization and diversification in the country's agriculture.

These crops are raised all over the country and throughout the year. In Assam, West Bengal and coastal regions of the Peninsula as many as three successive crops of rice are grown in a year. In irrigated areas of Punjab and Haryana the field is hardly vacant and wheat, gram, rice, bajra, and maize are the main crops [12].

Food crops have attracted highest attention of our planners, policy makers, administrators and agricultural scientists. It is in these crops, particularly in wheat, maize, bajra, jowar, and rice high yielding varieties have been developed and Green Revolution has set in. Whenever the production of these crops is adversely affected due to the vagaries of the weather famine and drought conditions develop which often leads to the political turmoil and instability.

Food crops include cereals and pulses amongst which rice, wheat, jowar, bajra, maize, barley, ragi (as cereals), gram and tur (as pulses) are important.

In India there are three major crop seasons (a) Kharif; (b) Rabi, and (c) Zaid

Kharif crops: (a) The Kharif crops are associated with the monsoon season. They are sown in the months of June and July and are harvested in autumn months i.e. in September and October; (b) Important Kharif crops are rice, jowar, maize, cotton, ragi, bajra, sugarcane and jute.

Rabi crops: The Rabi crops are sown in the period between October and December and harvested in April and May.

Important Rabi crops are wheat, peas, pulses, mustard and rapeseed.

Zaid crops; (a) Zaid crops are sown in the summer season; (b) Important Zaid crops are rice, maize, sunflower, vegetables and groundnut.

The major agricultural products can be broadly grouped into foods, fibers, fuels, and raw materials. Specific foods include cereals (grains), vegetables, fruits, oils, meats and spices. Fibers include cotton, wool, hemp, silk and flax. Raw materials include lumber and bamboo. Other useful materials are also produced by plants, such as resins, dyes, drugs, perfumes, biofuels and ornamental products such as cut flowers and nursery plants.

India is an agriculture-based country. Agriculture is the single most important contribution of Indian economy. About 70% of our population demands on agriculture [6]. In 2022-23 financial year, 18.30% of our National Income comes from agriculture [7]. More than half (59% in 2016) of country’s total workforce are working in agriculture and it’s supporting sectors like fisheries, forestry, etc. [6]. India’s agriculture includes rice, wheat, pulses, tea, coffee, fresh vegetables, fresh fruits, dry fruits, coconuts, major spices, milk, millets, cotton, jute, castor oil seed etc.

Agriculture has been the backbone of the Indian economy and it will continue to remain so for a long time. It has to support almost 17% of world population from 2.3% of world landmass and 4.2% of world’s water resources. The economic reforms, initiated in the country during the early 1990s, have put the economy on a higher growth trajectory. Annual growth rate in GDP has accelerated from below 6% during the initial years of reforms to more than 8% in recent years [8]. This happened mainly due to rapid growth in non-agriculture sector.

Agriculture is the main source of livelihood for over 80% of the small and marginal farmers in India [6]. Although, it employs about 52% of the labour force, it contributes to only 14.4% of GDP and 10.23% of all exports [9]. 80% of rural women, who constitute 41.8% of the agricultural work force, are amongst the least paid workers [10].

The agriculture of India remained under developed for a long time. It did not produce enough food. India had to buy food-grains from other countries, but the things are changing now. India is producing more food-grains than its needs. Some food-grains are being sent to other countries. Great improvements have been made in agriculture through the five-year plans. Green Revolution has been brought about in the agricultural field. Now our country is partially self-sufficient in food-grains. It is now in a position to export surplus food-grains and some other agricultural products to other countries.

Now India ranks first in the world in the production of tea and groundnuts. India ranks among the top five producers of many agricultural items like coffee, cotton, etc. India ranks 2 largest producers of wheat and rice in the world. Items like milk, many fresh fruits, spices, jute, millets, etc. where India is the largest producer. India is one the largest producer of Sugarcane in the world [11]. India is also a major exporter of rice and wheat to countries like Nepal, Bangladesh, Africa and others.

Although India has attained self-sufficiency in food staples, the productivity of its farms is below that of Brazil, the United States, France and other nations. Indian wheat farms, for example, produce about a third of the wheat per hectare per year compared to farms in France. Rice productivity in India was less than half that of China. Other staples productivity in India is similarly low. Indian total factor productivity growth remains below 2% per annum; in contrast, China's total factor productivity growths is about 6% per annum, even though China also has smallholding farmers. Several studies suggest India could eradicate its hunger and malnutrition and be a major source of food for the world by achieving productivity comparable with other countries.

Roy, A. et. al. (2015) analyzed Production and productivity of foodgrains for the period 1972-1973 to 2011-2012, which was divided into three decades, viz. 1982-1983 to 1991-1992, 1992-1993 to 2001-2002, 2002-2003 to 2011-12 [13]. Sharma, P. et. al. (2017) studied state-wise food security index (FSI) utilizing components related to food availability, food stability and food accessibility, and they explained the influence of climate change on FSI which is used as proxy for food security employing linear, non-linear and log-linear regression models and complied state-wise panel (during 1985-2009) of selected 13 Indian states [14]. The performance of agriculture sector including the dynamics of gross and net cropped area, cropping intensity, yield of major crops, cropping pattern, use of consumable and durable agricultural inputs, and institutional credit at state and all India level since the economic reforms and accelerate growth for the period of 2003-2004 to 2014-2015 have been studied by Khan, W. et. al. (2020) [15]. Madhukar, A. et. al. (2020) examining three major food crops — wheat, rice and maize across 29 Indian states and the four types of regression models were fitted on the annual yield data for the period 1967–2017 and Akaike information criterion (AIC) is used to best fit statistical models [16]. Kumar, A. et. al. (2015) estimates effects of various climatic, non-climatic and geographical factors on agricultural productivity in India for the period of 1980-2009 and they applied Cobb-Douglas production function to estimate the outputs; Prais Winsten models with panels corrected standard errors (PCSEs) estimations, indicated average maximum temperature appears positive on agriculture productivity of Rabi crops and negative impact on Kharif crops [17]. Samal, P. et. al. (2018) evaluates the growth experience of production and profitability in rice using secondary data for the period 1970-1971 to 2015-2016 and they observe that area growth in most of the states have either exhausted or declined rapidly and also, they analyzed the trends in costs and profits for the period 1980-81 to 2014-15 and it was revealed that cost of cultivation increased over years, but profit has not increased commensurately [18]. Kumar, A. (2013) et. al. focused on the impact of climate change on agricultural productivity in quantity terms, value of production in monetary terms and food security in India based on secondary data of 1980 to 2009 and they have used regression analysis to examine different climatic variables as well as used econometric model estimation for climatic, food security index and socio-economic variables [19]. Singh, A. K. (2021) examined the impact of climatic change on yield production and they used state wise data for the period of 1977-2014 in 15 Indian states and forecasted food-grain production for the years of 2040, 2060, 2080 and 2100. Regression coefficients of food-grain production and yield with climatic and non-climatic factors are estimated using Cobb-Douglas production function and marginal impact analysis method is used to examine the projected food-grain production [20]. Mathur A. S. et. al. (2006) studied growth of agriculture production trend in India for the period of 1993-1994 to 2003-2004 and analyzed various constraints which affects growth sector and they projected future growth in line with 11th five-year plan and highlighted role of public investment/govt. expenditure towards agricultural production in India [21]. Manogna, R. L. (2020) emphasized on analysis of the agricultural production of food grains in Indian states and also, they highlighted increasing agricultural production with efficient resource deployment and enhance production management and takes different variables effecting yield production for the period of 2006 to 2017 using Data Envelopment Analysis (DEA) and Malmquist Productivity Index (MPI) [22].

An analysis of the state-wise production of the major food crops so as to assess the difference in the production pattern with the geographical location would form an interesting research prospect. It will give us an idea about the how the topography of a place affects the production, followed by attempting to identify the underlying factor (s) responsible for the difference. This forms the broader objective of this dissertation.

The main objective of this study

(a) To assess the correlation structure existing among the state-wise production of 9 major food crops viz. Rice, Maize, Wheat, Bajra, Gram, Groundnut, Soya bean, Rapeseed & Mustard and Linseed for the year 2014-15; (b) To apply some dimensionality reduction technique with the aim of identifying the underlying sources of variation in the data; (c) To assess the homogeneity of the yield of staple and non-staple food crops across the different zones of India.

2. Materials and Methods

2.1 Materials

To fulfill the objectives of the study, the data considered is secondary in nature. It has been collected from the report entitled "Agricultural Statistics at A Glance 2015", published by Government of India; Ministry of Agriculture and Farmers Welfare; Department of Agriculture, Cooperation and Farmers Welfare; Directorate of Economics and Statistics.

The data on only the major food crops classified in the above-mentioned report has been considered viz. Rice, Maize, Wheat, Bajra, Gram, Groundnut, Soya bean, Rapeseed & Mustard and Linseed, for the year 2014-15, which is also the latest year for which the data is available.

2.2 Methods

The following section contains the methods used for attaining the objectives of the study:

2.2.1 Principal Component Analysis

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

The principal components can be used to visualize graphically the data. Also, it helps to identify the more important variables that explain the variations in the data.

PCA can be performed either on the correlation matrix or the covariance matrix. The covariance matrix is used when the variable scales are similar whereas the correlation matrixis used when the variables are measured on a different scale.

In linear algebra, an eigenvector or characteristic vector of a square matrix is a vector that does not change its direction under the associated linear transformation. In other words—if vis a vector that is not zero, then it is an eigenvector of a square matrix A if Av is a scalar multiple of v.

The general objectives of principal component analysis are- (a) Volume reduction, (b) Interpretation

Procedure:

Let the random vector $X = [x_1, x_2, \ldots, x_p]$ have the covariance matrix $\sum$ with the Eigen Values $\lambda_1 \geq \lambda_2 \geq \ldots \ldots \geq \lambda_p \geq 0$. Consider the linear combination,
$Y_1 = a_1 X = a_{11}X_1 + a_{12}X_2 + \ldots \ldots \ldots \ldots + a_{1P}X_P, Y_2 = a_2 X = a_{21}X_1 + a_{22}X_2 + \ldots \ldots \ldots \ldots + a_{2P}X_P :: Y$

Then,

$Var[Y_i] = a_i^{/}\sum a_i, i = 1,2, \ldots, p$ ------------------------(i)

$Cov[Y_i, Y_k] = a_i^{/}\sum a_k, k = 1,2, \ldots, p$---------------------(ii)

The principal components are those uncorrelated linear combinations, $Y_1, Y_2, \ldots \ldots, Y_P$ whose variances in (i) are as large as possible.

The 1st principal component is the linear combination with maximum variance i.e. it maximizes $Var[Y_1] = a_1^{/}\sum a_1$

We can define that,

1st principal component = linear combination $\boldsymbol{a_1^{/}X}$ that maximizes $V[\boldsymbol{a_1^{/}X}]$ subject to $\boldsymbol{a_1^{/}a_1 = 1}$

2nd principal component = linear combination $\boldsymbol{a_2^{/}X}$ that maximizes $\boldsymbol{V}[\boldsymbol{a_2^{/}X}]$ subject to $\boldsymbol{a_2^{/}a_2 = 1}$ **and** $\boldsymbol{Cov(a_1^{/}X, a_2^{/}X) = 0}$ and so on.

Suppose we have the covariance matrix $\sum$ with eigen value eigen vector pairs $(\boldsymbol{\lambda_1, e_1}), (\boldsymbol{\lambda_2, e_2}), \ldots \ldots, (\boldsymbol{\lambda_p, e_p})$, where, $\boldsymbol{\lambda_1 \geq \lambda_2 \geq \ldots \ldots \geq \lambda_p \geq 0}$

Let, $Y_1 = e_1^{/}X_1, Y_2 = e_2^{/}X_2, \ldots \ldots, Y_p = e_p^{/}X_p$ be the principal components then,

$$\boldsymbol{\sigma_{11} + \sigma_{22} + \ldots \ldots + \sigma_{pp} = \sum_{i=1}^{p} Var[X_i] = \lambda_1 + \lambda_2 + \ldots \ldots + \lambda_p}$$

The proportion of total variance due to the k^{th} principal component is:

$$\boldsymbol{\frac{\lambda_k}{\lambda_1 + \lambda_2 + \ldots \ldots + \lambda_p}, k = 1, 2, \ldots \ldots, p}$$

If most of the total population variance (say 85-90%) for large *p* can be attributed to the 1st 1,2 or 3 components, then these components can replace the original *p*-variables without much loss of information.

2.2.2 Kruskal-Wallis H Test

The Kruskal–Wallis test by ranks, Kruskal–Wallis H test (named after William Kruskal and W. Allen Wallis), or One-way ANOVA on ranks is a non-parametric method for testing whether samples originate from the same distribution. It is used for comparing two or more independent samples of equal or different sample sizes. The parametric equivalent of the Kruskal-Wallis test is the one-way analysis of variance (ANOVA). Since it is a non-parametric method, the Kruskal–Wallis test does not assume a normal distribution of the residuals, unlike the analogous one-way analysis of variance.

In case of several independent samples, Kruskal Wallis is used to test the null hypothesis that k independent samples are from the same population.

Computation of Kruskal Wallis H test:

(a) The scores are arranged into groups in ascending order of score value; (b) All the observations are ranked in a single series without regard to their groups. The smallest score value is replaced by rank 1. Tied scores receives the average of ranks available to those scores. The ranks are totaled within each group. The data would be arranged in the following manner;

Groups

$$[X_{11} \dots X_{1j} \dots X_{1k}, X_{21} \dots X_{2j} \dots X_{2k} \vdots X_{i1} \dots \vdots X_{n1} \vdots X_{ij} \dots \vdots X_{nj} \vdots X_{ik} \vdots X_{nk}]$$

Where,

X_{ij} is the data corresponding to the i^{th} observation of the j^{th} group, n is the number of observation in the j^{th} group ($j=1, 2, \dots, k$).

(c)The attributes of H approximate the distribution of Chi Square with k-1 degrees of freedom. The formula for obtaining the Kruskal Wallis test statistic can be written asfollows:

$$H = [\frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j}] - 3N(N+1)$$

Where,
N=Total number of observations
K=Number of groups
R_j= Sum of rank of the observations in the j^{th} group
n_j= Number of observation in the j^{th} group

3. Data Analysis

3.1 Graphical and Tabular Representation of Data:

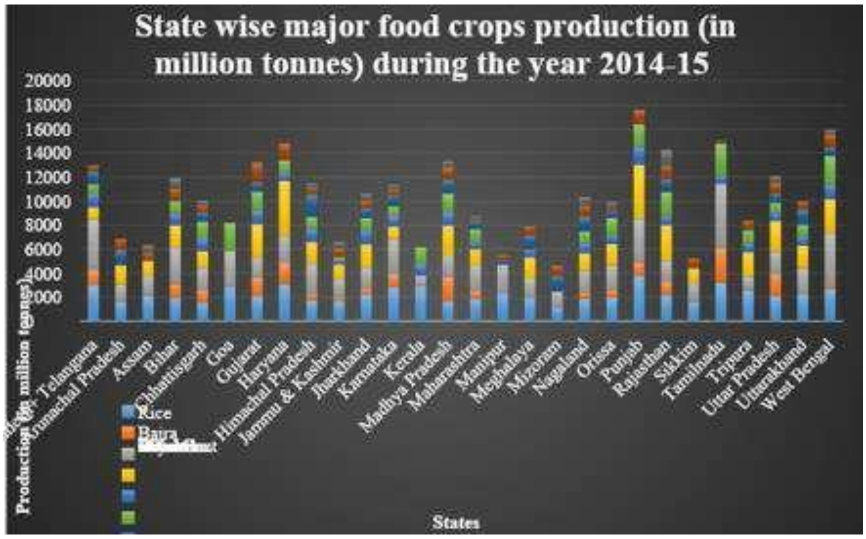


Figure 1

The above fig. 1 displays the state wise production of major food crops (in million tons) during the year 2014-15. From this graph we may have the following conclusion about the highest and the lowest crop producing states as below:

Table 1: highest and the lowest crop producing states as below

Crops	Productionlevel	State	Total Production
Rice	Highest	Punjab	3838
	Lowest	Mizoram	1181
Bajra	Highest	Tamil Nadu	2894
	Lowest	West Bangle	286
Maize	Highest	Tamilnadan	5360
	Lowest	Kerala	1015
Wheat	Highest	Haryana	4574
	Lowest	Andhra Pradesh	1000
Gram	Highest	Punjab	1500
	Lowest	J&K	371
Groundnut	Highest	Tamil Nadu	2697
	Lowest	J&K	274
R&M	Highest	Gujrat	1663
	Lowest	Maharashtra	200
Soya bean	Highest	Himachal	1883
	Lowest	Manipur	331
Linseed	Highest	Rajasthan	1285
	Lowest	Maharashtra	211

Overview: From Table.1 we have the same result as the graph in fig.1 shown before. By observing the graph and the table, we can say that Punjab is the highest crop producer state. The main food crops such as Rice, Wheat, Maize etc. are produced in Punjab the most. On the other hand, we have seen that Mizoram is the state where production of crops is very low.

4. Statistical Analysis of Data

(a) Covariance Matrix

Table 2: Covariance matrix of the nine-variable corresponding to the major food crops under consideration

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
X_1	364818.6	1326250.9	3930680	1829470.1	95108.54	242686.9	-152808	-11462.6	61678.36
X_2	1326251	5949710.2	3748791	2991480.5	147481	299319	-21119.7	84383.96	43859.43
X_3	393068.2	374879.6	1095281	6008.048	143521.3	440189.9	-79650.7	-47150.6	-42530.3
X_4	182947.1	299148.5	6008.048	1416088	275024.4	201111.7	16813.33	366279	83544.6
X_5	95108.54	147481	143521.3	275024.4	186059.4	174860.9	44093.05	49360.74	27040.57
X_6	242686.9	299319	440189.9	201111.7	174860.9	597640.8	-70718.2	-7250.74	31953.29
X_7	-152808	-21119.7	-79650.7	16813.33	44093.05	-70718.2	308463.2	27314.38	25039.87
X_8	-11462.6	84383.96	-47150.6	366279	49360.74	-7250.74	27314.38	158621.1	26291.64
X_9	-61678.4	43859.43	-42530.3	83544.6	27040.57	31953.29	25039.87	26291.64	118868.7

Footnote: Meaning of the symbols used in the table:

X_1 = Rice, X_2 = Bajra, X_3 = Maize, X_4 = Wheat, X_5 = Gram, X_6 = Groundnut, X_7 = Soya bean, X_8 = R&M, X_9 = Linseed

(b) Scatter Plot:

Following is the scatter plot matrix of the nine variables corresponding to the major food crops under consideration.

It can be seen from the scatterplot matrix that the variables corresponding to Soyabean, Linseed and Mustard (R & M) and Linseed are not quite significantly correlated with the other variables (with the only exceptions of the pair of Rice-Soyabean and Wheat-R & M).

(c) Principal Component Analysis

Through principal component analysis, by using the princomp () function in R, we have got the following results (Table 3)

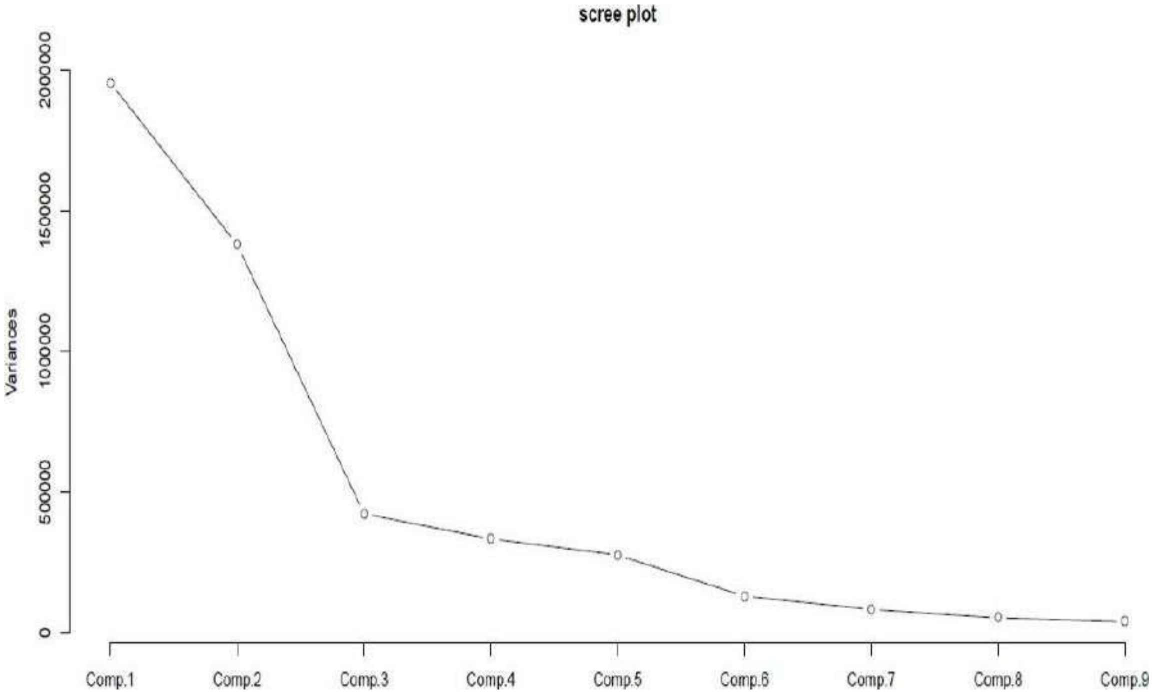
Importance Of Components	PC1	PC2	PC3	PC4	PC5
Proportion Of Variance	0.4188	0.2951	0.0905	0.07143	0.05916
Cumulative Proportion	0.4188	0.7139	0.8044	0.87581	0.93497

Here, PC stands for the word 'principal component'.

We observe that the 1st 5 principal components are able to explain more than 90% of the variation which exist in the data. Hence, we retain the 1st 5 principal components.

(d) Scree Plot:

Following figure shows the scree plot corresponding to the principal component analysis performed:



The above scree plot of the data justifies the results of principal component. The data are plotted as – the variances in y-axis and the PCs are in x-axis.

Table 3: The first 5 principal components which replace the original variables with little loss of information are shown below

	PC1	PC2	PC3	PC4	PC5
X_1	-0.28057412	0.16642878	-0.49617227	-0.06827129	0.020481
X_2	-0.3836957	0.07422663	0.60017038	-0.2138676	0.610281
X_3	-0.48861512	0.58508358	-0.0632901	0.558659	0.00391
X_4	-0.57015613	-0.71380962	-0.19210026	0.14589293	-0.02976
X_5	-0.20743503	-0.05373799	0.11987364	-0.01451981	-0.2727
X_6	-0.38669193	0.20886251	0.11106854	-0.61892419	-0.53844
X_7	0.05508183	-0.09897403	0.53112024	0.45566629	-0.49306
X_8	-0.11855795	-0.23106759	0.04982461	0.12542949	0.120917
X_9	-0.02414724	-0.07242579	0.20352721	-0.09784041	-0.06505

'+' sign indicates that the association of the variable with the principal component is positive whereas '-' sign indicates that the association of the variable with the principal component is negative.

Table 4: The proportion of contribution of each variable for the selected principal component are shown below

	PC1	**PC2**	**PC3**	**PC4**	**PC5**
X_1	0.111562281	0.07549103	0.20960769	0.029695163	0.009505748
X_2	0.152565628	0.03366873	0.25354163	0.093023485	0.283247936
X_3	0.194283838	0.26539018	0.02673687	0.242993365	0.001814692
X_4	0.22670629	0.32377949	0.08115265	0.063457342	0.013810193
X_5	0.08248061	0.02437521	0.05064055	0.006315512	0.126566185
X_6	0.153756995	0.0947387	0.04692084	0.269206208	0.249904559
X_7	0.021901716	0.04489399	0.22437144	0.19819583	0.228840849

X_8	0.047141182	0.10481079	0.02104838	0.054556598	0.056120646
X_9	0.009601461	0.03285188	0.08597995	0.042556496	0.030189194

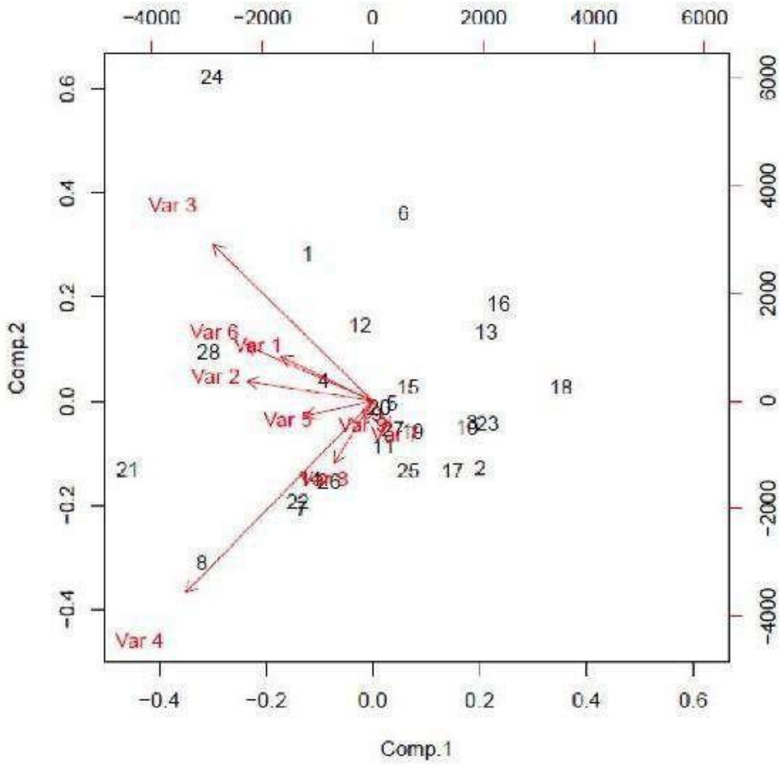
We see that variables X_3 and X_4 contribute most to the PC1, both negatively; variables X_3 and X_4 contribute most to the PC2, positively and negatively respectively; variables X_2 and X_7 contribute most to the PC3, both positively; variables X_3 and X_6 contribute most to the PC4, positively and negatively respectively and variables X_2 and X_6 contribute most to the PC5, positively and negatively respectively.

5. <u>Biplots and their Interpretation</u>

Biplots are a type of exploratory graph used in statistical analysis, which allows the information on both the variables and samples under consideration to be explained graphically. Therein, the observations (samples) are displayed as points whereas the variables are displayed as vectors.

(a) <u>Biplots for Examining Component 1</u>

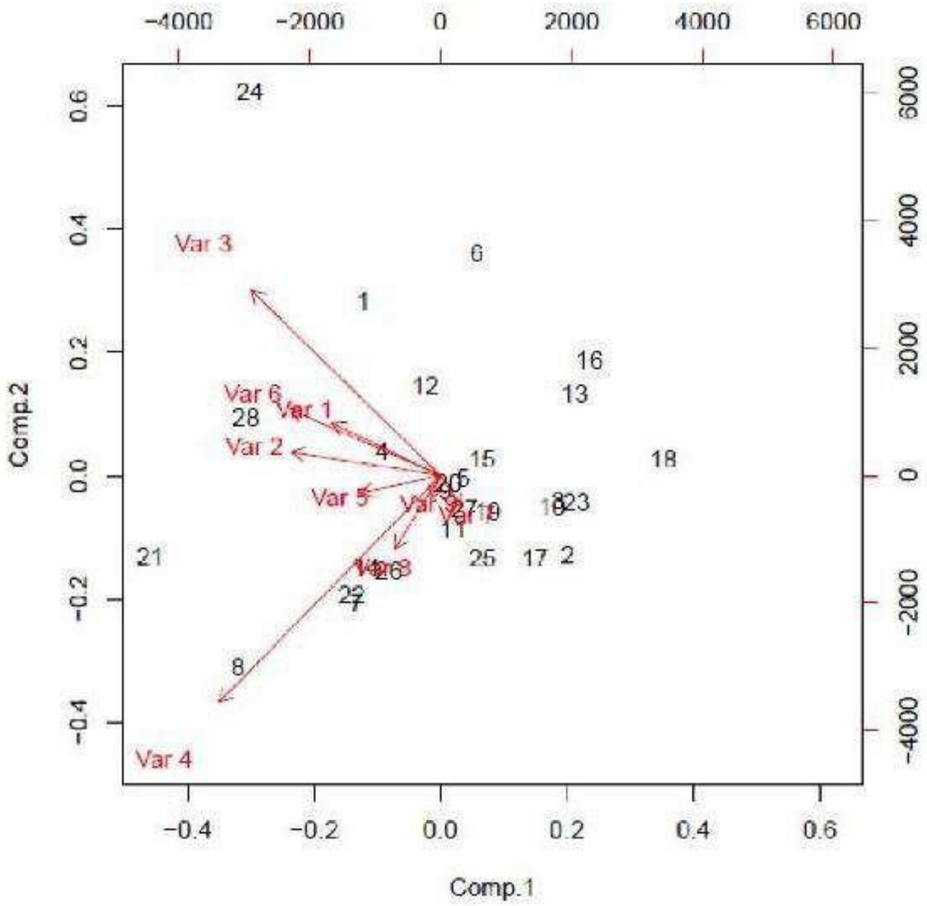
Following graph contains the biplot of the first principal component versus the second principal component:

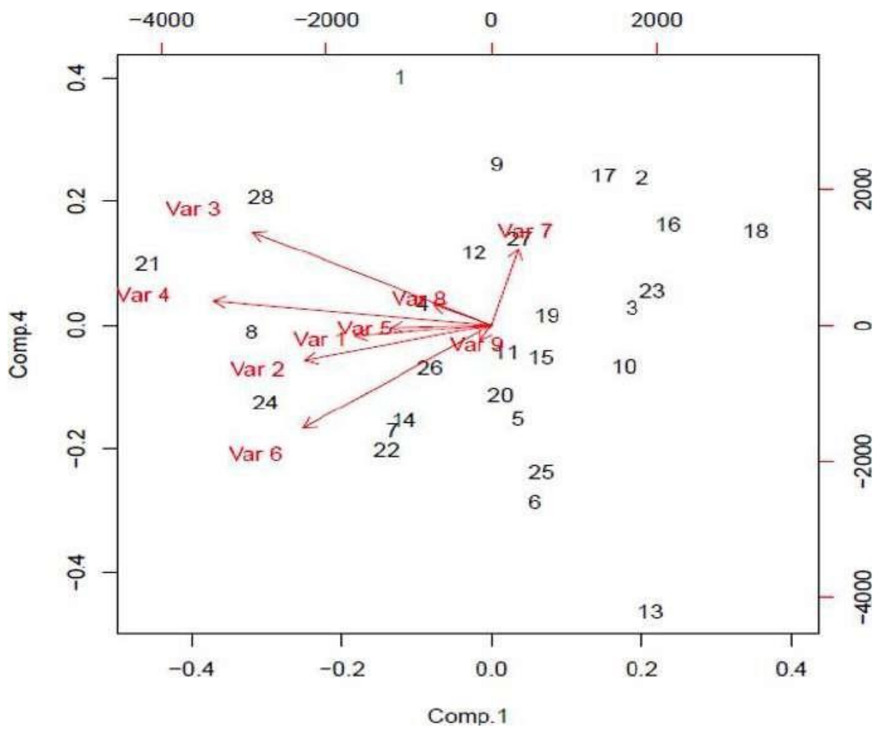


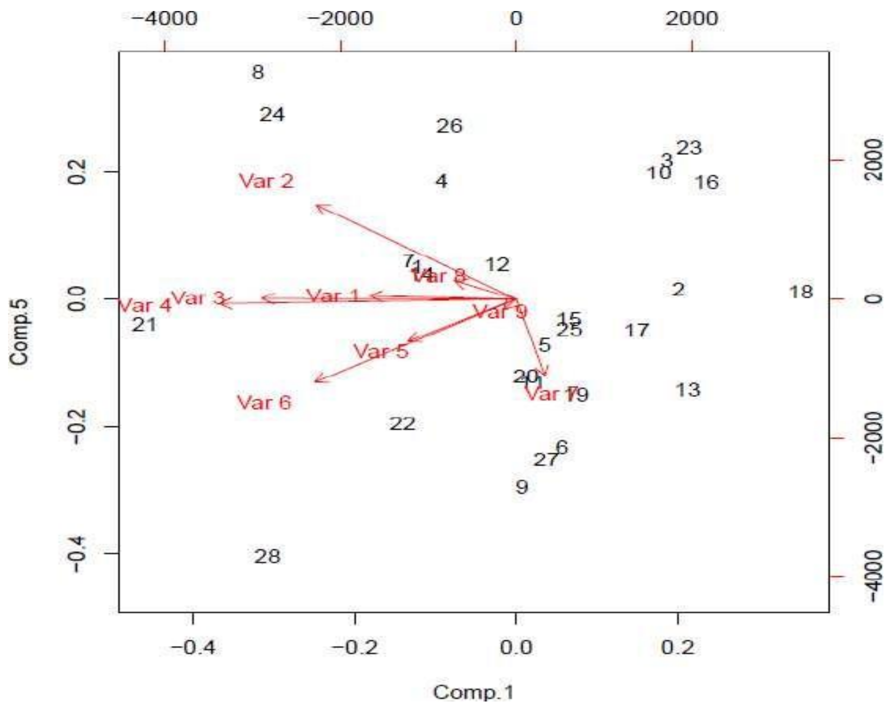
Here, the Left and the Bottom axis represent the loading of each variable on the principal components and the Right and the Top axis represent the score of each observation on the principal components under consideration.

Scoring highly on a single component simply means that the original variable values for these locations (states) are overwhelmingly explained by a single component.

Biplots of the first principal component against the third, fourth and fifth principal component are as follows:


Comp.1
Comp.2
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6

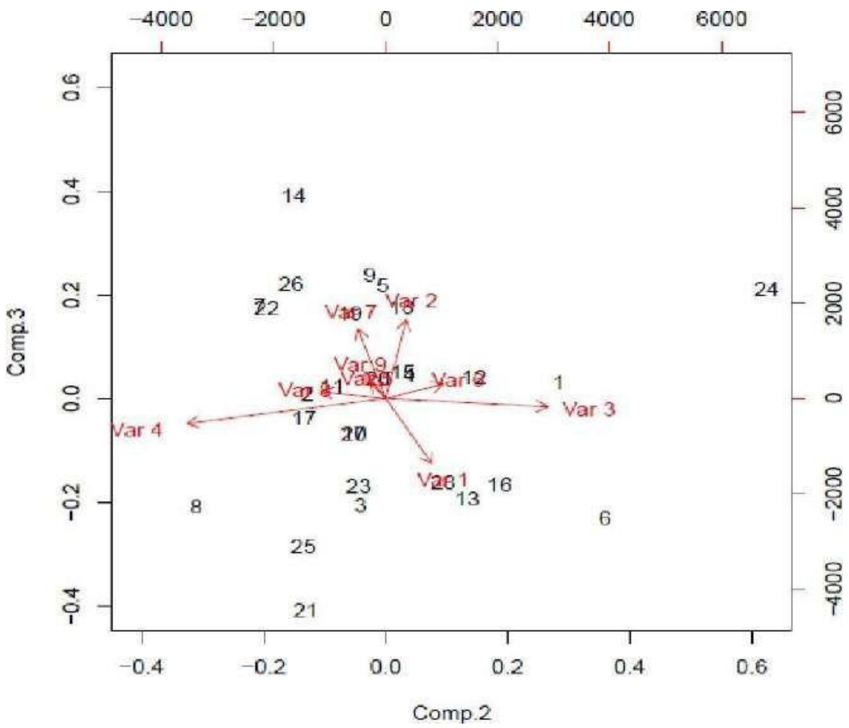

Comp.1
Comp.4
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6
Var 7
Var 8
Var 9

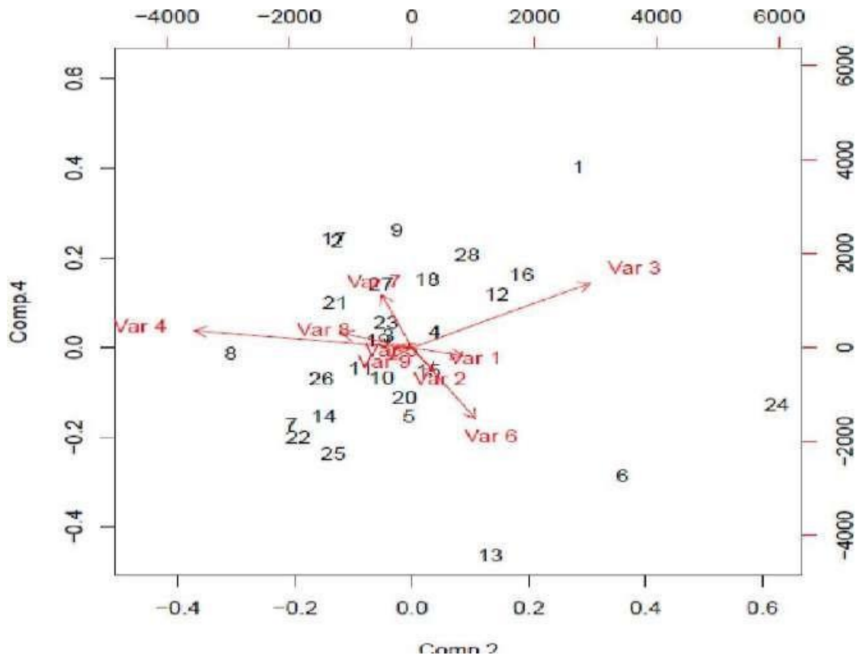


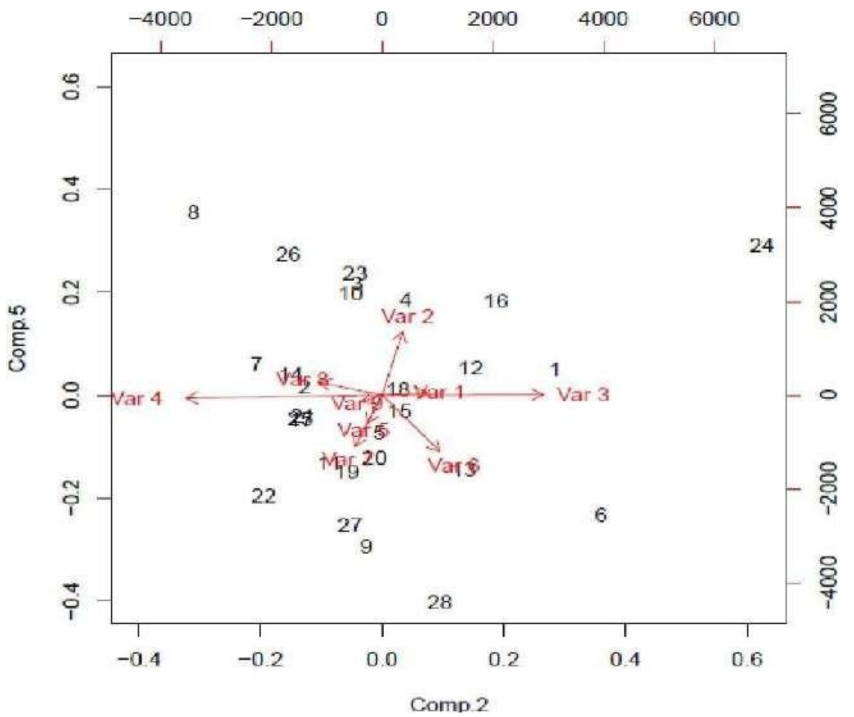
It can be seen from the biplots of PC1 versus all the other 4 PCs that the observations corresponding to states of Mizoram, Manipur, Arunachal Pradesh, Assam, Sikkim scores highly on the first component. It means that among all of the variables for these states, is more completely explained by a single component which is the first PC. Again, the variables maize and wheat load moderately highly but negatively associated with the first PC, both of which are staple food crops.

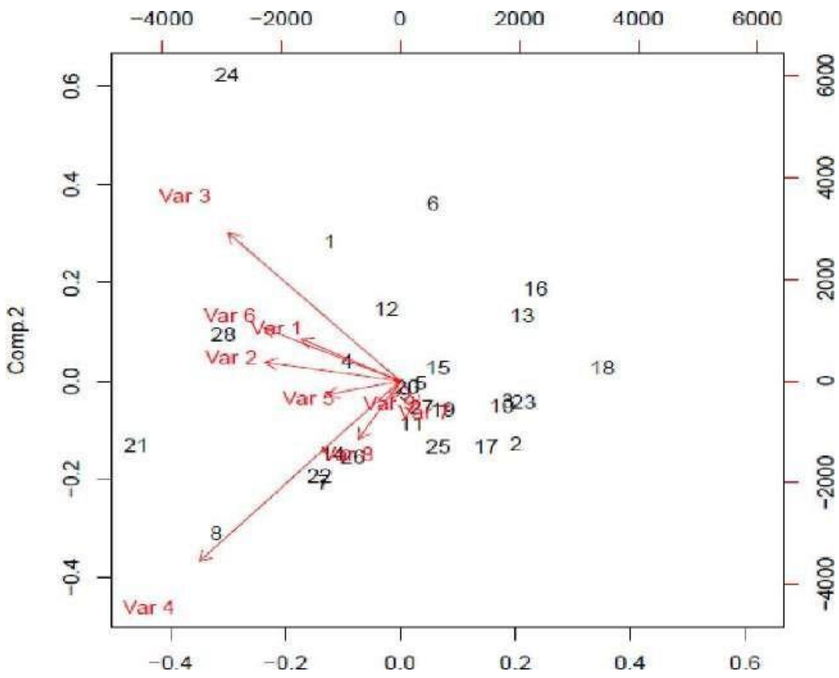
Hence, we can say that these 5 states (belonging to the Northeast region) are not good producers of maize and wheat.

(b) <u>Biplots for Examining Component 2</u>



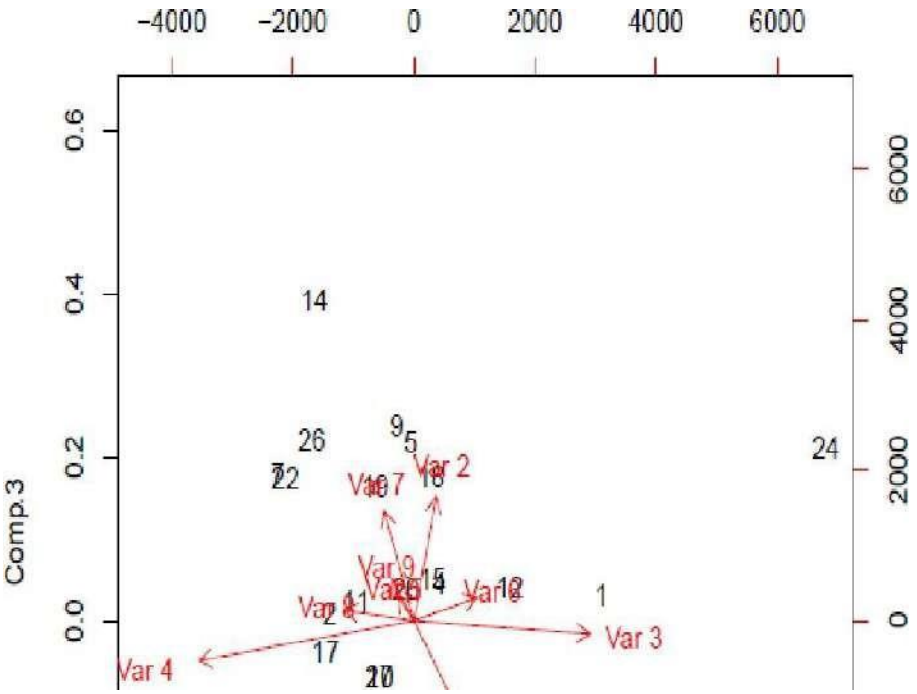

Comp.2
Comp.4
Var 1
Var 2
Var 3
Var 4
Var 6
Var 7
Var 8
Var 9

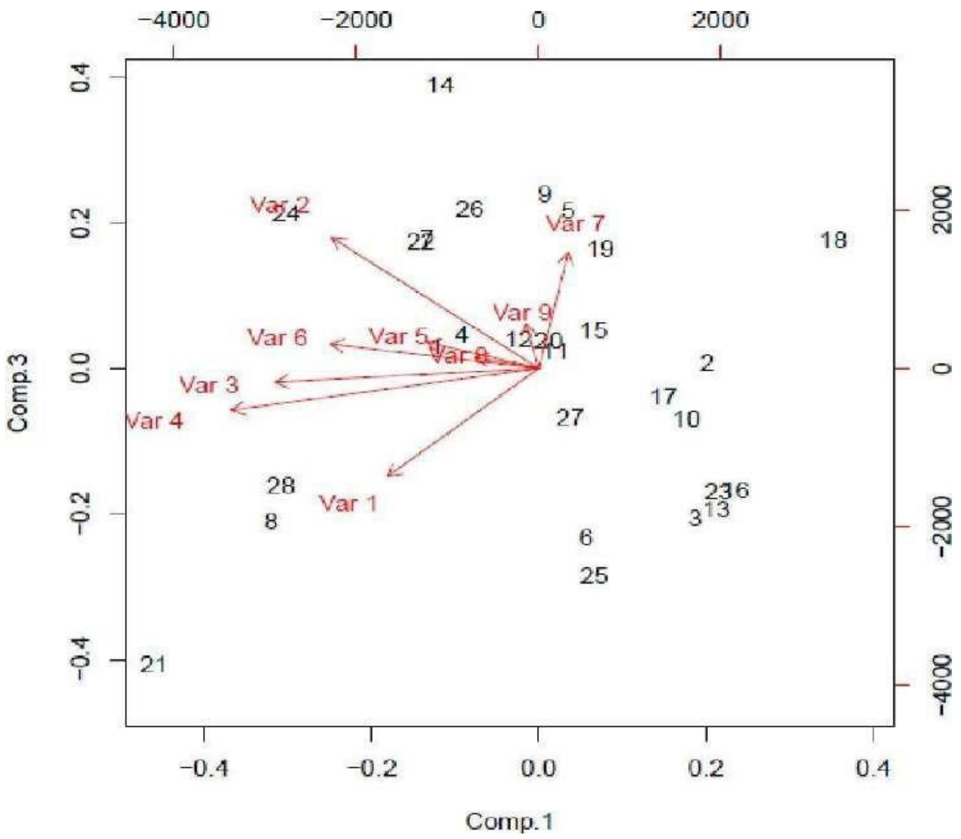

Comp.2
Comp.5
Var 1
Var 2
Var 3
Var 4
Var 6
Var 8
Var 9

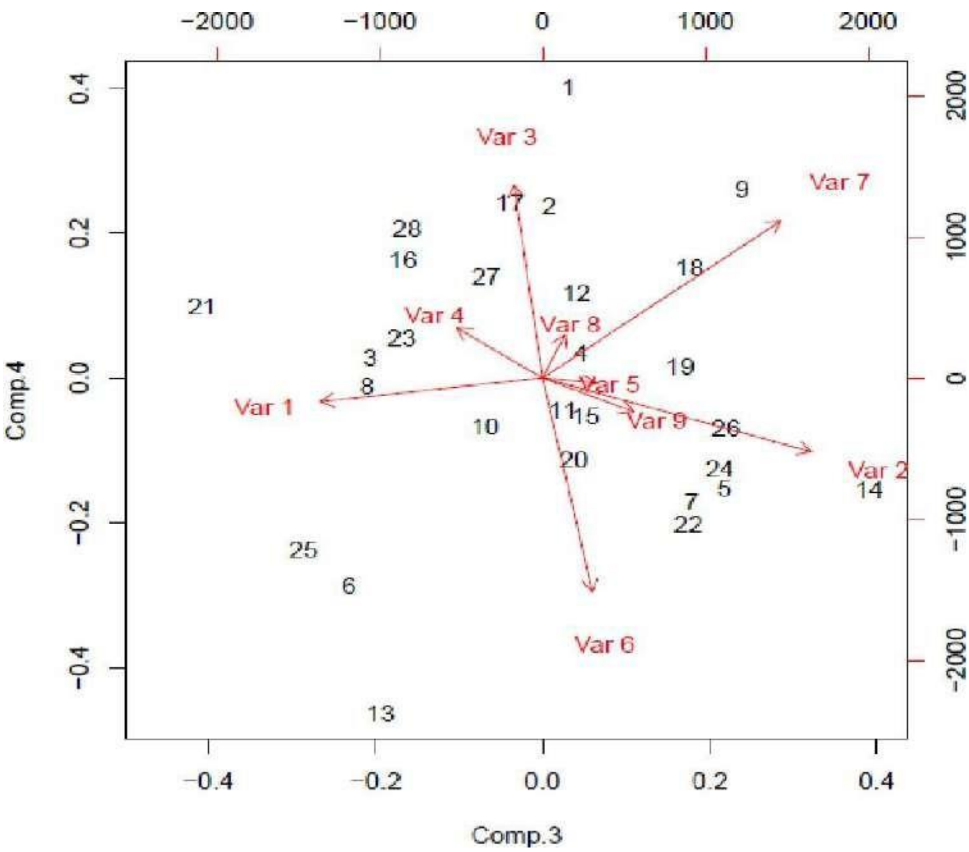


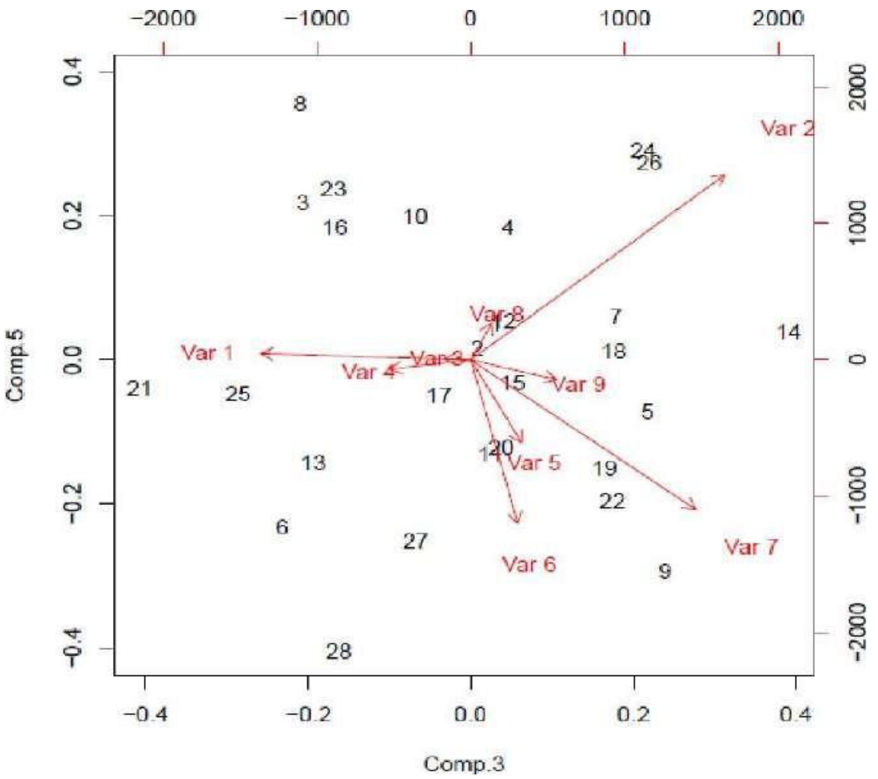
It can be seen from the biplots of PC2 versus all the other 4 PCs that the observation corresponding to the states of Tamil Nadu and Goa scores highly on the second component. It means that among all of the variables for Tamil Nadu and Goa is more completely explained by a single component which is the second PC. Again, the variables maize and wheat load moderately highly but positively and negatively associated respectively with the 2nd PC. Hence, we can say that Tamil Nadu and Goa are good producers of Maize but not Wheat.

(c) <u>Biplots for Examining Component 3</u>



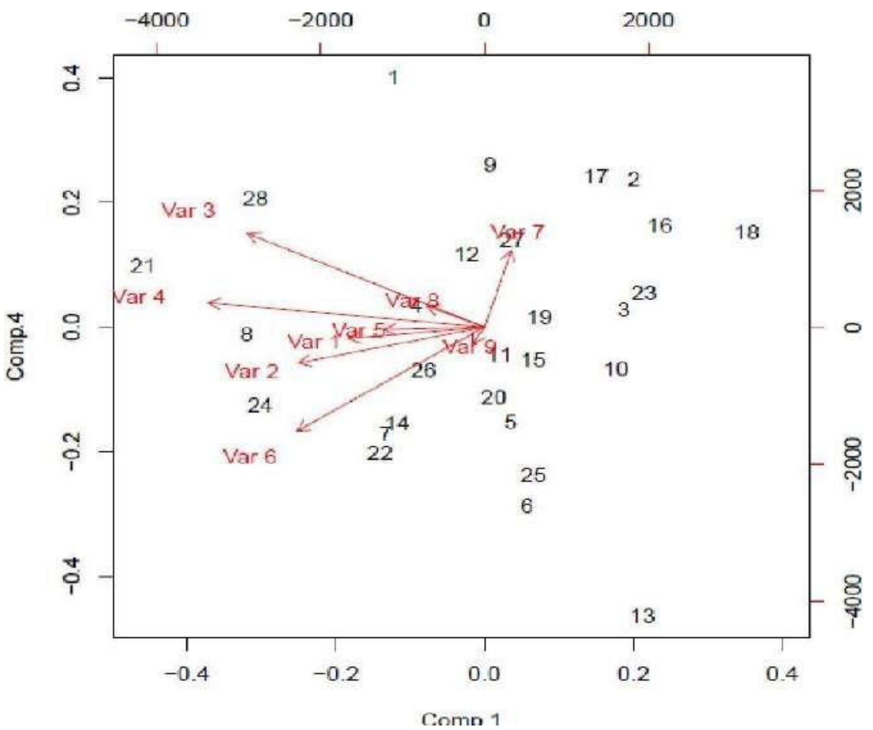

Comp.1
Comp.3
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6
Var 7
Var 9

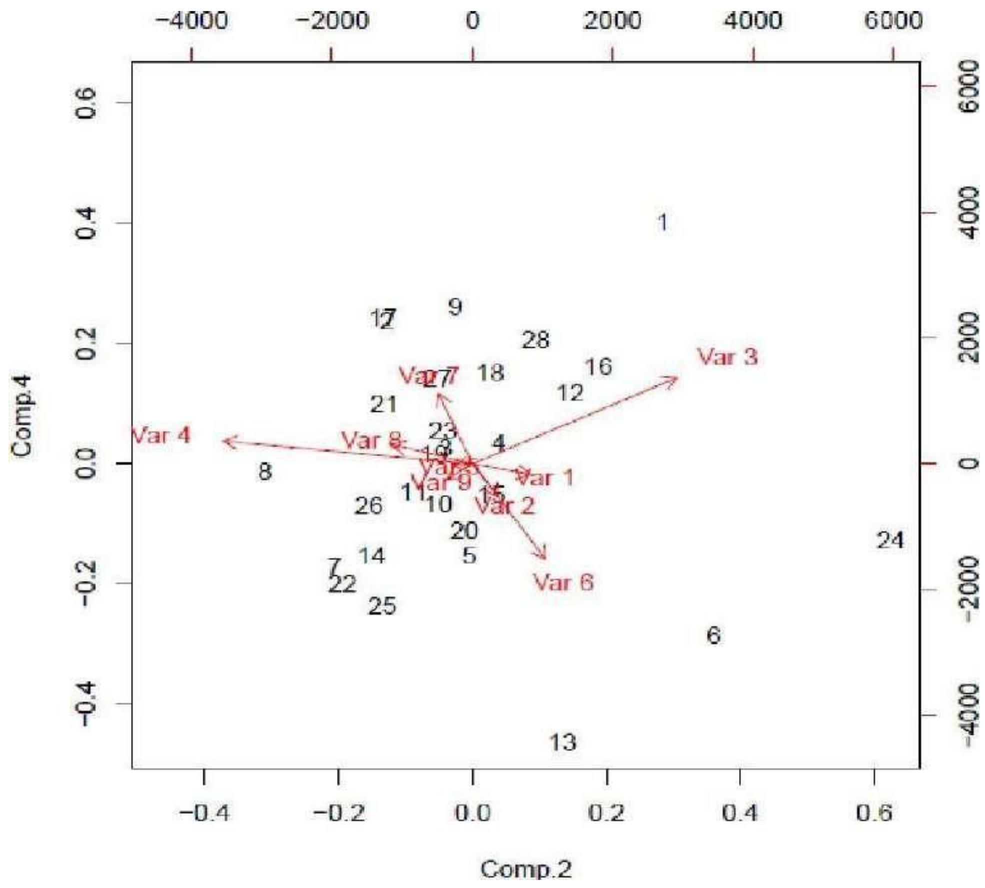

Comp.3
Comp.4
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6
Var 7
Var 8
Var 9

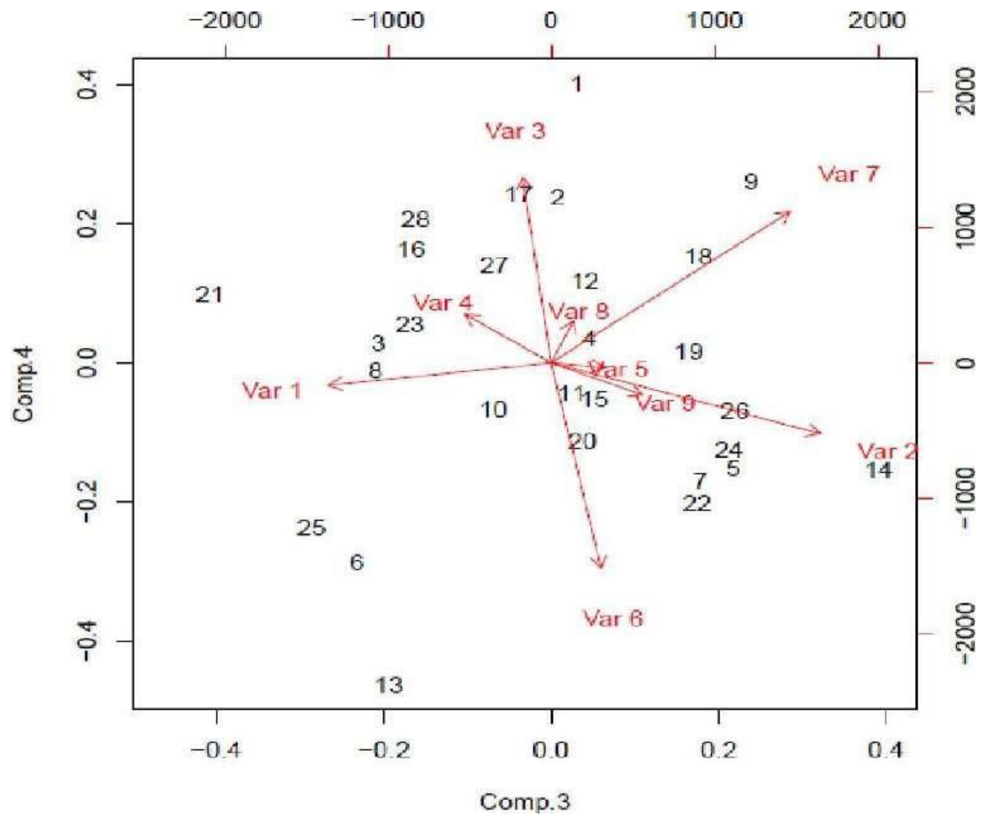


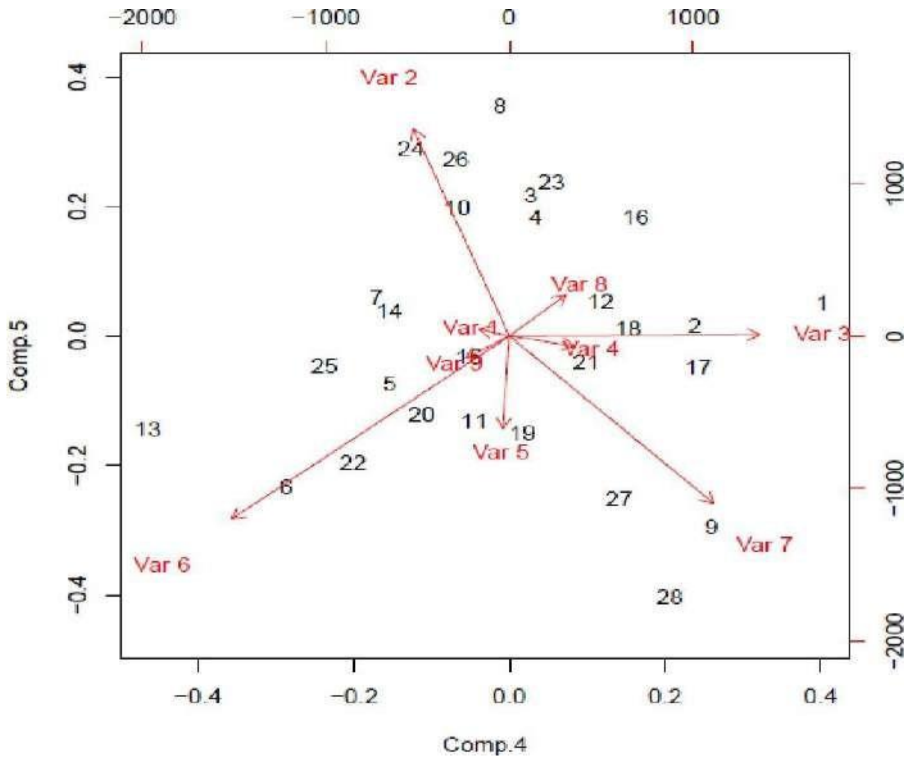
It can be seen from the biplots of PC3 versus all the other 4 PCs that the observations corresponding to states of Madhya Pradesh, Uttar Pradesh and Himachal Pradesh scores highly on the third component. It means that among all of the variables for Madhya Pradesh, Uttar Pradesh and Himachal Pradesh is more completely explained by a single component which is the third PC. Again, the variables bajra and soyabean load moderately highly, both positively associated with the 3rd PC. Hence, we can say that Madhya Pradesh, Uttar Pradesh and Himachal Pradesh are good producers of both bajra and Soyabean.

(d) <u>Biplots for Examining Component 4</u>



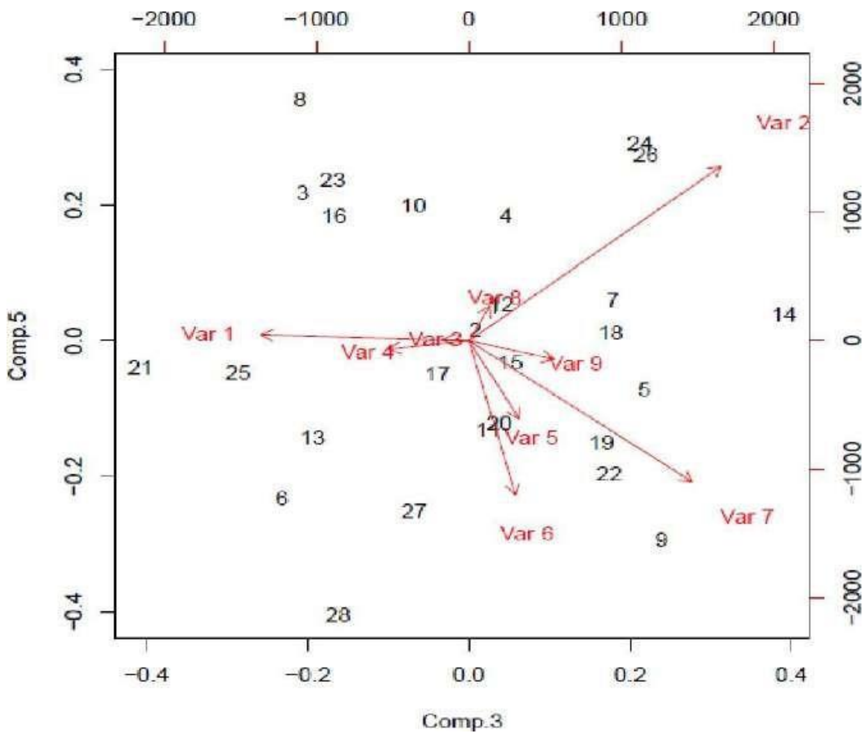

−4000
−2000
0
2000
4000
6000
Comp.4
Comp.2
−0.4
−0.2
0.0
0.2
0.4
0.6
Var 1
Var 2
Var 3
Var 4
Var 6
Var 8

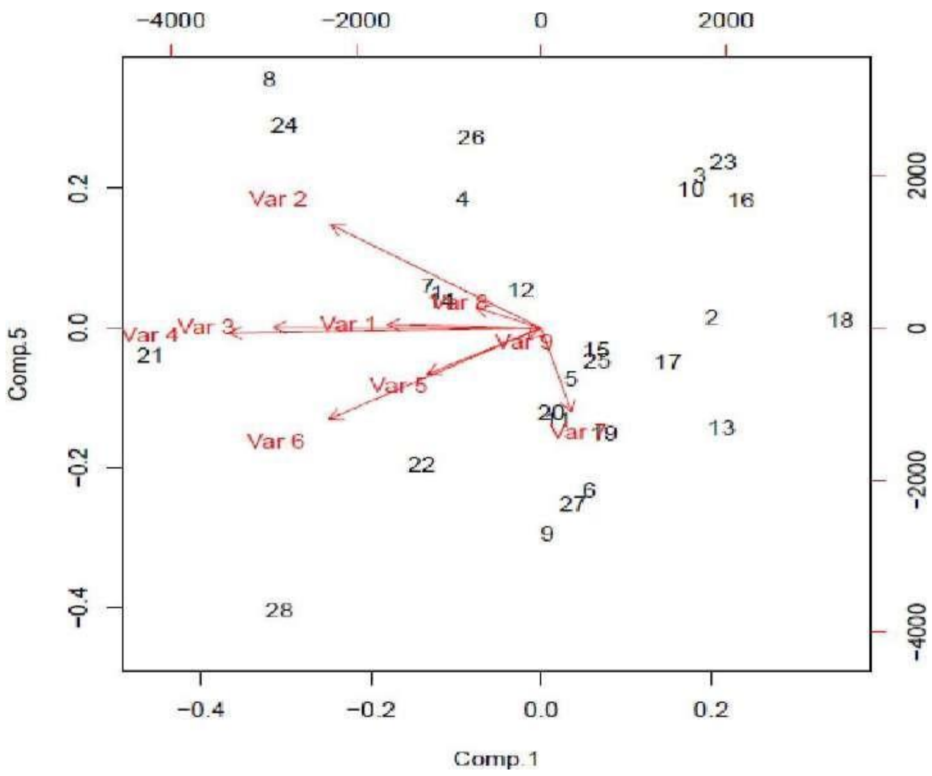

−2000
−1000
0
1000
2000
Comp.4
Comp.3
−0.4
−0.2
0.0
0.2
0.4
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6
Var 7
Var 8
Var 9

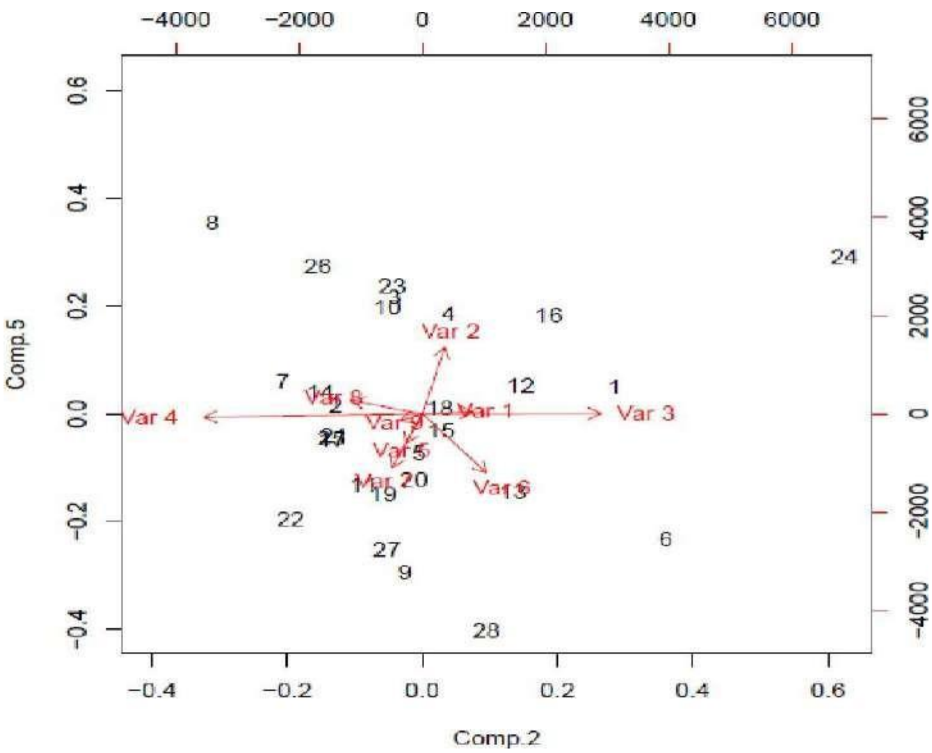


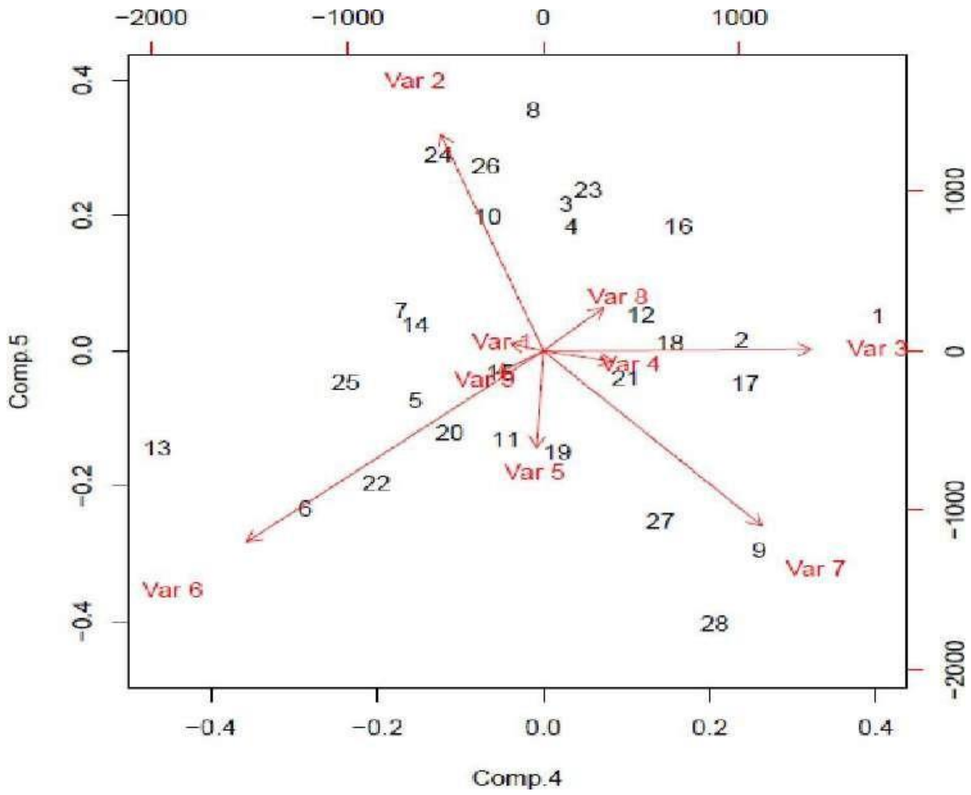
It can be seen from the biplots of PC4 versus all the other 4 PCs that the observations corresponding to state of Andhra Pradesh + Telangana, Himachal Pradesh, West Bengal, Meghalaya, Arunachal Pradesh scores highly on the fourth component. It means that among all of the variables for Andhra Pradesh + Telangana, Himachal Pradesh, West Bengal, Meghalaya, Arunachal Pradesh is more completely explained by a single component which is the fourth PC. Again, the variables maize and groundnut load moderately highly, but positively and negatively associated respectively with the 4th PC. Hence, we can say that Andhra Pradesh + Telangana, Himachal Pradesh, West Bengal, Meghalaya, Arunachal Pradesh are good producers of Maize but not groundnut.

(e) Biplots for Examining Component 5




Comp.5
Comp.1
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6
Var 7
Var 8
Var 9


Comp.5
Comp.2
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6
Var 9



It can be seen from the biplots of PC5 versus all the other 4 PCs that the observations corresponding to states of Haryana, Uttar Pradesh, Tamil Nadu and Sikkim scores highly on the fifth component. It means that among all of the variables for Haryana, Uttar Pradesh, Tamil Nadu and Sikkim is more completely explained by a single component which is the fifth PC. Again, the variables bajra, groundnut and soyabean load moderately highly, but positively, negatively and negatively associated respectively with the 5th PC. Hence, we cansay that Haryana, Uttar Pradesh, Tamil Nadu and Sikkim are good producers of bajra, but not groundnut.

(6) <u>Kruskal-Wallis Rank Sum Test</u>

(a) <u>For staple food crops</u>: Here, we consider the zone-wise yield of only the staple food crops among the major food cropsconsidered viz. rice, maize, wheat and bajra.

We intend to test the hypothesis

H_0: the zone wise yield of staple food crops for the 6 zones in India is homogeneous. Against

H_1: the zone wise yield of staple food crops for the 6 zones in India is not homogeneous.

<u>Results</u>: On performing the Kruskal-Wallis test in R, we find the p-value of the test to be $0.02174 < 0.05$. Hence, we reject our null hypothesis at 5% level of significance and conclude that the zone wise yieldof staple food crops for the 6 zones in India is not homogeneous.

(b) <u>For non-staple food crops</u>: Here, we consider the zone-wise yield of only the non-staple food crops among the major food cropsconsidered viz. Gram, Groundnut, Soya bean, R&M and Linseed.

We intend to test the hypothesis

H_0: the zone wise yield of non-staple food crops for the 6 zones in India is homogeneous. Against

H_1: the zone wise yield of non-staple food crops for the 6 zones in India is not homogeneous.

<u>Results</u>: On performing the Kruskal-Wallis test in R, we find the p-value of the test to be $0.01324 < 0.05$. Hence, we reject our null hypothesis at 5% level of significance and conclude that the zone wise yield of non-staple food crops for the 6 zones in India is not homogeneous.

7. Conclusion and Discussion

(a) The principal component analysis reveals that the nine original variables can be replace by the first five principal components, which are nothing but the linear combinations of the original variables that accounts for more than 90% variation present in the data; (b) The most contributing variables to the variation in the data have been found to be that corresponding to the crops bajra, maize, wheat, groundnut and soya bean. All of these crops but wheat is kharif crops; (c) The principal component analysis together with the graphical display of the data reveals that the north-eastern states are not good producers of wheat and maize whereas the states belonging to central India are good producers of bajra and Soya bean, which are kharif crops. Tamil Nadu is a leading producer of bajra and maize, both of which are again Kharif crops. Thus, they are the kharif crops which are causing most of the variation in the data; (d) Kruskal-Wallis test results show that zone wise yield of staple and non-staple food crops for the 6 zones in India is not homogeneous.

References

[1] Abellanosa, A. L., & Pava, H. M. (1980). Introduction to Crop Science. Central Mindanao University, Musuan, Bukidnon: Publications Office.
[2] Rubenstein, J. M. (2003). The Cultural Landscape: An Introduction to Human Geography (7th ed.).Pearson Education, Inc.
[3] Labour Code of the Philippines. Article 97(d), Chapter I, Title II.
[4] Black, H. C. (1990). Black's Law Dictionary: Definitions of the Terms and Phrases of American and English Jurisprudence (6th ed.). West Publishing Co.
[5] World Bank. (n.d.). Agricultural land (% of land area) – India. World Bank Open Data. https://data.worldbank.org/indicator/AG.LND.AGRI.ZS?locations=IN
[6] Food and Agriculture Organization of the United Nations. (n.d.). India at a Glance. https://www.fao.org/india/fao-in-india/india-at-a-glance/en/
[7] Ministry of Agriculture & Farmers Welfare. (n.d.). Contribution of Agricultural Sector in GDP. https://www.pib.gov.in/PressReleasePage.aspx?PRID=1909213
[8] International Monetary Fund. (n.d.). India. https://www.imf.org/en/Countries/IND
[9] Karnataka State Council for Science and Technology. (2015). The Pipeline Trencher. http://www.kscst.iisc.ernet.in/sppArchive/public/Abstract/038/7914.pdf
[10] NITI Aayog. (n.d.). Rural Women: Key to New India's Agrarian Revolution. https://www.niti.gov.in/rural-women-key-new-indias-agrarian-revolution
[11] Department of Agriculture, Cooperation and Farmers Welfare, Govt. of India. (2015). Agricultural Statistics at a Glance. www.agricoop.nic.in. Books Private Limited, New Delhi.
[12] [Website]. (n.d.). Essay on Important Food Crops of India. http://www.shareyouressays.com/120735/essay-on-important-food-crops-of-india-47-42-words
[13] Roy, A., Singh, N. U., Dkhar, D. S., Mohanty, A. K., Singh, S. B., & Tripathi, A. K. (2015). Food Security in North-East Region of India --- A State-wise Analysis. Agricultural Economics Research Review, 28 (Conference Number), 259-266. DOI: 10.5958/0974-0279.2015.00041.5
[14] Sharma, P., & Singh, A. K. (2017). Association of state-wise Food Security Index with Climatic Factors in India: Evidence from state-wise panel Data. Journal of Global Agriculture and Ecology, 6(3), 196-205.
[15] Khan, W., Ansari, S. A., & Ali, T. (2020). Economic reforms and agriculture in India: A state-wise analysis of the post-reforms period. J Public Affairs, 2020; e2390. DOI: https://doi.org/10.1002/pa.2390
[16] Madhukar, A., Kumar, V., & Dashora, K. (2020). Spatial and Temporal Trends in the Yields of Three Major Crops: Wheat, Rice, and Maize in India. Int. J. Plant Prod, 14, 187–207. DOI: https://doi.org/10.1007/s42106-019-00078-0
[17] Kumar, A., Sharma, P., & Joshi, S. (2015). Effects of Climatic Factors on Agricultural Productivity in India: A State-wise Panel Data Analysis. International Journal of Basic and Life Sciences, 3(1), ISSN (Online): 2320-513X.
[18] Samal, P., Jambhulkar, N. N., & Repalli, S. K. (2018). State-Wise Analysis of Growth in Production and Profitability of Rice in India. Indian Journal of Economics and Development, 4(3), 399-409. DOI: 10.5958/2322-0430.2018.00151.8
[19] Ajay Kumar and Pritee Sharma (2013). Impact of Climate Variation on Agricultural Productivity and Food Security in Rural India. Economics Discussion Papers, 43, Kiel Institute for the World Economy. DOI: https://dx.doi.org/10.2139/ssrn.4144089
[20] Singh, Ajay Kumar, & Jyoti, Bhim (2021). Projected Food-grain Production and Yield in India: An Evidence from State-wise Panel Data Investigation during 1977-2014. Journal of Agricultural Sciences, 16(1). URL: http://repo.lib.sab.ac.lk:8080/xmlui/handle/123456789/1059
[21] Mathur, Archana S., Das Surajit, & Sircar Subhalakshmi (2006). Status of Agriculture in India: Trends and Prospects. Economic and Political Weekly, 41(52), 5327-5336. DOI: http://www.jstor.org/stable/4419078
[22] Manogna, R. L., & Mishra, A. K. (2020). Agricultural production efficiency of Indian states: Evidence from data envelopment analysis. Int. J. Fin. Econ, 01–12. DOI: https://doi.org/10.1002/ijfe.2369