Journal of Scientific Research and Reports

Volume 30, Issue 12, Page 350-360, 2024; Article no.JSRR.128027
ISSN: 2320-0227

Classification of Mango Farmers in Malda District for better Developmental Policies: A Multivariate Clustering Approach

Bankim Sarkar [a], Mrittika Das [a], P K Sahu [a], Abdullah Mohammad Ghazi Al khatib [b] and Pradeep Mishra [c*]

[a] Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia, West Bengal, India.
[b] Department of Finance and Banking, Faculty of Economics, Damascus University, Damascus, Syrian Arab Republic.
[c] College of Agriculture, Rewa, JNKVV, India.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/jsrr/2024/v30i122679

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/128027

Original Research Article

Received: 08/10/2024
Accepted: 11/12/2024
Published: 16/12/2024

ABSTRACT

Mango is a bulky and nutritious fruit crop and it had high demand for it's delicious taste, soothing aroma, varietal diversity and wide adoptability to meet food and nutritional security. India is the largest producer of mango among that West Bengal is one of the contributors in the Indian mango

**Corresponding author: E-mail: pradeepjnkvv@gmail.com;*

Cite as: *Sarkar, Bankim, Mrittika Das, P K Sahu, Abdullah Mohammad Ghazi Al khatib, and Pradeep Mishra. 2024. "Classification of Mango Farmers in Malda District for Better Developmental Policies: A Multivariate Clustering Approach". Journal of Scientific Research and Reports 30 (12):350-60. https://doi.org/10.9734/jsrr/2024/v30i122679.*

basket. Malda district alone contributes more than 50 per cent of the mango production of the state from thirty one thousand hectares out of total eighty one thousand hectare mango area in West Bengal. In the present study an attempt has been made to critically analyse the status of mango farming by taking a random sample of 110 mango farmers from five selected blocks of Malda district adopting two stage sampling method with a specific objective to group the farmers in homogenous clusters to facilitate the developmental policy for different group of mango farmers. Among the parameters under study operational cost per unit of area under mango orchard is contributing maximum 26.67 per cent in cluster discrimination using Euclidian distance measure. Salient feature of the clustering analysis is the formation of five different groups by five farmers, thereby indicating the need for in depth study of these farmers so as to unearth the speciality of forming separate groups by these farmers. The present study thus, recommends for formation of cluster wise developmental policies for the betterment of the age-old mango farming in the district.

Keywords: Clustering; Euclidian distance; contribution; two stage sampling.

1. INTRODUCTION

Being an agrarian economy, India depends on agriculture and allied sectors for its development. In addition to food production, it is the most important source of income especially in the vast rural areas of the country. Despite of being sustainable in terms of food production there lack of self-sufficiency in providing nutritional security to every citizen of the nation. It is a matter of grave concern that one third of the world's malnourished child live in India (Global nutrition report,2018). Horticultural fruit crop ensures the nutritional security of the people to maintain good health. Mango is one of the most important fruits highly demanded for it's delicious taste, sweet fragrance, attracting colour, wide genetic variations and it's nutritional importance. India is the leading mango producing country across the world sharing 52 per cent of the global production whereas West Bengal ranks 8th, producing 4.25 per cent of the national production. Malda is the famous mango producing district of West Bengal, with 54 per cent share of the state's production. Around 4.25 lakh people are associated with the cultivation, business and marketing of mango and it's value-added products. The district has got G.I. tags on 3 different mango varieties viz. Lakhhonvog, Fazli, Khirsapati (www.ipiindia.gov.in,2022). Farmers are the provider of food grains as well as nutrition to the society and they keep us alive in this planet. They are the integral part of agri-based economy of a nation like India. If they remain in good condition, free from all problems in production, marketing, economic and social sectors, the entire population will be assured of food and nutritional security. To ensure their interest towards farming, it is a big challenge for the Government as well as for the agricultural scientists to take steps towards the age-old cultivation into a profitable enterprise. The first and most important condition for the sustenance of any agri-based economy is the continued upliftment of the situation of the farmers along with the augmentation of the production process. To have an overall idea about the present status of entire mango enterprise of the district and to evaluate it's sustainability, there is a tremendous need to study the current conditions of the growers. Sarkar et al*.* (2018) studied the mango growing community of Nadia district of West Bengal and reported the major problems of the farmers through ranking methods. Das et al. (2013) surveyed the mango growers of the English Bazaar block of Malda district and identified the socio-economic variables influencing the income, expenditure and yearly savings of the farmers. Lapple and Kelley (2013) studied decision of the farmers of Ireland towards adopting the organic method of cultivation techniques and found two distinct groups of the farmers one with moderate environmental consciousness and another one with high environmental conscious through cluster analysis and finally reported governmental incentives primarily focused on subsidy payments may not be enough to grow the organic industry. Nambi et al*.* (2015) classified the ripening period of mango in 5 different stages through hierarchical cluster analysis. Mandal et al .(2013) compared the use of clustering methods like k-means and k-mediod (PAM-partition around mediod) with various distance measures to categorise various mango cultivars according to their physical parameters and found K means provided maximum efficiency in case of manhattan distance where-as PAM in case of correlation distance. Thus, there is need to identify the important contributing factors/variables, their importance and grouping/clustering of cultivators into different groups so that appropriate policy measures and

developmental strategies could be taken up to uplift the mango cultivation and related activities in mango business. As such the present study was undertaken with an aim to find out optimum no. of homogenous farmers group based on their production and economic performance through applications of multivariate clustering technique and to throw light so that policy related issues could be considered by the policy makers to take some initiatives to uplift the mango production status of the district and run the entire mango-based economy of the district more smoothly to pull an international demand of the famous mangoes of the district.

2. MATERIALS AND METHODS

2.1 Materials

Selection of Location and Data Collection: Malda district of West Bengal has been chosen purposively for the present study. At first secondary data on area, production and productivity are collected from district horticultural office (www.ipiindia.gov.in).A two-stage sampling scheme is designed to reach farmers. At the first stage 5 blocks are selected using the sampling technique probability proportional to size without replacement (PPSWOR) and at the second stage 168 farmers were selected randomly using simple random sampling without replacement (SRSWOR) sampling scheme (Mukhopadhyay, 2022). Out of total 15 blocks in Malda district, 5 blocks viz. English Bazar, Ratua-2, Manikchak, Kaliachak-2, and Old Malda are selected on the basis of area data under mango cultivation for the year 2022 (Figs. 1 and 2). Thereafter, 168 farmers are selected from the 5 blocks through SRSWOR. From these selected168 farmers, 110 farmers are sorted, leaving aside the farmers (58 nos), who have leased out their orchards. An interview schedule was prepared for door-to-door survey to note down the responses from the farmers (Annexure 1) during the month of February to March, 2023 with the cooperation from Assistant Director of Agriculture (ADA), Department of Agriculture, Government of West Bengal, of the respective block. Most of the mango varieties are being characterised by alternate bearing habit; the growers have to face an on year and off year pattern in the production that truly impacts the whole mango-based economy. That is why any research on mango cultivation/economy needs to take care of at least two consecutive years data. As such this study is based on primary data of 2021 and 2022, two consecutive on and off year.

Fig. 1. Graphical view of the selectedblocks of Malda district

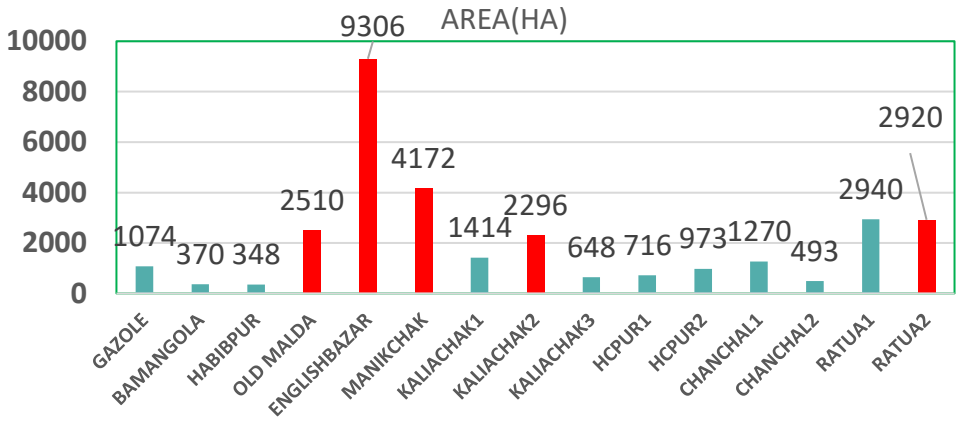


Fig. 2. Blockwise area under mango orchard in Malda District
Source- District Horticulture Office, Malda(2022)

2.2 Methods

An interview schedule is essentially a list containing a set of structured questions that have been created to serve as a guide for interviewers, researchers, and investigators in gathering information or data about a given topic or issue.

Among the analytical tools, different descriptive measures have been used. Descriptive statistics are the summarized form of the characteristic of a variable. It actually shows the data in a nutshell. It has a large difference from the inferential statistics because inferential statistics goes out of the data set.The different descriptive statistics used are-Arithmetic mean, Variance, $\sigma_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$, Standard Error of Mean, $S.E(Mean) = \frac{\sigma_X}{\sqrt{n}}$, Coefficient of variation, $CV = \frac{\sigma_X}{\bar{X}} \times 100$, Skewness measuring the lack of symmetry in frequency distribution curve of a variable and measured as $\gamma_1 = \sqrt{\beta_1} = \sqrt{\frac{\mu_3^2}{\mu_2^3}}$ where and μ_2 and μ_3 are 2nd and 3rd central moments respectively, kurtosis, peakiness or the flatness of the frequency distribution curve, measured in terms of $\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3$ where μ_4 and μ_2 are 4th and 2nd central moments respectively.

2.2.1 Cluster analysis

Cluster analysis is basically an exploratory multivariate data analysis technique where we try to group heterogenous multidimensional objects into homogenous groups using several distance measures. It is of two types.

a. Hierarchical Cluster Analysis
b. Non-Hierarchical Cluster Analysis

In hierarchical case the clusters have an inherent hierarchical structure in the formation and interpretation of result and visualized through a tree structure called 'Dendrogram'. In non-hierarchical cluster analysis, we do not have any hierarchy. If the no. of objects is large non-hierarchical cluster analysis is generally used.

In our study we have used K means non-hierarchical clustering technique. K means non-hierarchical clustering technique is accomplished in the following steps, 1) partitioning the items into k initial clusters randomly, 2) to calculate the centroids of each randomly chosen k clusters, 3) to reassign the items to the cluster whose centroid is nearest in distance, 4) repetition of step 3 until no more reassignment is required. To decide the optimum number of clusters, Elbow method is used. It is a method of drawing the curve of explained variance in Y axis and number of clusters in X axis. Each increase in k will add some gain in information and definitely the explained variance gets increased. We have to find that particular k after which the curve reaches towards plateau or that k corresponding to which an angle or elbow is created in the curve. The distance measure used is **Euclidean**

distance, given as $E_{ij}=\sqrt{\sum_{k=1}^{k}\left(X_{ik}-X_{jk}\right)^{2}}$ i and jare the two items for which distance is measured based on the k characters under consideration. Contribution of each and individual kcharacters towards clustering is measured by the individual share in squared Euclidean distances among all possible combinations of pair of clusters and the characteristics are ranked according to their contribution in measuring the squared Euclidean distances. Rank-1 is given to the highest mean difference and rank p is allotted to the lowest mean difference. Thereafter how many numbers of times each character is appearing in rank 1 among all possible combinations are recorded. The character appearing maximum no of times in rank 1 is considered as the most contributing character and the character appearing minimum no. of times in rank 1 is considered as the least contributing factor towards clustering (Singh and Chaudhury,2012).

3. RESULTS AND DISCUSSIONS

The study aims to group the farmers so that group sensitive action plan could be adopted for the development and progress of mango cultivation in Malda district and for the purpose multivariate clustering techniques is adopted with most important parameters of mango production system. In this study the number of variety of mangoes grown by the individual farmers (X_1), maximum (X_2) and minimum(X_3) age of the mango plants, plant density (X_4), average production per unit of land (X_5), operational cost per unit of land (X_6), average price per quintal of mango (X_7) and average return per unit of land (X_8) are used for cluster analysis using Euclidean distance measure and k-means cluster amalgamation technique. In k-means clustering technique the number of clusters have been decided by maximizing the ratio of between cluster sum of square to that of the total sum of squares using Elbow method as discussed already (Mukhopadhyay, 2022).

Before taking up the cluster analysis, it is pertinent to study the overall nature of the characters under consideration over the 110 farmers (Table 1 & Fig. 3) provides the description of the characters. Perusal of the table reveals that on an average, farmers are growing more than five varieties with a range of one to twelve having CV 41%. Among the 110 farmers minimum and maximum age of plantation varies between two years to hundred and fifty years, which clearly reflects the age-old tradition of mango cultivation in Malda district of West Bengal. Depending upon the nature of the varieties vis-à-vis the plant density, the production per unit area of mango orchard varies between 1.7 kilogram to 292.73 kilogram with an average of 43.89 kilogram per disimil associated with a CV of more than 96%. The study of Siddique and Sultana (2015) showed that the yield of mango of the district in 2010-11 was 7273 kg/ha which is equal to 29.38 kg/dismil. The present study proves that the yield of mango per unit area has increased significantly over the years. Wide variations in operational cost per disimil is revealed by the range of nil to Rs1485.The variations in price per quintal of mango is quite obvious because of varietal characteristics of mango fruit, so also the monetary return per unit of area under mango orchard. Thus, wide variations among the farmers with respect to all the characters justify the adoption of clustering approach to further study the mango growers of Malda district of West Bengal, India.

As has already been discussed, K means cluster analysis is taken up for 110 mango farmers with eight characters (Table 2). The cluster analysis has resulted in 13 clusters. The results of the cluster analysis have been presented in Table 2 to 5. Table 3 provides the inter-cluster distance among these 13 clusters. Maximum distance of 11.47 is recorded between the cluster 7 and cluster 13; whereas minimum of 2.01 between cluster 4 and cluster 9. Distance among the other clusters varies between 2.01 to 11.47. The distributions of farmers of different blocks in different clusters is presented in Table 4. From Table 4, it is found that maximum 41 number of farmers are allocated in cluster 5 followed by 31 in cluster 4, 10 in cluster 9, 7 in cluster 1, 6 in cluster 11, 2 each in cluster 6 and 12 and rest of the clusters are comprised of single farmer. If we consider the distributions of farmers of each block in different cluster group one can find that 30 farmers of Ratua-2 block are distributed over 9 clusters and 20 farmers of Manikchak block are distributed over 7 clusters. Similarly, 38 farmers of English bazar block and 9 farmers of Old Malda block are distributed in 6 clusters each. Thereby, maximum distribution quotient (number of clusters/number of farmers in the block) of 0.67 is recorded in Old Malda block against the minimum of 0.16 in English bazar block. Thus, the farmers of English bazar are more homogenous compared to the farmers of Old Malda. Though the study of Das and Mandal

(2020) revealed absence of significant difference among the blocks namely English Bazar, Ratua and Old Malda based on Index of performance measure, measured in terms of yield per acre, plant density, average age of plants and the various operational cost of production, the present study critically depicts the presence of block wise heterogeneity among farmers in their production and economical aspects which should be kept into consideration while framing the developmental policies for different blocks of the district (Table 4).

Character wise mean for different clusters is presented in Table 5. The maximum number of variety (8.4) is found in cluster 9 followed by 7 in cluster 10 whereas minimum of three varieties is found in cluster 7. So far about the average maximum and average minimum age of plants are concerned highest of 100 years are found in cluster 7. Though number of plants per unit area depends on the type of variety(ies) being grown by the farmers but in cluster 7 the minimum number of plants of 0.23 per dismal are grown by the farmers. So far about the average production per unit of land is concerned maximum is obtained in cluster 13 (292.73 kg per dismal) while the lowest 24.40 kg per dismal is found in cluster 5. Maximum operational cost of rupees 1484.85 is recorded in cluster 10 whereas minimum operational cost of rupees 136.36 found in cluster 3. Similarly, the farmers in group 2 have recorded maximum price of Rs2616.67 per quintal against the minimum of Rs 1422.46 in cluster 5. Maximum average return per dismalof land (Rs 4727.27) is recorded in cluster 13 while the minimum of rupees 316.84 is recorded in cluster 5. Thus, the 110 mango growers of Malda district are having some peculiarities as depicted fromi) the number of farmers in different clusters; out of 13 clusters, 5 are composed of single farmer in each case, ii) most importantly in cluster 7 two variables namely X_1 and X_4i.e., the no. of variety and no. of plant/dismal showing lowest values with only one farmer in the cluster, iii) similarly, cluster 13 is figuring the top most in two characters namely $X_5$and X_8 i.e., the average production and average return per unit of land with only one farmer in the group and so on. Thus, the cluster analysis reveals that these grouping of farmers in Malda district based on the above eight characters needs to be thoroughly studied before it is being utilized for further action plan.

It is also found from Table 5 that operational cost per dismil is the major characteristics that determines maximum (more than 26%) contribution towards clustering which reveals that operational cost is widely different from cluster to cluster. Operational cost is comprised of the expenditure due to the applications of fertilizer, hormones to deal with alternate bearing, insecticides, fungicides and nutritional supplements, irrigation and also for harvesting of mango and its marketing. Therefore, a sincere approach should be taken towards optimization of this operational cost while taking policies towards improvement of each and every cluster towards the augmentation of production system of the district.

Minimum age of the plant describes 22.22% contribution towards cluster differentiation. While maximum age of the plant contributes only 8.33% in clustering which reveals there exist more difference in minimum age of the plants between the 13 cluster of farmers than the maximum age of the plant. This result reveals that optimization of age of plants in orchards are needed. According to the report of National Horticulture Board,2023 (www.nhb.gov.in) the economic life of mango plant is6 to 35 years. Perusal of the Table 5 reveals that the mean of the maximum age of all the clusters is 39 years and some plants in almost 5 clusters (1^{st},6^{th},7^{th},9^{th} and 11^{th}) have crossed their economic life period. The study draws the attention towards replacement and rejuvenation of those aged orchards specially for the identified clusters.

The plant density and the productivity(production/dismil), two agronomic characters contribute 19.44% and 18.06% respectively in making 13 homogenous clusters of mango farmers. Substantial contribution of these two characters clearly indicates about need for optimization of the two parameters in mango gardens. The study of Das and Mandal (2020) showed the average plant density of the district as 39 plants/acre. But this study reveals an average plant density of 60 plants/acre of land Cluster 7 possess maximum aged orchard (100 years old) with minimum plant density/maximum spacing of 23 plant per acre and the cluster-3 possess minimum aged orchard (8 years) with maximum plant density/minimum spacing of 299 plant per acre. This result indicates the tendency of farmer to maximize yield at early age of the orchard by putting extra plant in same area of land. Therefore, optimization of planting density to enhance the performance of the farmers is needed.

Table 1. Description of the characters under study

	Minimum	Maximum	Mean	Std. Error	Std. Deviation	Skewness	Kurtosis	CV%
X_1	1.00	12.00	5.21	0.21	2.15	0.14	0.33	41.29
X_2	3.00	150.00	31.36	2.35	24.65	2.00	5.00	78.59
X_3	2.00	100.00	16.60	1.17	12.22	3.40	19.59	73.61
X_4	0.13	2.99	0.60	0.04	0.38	3.17	15.68	63.04
X_5	1.70	292.73	43.89	4.02	42.15	2.73	11.36	96.03
X_6	0.00	1484.85	344.91	22.91	240.29	1.83	6.06	69.67
X_7	450.00	3380.00	1799.92	54.28	569.28	0.42	-0.22	31.63
X_8	12.12	4727.27	794.57	79.66	835.48	2.14	5.39	105.15

Note(X_1-No of variety, X_2-Max age of the plant, X_3- Min age of the plant, X_4- No of plant/desimil, X_5-Avg production per desimil, X_6- Operational cost per desimil, X_7- Avg price per Q, X_8-Avg return per desimil)

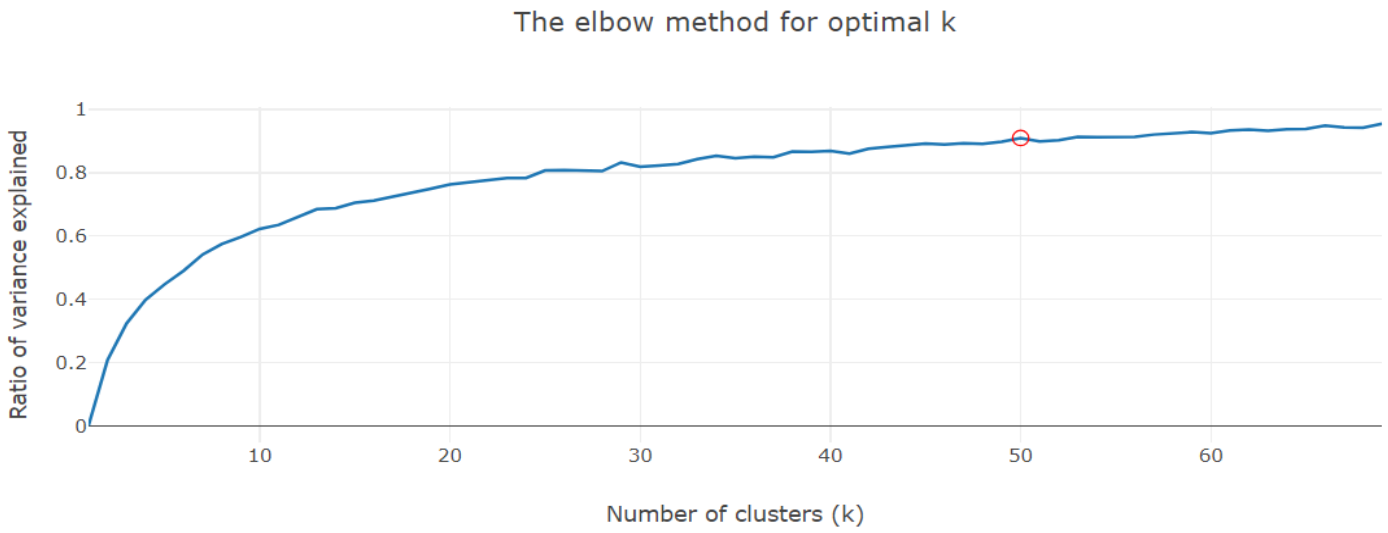


Fig. 3. Graphical presentation of the ratio of variance explained for different no. of clusters

Table 2. Change in variance explained due to change in number of clusters

No of cluster (k)	10	11	12	13	14	15	16
Variance explained	0.6307	0.6452	0.6652	0.6912	0.7024	0.7191	0.7261

Table 3. Cluster distances among the 13 clusters of mango farmers

Cluster	1	2	3	4	5	6	7	8	9	10	11	12	13
1		6.10	7.88	3.22	3.35	4.23	6.76	4.26	2.55	6.15	2.85	5.53	9.16
2	6.10		7.60	4.79	6.43	5.76	8.64	4.25	5.72	3.80	5.34	5.46	6.75
3	7.88	7.60		6.59	7.06	4.28	11.18	6.57	7.48	7.95	7.51	7.17	7.52
4	3.22	4.79	6.59		2.05	3.12	8.09	2.77	2.01	4.56	2.45	4.82	8.31
5	3.35	6.43	7.06	2.05		3.71	7.94	4.08	3.07	5.91	2.32	5.62	9.32
6	4.23	5.76	4.28	3.12	3.71		9.25	4.38	3.79	5.12	4.48	5.78	7.98
7	6.76	8.64	11.18	8.09	7.94	9.25		8.00	7.85	10.61	5.99	7.14	11.47
8	4.26	4.25	6.57	2.77	4.08	4.38	8.00		3.66	5.77	3.70	2.79	6.19
9	2.55	5.72	7.48	2.01	3.07	3.79	7.85	3.66		5.49	3.12	5.41	9.33
10	6.15	3.80	7.95	4.56	5.91	5.12	10.61	5.77	5.49		5.88	7.75	8.98
11	2.85	5.34	7.51	2.45	2.32	4.48	5.99	3.70	3.12	5.88		4.90	9.01
12	5.53	5.46	7.17	4.82	5.62	5.78	7.14	2.79	5.41	7.75	4.90		5.54
13	9.16	6.75	7.52	8.31	9.32	7.98	11.47	6.19	9.33	8.98	9.01	5.54	

Table 4. Block-cluster wise distribution of mango farmers

	English Bazar	Ratua-2	Manikchak	Kaliachak-2	Old-Malda	TOTAL
Cluster-1	3	2	2	0	0	7
Cluster-2	0	1	0	0	0	1
Cluster-3	0	0	1	0	0	1
Cluster-4	8	13	7	2	1	31
Cluster-5	18	5	6	8	4	41
Cluster-6	0	1	0	1	0	2
Cluster-7	0	0	0	0	1	1
Cluster-8	1	2	2	0	1	6
Cluster-9	5	4	1	0	0	10
Cluster10	0	1	0	0	0	1
Cluster-11	3	1	0	1	1	6
Cluster-12	0	0	1	0	1	2
Cluster-13	0	0	0	1	0	1
TOTAL	38	30	20	13	9	110
No. of cluster under each block	6	9	7	5	6	33
QUOTIENT	0.16	0.30	0.35	0.38	0.67	

Table 5. Average values of the characters over the clusters

Cluster	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	5.43	95.71	20.43	0.46	41.37	343.95	1559.55	684.74
2	6.00	30.00	30.00	1.11	95.96	1356.90	2616.67	2245.79
3	4.00	8.00	8.00	2.99	82.50	136.36	1875.00	1596.86
4	6.06	23.03	12.65	0.64	44.87	457.73	2028.91	842.42
5	3.90	18.76	14.80	0.47	24.40	200.57	1422.46	316.84
6	6.50	37.50	6.00	1.69	44.41	361.05	1386.61	671.62
7	3.00	100.00	100.00	0.23	57.27	204.55	2103.17	1136.36
8	5.50	29.00	16.17	0.68	110.26	428.69	2371.18	2646.29
9	8.40	57.00	13.05	0.49	27.42	311.04	2284.61	621.41
10	7.00	20.00	2.00	1.06	28.79	1484.85	1875.00	522.73

Cluster	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
11	3.33	40.00	35.00	0.43	38.15	388.30	1996.10	551.35
12	6.50	27.50	40.00	0.77	168.29	214.44	1978.57	3265.69
13	4.00	25.00	15.00	1.82	292.73	772.73	1607.14	4727.27
Average	5.36	39.35	24.08	0.99	81.26	512.40	1931.15	1525.34
%Contribution	2.78	8.33	22.22	19.44	18.06	26.39	0	11.11

Note (X_1-No of variety, X_2-Max age of the plant, X_3- Min age of the plant, X_4- No of plant/desimil, X_5-Avg production per desimil, X_6- Operational cost per desimil, X_7- Avg price per Q, X_8-Avg return per desimil)

The no. of variety contributes only 2.78% in clustering revealing that all cluster are growing closely 5 to 6 varieties as on average. The contribution of price is minimum in clustering which depicts that the price fluctuation is very less from cluster to cluster.

11.11% contribution in clustering shared by the return from mango production among the 13 groups clearly indicates wide variations in production vis-à-vis return from mango enterprise. The single farmer in cluster 13, growing 4 varieties with15 to 25 years oldorchard, is enjoying the maximum return and maximum profit with maximum productivity and with 3rd highest operational cost. These information in farming can be helpful in policy making purpose and also for the betterment of the mango farmings.

3.1 Policy Implication

The main objectives of the whole study was to facilitate the policy measures for the development of mango farmers based on their variability in farming putting them into different homogenous groups. From the study, it is suggested that the developmental / intervention policies should be taken based on the group characteristics of the respective cluster. It is also evident from the study that operational cost, minimum and maximum age of the plant, plant density, productivity and return are the major sources for which farmers are distributed among the different clusters. As such policies and programme should be suggested to address production process vis-à-vis farm economy taking in to consideration the above sources of variation.

4. SUMMARY AND CONCLUSION

Since independence, Indian agriculture has travelled from a food deficient country to food surplus country; therefore the need of the hour is to provide nutritional security to the mostly populated country in the world. Horticultural crops, in particular fruit crops like mango plays an important role in supplying nutritious food to the Indian populations, in addition to its economic importance. Among the districts in West Bengal, Malda is contributing 54 percent to the Bengal mango basket. So, mango cultivation has significant social and economical importance to the people of Malda. In this study, based on primary data obtained from five selected blocks of the district, the status of mango cultivations, its major economic parameters are critically examined. Altogether 110 mango growers are studied from five selected blocks viz. English bazar, Ratua-2, Manikchak, Kaliachak-2 and old Malda. The whole study has been undertaken to accomplish the objective of grouping the mango farmers in to homogenous clusters so as to facilitate the policy related issues with respect to development of mango farmers in the district. K-means clustering techniques using Euclidian distance measure reveals that 110 farmers could be grouped into 13 clusters. The peculiarity, as revealed by the clustering exercise is that in five clusters out of 13 clusters are consist of one farmer in each cluster. Thereby indicating that these five farmers have got certain extraordinary features which have placed them in different clusters. A thorough analysis is required so as to investigate whether these extraordinary features are good for the development or not; if found good these should be included in the policy measures so that other farmers are also benefited. It is also revealed from the study that among all the characters under consideration, operational cost per desimil is contributing maximum (26.39%) followed by minimum age of the plant (22.22%), plant density (19.44%), productivity (18.06%), return/dismal (11.11%), maximum age of the plant (8.33%) and so on. These characteristics should be focussed while making cluster wise policy.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of this manuscript.

DATA AVAILABILITY STATEMENT

On reasonable request, the corresponding author will provide data supporting the study's results. The raw data, cannot be made public for reasons of confidentiality and privacy. However, researchers who satisfy the requirements for access to confidential data can be given access to aggregated and anonymized data as well as the statistical analysis codes.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

Das. J. K., Pradhan, K., Majumder, G. and Ghoshal, R. (2013). Estimation of the attributes embedded with the economy of mango enterprise at Malda district of West Bangal, India. *Journal of Crop and Weed*, 9(2):39-43.

Das. S. K. and Mandal, S. (2020). Economics of production and marketing of mango crop a case study of Malda district, West Bengal. *Shodhganga: a reservoir of Indian theses @INFLIBNET.*

Lapple, D. and Kelley, H. (2013). Understanding the uptake of organic farming: Accounting for heterogeneities among Irish farmers. *Ecological Economics*,88: 11-19.

Mondal, B. and Paul Choudhury, J. (2013). A Comparative Study on K Means and PAM Algorithm using Physical Characters of Different Varieties of Mango in India. *International Journal of Computer Applications*, 78(5): 0975 – 8887.

Mukhopadhyay, P. (2022). Theory and methods of survey sampling, *PHI learning private limited*, Noida, Delhi, India.

Nambi, V. E., Thangavel, K. and Jesudas, D. M. (2015). Scientific classification of ripening period and development of colour grade chart for Indian mangoes (*Mangifera indica L*.) using multivariate cluster analysis. *Scientia Horticulturae*, 193: 90-98.

Sarkar, B., Mondal, S. and Basu, D. (2018). Problems and prospects of mango growers of Nadia district West Bengal. *Journal of Agricultural Engineering and Food Technology*, 5(2): 97-103.

Siddiqui, S. H., Aktar, N. and Sultana, C. (2015). A study on mango economy of Malda district, West Bengal. *National Geographical Journal of India*, 61(2): 129-140.

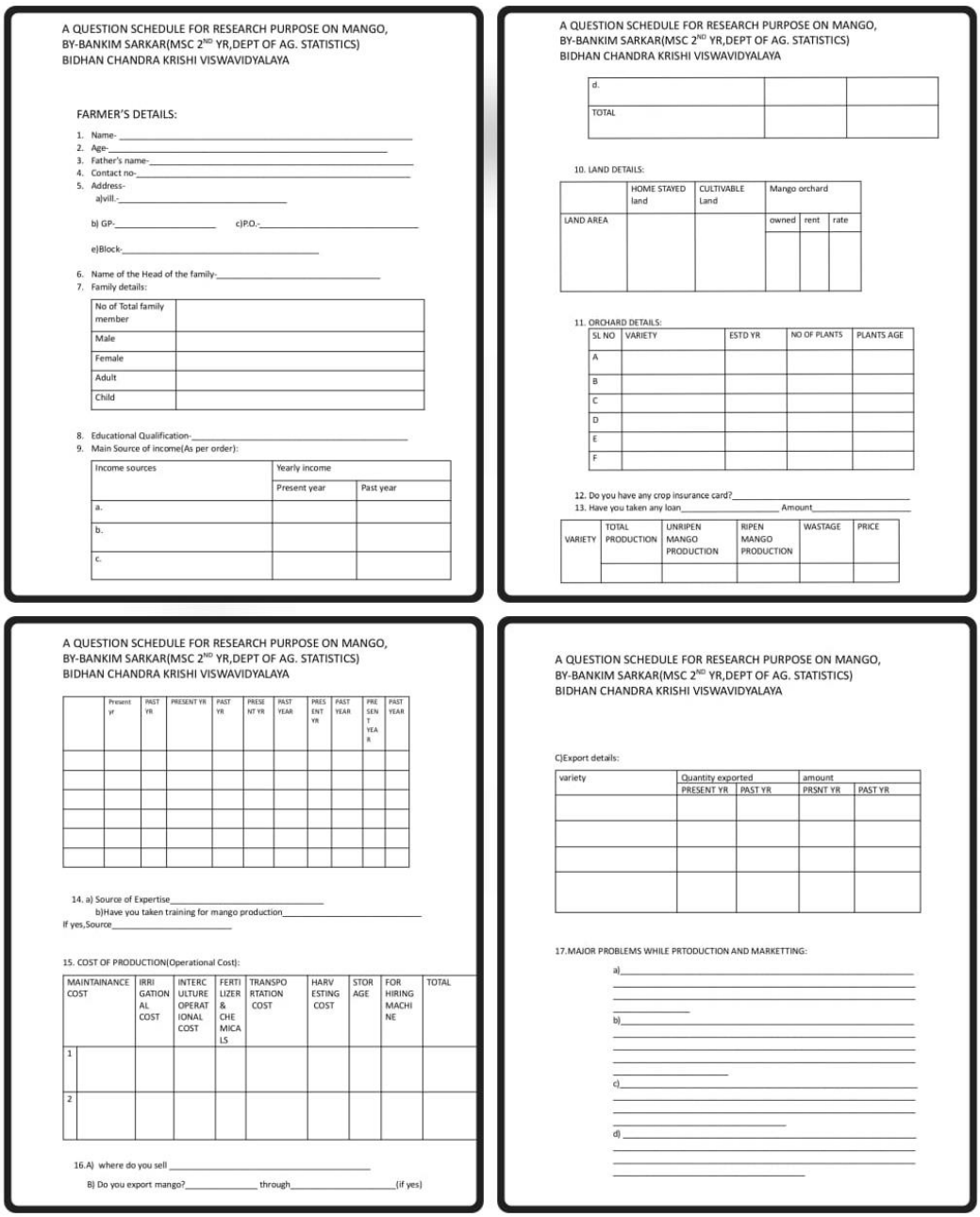
Singh, R. K. and Chaudhary, B.D. (2012). Biometrical methods in quantitative genetic Analysis, *Kalyani Publishers*, New Delhi, India.

www.globalnutritionreport.org last accessed on 15-04-2023

www.ipiindia.gov.in last accessed on 12-04-2023

www.nhb.gov.in last accessed on 14-04-2023

ANNEXURE-1



A QUESTION SCHEDULE FOR RESEARCH PURPOSE ON MANGO,
BY-BANKIM SARKAR(MSC 2ND YR,DEPT OF AG. STATISTICS)
BIDHAN CHANDRA KRISHI VISWAVIDYALAYA

FARMER'S DETAILS:

1. Name- _____
2. Age-_____
3. Father's name-_____
4. Contact no-_____
5. Address-
   a)vill.-_____
   b) GP-_____ c)P.O.-_____
   e)Block-_____
6. Name of the Head of the family-_____
7. Family details:

No of Total family member	
Male	
Female	
Adult	
Child	

8. Educational Qualification-_____
9. Main Source of income(As per order):

Income sources	Yearly income	
	Present year	Past year
a.		
b.		
c.		

A QUESTION SCHEDULE FOR RESEARCH PURPOSE ON MANGO,
BY-BANKIM SARKAR(MSC 2ND YR,DEPT OF AG. STATISTICS)
BIDHAN CHANDRA KRISHI VISWAVIDYALAYA

d.			
TOTAL			

10. LAND DETAILS:

	HOME STAYED land	CULTIVABLE Land	Mango orchard		
			owned	rent	rate
LAND AREA					

11. ORCHARD DETAILS:

SL NO	VARIETY	ESTD YR	NO OF PLANTS	PLANTS AGE
A				
B				
C				
D				
E				
F				

12. Do you have any crop insurance card?_____
13. Have you taken any loan_____ Amount_____

VARIETY	TOTAL PRODUCTION	UNRIPEN MANGO PRODUCTION	RIPEN MANGO PRODUCTION	WASTAGE	PRICE

A QUESTION SCHEDULE FOR RESEARCH PURPOSE ON MANGO,
BY-BANKIM SARKAR(MSC 2ND YR,DEPT OF AG. STATISTICS)
BIDHAN CHANDRA KRISHI VISWAVIDYALAYA

	Present yr	PAST YR	PRESENT YR	PAST YR	PRESENT YR	PAST YEAR	PRESENT YR	PAST YEAR	PRESENT YEAR	PAST YEAR

14. a) Source of Expertise_____
    b)Have you taken training for mango production_____
If yes,Source_____

15. COST OF PRODUCTION(Operational Cost):

	MAINTAINANCE COST	IRRIGATIONAL COST	INTERCULTURE OPERATIONAL COST	FERTILIZER & CHEMICALS	TRANSPORTATION COST	HARVESTING COST	STORAGE	FOR HIRING MACHINE	TOTAL
1									
2									

16.A) where do you sell _____
B) Do you export mango?_____ through_____(if yes)

A QUESTION SCHEDULE FOR RESEARCH PURPOSE ON MANGO,
BY-BANKIM SARKAR(MSC 2ND YR,DEPT OF AG. STATISTICS)
BIDHAN CHANDRA KRISHI VISWAVIDYALAYA

C)Export details:

variety	Quantity exported		amount	
	PRESENT YR	PAST YR	PRSNT YR	PAST YR

17.MAJOR PROBLEMS WHILE PRTODUCTION AND MARKETTING:

a)_____
b)_____
c)_____
d)_____

Fig. 4. Format of interview schedule

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/128027