



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
CENTRO MULTIDISCIPLINAR DE ANGICOS
CURSO INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA
TRABALHO DE CONCLUSÃO DE CURSO (2025.2)

MÉTODOS DE CARACTERIZAÇÃO DE GENÓTIPOS DE MELÃO POR CLUSTER
FUZZY¹

Hellen Vitória Araújo Costa²
Stefeson Bezerra de Melo³

RESUMO

O estudo caracterizou 21 genótipos de melão amarelo cultivados em Mossoró e Baraúna utilizando o método de clusterização Fuzzy C-Means (FCM) para agrupar os materiais com base em nove características dos frutos. Três classes foram formadas, mostrando que os genótipos se agrupam principalmente por características do fruto, e não apenas pelo ambiente de cultivo. As variáveis mais importantes na definição dos grupos foram o diâmetro longitudinal, o peso e o comprimento do fruto. A Classe 2 apresentou frutos maiores e mais doces, sendo a mais promissora para o mercado. O FCM mostrou-se eficiente para lidar com incertezas dos dados agrícolas e útil para apoiar programas de melhoramento genético do melão. **Palavras-chave:** Análise de agrupamento fuzzy, Qualidade de frutos, Caracterização morfológica

ABSTRACT

This study characterized 21 yellow melon genotypes cultivated in Mossoró and Baraúna using the Fuzzy C-Means (FCM) clustering method to group the materials based on nine fruit traits. Three classes were formed, indicating that the genotypes were grouped mainly according to fruit characteristics rather than solely by the cultivation environment. The most important variables in defining the groups were longitudinal diameter, fruit weight, and fruit length. Class 2 presented larger and sweeter fruits, making it the most promising for the market. The FCM method proved efficient in handling uncertainties in agricultural data and useful for supporting melon breeding programs. **Keywords:** Fuzzy clustering analysis; Fruit quality; Morphological characterization.

¹ Trabalho de Conclusão de Curso apresentado/defendido ao Curso de Graduação em Interdisciplinar em Ciência e Tecnologia, no dia 09 de dezembro de 2025, como requisito parcial para obtenção do título de Bacharel em Ciência e Tecnologia da Universidade Federal Rural do Semi-Árido (UFERSA).

² Orientanda do Curso de Graduação em Interdisciplinar em Ciência e Tecnologia, da Universidade Federal Rural do Semi-Árido (UFERSA).

³ Orientador do Curso de Graduação em Interdisciplinar em Ciência e Tecnologia, da Universidade Federal Rural do Semi-Árido (UFERSA). stefeson@ufersa.edu.br

1 INTRODUÇÃO

A lógica fuzzy é uma ferramenta flexível e robusta para a análise de cluster, especialmente em cenários com dados incertos ou sobreposição de grupos, como os encontrados na agricultura. Ao contrário dos métodos tradicionais, que atribuem cada ponto de dados a um único cluster, a lógica fuzzy permite que os pontos pertençam a múltiplos clusters simultaneamente, com diferentes graus de pertinência. Essa característica é fundamental em contextos em que as fronteiras entre os clusters não são claramente definidas, como em dados biológicos, ambientais e agrícolas.

A lógica fuzzy, em conjunto com a análise de cluster, possibilita a segmentação de áreas de cultivo de melão, considerando características do solo, clima, vegetação, produção dados morfológicos da planta. A inclusão desses dados permite o agrupamento de áreas com base na similaridade genética das plantas, otimizando o manejo específico para cada genótipo. A lógica fuzzy é utilizada para lidar com as incertezas inerentes aos dados, enquanto a análise de cluster identifica padrões e grupos homogêneos (Liu et al., 2013; Sannakki & Rajpurohit, 2011). A combinação dessas técnicas com algoritmos metaheurísticos como PSO (otimização por enxame de partículas) e GA (algoritmos genéticos) pode refinar ainda mais os agrupamentos, buscando soluções ótimas (Kuo et al., 2018). Isso resulta em um manejo mais eficiente, que leva em conta tanto as especificidades ambientais quanto as características genéticas das plantas (Onashoga et al., 2018).

A aplicação da lógica fuzzy em análise de cluster é evidente em diversas abordagens, como o algoritmo fuzzy c-means (FCM) e suas variações. O FCM emprega uma função de pertinência fuzzy para determinar o grau em que cada ponto de dados pertence a cada cluster, resultando em agrupamentos mais adaptáveis e flexíveis, adequados para conjuntos de dados complexos. O algoritmo FCM, conforme detalhado por Bezdek (1981), é uma técnica de agrupamento que atribui um grau de pertinência a cada ponto de dado, indicando sua associação com cada cluster. A utilização de conjuntos fuzzy possibilita o tratamento da incerteza, um aspecto comum em dados do mundo real, onde a filiação a um cluster nem sempre é binária.

O FCM pode ser aprimorado com o uso de funções kernel e conjuntos fuzzy intuicionísticos para um melhor desempenho em dados não lineares, conforme proposto por Kuo et al. (2018). No contexto agrícola, a lógica fuzzy tem se mostrado valiosa em diversas aplicações, como na modelagem de sistemas de irrigação (Giusti e Marsili-Libelli, 2015), no desenvolvimento de sistemas de suporte à decisão para a seleção de solo (Onashoga et

al., 2018), e na análise do impacto das mudanças climáticas na produção agrícola (Shahjalal et al., 2021). Estas aplicações demonstram a versatilidade da lógica fuzzy em lidar com a complexidade e as incertezas inerentes aos dados agrícolas.

Ao agrupar regiões com características similares, a lógica fuzzy pode auxiliar na otimização de práticas de manejo, como irrigação e fertilização, e na identificação de áreas com potencial de produção diferenciado, conforme observado por Giusti e Marsili-Libelli (2015) e Onashoga et al. (2018).

Ao empregar a análise de cluster fuzzy, este experimento visa identificar e caracterizar os 21 genótipos de melão com base em 9 variáveis em dois ambientes, permitindo um melhor entendimento das particularidades de cada genótipo em dentro de cada local.

2 METODOLOGIA

2.1 Localização e coleta de dados

Foram realizados dois experimentos em condições de campo em épocas e locais diferentes com 21 genótipos de melão, conforme descritos na tabela 1. O primeiro foi realizado no município de Mossoró – RN (5° 11'S, 37° 21' W, 23 m) e o segundo no município de Baraúna - RN (5° 05'S, 37° 38' W, 95 m), nos períodos de julho a setembro de 2017 e novembro a janeiro de 2018, respectivamente.

Durante o período de condução dos experimentos, nos dois municípios os valores médios de temperatura máxima, temperatura mínima, umidade relativa e precipitação pluviométrica foram: Mossoró (T_{máx} = 28,47 °C, T_{mín} = 27,05 °C, UR = 59,12 % e PP = 0,0 mm); Baraúna (T_{máx} = 28,82 °C, T_{mín} = 27,59 °C, UR = 65,71 % e PP = 0,06 mm). Ambos os ensaios foram conduzidos em delineamento em blocos casualizados com 21 tratamentos, 3 repetições e 10 plantas por parcela. Os tratamentos utilizados foram acessos de melão amarelo da Coleção de Germoplasma de Meloeiro da Universidade Federal Rural do Semi-Árido – UFERSA

Para obtenção das plântulas, os acessos foram semeados em bandejas de poliestireno (200 células) preenchidas com substrato comercial (Tropstrato[®]). Doze dias após a semeadura, as mudas foram transplantadas para o campo num espaçamento 0,3 m entre plantas e 2,00 m entre fileiras. Para condução em campo, o manejo foi semelhante ao comercial e nenhum inseticida foi aplicado para controle de pragas e a obtenção dos dados foram realizadas em plantas adultas coletando as seguintes variáveis: número de frutos por planta (NFP), peso em kg, diâmetro longitudinal (DL) em cm, diâmetro transversal (DT) em

cm, comprimento longitudinal (CL) em cm, índice de formato (IF), espessura de polpa (EP) em cm, firmeza de polpa (FP), sólidos solúveis (SS) em °brix.

Tabela 1. Genótipos de melão amarelo avaliados em ensaios de campo nos municípios de Mossoró-RN (Mos) e Baraúna-RN (Bar) nos anos de 2017 e 2018

Tratamentos	
Goldex	MRL-14
MRL-03	MRL-15
MRL-04	MRL-16
MRL-05	MRL-17
MRL-06	MRL-18
MRL-07	MRL-19
MRL-08	MRL-22
MRL-09	MRL-23
MRL-11	Tantalo
MRL-12	Titannium
MRL-13	

2.2 Análise de Cluster utilizando Fuzzy C-Means

A análise de cluster é uma técnica de mineração de dados usada para agrupar instâncias em grupos com base em suas similaridades. O algoritmo Fuzzy C-Means (FCM) é um método popular de cluster que permite que cada ponto de dados pertença a múltiplos clusters, com graus de pertinência distintos, o que o diferencia de algoritmos como o K-means. O FCM é considerado uma técnica de clustering flexível e robusta, especialmente em cenários com dados incertos ou sobreposição de grupos (Dalmaijer, Nord, Astle, 2022).

O FCM opera com uma matriz de valores de pertinência, indicando a força da associação de cada ponto de dados a um cluster específico. O objetivo do FCM é minimizar uma função de custo, considerando tanto a distância entre os pontos de dados e os centros dos clusters quanto os graus de pertinência atribuídos a cada ponto. Essa função de custo conforme (Bezdek, 1981) é expressa da seguinte maneira:

$$J(U, V) = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m \|x_j - v_i\|^2$$

em que:

$J(U,V)$: Representa a função objetivo a ser minimizada, indicando o custo do agrupamento.

C : número total de clusters.

N : número total de pontos de dados no conjunto.

u_{ij} : grau de pertencimento do ponto x_j ao cluster i .

m : parâmetro de fuzzificação, que controla o grau de sobreposição entre os clusters ($m > 1$).

x_j : vetor que representa um ponto de dados.

v_i : vetor do centro do cluster i .

$\| \cdot \|$: representa a norma

O algoritmo FCM segue um procedimento iterativo que envolve as seguintes etapas principais:

Inicialização: Os centros dos clusters (classes) são definidos aleatoriamente, e a matriz de pertencimento é gerada.

Cálculo dos centros dos clusters: Os novos centros dos clusters são calculados com base na média ponderada dos pontos de dados, considerando os graus de pertencimento.

Atualização da matriz de pertencimento: Os graus de pertencimento são recalculados para refletir as novas posições dos centros dos clusters.

Verificação de convergência: O processo é repetido até que a mudança na função de custo seja menor que um limite pré-definido.

O algoritmo FCM é amplamente utilizado em diversas áreas devido à sua flexibilidade e capacidade de lidar com incertezas nos dados. Em comparação com o K-means, o FCM oferece uma abordagem mais robusta, já que cada elemento pode pertencer a todos os grupos com um certo grau de pertinência.

O FCM também pode ser aprimorado com o uso de funções kernel e conjuntos fuzzy intuicionísticos para um melhor desempenho em dados não lineares (Kuo et al., 2018). E para avaliar a qualidade dos cluster foi utilizado o coeficiente de Silhueta, que é uma medida que avalia a qualidade do agrupamento, indicando o quão bem cada ponto se encaixa em seu cluster em comparação com outros clusters (Rousseeuw, 1987). Todas as análises dos dados foram realizadas na linguagem de programação Python 3.0 utilizando o ambiente computacional Jupyter Notebook.

3. RESULTADOS E DISCUSSOES

Inicialmente foi realizada um teste de normalidade e todas as nove variáveis foram normais 5% de significância pelo de kolmogorov-smirnov. E uma matrix de correlação de Pearson foi implementada para observar a relação entre as variáveis.

A correlação na tabela 2, sugere que a maioria das variáveis apresenta correlações moderadas a fortes entre si, especialmente aquelas que estão relacionadas ao peso e diâmetro do fruto. O número de frutos por planta (NFP) tem uma correlação muito fraca com as demais variáveis, o que indica não estar influenciado de forma significativa pelas variáveis analisadas neste estudo.

Tabela 2 - Matriz de correlação de Pearson de genótipos de melão amarelo avaliados em ensaios de campo nos municípios de Mossoró-RN e Baraúna-RN nos anos de 2017 e 2018

	NFP	PESO	DL	DT	CL	IF	EP	FP	SS
NFP	1,00								
PESO	-0,08	1,00							
DL	-0,14	0,90	1,00						
DT	-0,03	0,97	0,79	1,00					
CL	-0,07	0,82	0,92	0,74	1,00				
IF	-0,18	0,11	0,52	-0,11	0,47	1,00			
EP	0,03	0,89	0,80	0,88	0,74	0,08	1,00		
FP	-0,03	0,45	0,57	0,40	0,50	0,37	0,45	1,00	
SS	0,06	0,34	0,39	0,28	0,37	0,24	0,28	0,13	1,00

(NFP) = número de frutos por planta peso, (DL) = diâmetro longitudinal, (DT) = diâmetro transversal, (CL) = comprimento longitudinal, (IF) = índice de formato, (EP) = espessura de polpa, (FP) = firmeza de polpa, (SS) = sólidos solúveis (SS). Números destacados em vermelho representam correlação significativas a 5% de probabilidade

Na tabela os resultado do algoritmo da FCM, e observa-se que a Classe 0 Apresenta uma leve predominância de genótipos de avaliados em Mossoró, já na Classe 1 há uma maior concentração de genótipos avaliados em Baraúna. Para a Classe 2 predominância de genótipos avaliados em Mossoró.

A classificação pelo método FCM resultou em uma distribuição onde algumas classes podem ter uma concentração maior de genótipos de um determinado local, mas sem uma exclusividade. Alguns genótipos, como MRL-23, aparecem em mais de uma classe, indicando que eles podem apresentar características que os colocam em uma zona de transição entre as classes.

Tabela 3 - Classes de genótipos de melão amarelo construídos pelo algoritmo fuzzy c-means (FCM)

Classes	Genótipos	Mossoró	Baraúna	Total
0	Bar-Goldex, Bar-MRL-05, Bar-MRL-06, Bar-MRL-17, Bar-MRL-18, Mos-Goldex, Mos-MRL-06, Mos-MRL-13, Mos-MRL-17, Mos-MRL-18, Mos-MRL-19, Mos-MRL-23	7	5	12
1	Bar-MRL-03, Bar-MRL-04, Bar-MRL-09, Bar-MRL-11, Bar-MRL-12, Bar-MRL-13, Bar-MRL-14, Bar-MRL-19, Bar-MRL-22, Bar-MRL-23, Mos-MRL-04, Mos-MRL-11, Mos-MRL-12, Mos-MRL-14, Mos-MRL-22	5	10	15
2	Bar-MRL-07, Bar-MRL-08, Bar-MRL-15, Bar-MRL-16, Bar-Tantalo, Bar-Titannium, Mos-MRL-03, Mos-MRL-05, Mos-MRL-07, Mos-MRL-08, Mos-MRL-09, Mos-MRL-15, Mos-MRL-16, Mos-Tantalo, Mos-Titannium	9	6	15
Total		21	21	42

Outros, como Bar-Titannium e Mos-Titannium, estão agrupados na Classe 2, sugerindo que compartilham características mais distintas dos outros genótipos.

A distribuição dos genótipos nas classes, considerando os locais, demonstra que a classificação pelo método FCM não separa os genótipos exclusivamente por local de origem, mas por características intrínsecas que podem estar relacionadas ou não com o local de cultivo.

A análise de agrupamento de genótipos é relevante para o melhoramento genético. Ao classificar os genótipos em grupos distintos, os melhoristas podem focar nas características que melhor diferenciam cada classe e usar essas diferenças para aprimorar os genótipos. A lógica fuzzy, em conjunto com a análise de cluster, possibilita a segmentação de áreas de cultivo de melão, considerando características do solo, clima, vegetação e produção (Gavioli et al., 2022).

Este método é útil quando os dados apresentam sobreposição ou transições graduais entre classes, como é comum em características de plantas. As características utilizadas para

essa classificação, como diâmetro longitudinal (DL), peso do fruto (PESO), comprimento longitudinal (CL), diâmetro transversal (DT), sólidos solúveis (SS), espessura da polpa (EP), índice de formato (IF), e firmeza da polpa (FP), são relevantes para o melhoramento do melão.

Estudos como os de Silva et al. (2005) ressaltam que a seleção de genótipos de melão deve considerar múltiplos fatores, incluindo a produtividade, o mercado, a resistência a pragas e doenças, e as características de pós-colheita, além dos sólidos solúveis.

A classificação dos genótipos de melão por FCM revela uma distribuição relevante entre os locais de Mossoró e Baraúna, com cada classe apresentando uma mistura de genótipos de diferentes origens, o que destaca a complexidade das características avaliadas e a importância de abordagens como o FCM para uma análise mais precisa.

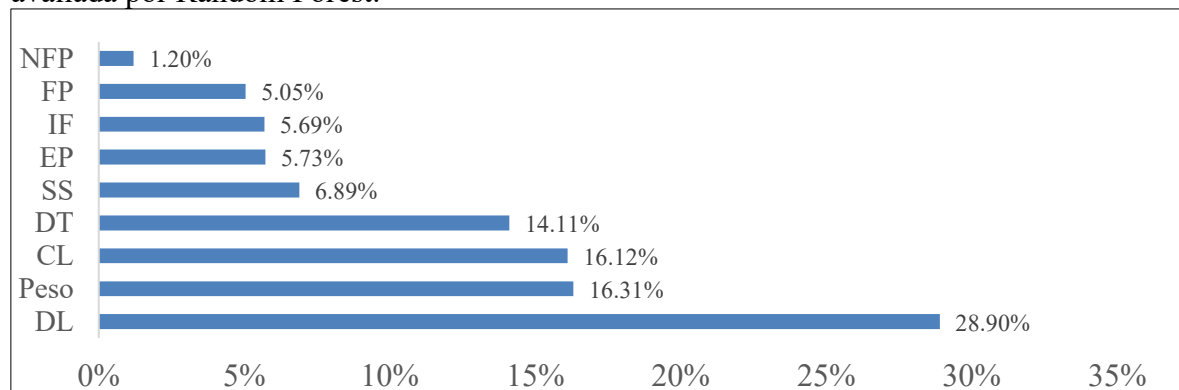
Em relação à qualidade do fruto, é crucial considerar os teores de sólidos solúveis (SS). A metodologia proposta por Gorgatti Netto et al. (1994) sugere a medição do SS na região equatorial do fruto, onde há maior concentração de açúcares. A análise de SS para a maioria dos genótipos cultivados em Mossoró, Carnaubais e alguns em Alto do Rodrigues não se enquadrou na escala de melões aptos à comercialização (Gorgatti Netto et al., 1994).

A literatura reforça a importância da avaliação das características do fruto no melhoramento do melão. Carmo et al. (2015) destacam que, para o sucesso do melhoramento, é essencial considerar aspectos como produtividade, mercado, comercialização, resistência a pragas e doenças, conservação pós-colheita e sólidos solúveis.

Além disso, os resultados obtidos por Yuri et al. (2020) demonstraram que a produtividade total e comercial do melão varia entre genótipos, com alguns se destacando em condições de verão e outros em condições de inverno.

Foi realizada junto a FCM uma análise de Random forest para verificar quais variáveis tiveram mais impacto na formação das classes, e o resultado pode ser visto na figura 1.

Figura 1 - Contribuição percentual das Variáveis na FCM para genótipos de melão amarelo, avaliada por Random Forest.



Com base na análise da importância das variáveis para a classificação de genótipos de melão amarelo, utilizando o algoritmo Random Forest nas variáveis, observa-se que as dimensões do fruto, especialmente o diâmetro longitudinal (DL), são os fatores mais relevantes para distinguir os genótipos.

A análise revela que, diâmetro longitudinal (DL) apresenta a maior importância (28,90%), seguido pelo peso do fruto (PESO) (16,30%), comprimento longitudinal (CL) (16,11%) e diâmetro transversal (DT) (14,11%). Essas variáveis são cruciais para a classificação dos genótipos, refletindo a importância das dimensões do fruto na diferenciação.

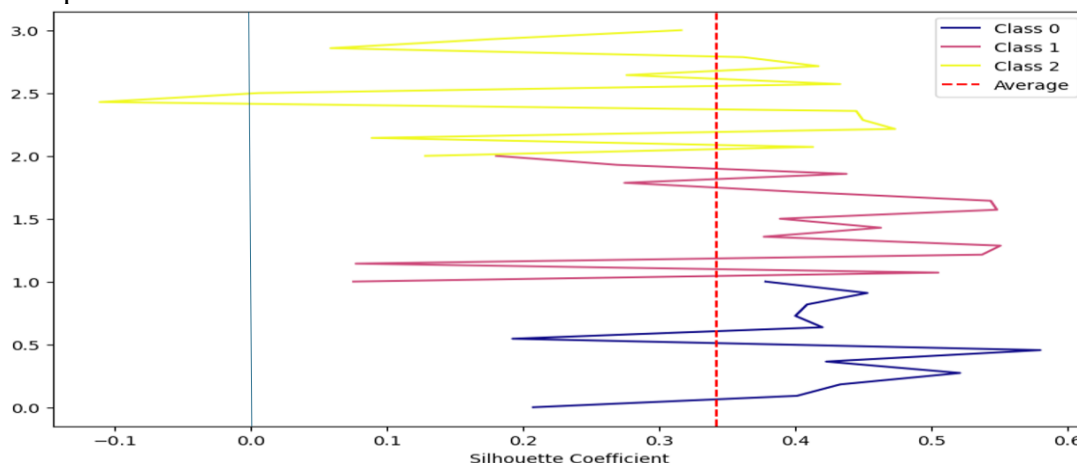
O teor de sólidos solúveis (SS), que indica o grau de doçura do fruto, também apresenta uma importância considerável (6,89%) na classificação, demonstrando ser um critério relevante na avaliação da qualidade do melão.

Outras características, como a espessura da polpa (EP), o índice de formato (IF) e a firmeza da polpa (FP), possuem uma importância menor (5,73%, 5,69% e 5,05%, respectivamente), embora contribuam para a caracterização dos genótipos.

O número de frutos por planta (NFP) é a variável que apresenta a menor importância (1,20%) para a classificação, sugerindo que, apesar de influenciar a produtividade, sua variação não está fortemente correlacionada com a distinção dos grupos de genótipos obtidos via FCM.

As dimensões do fruto, especialmente o diâmetro longitudinal (DL), e o teor de sólidos solúveis (SS) são os atributos que mais influenciam essa classificação. Esses resultados podem auxiliar no melhoramento genético do melão, visando a seleção de genótipos com características desejáveis para produtores e consumidores, para analisar as classes foi utilizado o coeficiente de Silhueta, que pode ser visto na figura 2.

Figura 2 - Contribuição percentual das Variáveis na FCM para genótipos de melão amarelo, avaliada por Random Forest



O eixo X (Silhouette Coefficient) representa o coeficiente de silhueta, que varia de -1 a 1. Sendo que valores próximos a 1 indicam que os pontos estão bem agrupados, enquanto valores próximos a -1 sugerem que os pontos podem estar mal agrupados. O eixo vertical (Y) representa as diferentes classes (Class 0, Class 1 e Class 2) e a média geral dos coeficientes de silhueta no tracejado em vermelho com valor igual a 0,34.

A Classe 0 em azul, tem valores que variam entre 0,19 a 0,58, com a maioria dos pontos concentrados em torno de 0,40. Isso sugere que os genótipos dessa classe têm uma separação moderada em relação aos outros clusters, indicando que eles estão razoavelmente bem agrupados, mas ainda pode haver algumas sobreposições.

A Classe 1 em vermelho apresenta coeficientes de silhueta que variam de 0,08 a 0,55, com uma média de 0,38. Isso indica que os genótipos dessa classe estão bem agrupados, mas ainda há espaço para melhorias na separação.

A Classe 2 em amarelo estão os valores mais, variando de -0,11 a 0,47, com média igual a 0,26, sugerindo que os genótipos dessa classe estão bem agrupados e se distanciam dos outros clusters, indicando uma boa definição e separação para a maioria.

A análise do coeficiente de silhueta fornece informações valiosas sobre a qualidade do agrupamento dos genótipos de melão amarelo. A Class 2 se destaca como a mais bem definida, enquanto as Classes 0 e 1 apresentam uma separação moderada. Esses resultados podem orientar futuras análises e decisões sobre o melhoramento genético, ajudando a identificar quais genótipos são mais homogêneos e adaptáveis às condições de cultivo.

Assim a análise de clusterização FCM que resultou na formação de três classes distintas de genótipos de melão pode ser observada na tabela 4, com os valores médias de cada variável dentro de suas respectivas classes.

A Classe 0 apresentou médias de 1,02 para número de frutos por planta (NFP), 1,87 para peso do fruto (PESO), 17,88 para a dimensão longitudinal do fruto (DL), 14,35 para a dimensão transversal do fruto (DT), 11,87 para a dimensão do comprimento do fruto (CL), 1,25 para o índice de formato (IF), 4,62 para a espessura da polpa (EP), 3,78 para a firmeza da polpa (FP), 11,16 para sólidos solúveis (SS) e 0,40 para o escore de silhueta.

Tabela 4 - Estatísticas descritivas dos genótipos de melão amarelo das classes 0, 1 e 2 construídos pelo algoritmo fuzzy c-means (FCM)

Classe	Estatísticas descritivas	NFP	PESO	DL	DT	CL	IF	EP	FP	SS	Silhouette Score
Classe 0	Médias	1,02	1,87	17,88	14,35	11,87	1,25	4,62	3,78	11,16	0,40
	Desvio-padrão	0,25	0,23	0,79	0,66	0,65	0,04	0,22	0,73	0,80	0,11
	C.V.(%)	24,21	12,41	4,41	4,58	5,45	3,23	4,72	19,25	7,13	27,32
Classe 1	Médias	1,03	1,31	15,93	12,71	10,25	1,25	4,12	2,78	11,33	0,38
	Desvio-padrão	0,31	0,18	0,85	0,65	0,85	0,06	0,25	0,44	1,22	0,17
	C.V.(%)	30,09	13,69	5,36	5,08	8,30	5,04	6,10	16,02	10,80	44,02
Classe 2	Média	0,96	2,06	19,38	14,57	13,10	1,33	4,76	4,35	12,59	0,26
	Desvio-padrão	0,18	0,29	0,73	0,91	0,82	0,08	0,27	0,90	1,25	0,19
	C.V.(%)	18,51	14,05	3,75	6,22	6,25	6,08	5,73	20,73	9,94	72,27

A Classe 1 apresentou médias de 1,03 para NFP, 1,31 para PESO, 15,93 para DL, 12,71 para DT, 10,25 para CL, 1,25 para IF, 4,12 para EP, 2,78 para FP, 11,33 para SS e 0,38 para o escore de silhueta.

A Classe 2 apresentou médias de 0,96 para NFP, 2,06 para PESO, 19,38 para DL, 14,57 para DT, 13,10 para CL, 1,33 para IF, 4,76 para EP, 4,35 para FP, 12,59 para SS e 0,26 para o escore de silhueta.

Os resultados mostram que as classes diferem em diversas características de fruto. A Classe 2 se destacou por apresentar os maiores valores médios para peso do fruto (2,06 kg) e sólidos solúveis (12,59 °Brix), indicando frutos maiores e mais doces, o que pode ser desejável para o mercado (CEAGESP, 2016). Em estudos de melão minimamente processado, é essencial avaliar parâmetros como sólidos solúveis (SS) para garantir a qualidade do produto (Alves et al., 2019).

Por outro lado, a Classe 1 apresentou o menor peso médio do fruto (1,31 kg). Os coeficientes de variação (CV) foram variáveis entre os caracteres, com o escore de silhueta apresentando maior variabilidade em todas as classes.

A análise do escore de silhueta revelou valores baixos para todas as classes (variando de 0,26 a 0,40), indicando que a separação entre os grupos não é totalmente nítida, sugerindo possíveis sobreposições entre as classes. É importante considerar que a clusterização de dados biológicos, como genótipos de plantas, pode apresentar certa complexidade devido à interação entre fatores genéticos e ambientais (Nunes et al., 2006; Nunes et al., 2011).

4. CONCLUSÃO

A técnica de clusterização Fuzzy C-Means (FCM) demonstrou ser uma abordagem eficaz para a classificação de genótipos de melão. As dimensões do fruto, o peso e o teor de sólidos solúveis foram identificados como os atributos mais significativos para a classificação.

A metodologia FCM se mostrou particularmente adequada para lidar com a incerteza inerente aos dados agrícolas, permitindo a criação de agrupamentos flexíveis e adaptáveis.

Os resultados obtidos neste estudo oferecem subsídios para programas de melhoramento genético, possibilitando a seleção de genótipos com características mais desejáveis, tanto para produtores quanto para consumidores. Além disso, o uso dessa técnica pode ser uma ferramenta para otimizar a produção e a qualidade dos frutos.

REFERÊNCIAS

ALVES, A. M. A. **Conservação de melão ‘Cantaloupe’ minimamente processado com diferentes recobrimentos**. 2019. Dissertação (Mestrado) – Universidade Federal de Campina Grande, Pombal, PB, 2019.

BEZDEK, J. C. **Pattern recognition with fuzzy objective function algorithms**. New York: Plenum Press, 1981.

CARMO, I.L.G.S.; SILVA, E.S.; MONTEIRO NETO, J.L.L.; TRASSATO, L.B.; MEDEIROS, R.D.; PORTO, D.S. **Desempenho agrônômico de cultivares de melancia no cerrado de Boa vista, Roraima**. Revista Agroambiente On-line, v. 9, n. 3, p. 268–274, jul.-set. 2015. DOI: 10.18227/1982-8470ragro.v9i3.2531. Disponível em: <https://doi.org/10.18227/1982-8470ragro.v9i3.2531>

CEAGESP. **A medida da doçura das frutas**. Cartilha Técnica 08. São Paulo: Companhia de Entrepósitos e Armazéns Gerais de São Paulo (CEAGESP), 2016.

DALMAIJER, E. S.; NORD, C. L.; ASTLE, D. E. **Statistical power for cluster analysis**. BMC Bioinformatics, v. 23, art. 205, 2022. DOI: 10.1186/s12859-022-04784-3. Disponível em: <https://doi.org/10.1186/s12859-022-04675-1>

DUNN, J. C. **A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters**. Journal of Cybernetics, v. 3, n. 3, p. 32-57, 1973. DOI: 10.1080/01969727308546046. Disponível em: <https://doi.org/10.1080/01969727308546046>

FERREIRA, R. L. **Cobertura do solo e métodos de plantio na produção do melão**. Horticultura Brasileira, São Pedro, v. 18, suplemento, p. 517-518, 2000.

GAVIOLI, A.; DE SOUZA, E. G.; BAZZI, C. L.; SCHENATTO, K.; BETZEK, N. M. **Identification of management zones in precision agriculture: An evaluation of alternative cluster analysis methods**. Biosystems Engineering, v. 181, p. 86-102, 2019. DOI: 10.1016/j.biosystemseng.2019.02.019. Disponível em: <https://doi.org/10.1016/j.biosystemseng.2019.02.019>

GIUSTI, E.; MARSILI-LIBELLI, S. **A Fuzzy Decision Support System for irrigation and water**. Environmental Modelling & Software, v. 63, p. 73-86, 2015. DOI: 10.1016/j.envsoft.2014.09.020. Disponível em: <https://doi.org/10.1016/j.envsoft.2014.09.020>

GORGATTI NETTO, A.; GAYET, J. P.; BLEINROTH, E. W.; MATALLO, M.; GARCIA, E. E. C.; GARCIA, A. E.; ARDITO, E. F. G.; BORDIN, M. R. **Melão para exportação: procedimentos de colheita e pós-colheita**. Brasília: EMBRAPA-SPI, 1994. 37 p. (FRUPEX. Publicações Técnicas, 6).

KUO, R. J.; LIN, T. C.; ZULVIA, F. E.; TSAI, C. Y. **A hybrid metaheuristic and kernel intuitionistic fuzzy c-means algorithm for cluster analysis**. Applied Soft Computing, v. 67, p. 299-308, 2018. DOI: 10.1016/j.asoc.2018.02.039. Disponível em: <https://doi.org/10.1016/j.asoc.2018.02.039>

LIU, Y.; JIAO, L.; LIU, Y.; HE, J. **A self-adapting fuzzy inference system for the evaluation of agricultural land**. Environmental Modelling & Software, v. 40, p. 226-234, 2013. DOI: 10.1016/j.envsoft.2012.09.013. Disponível em: <https://doi.org/10.1016/j.envsoft.2012.09.013>

MENEZES, J. B.; GOMES JR., J.; ARAÚJO NETO, S. E.; SIMÕES, A. N. **Armazenamento de dois genótipos de melão-amarelo sob condições ambiente**. Horticultura Brasileira, Jaboticabal, v. 19, n. 1, p. 42-49, 2001.

NAYAK, J.; NAIK, B.; BEHERA, H. S.; ABRAHAM, A. **Hybrid chemical reaction based metaheuristic with fuzzy c-means algorithm for optimal cluster analysis**. Expert Systems with Applications, v. 79, p. 282-295, 2017. DOI: 10.1016/j.eswa.2017.02.037. Disponível em: <https://doi.org/10.1016/j.eswa.2017.02.037>

NUNES, G. H. S.; MADEIROS, A. E. S.; GRANGEIRO, L. C.; SANTOS, G. M.; SALES JÚNIOR, R. **Estabilidade fenotípica de híbridos de melão amarelo avaliados no Pólo Agrícola Mossoró-Assu**. Pesquisa Agropecuária Brasileira, Brasília, v. 41, n. 9, p. 1369-1376, 2006.

NUNES, G. H. S.; ANDRADE NETO, R. C.; COSTA FILHO, J. H.; MELO, S. B. **Influência de variáveis ambientais sobre a interação genótipos x ambientes em meloeiro**. Revista Brasileira de Fruticultura, v. 33, p. 1194-1199, 2011.

ONASHOGA, A.; OJESANMI, O.; JOHNSON, F.; AYO, F. E. **A fuzzy-based decision support system for soil selection in olericulture**. Journal of Agricultural Informatics, v. 9, n. 3, p. 65-77, 2018. DOI: 10.17700/jai.2018.9.3.480. Disponível em: <https://doi.org/10.17700/jai.2018.9.3.480>

PEREIRA, M. T. L.; SILVA, M. M. S.; MEDEIROS, A. E. P.; SILVA JÚNIOR, A. F. **Modelagem matemática da secagem de casca de melão**. Brazilian Journal of Development, v. 7, n. 6, p. 54935-54947, 2021. DOI: 10.34117/bjdv7n6-070. Disponível em: <https://doi.org/10.34117/bjdv7n6-070>

ROUSSEEUW, P. J. **Silhouettes: A graphical aid to the interpretation and validation of cluster analysis**. Journal of Computational and Applied Mathematics, v. 20, p. 53-65, 1987. DOI: 10.1016/0377-0427(87)90125-7. Disponível em: [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)

SANNAKKI, S. S.; RAJPUROHIT, V. S.; KUMAR, A. R. **A Survey on Applications of Fuzzy Logic in Agriculture**. Journal of Computer Applications, v. 4, n. 1, p. 8-11, 2011.

SHAHJALAL, M.; HUDA, M. N.; ZAMAN, M. A.; ISLAM, M. S.; AKTER, S. **Fuzzy Logic Approach for Identifying the Effects of Climate**. International Journal of Agricultural Economics, v. 6, n. 4, p. 181-192, 2021. DOI: 10.11648/j.ijae.20210604.15. Disponível em: <https://doi.org/10.11648/j.ijae.20210604.15>

SILVA, L. A.; INNECCO, R.; COSTA, J. T. A.; MELO, F. I. O.; MALUF, W. R.; PEDROSA, J. F. **Estudo de aspectos quantitativos e qualitativos de frutos de genótipos de melão**. Revista Ciência Agronômica, Fortaleza, v. 36, n. 3, p. 310-315, 2005.

YURI, J. E.; RESENDE, G. M.; COSTA, N. D. **Características produtivas de genótipos de melão amarelo e pele de sapo em duas épocas de plantio no submédio do vale do São Francisco**. Revista Brasileira de Agricultura Irrigada, v. 14, n. 2, p. 3897-3905, 2020.