

Original Research Article

<https://doi.org/10.20546/ijcmas.2019.801.335>

## Food Grain Production Index Using Principal Component Analysis in Karnataka State, India

N.L. Pavithra\*, K.V. Ashalatha, J. Megha, G.R. Manjunath and Siddu Hanabar

*Department of Agricultural Statistics, College of Agriculture, UAS, Dharwad-580005, Karnataka, India*

*\*Corresponding author*

### ABSTRACT

Karnataka State has a typical composition having a large share of its area under highly diversified agricultural crops, higher growth in agriculture assumes great importance and is a matter of concern for policy planners and research scholars in recent times. A composite index is a grouping of equities, indexes or other factors combined in a standardized way, providing a useful statistical measure of overall food performance over time, and it is also known simply as a "composite". Usually, a composite index has a large number of factors that are averaged together to form a product representative of an overall food sector. Indicators are useful for determining trends and drawing conclusions for particular issues in policy analysis. They can also be helpful in making policy and in monitoring performance. When several indicators are compiled into a single index using a specific technique, then a composite indicator is formed. The composite indicator measures multi-dimensional concepts, which cannot be explained by a single indicator. Here, Food grain Production Index (FgPI) has been constructed using Principal Component Analysis (PCA) for 30 districts of Karnataka, India. In present study, the indicators like production of tur, production of paddy, production of total pulses and production of total cereals have been taken.

#### Keywords

Composite Indicator,  
Food grain production  
index, Principal  
component analysis

#### Article Info

**Accepted:**  
26 December 2018  
**Available Online:**  
10 January 2019

### Introduction

Karnataka State has a typical composition of having a large share of its area under severe climatic constraints with a highly diversified agricultural sector. Karnataka is the largest producer of coarse cereals, coffee, raw silk and tomatoes among the states in India. The main crops grown here are rice, ragi, jowar, maize, and pulses (Tur and gram) besides oilseeds and number of cash crops. The state of Karnataka is blessed with varied agro-

climatic conditions which permits the farmers of the state to cultivate not only a variety of crops in a season but also a number of crops like cereals, pulses, oilseeds, commercial crops and horticultural crops across different seasons of the year. Agriculture in Karnataka has occupied around 12.31 million hectares of land, this comes to 64.6 per cent of the total area. The state is one of the major producers of paddy among all other states in India. Karnataka has large rainfed areas next only to Rajasthan as the future of agriculture growth

in the state depends on this factor which accounts for more than 75 per cent of the cropped area. The share of agriculture in the state GDP is around 16 per cent which is higher than the current national average of all the states in India. Karnataka is the state to come up with a separate agriculture budget. In this paper, Foodgrain Production Index (FgPI) has been constructed using PCA with production indicator for 30 districts of Karnataka State, India.

**Materials and Methods**

The data on food grain production has been taken from the secondary source (Directorate of Economics and Statistics, Govt. of Karnataka, Bangalore and Karnataka at a Glance: Government of Karnataka) for all the 30 districts of Karnataka for the year 2014-15. District wise data on production of paddy, production of pigeonpea, production of total pulses and production of total cereals has been analyzed. 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 (or sometimes, principal modes of variation).

The number of principal components is less than or equal to the smaller of the number of original variables or the number of observations. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables.

Principal component analysis was used for data reduction technique as well as for the solution of multicollinearity. Principal component analysis was employed, with a view to aggregate the performance indicators into a few groups of factors. This technique was used by many researchers for grouping the factors and is the oldest and the best known technique of multivariate analysis. The principal components will be utilized in the construction of composite index (Kumar *et al.*, 2013).

Maximum Likelihood Estimate (M.L.E.) of variance-covariance matrix ( $\Sigma$ ) of the given data set will be estimated by

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' \quad \dots (i)$$

Where,

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_q \end{bmatrix},$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

And n is total number of districts.

Then Correlation Matrix (CM) was obtained using above variance-covariance matrix as

$$CM = (\sqrt{V})^{-1} \hat{\Sigma} (\sqrt{V})^{-1} \quad (ii)$$

Where,

V = Diagonal matrix obtained from variance-covariance matrix and

$\hat{\Sigma}$  = M. L.E. of variance-covariance matrix.

Let  $x_1, x_2, \dots$ , are variables under study, then first principal component may be defined as

$$Z_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p$$

Such that variance of  $z_1$  is as large as possible subject to the condition that

$$a_{11}^2 + a_{12}^2 + \dots + a_{1p}^2 = 1$$

This constraint is introduced because if this is not done, then  $\text{Var}(z_1)$  could be increased simply by multiplying any  $a_{1j}$ 's by a constant factor. The second principal component is defined as

$$Z_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p$$

Such that  $\text{Var}(z_2)$  is as large as possible next to  $\text{Var}(z_1)$  subject to the constraint that

$$a_{12}^2 + a_{22}^2 + \dots + a_{2k}^2 = 1 \text{ and } \text{cov}(z_1, z_2) = 0 \text{ and so on}$$

It is quite likely that first few principal components account for most of the variability in the original data. If so, these few principal components can then replace the initial  $p$  variables in subsequent analysis, thus, reducing the effective dimensionality of the problem.

It is a mathematical technique, which does not require user to specify the statistical model or assumption about distribution of original varieties. It may also be mentioned that principal components are artificial variables and often it is not possible to assign physical meaning to them. Next step is to obtain principal components using eigen vectors of the estimated correlation matrix and standardized values of variables. The principal components will be obtained by using the formula given below.

$$P_1 = a_{11}Z_1 + a_{12}Z_2 + \dots + a_{1Q}Z_Q$$

$$P_2 = a_{21}Z_1 + a_{22}Z_2 + \dots + a_{2Q}Z_Q$$

·  
·  
·

$$P_Q = a_{Q1}Z_1 + a_{Q2}Z_2 + \dots + a_{QQ}Z_Q$$

Where,

$P_{qs}$ :  $q^{\text{th}}$  principal components

$Z_{qs}$ : standardized values of  $q^{\text{th}}$  variable

$a_{kq}$ : element belonging to  $k^{\text{th}}$  eigenvector and for  $q^{\text{th}}$  variable,

$$k = 1, 2, \dots, Q; q = 1, 2, \dots, Q.$$

Now, the composite index will be constructed using the obtained eigen values of variables and principal components as under:

$$CI_i = \frac{\lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_Q P_Q}{\sum_{j=1}^Q \lambda_j} \quad \text{(iii)}$$

Where,

$CI_i$  = Composite index for  $i^{\text{th}}$  district

$\lambda_j$ s are eigen values

$P_{qs}$  =  $q^{\text{th}}$  principal components,  $i=1, 2, \dots, N$ ;  $j=1, 2, \dots, Q$ .

Further, the composite index of each district will be normalized by using the following formula:

$$CI_{Ni} = \frac{CI_i - \min(CI)}{\max(CI) - \min(CI)} \quad \text{(iv)}$$

Where,

$CI_{Ni}$  = Normalized value of composite index of  $i^{th}$  district

min (CI) = Minimum value of composite index

max (CI) = Maximum value of composite index among all.

### Variance Inflation Factor

Regression analysis was performed and Variance Inflation Factor (VIF) for each variable was obtained to detect multicollinearity by regressing one variable to other remaining variables. The Variance Inflation Factor for  $j^{th}$  variable can be obtained as under

$$VIF_j = \frac{1}{1 - R_j^2}$$

Where,

$VIF_j$  is Variance Inflation Factor for  $j^{th}$  variable and

Coefficients of determination ( $R_j^2$ ) will be obtained by regressing  $j^{th}$  variable on other variable(s).

### Results and Discussion

Correlation coefficient between production of food grains in Karnataka state were calculated which is presented in Table 1 in the form of correlation matrix.

From the Table 1 it is clear that the total cereals were highly significant and positively correlated with paddy and pigeonpea was significant and positively correlated with total pulses for food grain production. Paddy was negatively correlated and non-significant with

pigeonpea and total pulses and other variables were non-significant. An attempt on correlating the paddy was highly significant and positively correlated with total cereals *i.e.* correlation coefficient ( $r$ ) was found out to be 1.00 (positively correlated) indicating that the increase in production of paddy results in significant increase in production of food grains. Paddy is one of the major contributing to totals cereals so there was a perfect correlation between paddy and total cereals. Pigeonpea was significant and positively correlated with total pulses for food grain production with a correlation coefficient  $r=0.969$ . Other variables were non-significant.

In statistics, multicollinearity is a phenomenon in which two or more predictor variable in the regression model was highly correlated and it affects calculations regarding individual predictors. Variance Inflation Factor (VIF) above 5 indicates multicollinearity.

The results of regression analysis along with VIFs are given in Table 2. It was concluded that the linear relationship among variables were highly significant. It was also concluded that the variable total cereals had serious multicollinearity as VIF for total cereals were greater than 10.

The variables paddy, pigeonpea and total pulses showed very little multicollinearity. Overall, it was concluded that there was multicollinearity among variables. Thus, the composite index was constructed using principal component analysis to overcome the problem of multicollinearity.

The composite indices of food grains production was worked out for districts of Karnataka state for the study period of 1990 to 2015. The districts were ranked on the basis of composite indices. The composite indices for food grains production along with the rank of the districts were presented in the Table 3.

**Table.1** Correlation between food grains

Crops	Paddy	Total cereals	Pigeonpea	Total pulse
Paddy	1.000	1.000**	-0.200	-0.121
Total cereals	1.000**	1.000	-0.200	-0.121
Pigeonpea	-0.200	-0.200	1.000	0.969**
Total pulse	-0.121	-0.121	0.969**	1.000

Note: \*\* significant at 1 %

**Table.2** Detection of multicollinearity

Model	Dependent variable	Independent variable	Significant value	R <sup>2</sup> value	VIF
1	Paddy	Total cereals, pigeonpea and total pulses	<0.0001	1.000	1.143
2	Total cereals	Pigeonpea, total pulses and paddy	<0.0001	1.000	18.603
3	Pigeonpea	Total pulses, total cereals and paddy	<0.0001	0.973	1.015
4	Total pulses	Pigeonpea, total cereals and paddy	<0.0001	0.972	1.042

**Table.3** Values of Composite Indices (C.I) of food grains production for different districts along with the ranks

Districts	CI	Rank
Raichur	1.00	1
Davangere	0.84	2
Bellary	0.77	3
Shivamoga	0.73	4
Mysore	0.68	5
Koppal	0.57	6
Mandya	0.54	7
Uttara Kannada	0.43	8
Hassan	0.42	9
Dakshina Kannada	0.41	10
Udupi	0.39	11
Belagavi	0.39	12
Haveri	0.38	13
Yadgiri	0.37	14
Chickmagalur	0.36	15
Kodagu	0.33	16
Chamarajanagara	0.28	17
Tumkur	0.26	18
Dharwad	0.23	19
Ramanagara	0.23	20
Chickballapur	0.22	21
Bangalore Urban	0.22	22
Bangalore Rural	0.21	23
Chitradurga	0.21	24
Gadag	0.21	25
Kolar	0.21	26
Bagalkot	0.20	27
Bidar	0.18	28
Vijayapura	0.11	29
Kalaburgi	0.00	30

During the study period, indicators based on production of food grains, the district Raichur was on the first position and the district Kalaburgi was on the last position. The composite indices varied from 0.0000 to 1.000. Four most food grains producing districts were Raichur, Davangere, Bellary and Shivamoga. Four least food grains producing districts were Bagalkot, Bidar, Vijayapura and Kalaburgi.

## References

- Anonymous, 2015, Annu. Rep. (2014-15). Department of Planning, Programme Monitoring and Statistics, Karnataka, p. 849.
- Kannan, E., Kumar, P., Vishnu, K. and Abraham, H., 2013, Assessment of pre and post-harvest losses of rice and red gram in Karnataka. *Res. Rep.*, pp. 46-52.
- Kumar, M., Ahmad, T., Rai, A. and Sahoo, P. M., 2013, Methodology for construction of composite Index. *Int. J. Agric. Stat. Sci.*, 9(2): 639-647.
- Nethravathi, A. P. and Yeledhalli, R. A., 2016, Growth and instability in area, production and productivity of different crops in Bengaluru division. *Int. J. Agric. Environ. Biotechnol.*, 9(4): 599-611.
- Prem, N., Sharma, S. D., Rai, S. C. and Bhatia, V. K., 2009, Inter-district variation of socio-economic development in Andhra Pradesh. *J. Indian Soc. Agric. Stat.*, 63(1): 35-42.
- Pushpa, M. S., 2007, An econometric analysis of demand and supply response of pulses in India. *Karnataka J. Agric. Sci.*, 20(3): 545-550.
- Sarawathi, P. A., Basavaraja, H., Kunnal, L. B., Mahajanshetti, S. B. and Bhat, A. R. S., 2012, Growth in area production and productivity of major crops in Karnataka. *Karnataka J. Agric. Sci.*, 25(4): 431-436.

## How to cite this article:

Pavithra, N.L., K.V. Ashalatha, J. Megha, G.R. Manjunath and Siddu Hanabar. 2019. Food Grain Production Index Using Principal Component Analysis in Karnataka State, India. *Int.J.Curr.Microbiol.App.Sci.* 8(01): 3138-3143. doi: <https://doi.org/10.20546/ijcmas.2019.801.335>