PLANT BREEDING

A novel fuzzy approach to identify the phenotypic adaptability of common bean lines

Vinícius Quintão Carneiro[1]*, **Jussara Mencalha[2]**, **Isabela de Castro Sant'anna[3]**, **Gabi Nunes Silva[4]**, **Júlio Augusto de Castro Miguel[1]**, **Pedro Crescêncio Souza Carneiro[5]**, **Moysés Nascimento[6]** and **Cosme Damião Cruz[5]**

[1]Departamento de Biologia, Universidade Federal de Lavras, Av. Central, s/n, Campus Universitário, 37200-000, Lavras, Minas Gerais, Brazil. [2]Departamento de Fitotecnia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. [3] Centro de Seringueira e Sistemas Agroflorestais, Instituto Agronômico de Campinas, Votuporanga, São Paulo, Brazil. [4]Departamento Acadêmico de Matemática e Estatística, Universidade Federal de Rondônia, Ji-Paraná, Rondonia, Brazil. [5]Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. [6]Departamento de Estatística, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. *Author for correspondence. E-mail: vinicius.carneiro@ufla.br

ABSTRACT. The genotype by environment interaction is the main factor that influences the response of evaluated genotypes in trials of value for cultivation and use. Adaptability and stability analyses are fundamental to understanding the performance of genotypes in a growing region. Some of these methodologies incorporate previous information for recommending an extra group of genotypes denominated as specific ideotypes under certain cultivation conditions. Based on this strategy, the centroid method and its modifications have been widely used due to the simplicity of classification of the evaluated genotypes. However, these methodologies present problems in identifying adaptability patterns of some genotypes. Artificial intelligence techniques, such as fuzzy C-means, can be an alternative to reduce these difficulties, since they use, in addition to distance information between genotypes, memberships (measures quantifying how much an observation belongs to a particular class) to increase discriminatory power. Therefore, our aim was to propose and evaluate the phenotypic adaptability method by fuzzy clustering to assist cultivar recommendations. The adaptation of the fuzzy C-Means method to classify the genotypes was implemented in BioFuzzy software. The grain yield data of black common bean genotypes were used to evaluate the potential of the method. The results obtained by this method were compared with those obtained by the centroid method. The phenotypic adaptability method by fuzzy clustering was effective in identifying the adaptability patterns of common bean genotypes. Moreover, the discriminatory power was higher than that observed with the centroid method.

Keywords: artificial intelligence; fuzzy logic; plant breeding.

Introduction

The evaluation of elite genotypes in trials of value for cultivation and use (VCU) is essential to identify potential genotypes for recommended to growers. To verify how these genotypes respond to environmental variations, these trials are performed in a set of environments that represent the growing regions of each crop.

The genotype by environment interaction (GE interaction) is the main factor that influences the responses of the evaluated genotypes in VCU trials (Malosetti, Ribaut, & Van Eeuwijk, 2013; Van Eeuwijk, Bustos-Korts, & Malosetti, 2016). In addition to studies about the nature of GE interactions (Cruz & Castoldi, 1991), adaptability and stability analyses are fundamental to understanding genotype performance in a growing region.

These analyses can be based on different statistical principles, such as regression (Eberhart & Russell, 1966; Nascimento et al., 2011), mixed model (Resende, 2004), principal component (Yan, Hunt, Sheng, & Szlavnics, 2000) or factor analyses (Murakami & Cruz, 2004). For cultivar recommendation purposes, these analyses involve classification procedures in which the genotypes are allocated into classes of adaptability and stability according to their performances in the evaluated environments. Although these methods are widely used, some present a high number of parameters to be interpreted, which makes cultivar recommendations even more complex (Cruz, Regazzi, & Carneiro, 2012).

Among the statistical approaches, there are adaptability and stability methodologies (Lin & Binns, 1988; Rocha, Muro-Abad, Araújo, & Cruz, 2005; Vasconcelos, Reis, Cruz, Sediyama, & Scapim, 2011; Nascimento

et al., 2015) that incorporate in the analysis previous information for recommending an extra group of genotypes denominated as specific ideotypes under certain cultivation conditions. Based on this strategy, the centroid method (Rocha et al., 2005) and its modifications (Vasconcelos et al., 2011; Nascimento et al., 2015) have been widely used by breeding programs due to the simplicity of classification of the evaluated genotypes.

This classification is based on the Cartesian distance of the evaluated genotypes to four pre-established ideotypes. Despite providing practical results, these methods present difficulties in identifying the adaptability of genotypes with similar proximity to two or more ideotypes (Vasconcelos et al., 2011).

Artificial intelligence techniques such as artificial neural networks (ANNs) have also been used in cultivar recommendations (Barroso, Nascimento, Nascimento, Silva, & Ferreira, 2013; Nascimento et al., 2013; Teodoro et al., 2015). These methodologies present recognized potential for solving classificatory and predictive problems (Silva, Tomaz, Sant'Anna, Nascimento, & Bhering, 2014; Sant'Anna et al., 2015; Glória et al., 2016; Carneiro et al., 2017). This approach applied to classificatory problems requires prior knowledge of the classifications of the set of observations in the training process (Ma, Zhang, & Wang, 2014; Singh, Ganapathysubramanian, Singh, & Sarkar, 2016). Therefore, for using ANNs in cultivar recommendations, which in statistical approaches is a classificatory problem, it is necessary to know the adaptability and/or stability of a dataset for training the ANN. Despite their efficiency, these techniques require high processing time and computational demand.

Unsupervised artificial intelligence techniques have not yet been used to assist in recommending cultivars. These techniques dispense previous knowledge about the classification of observations and require the algorithm itself to find unknown patterns in the data (Schmidhuber, 2015). This is the principle of clustering techniques such as k-means (Macqueen, 1967), fuzzy C-means (FCM) (Bezdek, Ehrlich, & Full, 1984) and self-organizing maps (Kohonen, 1982) used to organize and categorize data (Jang, Sun, & Mizutani, 2012).

FCM can be an alternative to reduce the difficulties of identifying grouping patterns regarding phenotypic adaptability, since it uses, in addition to distance information between genotypes, memberships (measures that quantify how much an observation belongs to a particular class) to increase discriminatory power. Therefore, our aim was to propose and evaluate the phenotypic adaptability method by fuzzy clustering to assist in cultivar recommendation.

Material and methods

The grain yield data (kg ha^{-1}) of black common bean genotypes described in detail by Oliveira, Carneiro, Carneiro, and Cruz (2006) were used to evaluate the potential of the phenotypic adaptability method by fuzzy clustering. Succinctly, 18 lines and the two cultivars Ouro Negro and Valente (Table 1) were evaluated in 12 experiments conducted in a randomized block design with three replications. These experiments were carried out in six cities of Minas Gerais State (Viçosa, Coimbra, Ponte Nova, Leopoldina, Florestal and Capinópolis) in the fall-winter seasons in 2002 and in the dry and fall-winter seasons in 2003.

Table 1. Identification of black common bean genotypes.

Identification	Genotypes	Identification	Genotypes
1	VP1	11	VP11
2	VP2	12	VP12
3	VP3	13	VP13
4	VP4	14	Vi 5700
5	VP5	15	Vi 5500
6	VP6	16	Vi 7800
7	VP7	17	CNFP 9346
8	VP8	18	CNFP 7988
9	VP9	19	Ouro Negro
10	VP10	20	Valente

The data were submitted to individual analysis of variance, which was performed assuming fixed effects for genotypes. The grouping of means was performed according to the Scott-Knott test (Scott & Knott, 1974) for the experiments in which a significant effect of genotypes was verified.

After individual analysis of variance, the maximum F test was performed to verify the homogeneity of the residual variances of the experiments. The residual variances of the experiments were considered

homogeneous when the ratio between the greatest and the smallest residual variance did not exceed the value 7 (Banzatto & Kronka, 1995). Subsequently, a joint analysis of variance was performed in which the following statistical model was adopted:

$$y_{ijk} = \mu + G_i + B/E_{jk} + E_k + GE_{ik} + \varepsilon_{ijk},$$

in which y_{ijk} is the observed value of ith genotype evaluated in the kth environment and in j^{th} block; μ is the general mean; G_i is the fixed effect of the ith genotype (i = 1, 2, 3..., 20); B/E_{jk} is the random effect of the jth block in the kth environment; E_{ik} is the random effect of the kth environment; GE_{ik} is the random effect of genotype and environment interactions between the ith genotype and the kth environment; and ε_{ijk} is the experimental error associated with y_{ijk}.

Detailed studies about the phenotypic adaptability of genotypes belonging to the evaluated environments were realized by the fuzzy clustering method, as detailed below.

Phenotypic adaptability by fuzzy clustering

Phenotypic adaptability by the fuzzy clustering method is based on the FCM approach (Bezdek et al., 1984), which is a technique that aims to group observations into a number of predetermined clusters. The allocation of individuals into each cluster is performed by minimizing the following objective function:

$$J(U, c_1, \ldots, c_c, \lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{c} \left(\sum_{j}^{n} u_{ij}^m \| c_i - x_j \|^2 \right) + \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{c} u_{ij} - 1 \right) \tag{1}$$

$$\left(u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{id}}{d_{kj}} \right)^{2/(m-1)}} \right)$$

$$\left(c_i = \frac{\sum_{j}^{n} u_{ij}^m x_j}{\sum_{j}^{n} u_{ij}^m} \right);$$

$$(\sum_{i=1}^{c} u_{ij} = 1, \quad \forall j = 1, \ldots, n),$$

where U is the membership matrix of *j* observations in *i* predefined clusters. c_i is the centroid of the ith cluster, $m \in [1, \infty]$ is the fuzzy weighting exponent and λ_j are the Lagrange multipliers for the *n* constraints.

The objective function minimization (1) is performed by an iterative process, according to the following steps: i) initialize the U matrix with random values between zero and one so that the restriction $\sum_{i=1}^{c} u_{ij} = 1$ is respected; ii) obtain the centroids of clusters defined a priori; iii) obtain the objective function according to equation 1; and iv) initialize the iterative process and define the stop criterion, which is determined between two consecutive iterations, and obtain the new U matrix.

To apply the FCM algorithm to identify phenotypic adaptability patterns, 1,000 ideotypes were added to the original data set of the average values of the genotypes in each environment. These ideotypes were partitioned into four classes (wide adaptability, adaptability to favorable environments, adaptability to unfavorable environments and unadapted) according Rocha et al. (2005). That is, 250 ideotypes with same average value were denoted for each class. The ideotypes for wide adaptability (cluster I) were defined by the maximum values in all environments evaluated, while unadapted ideotypes presented the minimum values in all environments. The ideotypes for adaptability to favorable environments presented maximum values in favorable environments and minimum values in unfavorable environments. The ideotypes of adaptability to unfavorable environments presented opposite values.

The mean values of genotypes and ideotypes in each environment, that is, a matrix of dimensions M(1000+20)x12, were adopted to verify the phenotypic adaptability patterns by the fuzzy clustering method. The influence of the fuzzy weighting exponent value (*m*) on the membership of each genotype in clusters of phenotypic adaptability was also evaluated. The values considered for this exponent were 1.5 and 2.0.

The classification of each genotype was performed considering the greater membership observed among the four clusters established a priori. After identifying in which cluster each genotype presented greater pertinence, this result was compared with the classification of the already known ideotypes, which allowed the identification of the patterns of phenotypic adaptability.

The results obtained by FCM approach were compared with those obtained by the centroid method (Rocha et al., 2005). This method is based on comparisons of cartesian distances between genotypes and four

ideotypes, which were created from experimental data. These ideotypes were partitioned into four classes: wide adaptability, adaptability to favorable environments, adaptability to unfavorable environments and unadapted. The classification of environments into favorable or unfavorable was made by the environmental index proposed by Finlay and Wilkinson (1963).

Only one ideotype for each class was used in this analyse. The average values of each ideotype was equal those to FCM approach, respectively. The classification of each genotype was performed considering the greater spatial probability observed among the four classes established a priori. The spatial probability of each genotype to reference ideotypes was calculated by the following formula:

$$P_{ik} = \frac{\left(\frac{1}{d_{ik}}\right)}{\sum_{k=1}^{4}\frac{1}{d_{ik}}},$$

where P_{ik}: spatial probability that genotype i is similar to ideotype k; d_{ik}: distance from genotype i to ideotype k.

The variance analyses and the centroid method were performed in GENES software (Cruz, 2016). Phenotypic adaptability by the fuzzy clustering method was developed with the aid of MATLAB software's fuzzy logic toolbox and implemented in BioFuzzy software, available at https://github.com/VQCarneiro/BioFuzzy.

Results

There was a significant effect of the genotypes in all experiments ($p < 0.01$), except for the experiment in Ponte Nova during the dry season of 2003 (Table 2). The coefficients of variation were lower than 20%, indicating high experimental accuracy and reliability of the results. The overall mean was 2,389 kg ha^{-1}; however, variation between 1,436 and 3,665 kg ha^{-1} was observed for the experiments in Viçosa and Coimbra during the fall-winter season of 2003.

Table 2. Grain yield (kg ha^{-1}) of the black common bean genotypes (GEN) evaluated in the fall-winter/2002, dry/2003 and fall-winter/2003 seasons in the municipalities of Viçosa (VI), Coimbra (CB), Ponte Nova (PN), Leopoldina (LP), Florestal (FT), and Capinópolis (CP) in Minas Gerais State, Brazil.

GEN	Fall - Winter Season/2002				Dry Season/2003				Fall - Winter Season/2003			
	VI	CB	PN	LP	VI	CB	PN	FT[5]	VI	CB	PN	CP
VP1	1863 c*	3061 a	2588 d	2171 a	3475 a	1656 f	2701	712 f	1230 d	3891 b	2850 d	2858 d
VP2	1483 h	2250 b	2485 d	1982 b	3005 b	1774 f	3081	1551 c	1623 c	3322 c	2585 h	2508 g
VP3	2133 b	2955 a	2878 b	2043 b	2860 b	2065 d	2704	1052 e	2215 a	4047 b	2760 e	2480 g
VP4	1839 c	2782 a	2429 e	1995 b	2689 b	1897 e	2203	1571 c	1612 c	3367 c	2695 f	2464 g
VP5	1741 d	2411 b	2790 b	2209 a	3326 a	2017 e	2828	2479 a	1495 c	3403 c	2902 c	3058 b
VP6	1782 d	2548 b	2836 b	2009 b	3007 b	1364 g	2837	1632 c	1410 d	3669 c	2817 e	2436 h
VP7	1491 h	2176 b	2111 g	1811 b	2774 b	2516 b	2886	2476 a	1565 c	3531 c	2496 i	2564 f
VP8	1544 g	2370 b	2668 c	2285 a	3164 a	2099 d	2853	1566 c	1346 d	3516 c	2581 h	2619 e
VP9	1416 i	2186 b	2431 e	1795 b	3364 a	2705 a	2912	1969 b	1027 e	3494 c	2065 m	2480 g
VP10	1481 h	2243 b	2390 e	1796 b	2746 b	2251 c	2572	1412 d	1368 d	3548 c	2358 k	2181 j
VP11	1475 h	2365 b	2375 e	1943 b	2983 b	2471 b	2577	1719 c	1343 d	3879 b	2345 k	2664 e
VP12	1424 i	2302 b	2210 f	1928 b	3279 a	2475 b	2452	1359 d	1247 d	3596 c	2242 l	2430 h
VP13	1635 f	2376 b	2349 e	2161 a	3098 a	1974 e	2798	1469 d	1353 d	3669 c	2436 j	2597 f
Vi 5700	1861 c	3110 a	2563 d	2375 a	3488 a	2586 a	2753	1689 c	1626 c	3856 b	3000 b	2969 c
Vi 5500	1687 e	3044 a	2820 b	2421 a	2977 b	2110 d	3307	1992 b	1843 b	3902 b	3050 a	2358 i
Vi 7800	1899 c	2735 a	2403 e	1892 b	2535 b	2182 c	2724	1452 d	1679 c	3445 c	2641 g	2419 h
CNFP 9346	1308 j	2288 b	2214 f	1456 b	3261 a	2177 c	2963	2051 b	1253 d	3478 c	2486 i	2447 g
CNFP 7988	1087 k	2311 b	2248 f	1710 b	3307 a	1972 e	2405	639 f	1026 e	3097 c	2085 m	2541 f
Ouro Negro	2260 a	3006 a	3230 a	2096 a	3110 a	2217 c	3094	1659 c	1937 b	4405 a	2804 e	3285 a
Valente	1531 g	3049 a	2376 e	2009 b	3504 a	2389 b	2702	1686 c	530 f	4179 a	2785 e	2680 e
Means	1647	2578	2520	2004	3098	2145	2768	1607	1436	3665	2599	2602
Ej [†]	-742	189	131	-385	709	-244	379	-782	-953	1276	210	213
CV (%)	12	8	12	10	10	14	14	17	20	6	11	11
p-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	>0.05	<0.01	<0.01	<0.01	<0.01	<0.01

[†]E$_j$ – Environmental index; *Values followed by the same letter in the column belong to the same group, Scott and Knott test for significance at the 5% probability level.

The experiments carried out in the fall-winter season of 2002 in Viçosa and Leopoldina presented averages lower than the overall mean (2,389 kg ha^{-1}). Therefore, these environments were considered representative of unfavorable environments due to their negative environmental indexes (Table 2). The experiments installed in Coimbra and Florestal during the dry season of 2003 and in Viçosa during the fall-winter season of 2003 were also classified as unfavorable environments. The others were considered representative of favorable environments since they presented averages higher than the overall average and positive environmental indexes (Table 2).

The cultivars Ouro Negro and Valente were allocated by the Scott Knott test (Scott & Knott, 1974) as belonging to the group of the highest-yielding genotypes for 58 and 30% of the environments, respectively (Table 2). While Ouro Negro stood out in both favorable and unfavorable environments, the Valente cultivar excelled in favorable environments. Higher averages (4,405 and 4,179 kg ha^{-1}) were observed in Coimbra for these cultivars during the fall-winter season of 2003, where the highest environmental average was observed. The lines Vi 5700 and Vi 5500 presented averages as high as the highest-yielding genotypes in 30 and 25% of the experiments, respectively. These lines also stood out in both classes of environments. The line CNFP 9346, despite having one of the highest yields in the experiment during the dry season of 2003 in Viçosa, presented low averages for the other regions.

The residual variances of the experiments were considered homogeneous (Tables 1 and 3). The joint analysis of the experiments showed that the effects of genotypes and environments were significant (p < 0.05) (Table 3), as described by Oliveira et al. (2006). In addition, there was a significant interaction between genotypes and environments. The coefficient of variation was 11.64%, that is, a low value for this characteristic.

Table 3. Summary of joint variance analysis of black common bean genotypes evaluated for grain yield (kg ha^{-1}) in 12 environments in Minas Gerais State.

Sources of Variation	DF	Mean Square
Block/E	24	427258.22
Environments (E)	11	25596941.77[*]
Genotypes (G)	19	1019567.40[*]
GE interaction	209	230746.91[**]
Error	456	77373.15
Mean		2389.49
CV (%)		11.64

[**], [*]Significant according to an F-test at the 0.01 and 0.05 probability levels, respectively; DF: degrees of freedom; CV: coefficient of variation.

Using the centroid method, it was verified that the lines VP3 (29,1%), VP5 (31,9%), Vi 5700 (36,5%), Vi 5500 (35,3%) and the cultivar Ouro Negro (44,6%) presented wide adaptability (Table 4). Seven lines were classified as adapted to unfavorable environments (Table 4). Line VP7 stood out among the others in this cluster with a probability of 36.7%. Line VP1 and the Valente cultivar presented phenotypic adaptability to favorable environments with probabilities equal to 33.8 and 29.7%, respectively.

All genotypes showed probabilities by the centroid method (Rocha et al., 2005), ranging between 15.0 and 44.6% for all clusters of phenotypic adaptability. This variation was observed for the cultivar Ouro Negro, with a 44.6% probability in cluster I and 15% probability in cluster IV. However, most of the genotypes presented very close probabilities between the two clusters, especially VP 6 (26.2%) and CNFP 9346 (27.4%). The difference between the probabilities of CNFP 9346 in clusters III and IV was less than 0.1% (Table 4). A similar difference was also observed when considering the probabilities of line VP 6 in clusters II and IV.

The phenotypic adaptability pattern of black common bean genotypes regardless of the *m* value adopted in the FCM methodology presented 100% similarity to that obtained with the centroid method (Table 4). The cultivar Ouro Negro presented better performance in cluster I, with membership equal to 91.2 and 65.7% when adopting m values of 1.5 and 2, respectively (Table 4). Considering these *m* values, line VP 1 was the most noteworthy in cluster II, with membership equal to 63 and 43.5%. On the other hand, line VP 7 was highlighted in cluster III, with membership equal to 74.4 and 50.1% considering these same *m* values. The lines VP 8, VP 9, VP 11, and CNFP 9346 and cultivar Valente did not present a predominant phenotypic adaptability pattern when adopting *m* equal to 1.5 since they did not have membership of more than 50% in a cluster. For these lines, the second largest membership occurred in cluster IV. However, the second largest membership of cultivar Valente occurred in cluster I.

The memberships of the genotypes in the predominant clusters were higher for the fuzzy clustering method than for the spatial probabilities in centroid method (Table 4). It was also observed that for *m* equal to 1.5, the membership values were higher than the membership values considering *m* equal to 2 (Table 4).

These facts could be verified by the response of the cultivar Ouro Negro, which presented the greatest discrepancy between the memberships in clusters I and IV. Adopting m values equal to 1.5 and 2, the membership ranged from 1.1 to 91.2% and from 7.4 to 65.7%, respectively. In addition, adopting m equal to 1.5 provides a better discrimination of genotypes within the same group.

Table 4. Memberships, spatial probabilities and phenotypic adaptability classification of black common bean genotypes according to the centroid method (C_0) and FCM (C_1) with m values equal to 1.5 or 2.0.

Genotypes	I			II			III			IV			Class[†]
	C_0	C_1		C_0	C_1		C_0	C_1		C_0	C_1		
		1.5	2.0		1.5	2.0		1.5	2.0		1.5	2.0	
VP1	22.5	12.4	19.3	33.8	63.0	43.5	19.1	6.5	13.9	24.7	18.1	23.3	II
VP2	22.4	15.1	19.8	22.8	16.1	20.6	27.0	32.5	28.9	27.8	36.2	30.6	IV
VP3	29.1*	43.0	33.6	23.5	18.0	21.8	25.7	26.1	26.1	21.6	12.9	18.5	I
VP4	21.6	11.6	18.1	20.2	8.7	15.7	31.1	50.4	37.5	27.2	29.3	28.7	III
VP5	31.9	52.8	39.1	20.9	9.5	16.7	27.7	30.2	29.5	19.6	7.4	14.7	I
VP6	23.8	20.3	22.6	26.2	29.4	27.3	23.8	20.5	22.7	26.2	29.9	27.4	IV
VP7	23.9	12.9	21.0	17.8	3.9	11.7	36.7	74.4	50.1	21.6	8.7	17.3	III
VP8	24.5	22.6	23.9	23.4	18.6	21.8	26.8	32.6	28.6	25.4	26.2	25.7	III
VP9	23.3	17.1	21.4	21.4	12.1	18.0	29.4	44.4	34.2	25.9	26.3	26.4	III
VP10	19.2	6.6	14.0	20.1	7.9	15.4	28.7	33.7	31.5	32.0	51.8	39.1	IV
VP11	23.8	18.8	22.3	21.5	12.4	18.2	29.3	44.2	34.0	25.4	24.5	25.5	III
VP12	20.6	9.7	16.4	21.4	11.2	17.7	27.9	33.7	30.4	30.1	45.4	35.4	IV
VP13	22.6	15.8	20.2	23.0	16.8	20.9	26.8	32.0	28.6	27.6	35.5	30.2	IV
Vi 5700	36.5	73.9	49.7	21.9	9.3	17.8	23.6	12.7	20.7	17.9	4.1	11.8	I
Vi 5500	35.3	69.3	47.0	21.7	9.5	17.6	24.7	16.3	22.8	18.3	4.9	12.6	I
Vi 7800	22.8	14.7	20.2	20.4	9.3	16.2	31.1	52.0	37.8	25.7	24.0	25.8	III
CNFP 9346	22.6	15.7	20.2	22.6	15.5	20.2	27.4	34.6	29.8	27.4	34.3	29.8	III
CNFP 7988	15.1	1.2	7.5	21.5	4.8	15.2	19.1	3.0	12.0	44.4	91.0	65.4	IV
Ouro Negro	44.6	91.2	65.7	21.6	4.8	15.3	18.8	2.8	11.7	15.0	1.1	7.4	I
Valente	25.3	23.9	25.1	29.7	45.2	34.7	21.3	12.1	17.9	23.8	18.7	22.3	II

†: adaptability classification of genotypes by the centroid method (C_0) and FCM (C_1); I – Wide adaptability; II – Adaptability to favorable environments; III – Adaptability to unfavorable environments; IV – Unadapted. *: Greater spatial probability and membership observed among the four clusters established a priori in Centroid Method (C_0) and FCM (C_1) with m values equal to 1.5 or 2.0.

Adopting the centroid method, the ideotype of cluster I was characterized by having, in all environments, maximum values of grain yield ranging from 2,215 to 4,405 kg ha^{-1}. On the other hand, the ideotype of cluster IV presented minimum values for this characteristic ranging between 530 and 3,097 kg ha^{-1}. In the FCM method, the average values of cluster I ranged from 2,207 to 4,398 kg ha^{-1} and from 2,211 to 4,401 kg ha^{-1} when using m equal to 1.5 and 2.0, respectively (Table 5). For these values of m, in the average values of cluster IV, there was variation from 540 to 3,102 kg ha^{-1} and from 535 to 3,099 kg ha^{-1}.

Table 5. Ideotype average values (ID) of phenotypic adaptability clusters (I – wide adaptability; II – adaptability to favorable environments; III – adaptability to unfavorable environments; IV – unadapted) obtained by the centroid method and FCM.

ID	Fall-Winter/2002				Dry Season/2003				Fall-Winter/2003			
	VI[1]	CB[2]	PN[3]	LP[4]	VI	CB	PN	FT[5]	VI	CB	PN	CP[6]
I	2260	3110	3230	2421	3504	2705	3307	2479	2215	4405	3050	3285
II	1087	3110	3230	1456	3504	1364	3307	639	530	4405	3050	3285
III	2260	2176	2111	2421	2535	2705	2203	2479	2215	3097	2065	2181
IV	1087	2176	2111	1456	2535	1364	2203	639	530	3097	2065	2181
FCM (m = 1,5)												
I	2254	3106	3224	2417	3499	2698	3302	2469	2207	4398	3047	3278
II	1091	3108	3226	1460	3502	1368	3303	644	535	4401	3048	3281
III	2251	2180	2115	2414	2541	2698	2211	2469	2204	3104	2072	2186
IV	1092	2179	2114	1461	2542	1373	2209	648	540	3102	2069	2185
FCM (m = 2,0)												
I	2257	3108	3227	2419	3501	2701	3304	2474	2211	4401	3048	3282
II	1089	3109	3227	1458	3503	1366	3305	642	533	4403	3049	3283
III	2256	2178	2113	2418	2538	2702	2207	2474	2210	3100	2068	2183
IV	1089	2178	2113	1459	2539	1369	2206	643	535	3099	2067	2183

[1]VI – Viçosa; [2]CB – Coimbra; [3]PN – Ponte Nova; [4]LP – Leopoldina; [5]FT – Florestal; [6]CP – Capinópolis.

Discussion

The black common bean genotypes presented a differential response to environmental variations, as emphasized by the response of the cultivar Valente, which presented the highest average in Viçosa during the dry season of 2003 and the lowest mean for the same site and year during the fall-winter season. This result corroborates those of Oliveira et al. (2006), which affirms that the season is the condition that contributes most to the genotype and environment interaction. Therefore, detailed studies using the phenotypic adaptability method by fuzzy clustering and the centroid method (Rocha et al., 2005) were performed to identify the phenotypic adaptability pattern of the genotypes evaluated.

The phenotypic adaptability method by fuzzy clustering presented a similarity of 100% in the phenotypic adaptability pattern when compared to that obtained with the centroid method. However, the fuzzy approach provided higher discriminatory power of the genotypes, since the memberships obtained in the four clusters demonstrated a higher discrepancy in relation to spatial probabilities obtained by the centroid method. Studies have reported that the spatial probabilities obtained by the centroid method are very close to each other, which makes it difficult to recommend cultivars (Amorin, Souza, Silveira, Nascimento, & Ferreira, 2011; Batista, Hamawaki, Sousa, Nogueira, & Hamawaki, 2015). Therefore, it was verified that this proposed method allowed the identification of the phenotypic adaptability pattern of black common bean genotypes with high discriminatory power; consequently, this method aids in cultivar recommendation in breeding programs.

The concept of adaptability based on ideotypes given by the fuzzy clustering method is similar to that proposed by Rocha et al. (2005). However, these methodologies differ in some particularities related to the processing of genotype information. While the centroid method adopts only four ideotypes that represent the phenotypic adaptability pattern (Rocha et al., 2005), the fuzzy clustering method adopts 250 ideotypes for each of these patterns. It is highlighted in this work that the use of 1,000 ideotypes was effective in identifying the phenotypic adaptability of the 20 genotypes.

The FCM, despite presenting a similar classification to the centroid method, provided more discrepant membership values in the clusters. This high discrepancy made it possible to verify that there were genotypes classified as recommendable (wide or specific adaptability) that still had high membership values in the cluster of the unadapted genotypes. These genotypes presented a predominant membership below 50% in the recommended clusters (wide or specific adaptability) and the second largest membership in the unadapted cluster. Therefore, we recommend that lines such as VP 8, VP 9, VP 11, and CNFP 9346 be considered unadapted since they presented predominant membership in the cluster of genotypes adapted to unfavorable environments at a level below 50% and the second highest membership in the cluster of unadapted genotypes. According to Scott-Knott's grouping pattern, in at least 50% of the evaluated environments, these lines were allocated in groups with low means (Table 2). That is, the classification of these lines as unadaptable to unfavorable environments as suggested by the centroid method constitutes a risk for cultivar recommendation.

The value of m equal to 1.5 provided a greater discrimination of the genotypes by their respective memberships. Under these conditions, the best discrimination of the Vi5700, Vi5500, and Ouro Negro cultivars was observed in the widely adaptable cluster, as the membership of these genotypes in this cluster was equal to 73.9, 69.3, and 91.2%, while when adopting m equal to 2, the membership was 49.7, 47, and 65.7%, respectively (Table 4). The literature reports that the increase in the m value provides a greater fuzzy effect in the grouping by the FCM approach (Pimentel & Souza, 2013).

Despite its great potential, fuzzy logic has not been sufficiently explored by breeding programs. This approach has been used as a principle in grouping techniques such as FCM (Rundo et al., 2017) and in fuzzy decision-making systems (Mardani, Jusoh, & Zavadskas, 2015; Carneiro et al., 2018), which have not yet been explored in plant breeding. The FCM method allows us to observe the groupings and better discriminates them through memberships (Pimentel & Souza, 2013), which is not possible in other fuzzy techniques, such as k-means and self-organizing maps. Thus, it was verified with this work that fuzzy logic, especially the phenotypic adaptability method by fuzzy clustering, presents high potential to be adopted by breeding programs in genotype selection studies and especially in the recommendation of cultivars.

Conclusion

The phenotypic adaptability method by fuzzy clustering was effective in identifying the adaptability patterns of black common bean genotypes. Moreover, the discriminatory power of this method obtained by membership in each cluster was higher than that observed in the results by spatial probability of the centroid method.

References

Amorin, B. S., Souza, G. I., Silveira, M. A., Nascimento, I. R., & Ferreira, T. A. (2011). Adaptabilidade fenotípica de genótipos de batata-doce oriundos de sementes botânicas na região Sul do Estado do Tocantins. *Revista Brasileira de Tecnologia Aplicada nas Ciências Agrárias*, *4*(3), 31-39. DOI: https://doi.org/10.5777/paet.v4i3.1499

Banzatto, D.A.; & Kronka, S. N. (1995). *Experimentação agrícola* (3. ed.). Jaboticabal, SP: FUNEP.

Barroso, L. M. A., Nascimento, M., Nascimento, A. C. C., Silva, F. F., & Ferreira, R. P. (2013). Uso do método de Eberhart e Russell como informação a priori para aplicação de redes neurais artificiais e análise discriminante visando a classificação de genótipos de alfafa quanto à adaptabilidade e estabilidade fenotípica. *Revista Brasileira de Biomedicina, 31*(2), 176-188.

Batista, R. O., Hamawaki, R. L., Sousa, L. B., Nogueira, A. P. O., & Hamawaki, O. T. (2015). Adaptability and stability of soybean genotypes in off-season cultivation. *Genetics and Molecular Research, 14*(3), 9633-9645. DOI: https://doi.org/10.4238/2015.August.14.26

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*(2-3), 191-203. DOI: https://doi.org/10.1016/0098-3004(84)90020-7

Carneiro, V. Q., Silva, G. N., Cruz, C. D., Carneiro, P. C. S., Nascimento, M., & Carneiro, J. E. S. (2017). Artificial neural networks as auxiliary tools for the improvement of bean plant architecture. *Genetics and Molecular Research*, *16*(2), 1-12. DOI: https://doi.org/10.4238/gmr16029500

Carneiro, V. Q., Prado, A. L., Cruz, C. D., Carneiro, P. C. S., Nascimento, M., & Carneiro, J. E. S. (2018). Fuzzy control systems for decision-making in cultivars recommendation. *Acta Scientiarum. Agronomy, 40*(1), 1-8. DOI: https://doi.org/10.4025/actasciagron.v40i1.39314

Cruz, C. D. (2016). Genes Software – extended and integrated with the R, Matlab and Selegen. *Acta Scientiarum. Agronomy, 38*(4), 547-552. DOI: https://doi.org/10.4025/actasciagron.v38i4.32629

Cruz, C. D., & Castoldi, F. L. (1991). Simple and complex decomposition in parts of the genotypes x environments interaction. *Revista Ceres*, *38*, 422-430.

Cruz, C. D., Regazzi, A. J., & Carneiro, P. C. S. (2012). *Modelos biométricos aplicados ao melhoramento genético* (4. ed). Viçosa, MG: Editora UFV.

Eberhart, S. A., & Russell, W. A. (1966). Stability Parameters for Comparing Varieties. *Crop Science*, *6*(1), 36-40. DOI: https://doi.org/10.2135/cropsci1966.0011183X000600010011x

Finlay, K., & Wilkinson, G. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research, 14*(6), 742-754. DOI: https://doi.org/10.1071/AR9630742

Glória, L. S., Cruz, C. D., Vieira, R. A. M., Resende, M. D. V., Lopes, P. S., Otávio, H. G. B., & Silva, F. F. (2016). Accessing marker effects and heritability estimates from genome prediction by Bayesian regularized neural networks. *Livestock Science, 191*, 91-96. DOI: https://doi.org/10.1016/j.livsci.2016.07.015

Jang, J. S. R., Sun, C. T., & Mizutani, E. (2012). *Neuro-fuzzy and soft computing - A computacional approach to learning and machine intelligence*. New Delhi, IN: PHI Learning Private Limited.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*, 59-69. DOI: https://doi.org/10.1007/BF00337288

Lin, C. S., & Binns, M. R. (1988). A superiority measure of cultivar performance for cultivar × location data. *Canadian Journal of Plant Science, 68*(1), 193-198. DOI: https://doi.org/10.4141/cjps88-018

Ma, C., Zhang, H. H., & Wang, X. (2014). Machine learning for Big Data analytics in plants. *Trends in Plant Science, 19*(12), 798-808. DOI: https://doi.org/10.1016/j.tplants.2014.08.004

Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. In M. L. C. Lucien, & J. Neyman (Eds.), *Statistics* (p. 281-297). Berkeley, CA: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability; University of California.

Malosetti, M., Ribaut, J. M., & Van Eeuwijk, F. A. (2013). The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology, 4*(44), 1-17. DOI: https://doi.org/10.3389/fphys.2013.00044

Mardani, A., Jusoh, A., & Zavadskas, E. K. (2015). Fuzzy multiple criteria decision-making techniques and applications - Two decades review from 1994 to 2014. *Expert Systems with Applications, 42*(8), 4126-4148. DOI: https://doi.org/10.1016/j.eswa.2015.01.003

Murakami, D. M., & Cruz, C. D. (2004). Proposal of methodologies for environment stratification and analysis of genotype adaptability. *Crop Breeding and Applied Biotechnology, 4*, 7–11.

Nascimento, M., Silva, F. F., Sáfadi, T., Nascimento, A. C. C., Ferreira, R. P., & Cruz, C. D. (2011). Abordagem bayesiana para avaliação da adaptabilidade e estabilidade de genótipos de alfafa. *Pesquisa Agropecuária Brasileira, 46*(1), 26-32. DOI: https://doi.org/10.1590/S0100-204X2011000100004

Nascimento, M., Peternelli, L. A., Cruz, C. D., Campana, A. C. M., Ferreira, R. P., Bhering, L. L., & Salgado, C. C. (2013). Artificial neural networks for adaptability and stability evaluation in alfalfa genotypes. *Crop Breeding and Applied Biotechnology, 13*(2), 152-156.

Nascimento, M., Ferreira, A., Nascimento, A. C. C., Silva, F. F., Ferreira, R. P., & Cruz, C. D. (2015). Multiple centroid method to evaluate the adaptability of alfalfa genotypes. *Revista Ceres, 62*(1), 30-36. DOI: https://doi.org/10.1590/0034-737X201562010004

Oliveira, G. V., Carneiro, P. C. S., Carneiro, J. E. S., & Cruz, C. D. (2006). Adaptability and stability of common bean in Minas Gerais State, Brazil. *Pesquisa Agropecuária Brasileira, 41*(2), 257-265. DOI: https://doi.org/10.1590/S0100-204X2006000200010

Pimentel, B. A., & Souza, R. M. C. R. (2013). A multivariate fuzzy c-means method. *Applied Soft Computing, 13*(4), 1592-1607. DOI: https://doi.org/10.1016/j.asoc.2012.12.024

Resende, M. D. V. (2004). *Métodos estatísticos ótimos na análise de experimentos de campo*. Colombo, PR: Embrapa Florestas.

Rocha, R. B., Muro-Abad, J. I., Araújo, E. F., & Cruz, C. D. (2005). Avaliação do método centróide para estudo de adaptabilidade ao ambiente de clones de *Eucalyptus grandis. Ciência Florestal, 15*(3), 255-266.

Rundo, L., Militello, C., Russo, G., Garufi, A., Vitabile, S., Giliardi, M. C., & Mauri, A. (2017). Automated prostate gland segmentation based on an unsupervised fuzzy C-means clustering technique using multispectral T1w and T2w MR imaging. *Information, 8*(49), 1-28. DOI: https://doi.org/10.3390/info8020049

Sant'Anna, I. C., Tomaz, R. S., Silva, G. N., Nascimento, M., Bhering, L. L., & Cruz, C. D. (2015). Superiority of artificial neural networks for a genetic classification procedure. *Genetics and Molecular Research, 14*(3), 9898-9906. DOI: https://doi.org/10.4238/2015.August.19.24

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61*, 85-117. DOI: https://doi.org/10.1016/j.neunet.2014.09.003

Scott, A. J., & Knott, M. (1974). A cluster analysis method for grouping means in the analysis of variance. *Biometrics, 30*(3), 507-512. DOI: https://doi.org/10.2307/2529204

Silva, G. N., Tomaz, R. S., Sant'Anna, I. C., Nascimento, M., & Bhering, L. L. (2014). Neural networks for predicting breeding values and genetic gains. *Scientia Agricola, 71*(6), 494-498. DOI: https://doi.org/10.1590/0103-9016-2014-0057

Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science, 21*(2), 110-124. DOI: https://doi.org/10.1016/j.tplants.2015.10.015

Teodoro, P. E., Barroso, L. M. A., Nascimento, M., Torres, F. E., Sagrilo, E., Santos, A., & Ribeiro, L. P. (2015). Redes neurais artificiais para identificar genótipos de feijão-caupi semiprostrado com alta adaptabilidade e estabilidade fenótipicas. *Pesquisa Agropecuária Brasileira, 50*(11), 1054-1060. DOI: https://doi.org/10.1590/S0100-204X2015001100008

Van Eeuwijk, F. A., Bustos-Korts, D. V., & Malosetti, M. (2016). What should students in plant breeding know about the statistical aspects of genotype × environment interactions? *Crop Science, 56*(5), 2119-2140. DOI: https://doi.org/10.2135/cropsci2015.06.0375

Vasconcelos, E. S., Reis, M. S., Cruz, C. D., Sediyama, T., & Scapim, C. A. (2011). Integrated method for adaptability and phenotypic stability analysis. *Acta Scientiarum. Agronomy, 33*(2), 251-257. DOI: https://doi.org/10.4025/actasciagron.v33i2.8272

Yan, W., Hunt, L. A., Sheng, Q., & Szlavnics, Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science, 40*(3), 597-605. DOI: https://doi.org/10.2135/cropsci2000.403597x.