Title: Principal dimensions of regional agricultural &socio-economic disparities in Haryana

Authors: Ekta Hooda*, B K Hooda, Veena Manocha& Nitin Tanwar*$

Department of Mathematics, Statistics & Physics**,** CCS HAU-Hisar (Haryana)

*Ph. D.StudentDepartment of Mathematics, Statistics & Physics**,** CCS HAU-Hisar (Haryana)

$Corresponding author E-mail: nitintanwar5@gmail.com; Mob: 9268897969

Abstract

The present study is an attempt towards identification of principal agricultural and socio-economic dimensions in Haryana using principal component analysis (PCA) and canonical correlation analysis techniques. Principal component analysis transforms the original set of variables into a smaller set of linear combinations that account for most of the variation of the original data whereas canonical correlation analysis determines pairs of canonical variates which are orthogonal linear combinations of the variables within each set that best explain the variability both within and between sets. Canonical correlation analysis also identifies and measures the strength of relationships between two vectors of variables measured on the same individuals. The study was conducted for three periods i.e. 1991-92, 2001-02 and 2011-12. The district was considered as the unit of analysis and analysis is based on 19 indicators from the agriculture sector and 9 indicators from the socio-economic sector. The first principal component (PC) of agriculture sector represents the overall level of agriculture and livestock with 42.07, 28.71, and 28.01 per cent of the total variation in periods 1991-92, 2001-02 and 2011-12. Whereas, the first PC of socio-economic sector extracted 43.2, 42.6 and 56.6 per cent variation for the periods 1991-92, 2001-02 and 2011-12, respectively. Population density per sq km, number of vehicles on road/lakh population and number of cooperative societies/lakh population have been most important variables for the first principal component from the socio-economic sector in the periods 1991-92 and 2001-02. However, infant mortality rate, number of vehicles on road/lakh population and main workers as percentage of total population has observed to be the most important indicators during 2011-12. Canonical analysis of first two PCs from each of the agriculture and socio-economic sector indicated that the dimensions represented by the second principal component of agriculture sector and first principal component of socio-economic sector established a strong association between the agriculture and the socio-economic sectors. The significant canonical correlation between the vectors represented by first two PCs of agricultural and socio-economic sectors suggest that developments in socio-economic sectors and agricultural sectors go together. That is socio-economic development in Haryana can be achieved through development in agriculture.

Key Words: *Principal Component Analysis, Canonical Correlation Analysis, Principal Dimensions.*

1. Introduction

Regional disparity means divergence or inequality of phenomena or processes having specific territorial allocation and occurring at least in two entities of the territorial structure. It refers to differences between economic performance and welfare between different regions. Regional

disparities are manifested in different conditions of life as well as in unequal economic and development potential.Regional disparities in economic development are one of the common features found in all over the world today. India is facing the problem of severe regional disparity and the indicators of such disparities are reflected by the factors like per capita income, the proportion of population living below the poverty line, the percentage of urban population, percentage of working population engaged in agriculture, the percentage of workers engaged in industries, infrastructural development etc.

Economic planning of a country is aimed at bringing about a balanced regional development and reduction in regional disparities in the pace of development. Since independence, India has implemented many developmental programmes to enhance the quality of life of people by providing basic necessities for effective improvement in their social and economic well-being. The literacy level, housing conditions and overall quality of life of the masses has considerably improved after independence. However, disparities in the level of development can still be observed at districts and state levels with certain areas went ahead leaving other lagged behind.

The process of development cannot be captured fully by any single indicator. Also, a number of indicators analyzed individually do not provide an easily comprehensible picture of the true development patterns. Arbitrary selection of a large number of indicators from various sectors has now become a routine practice for inter-regional disparity analysis. This usually happens because the development analysts are perhaps, uncertain about the relative importance of indicators in regional discrimination and hence consider as many as possible to make sure that all the important ones are included. Inclusion of some of the indicators may be totally irrelevant to the classification of regions and may mask any genuine pattern that exists in the present data. Further, the development indicators within and across sectors are linked together and additional information supplied by one indicator independently of the others may be negligible. [1] Studied developmental disparities in districts of Haryana according to their level of development. The study utilized data over three points of time, viz. 1991-92, 2001-02, and 2011-12. Assessment of development in agricultural, industrial, infrastructural and socio-economic sectors has been studied using composite indices based on forty indicators. The districts of Ambala, Faridabad and Gurgaon ranked first in overall development in 1991-92, 2001-02 and 2011-12, respectively, whereas Mahendragarh ranked last in 1991-92 and 2001-02 and the newly formed district Mewat in 2011-12.

Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA) are important multivariate techniques frequently used in biological and social sciences. PCA transforms the original set of variables into a smaller set of linear combinations that account for

most of the variation of the original data whereas CCA determines pairs of canonical variates which are orthogonal linear combinations of the variables within each set that best explain the variability both within and between sets. CCA also identifies and measures the strength of relationship between two vectors of variables measured on the same individuals.

Many studies have been conducted in past to classify regions in Haryana and other states of India based on indicators from agriculture, socio-economic, infrastructure and industrial sectors. [2] Used structural equation modeling approach for classification of regions in Haryana at block level. [3] Analysed the spatial patterns of regional disparities and variations in the levels of agricultural development among the districts of the state of Uttar Pradesh. [4] Measured the levels of agricultural development for the state of Odisha where the study revealed that 7 out of 30 districts of Odisha came under the category of backward regions, showing that large regional disparities exist in levels of agricultural development in the state. [5] Identified the evolutionary path of a number of local systems in a Mediterranean country vulnerable to soil degradation in the last decades. Multivariate techniques, principal component analysis and canonical correlation analysis were used to evaluate the socio-ecological conditions and to estimate rapidity of change of local systems by considering 6 bio-physical factors predisposing soil to degradation and 23 socio-economic indicators over fifty years (1960–2010). [6] Made an attempt to capture the recent dynamics of development of districts of Eastern Uttar Pradesh in respect of three sectors- Agriculture, Social and Infrastructure. Technique of composite indices suggested by [7] was used in addition to principal component and factor analysis. Ranking seems to very close to ground reality and provides useful information for further planning and corrective measures for future development of Eastern Uttar Pradesh's Districts. [8] Ranked the districts of Himachal Pradesh on the basis of their levels of development obtained with the help of optimum combination of 35 indicators related to agriculture, social and industrial sectors. The district wise data in respect of the indicators published by Himachal Pradesh government for the year 2014-15 were used for all 12 districts in Himachal Pradesh for the study. Principal component analysis and factor analysis were used for ranking the districts.

Haryana is one of the 29 states in India, situated in North India. It was carved out of the former state of East Punjab on 1 November 1966 on a linguistic basis. It is a small state in the Indian Union with a geographical area of about 44,212 km^2. As of 2011 census of India, the state is 18^{th} largest by population with 25,353,081 inhabitants. It is predominantly an agrarian state with 65.12% of the population living in rural areas. Agriculture plays an important role in the socio-economic development o the state. Therefore, the present study has been planned to

identify the principal dimensions of regional disparities in agriculture and socio-economic sectors in Haryana, which play a major role in developmental disparities in various districts of Haryana. PCA has been used to identify principal dimensions for the agriculture and socio-economic sectors for the periods 1991-92, 2001-02 and 2011-12. Principal Canonical Correlation Analysis (PCCA) proposed by [9] has been used to study the association between agricultural and socio-economic dimensions in Haryana.

2. Materials and Methods

An individual district of the state of Haryana has been considered as the unit of analysis. The necessary data on agriculture and socio-economic sectors have been collected from the various issues of Statistical Abstracts published by government of Haryana. The study utilized district-wise data of Haryana for the three points of time, i.e. Period-I: 1991-92, Period-II: 2001-02 and Period-II: 2011-12 with sector-wise indicators given below:

Indicators for Agriculture Sector

AG1: Percentage of gross area sown under foodgrain to total cropped area
AG2: Irrigation intensity
AG3: Percentage of gross area sown under commercial crops to total cropped area
AG4: Gross value from agriculture/ha at current prices (in Rs.)
AG5: Gross value of agriculture output per capita (rural) at current prices (in Rs.)
AG6: Percentage of area under HYV of wheat to total cropped area
AG7: Productivity of cereals (t/ha)
AG8: Productivity of pulses (t/ha)
AG9: Productivity of oilseeds (t/ha)
AG10: Number of regulated markets
AG11: Percentage of agriculture workers to total work force
AG12: Cropping intensity
AG13: Average annual rainfall (mm)
AG14: Number of tractors/000ha of gross cropped area
AG15: Tubewells&pumpssets/000ha of gross cropped area
AG16: Fertilizer consumption (in kg) in terms of nutrients /ha of gross cropped area
AG17: Cattle per sq km
AG18: Buffalo per sq km
AG19: Poultry per sq km

Indicators for Socio-Economic Sector

SE1: Main workers as % of total population
SE2: Literacy (%)
SE3: Female literacy (%)
SE4: Population density per sq km
SE5: Infant mortality rate
SE6: Number of registered motor vehicles/lakh population

SE7: Number of vehicles on road/lakh population
SE8: Number of cooperative societies/lakh population
SE9: Urban population (%)

Selection of Principal Dimensions (Indicators) Using Principal Components

Principal Component Analysis (PCA) transforms the original set of variables into a smaller set of linear combinations that account for most of the variation of the original data. The standard PCA results guarantee that retaining first few principal components with the largest associated variance produces the subset oflinear combinations of the original variables which, best approximates original data. The first principal component (PC1) is that weighted linear combination of the observed variables which accounts for the largest amount of the total variation in the data. The second principal component (PC2) is the weighted linear combination which is uncorrelated with PC1 and accounts for the maximum amount of the remaining variation in data and so on.

Principal Canonical Correlation Analysis

Canonical correlation analysis (CCA) is frequently used to analyze association between two vectors/sets of variables [10]. In most applications one vector(X: $p_1 \times 1$) is called the set of predictors and the other vector(Y: $p_2 \times 1$) is called the set of criterion or response variables. The idea of canonical correlation is to find two linear composites, one for X and one for Y, such that their correlation is maximum. The resulting correlation is called the first canonical correlation and the pair of linear combinations as the first canonical variate pair. In this context canonical correlation looks like PCA where k independent components are extracted which are linear combinations of the original variables and these k components explain maximum variation in the original data set. The procedure is continued until two new coordinate systems are specified completely. In practice, a maximum of q = min (p_1, p_2) canonical variates pairs can be extracted, where p_1 and p_2 represent the number of variables in the sets X and Y respectively. Canonical correlation analysis may be performed either using a joint covariance matrix (S) or joint correlation matrix (R) for the vectors X and Y depending upon the measurements considered in subsectors X and Y.

If

$$R = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} \quad (1)$$

is the joint correlation matrix of X and Y, then canonical correlations between X and Y can be found by solving the eigenvalue equations

$$R_{xx}^{-1} R_{xy} R_{yy}^{-1} R_{yx}\, a = \lambda_a \quad (2)$$

And

$$R_{yy}^{-1}R_{yx}R_{xx}^{-1}R_{xy}\,b = \lambda_b \tag{3}$$

Where the eigenvalues λ are the squared canonical correlations and the eigenvectors aandb are the eigenvectors of the two matrices. The largest eigen value is the square of the first canonical correlation. In practice, only one of the eigenvalue equations needs to be solved since the solutions are related by

$$a = \frac{1}{\sqrt{\lambda}} R_{xx}^{-1}R_{xy}b \tag{4}$$

And

$$b = \frac{1}{\sqrt{\lambda}} R_{yy}^{-1}R_{yx}a \tag{5}$$

Then, U = a'Xand V = b'Y is the canonical variate pair. The significance of $q_1 < q = \min(p_1, p_2)$ canonical variate pair may be judged by the test statistics

$$\chi_{k1}^2 = -\left[(n-1) - (p_1 + p_2 + 1)/2\right]\ln\Lambda_{k1} \tag{6}$$

This follows a Chi-square distribution [11] with degree of freedom $(p_1\text{-}k_1)$ $(p_2\text{-}k_1)$. Where,

$$\Lambda_{k1} = \prod_{i=k1+1}^{k}(1-\lambda_{(i)}),\ k_1 < k \tag{7}$$

And $\lambda_{(i)}$ is the i^{th} eigenvalue of

$$R_{yy}^{-1}R_{yx}R_{xx}^{-1}R_{xy} \tag{8}$$

[7] Proposed a modified method of canonical correlation analysis and called it principal canonical correlation analysis. PCCA is canonical correlation analysis of two sets of principal component scores. A separate PCA is performed for each set and component scores are computed. PCCA then uses these PC scores instead of the original random vectors. PCA transforms the given data of correlated variables into a new data set of uncorrelated PC scores and these scores are derived from the original variables that retain a certain percentage of the inherent variability. Also, each PC score accounts for a decreasing proportion of the total variance inherent in the data. [12] Pointed out that the interpretation of principal components is easier than the canonical variates. Therefore, it is assumed that PCCA has some merit, because PC scores descend in order of the amount of information that they contain. Thus, by using only selected PC scores, it will be easier to interpret the CCA. Therefore, comparing CCA with PCA, the canonical correlation of first two principal components is more useful for study of the relation between the sets of variables.

3. Results and Discussion

3.1 Principal Dimensions of Agriculture in Haryana

Period-wise PCA was performed with correlation matrix as input. The eigenvalues and the percentage of variation explained by first six PCs for 19 indicators of the Agriculture sector are presented in Table-1. The first 6 PCs explained 90.4, 87.4 and 84.7 percent variation of the data sets for the periods 1991-92, 2001-02 and 2011-12, respectively. The first two PCs explained 59.6, 53.0 and 48.9 percent variation of the data set in periods 1991-92, 2001-02 and 2011-12, respectively and hence, can be considered as principal dimensions for the 19 indicators of the agriculture sector. The corresponding principal component loadings have been presented in Table-2.

First principal component for agriculture sector (AGPC1) explained 42.07, 28.71, and 28.01 per cent of the total variation in periods 1991-92, 2001-02 and 2011-12, respectively. Loading pattern for the first principal component indicate that the most important indicatorsfor this dimension are AG4 (Gross value from agriculture/ha at current prices), AG7 (Productivity of cereals), AG16 (Fertilizer consumption in terms of nutrients/ha of gross cropped area), and AG17 (Cattle per sq km). Forthe period 1991-92, loadings for AG4, AG7, AG16 and AG17 are 0.93, 0.78, 0.81 and 0.82 respectively.The first component is also influenced by AG1 (gross value from agriculture/ha at current prices), AG2, AG3, AG6, AG7, AG8, AG12, AG14 and AG18. All these variables have positive association with PC1 except AG3 (percentage of gross area sown under commercial crops to total cropped area) which is having a negative association. Similar loading patterns have also been observed for this component during the periods 2001-02 and 2011-12. Thus, AGPC1 gives the overall level of agriculture and livestock in all the periods and can be considered as principal dimensions of agriculture sector.

The second principal component for the agriculture sector (AGPC2) explained 17.48, 24.24 and 20.86 percent of the total variability for the periods 1991-92, 2001-02 and 2011-12, respectively. For the period 1991-92 the most important indicator for AGPC2 is AG11 (percentage of agriculture workforce to total workforce) with loading 0.85 followed by AG13 (average annual rainfall) with loading -0.73.Other important variables for AGPC2 are AG5 (Gross value of agriculture output per capita (rural) at current prices), AG15 (Tube wells & pump sets/000ha of gross cropped area) and AG19 (Poultry per sq km). AG5 and AG11 have positive influence with loadings 0.63 and 0.85 whereas AG13, AG15 and AG19 have negative influence with loadings -0.73, -0.64 and -0.63 respectively. AGPC2 is a contrast between indicators with positive and negative loadings. By and large, similar loading patterns have also been observed for the PCs of agriculture sector for the periods 2001-02 and 2011-12. AG4, AG7 and AG16 have been observed to be the most influential variables for the principal component in period 2001-02 whereas AG4, AG7 and AG17 in period 2011-12. AG11 has

been observed to be the most important variable for the second principal component (AGPC2) in all of the periods for the agriculture sector. The other components have no clear cut loadings pattern and are of lesser importance.

Scatter plot for the period 2011-12 for first two principal components of agricultural sector indicators has been presented in figure 1. It indicates that the districts Karnal, Kurukshetra, Panipat and Yamuna nagar have high principal component scores value for AGPC1 and AGPC2. On the other hand, the districts Mewat, Mahendragarh and Bhiwani have low scores for both of the principal components of agricultural sector indicators. The developmental disparities indicated by AGPC1 and AGPC2 in figure 1 are in accordance with the disparities reflected by composite index of development in [1].

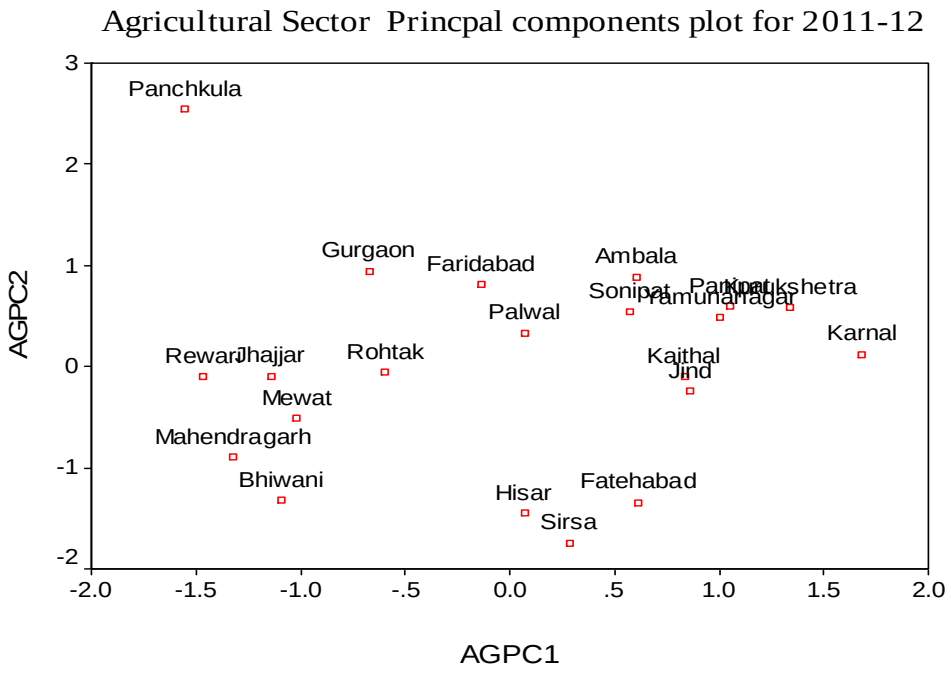

Figure 1 (Source: Own research)

3.2 Principal Socio-Economic Dimensions in Haryana

The first three PCs (Table-1) for the socio-economic sector explained nearly 85.90, 76.74 and 81.82 percent variability of the data set having 9 indicators for the periods 1991-92, 2001-02 and 2011-12, respectively. The first PC explained 43.2, 42.6 and 56.6 percent variation for the periods 1991-92, 2001-02 and 2011-12, respectively. Loadings for the socio-economic sectors have been presented in Tables-3. It was observed that SE4 (population density per sq km), SE7 (number of vehicles on road/lakh population) and SE8 (number of cooperative societies/lakh population) are the most important variables for the first principal component from the socio-

economic sector (SEPC1) in the period 1991-92. Loadings for SE4, SE7 and SE8 during this period are 0.85, 0.86 and 0.82 respectively. Similar loading pattern have also been observed for the period 2001-02 with loading of 0.78, 0.81 and 0.84 for the indicators SE4, SE7 and SE8 respectively. The indicators SE1 (main workers as percentage of total workers), SE3 (female literacy percentage) and SE5 (infant mortality rate) also have high component loadings for first two periods justifying SEPC1 to be a principal socio-economic dimension during these periods. The other three indicators, viz. SE2 (literacy percentage), SE5 (infant mortality rate) and SE6 (number of registered motor vehicles per lakh population) have high loadings on second principal component from the socio-economic sector (SEPC2) in 1991-92 which explains 32.23 percent of variability in the data. Only two indicators SE2 and SE6 have high loadings on SEPC2 for the 2001-02 while SE9 (urban population percentage) has highest loading on SEPC3 for the period 2001-02.

The second PC is determined by the indicators SE6 (number of registered motor vehicles/lakh population) and SE2 (percent literacy) in both the periods. The loading pattern is slightly different for the period 2011-12, where the most important variable for the principal dimension (SEPC1) is SE5 (infant mortality rate) followed by SE7 and SE1. Except SE9, all the indicators have high positive loading on first principal component and SEPC1 can be considered as the overall dimension of socio-economic development. Since SE9 (urban population percentage) has high loading on SEPC3 for both periods 2001-02 and 2011-12, it indicates that urban population percentage is defining a separate socio-economic dimension in these periods.

Similarly, Scatter plot for the period 2011-12 for first two principal components of socio-economic sector indicators shows high scores on principal components and has been presented in figure 2. This plot indicates that Faridabad is the most developed district having high scores on both the principal components while Gurgaon and Panchkula districts have high scores on SEPC1 only. On the other hand districts Mewat, Palwal, Bhiwani and Sirsa have low principal component scores and having classified as less developed districts on the socio-economic front. This is also in accordance with developmental ranking based on composite indices of development in [1].

Thus AGPC1 and SEPC1 are respective the principal dimensions for assessing disparities in agricultural and socio-economic sectors.

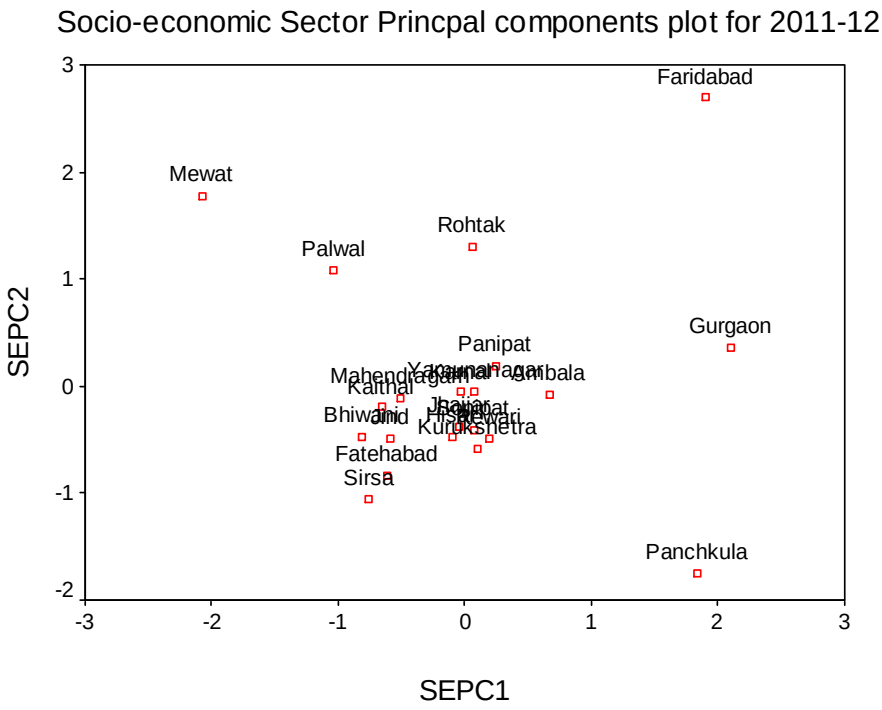

Figure 2 (Source: Own research)

3.3Association between Socio-Economic and Agricultural Dimensions

Canonical correlation analysis is used to find linear combinations of the variables in the sets of variables having maximum correlation. These combinations are the first coordinates in the new system and represent principal dimensions from the two sets of variables. Second pair of linear combinations is then obtained such that it has maximum correlation and is uncorrelated with the first linear combination. The procedure is continued until two new coordinate systems are specified completely. The first two principal components of agriculture and socio-economic sectors explained 75.45, 61.78 and 71.76 percent of the total variation in the data for the periods 1991-92, 2001-02 and 2011-12 respectively. Therefore, canonical correlation analysis [12] has been performed using first two principal components from each of agriculture and socio-economic sectors to examine the association between the principal dimensions of agriculture and socio-economic sectors. Period-wise canonical correlations, canonical loadings and cross-loadings based on the first and second principal components for Agriculture and Socio-Economic sectors have been summarized in Table-4. The p-values in the Table-4 indicate that both first and second canonical variate pairs are significantly correlated for 1991-92 data, whereas only first canonical variate pair has significant canonical correlations for the period 2001-02 and 2011-12 with respective canonical correlations 0.785 and 0.701.

The first canonical correlation for 1991-92 is 0.843 and the corresponding canonical variates are:

$$U_1 = -0.135AGPC1 + 0.991AGPC2 \quad (9)$$

And

$$V_1 = -0.869SEPC1 + 0.495SEPC2 \quad (10)$$

These variates are mainly determined by the dimension represented by second principal component for agriculture sector and first component that of socio-economic sector indicators. The second canonical variate pair for this period is:

$$U_2 = 0.991AGPC1 + 0.135AGPC2 \quad (11)$$

And

$$V_2 = 0.495SEPC1 + 0.869SEPC2 \quad (12)$$

In this case, the variable U_1 is mainly determined by the first PC of agriculture sector with loading 0.991 while U_2 has high loadings for both PCs of scio-economic sector with loadings 0.495 and 0.869.A similar loading pattern has also been observed for 2001-02 and 2011-12 for the first canonical variate pair. Loading pattern indicate that the principal dimension represented by U_1 is dominated by the second principal component of the agriculture sector indicators, whereas, the principal dimension represented by V_1 is dominated by the first PC of the socio-economic sector indicators. The most important indicators for AGPC2 for the period 1991-92 are AG11 (percentage of agriculture workforce to total workforce), AG13 (average annual rainfall), AG5 (Gross value of agriculture output per capita (rural) at current prices), AG15 (Tube wells& pump sets/000ha of gross cropped area) and AG19 (Poultry per sq km). AGPC2 forms a contrast indicators represented by AG5 and AG11 with that of AG13, AG15 and AG19. The indicators represented by AG4, AG7 and AG16 have been observed to be the most influential variables in period 2001-02 whereas AG4, AG7 and AG17 in period 2011-12. The most important variable for the principal dimension (SEPC1) is SE5 (infant mortality rate) followed by SE7 (Number of vehicles on road/lakh population) and SE1 (main workers as percentage of total workers). These two dimensions AGPC2 and SEPC1 establish a strong association between the agriculture and the socio-economic sectors. The significant canonical correlation between the vectors represented by first two PCs of agricultural and socio-economic sectors suggest that developments in socio-economic sectors and agricultural sectors go together. That is socio-economic development in Haryana can be achieved through development in agriculture.

Conclusion: The study indicates that there are interdistricts disparities in Haryana with reference to various dimensions of agriculture and socio-economic sectors. The indicators used in the study showed that some districts are highly developed in agriculture sector and some are highly developed in socio economic sectors. First two PC plot provide an excellent view of agricultural and social sector disparities which resembles the real picture of districts in reference to agricultural and social sector in Haryana. The results of the agriculrual dimension indicate that the districts Karnal, Kurukshetra, Panipat and Yamuna nagar have high principal component scores value for first and second principal component for the agriculture sector while, the districts Mewat, Mahendragarh and Bhiwani have low scores for both of the principal components of agricultural sector indicators. Similarly, socio-economic dimensions indicated that Faridabad district is the most developed district having high scores on both the principal components while Gurgaon and Panchkula districts have high scores on first principal component of the socio-economic sector only. On the other hand districts Mewat, Palwal, Bhiwani and Sirsa have low principal component scores and having classified as less developed districts on the socio-economic front. Canonical analysis of first two PCs from each of the agriculture and socio-economic sector indicated that the dimensions represented by the second principal component of agriculture sector and first principal component of socio-economic sector established a strong association between the agriculture and the socio-economic sectors.

References

1. Hooda E, Hooda BK, Manocha V. Dynamics of inter-district developmental disparities in Haryana. Journal of Applied and Natural Science. 2017;9(2): 983 – 991.
2. Sheoran OP. Structural Equation Modelling with Latent Variables for Assessment of Regional Development in Haryana. Unpublished PhD thesis, CCS HAU Hisar (Haryana);2010.
3. Shafiqullah. Impact of Regional Disparities on Agricultural Development in Uttar Pradesh - A Geographical Analysis. Global Journa Human Social Science. 2013;13(5):36-46.
4. Jena D. Agricultural Development Disparities in Odisha: A Statistical Study. American Review of Mathematics and Statistics. 2014;2(1): 45-53.
5. Salvati L, Zitti M, Carlucci M. Territorial systems, Regional disparities and Sustainability: Economic structure and soil degradation in Italy. Sustainability. 2014;6:3086-3104.

6. Tanwar N, Tyagi S, Sisodia BVS, Hooda BK. Dynamics of socio-economic development of districts of eastern Uttar Pradesh. Journal of Applied and Natural Science. 2016;8(1): 5.
7. Narain P, Rai SC, Shanti S. Statistical Evaluation of Development on Socio-economic Front. Journal of Indian Society of Agricultural Statistics. 1991;43:329-345.
8. Tripathi RK, Tanwar N. Inter-District Disparities in the Socio-Economic Development of HimachalPradesh. Indian Journal of Economics and Development. 2017;13:13-18.
9. Sugiayama T. and Takeda A. On the characterization of principal canonical correlation analysis,Bulletin of the International Statistical Institute, 52nd Session, Finland. 1999.
10. Anderson TW. An introduction to multivariate analysis, 3rd edition, John Wiley & Sons, New York;2003.
11. Bartlett MS. The effect of standardization on an approximation in factor analysis. Biometrika. 1951;38: 337-344.
12. Sugiyama T, Ogura T, Sakaori F, Yamada T. Distributions and the bootstrap method of some statistics in principal canonical correlation analysis. Journal of Japan Statistical Society. 2007;37(2), 239-251.

13. Table-1:-Principal component analysis of Agriculture and Socio-Economic Sector indicators for periods 1991-92, 2001-02 and 2011-12.

PC	**Period**								
	1991-92			2001-02			2011-12		
	Eigen value	Variation (%)	Cum. Variation (%)	Eigen value	Variation (%)	Cum. Variation (%)	Eigen value	Variation (%)	Cum. Variation (%)
	Agriculture Sector								
1	7.99	42.07	42.07	5.45	28.71	28.71	5.32	28.01	28.01
2	3.32	17.48	59.55	4.60	24.24	52.96	3.96	20.86	48.88
3	2.35	12.38	71.93	2.27	11.95	64.92	2.68	14.11	62.99
4	1.42	7.51	79.44	1.75	9.23	74.15	1.72	9.07	72.07
5	1.12	5.93	85.38	1.39	7.36	81.51	1.46	7.70	79.78
6	.94	4.98	90.36	1.14	6.02	87.54	.94	4.95	84.73
	Socio-Economic Sector								
1	3.89	43.22	43.22	3.83	42.63	42.63	5.09	56.57	56.57
2	2.90	32.23	75.45	1.72	19.15	61.78	1.36	15.19	71.76
3	.94	10.44	85.90	1.34	14.95	76.74	.90	10.05	81.82

Table-2:- Loadings for PCs' of Agriculture sector indicators for periods1991-92, 2001-02 and 2011-12

Indicator	Components Loading for Agricultural Indicators														
	Period-II(1991-92)					Period-II(2001-02)					Period-II(2011-12)				
	PC1	PC2	PC3	PC4	PC5	PC1	PC2	PC3	PC4	PC5	PC1	PC2	PC3	PC4	PC5
AG1	0.70	-0.38	-0.26	-0.41	0.26	0.57	-0.61	-0.32	0.19	-0.11	0.46	0.69	0.24	-0.01	-0.26
AG2	0.74	0.06	0.42	-0.24	-0.27	0.53	0.31	0.04	-0.08	-0.64	0.46	-0.13	0.10	0.76	-0.07
AG3	-0.73	0.46	0.20	0.35	-0.22	-0.54	0.65	0.29	-0.14	0.19	-0.45	-0.76	-0.27	-0.22	0.22
AG4	0.93	0.14	0.09	0.17	-0.19	0.95	-0.22	-0.03	0.05	0.08	0.82	0.30	0.09	0.22	0.10
AG5	0.60	0.63	0.19	0.04	-0.23	0.76	0.52	0.22	0.02	0.19	0.65	-0.56	-0.17	0.15	0.33
AG6	0.76	0.12	0.14	0.11	0.44	0.13	0.29	0.15	-0.26	0.67	0.68	-0.32	-0.05	-0.47	-0.26
AG7	0.78	0.37	0.17	0.35	0.10	0.84	0.21	0.24	-0.15	0.07	0.84	-0.14	0.09	0.06	0.42
AG8	0.75	0.01	-0.34	0.13	-0.26	0.00	-0.11	-0.14	0.74	0.25	-0.16	-0.26	0.73	0.02	0.13
AG9	0.51	0.05	-0.61	-0.28	0.31	-0.24	0.56	-0.59	-0.17	-0.18	-0.21	-0.05	0.63	0.40	0.39
AG10	0.08	0.23	0.80	-0.20	0.35	0.55	0.23	0.39	-0.32	-0.36	0.69	-0.29	-0.46	-0.11	-0.04
AG11	0.17	0.85	-0.04	-0.06	0.37	-0.08	0.85	-0.24	0.36	0.14	0.15	-0.81	-0.23	0.04	-0.33
AG12	0.76	0.41	0.03	-0.31	-0.30	0.53	0.43	0.03	0.44	-0.26	0.53	-0.21	-0.24	0.40	-0.25
AG13	0.14	-0.73	0.53	0.10	0.10	-0.14	-0.77	0.46	0.19	0.09	0.15	0.61	-0.37	-0.11	-0.01
AG14	0.61	-0.25	-0.03	0.57	0.28	0.45	-0.38	-0.33	-0.46	0.20	-0.34	0.66	-0.41	0.32	0.21
AG15	0.49	-0.64	-0.35	0.33	-0.12	0.49	-0.52	-0.50	-0.05	0.14	0.24	0.55	0.38	-0.50	0.15
AG16	0.81	0.10	-0.02	0.36	0.02	0.89	-0.28	-0.10	-0.04	0.20	0.77	0.11	0.13	-0.18	0.06
AG17	0.82	-0.11	0.22	-0.14	-0.11	0.47	0.41	0.65	0.05	0.14	0.81	0.13	-0.02	-0.21	0.40
AG18	0.75	-0.17	-0.16	-0.26	-0.15	0.46	0.57	-0.29	0.46	0.09	0.39	0.10	0.57	0.09	-0.64
AG19	0.23	-0.63	0.62	-0.11	-0.06	-0.14	-0.65	0.55	0.38	-0.13	0.05	0.03	0.61	-0.68	0.18

Table-3Component Loadings for PCs' of Socio- Economic Indicators for periods 1991-92, 2001-02 and 2011-12

Indicator	Components Loading for Socio- Economic Indicators								
	Period-I(1991-92)			Period-II(2001-02)			Period-III(2011-12)		
	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
SE1	0.78	0.49	0.01	0.65	0.45	-0.26	0.88	0.21	-0.15
SE2	-0.02	0.85	-0.06	-0.39	0.69	0.17	0.63	-0.57	0.24
SE3	0.72	-0.57	0.28	0.75	-0.21	0.51	0.83	-0.14	0.16
SE4	0.85	-0.29	0.38	0.78	0.01	0.49	0.87	-0.16	0.20
SE5	0.76	0.56	0.02	0.75	0.39	-0.31	0.94	0.04	0.03
SE6	-0.04	0.70	0.66	-0.38	0.83	0.24	0.60	-0.36	-0.04
SE7	0.86	0.24	-0.28	0.81	0.36	-0.06	0.91	0.32	0.04
SE8	0.82	-0.19	-0.40	0.84	-0.18	-0.36	0.60	0.71	-0.27
SE9	0.20	-0.79	0.21	0.19	-0.02	0.69	-0.25	0.47	0.83

Table-4 Principal Canonical Correlation Analysis of agriculture and socio-economic sectorsfor periods 1991-92, 2001-02 and 2011-12

Variables	Period					
	1991-92		2001-02		2011-12	
Agriculture Sector	U_1	U_2	U_1	U_2	U_1	U_2
AGPC1	-0.135	0.991	0.028	1.000	-0.151	0.989
AGPC2	0.991	0.135	-1.000	0.028	0.989	0.151
Socio-Economic Sector	V_1	V_2	V_1	V_2	V_1	V_2
SEPC1	-0.869	0.495	0.998	-0.067	0.993	0.117
SEPC2	0.495	0.869	0.067	0.998	0.117	-0.993
Canonical correlation	0.843**	0.580*	0.785**	0.141	0.701*	0.039
p-value	< 0.001	0.023	0.004	0.577	0.019	0.870

*: significant at 5% level; **: significant at 1% level.