



ISSN: 2447-3359

## REVISTA DE GEOCIÊNCIAS DO NORDESTE

*Northeast Geosciences Journal*

v. 11, nº 1 (2025)

<https://doi.org/10.21680/2447-3359.2025v11n1ID37229>



# Um novo sistema de classificação de solos para a província do quadrilátero ferrífero usando análise estatística multivariada

*A new soil classification system for the quadrilátero ferrífero province using multivariate statistical analysis*

Gustavo Henrique Nogueira<sup>1</sup>; Leonardo Alberto Sala<sup>2</sup>; José Matheus Vieira Matos<sup>3</sup>; Tatiana Barreto dos Santos<sup>4</sup>

<sup>1</sup> Departamento de Engenharia de Minas, Universidade Federal de Ouro Preto. Morro do Cruzeiro Campus, s/n – Bauxita, CEP: 35400-000, Ouro Preto, Minas Gerais, Brazil. Email: ghnogueira89@gmail.com

**ORCID:** <https://orcid.org/0000-0003-2203-727X>

<sup>2</sup> Mining Engineer at Vale S/A, Geomechanical and Hydrogeological Management, Mina Águas Claras, CEP: 34000-000, Nova Lima, Minas Gerais, Brazil. Email: leonardo.sala@aluno.ufop.edu.br

**ORCID:** <https://orcid.org/0000-0002-6052-8427>

<sup>3</sup> Departamento de Engenharia de Minas, Universidade Federal de Ouro Preto. Morro do Cruzeiro Campus, s/n – Bauxita, CEP: 35400-000, Ouro Preto, Minas Gerais, Brazil. Email: jose.matos@aluno.ufop.edu.br

**ORCID:** <https://orcid.org/0000-0002-3700-166X>

<sup>4</sup> Departamento de Engenharia de Minas, Universidade Federal de Ouro Preto. Morro do Cruzeiro Campus, s/n – Bauxita, CEP: 35400-000, Ouro Preto, Minas Gerais, Brazil. Email: tatiana.santos@ufop.edu.br

**ORCID:** <https://orcid.org/0000-0001-5484-6675>

**Resumo:** O Sistema Unificado de Classificação de Solos (USCS) é amplamente utilizado na engenharia geotécnica, mas tem limitações na classificação de solos tropicais, devido a características como a presença de saprolitos ou lateritas. Este estudo propõe um sistema regional de classificação de solos para a província mineral do Quadrilátero Ferrífero, no Brasil, com o objetivo de aprimorar a compreensão do comportamento dos solos na região. Foram analisadas 101 amostras de solo, tanto naturais quanto compactadas, incluindo variáveis como ângulo de atrito efetivo, coesão, índice de plasticidade, gravidade específica das partículas e teor de finos. Utilizou-se a Análise de Componentes Principais (PCA) e a clusterização *k-means* para desenvolver o sistema de classificação proposto. Este categoriza os solos em três classes distintas (A, B e C) com base em seus parâmetros geotécnicos. A análise mostrou que o sistema proposto supera o USCS na diferenciação do comportamento dos solos na província do Quadrilátero Ferrífero, estabelecendo um gráfico de classificação que explica 81,68% da variabilidade dos parâmetros analisados. Comparado ao USCS, o novo sistema oferece uma ferramenta mais precisa para prever o comportamento dos solos, sendo útil em engenharia de fundações, projetos de escavação e outras aplicações geotécnicas na região.

**Palavras-chave:** Sistema de classificação de solos; Quadrilátero Ferrífero; Estatística multivariada.

**Abstract:** The Unified Soil Classification System (USCS) is widely used in geotechnical engineering, but it has limitations in classifying tropical soils due to characteristics such as the presence of saprolites or laterites. This study proposes a regional soil classification system for the mineral province of the Quadrilátero Ferrífero, in Brazil, aiming to enhance the understanding of soil behavior in the region. A total of 101 soil samples, both natural and compacted, were analyzed, including variables such as effective friction angle, cohesion, plasticity index, specific gravity of particles, and fines content. Principal Component Analysis (PCA) and *k-means* clustering were used to develop the proposed classification system. This categorizes the soils into three distinct classes (A, B, and C) based on their geotechnical parameters. The analysis showed that the proposed system outperforms the USCS in differentiating soil behavior in the Quadrilátero Ferrífero province, establishing a classification chart that explains 81.68% of the variability of the analyzed parameters. Compared to the USCS, the new system provides a more accurate tool for predicting soil behavior, being useful in foundation engineering, excavation projects, and other geotechnical applications in the region.

**Keywords:** Soil classification system; Quadrilátero Ferrífero; Multivariate statistical.

## 1. Introdução

Os sistemas de classificação de solos são ferramentas essenciais na engenharia geotécnica para caracterizar o comportamento dos solos. Esses sistemas, incluindo o amplamente utilizado Sistema Unificado de Classificação de Solos (USCS), auxiliam na previsão do desempenho do solo para diversas aplicações em terraplenagem e construção civil. No entanto, os sistemas tradicionais, conforme observado por Fookes (1994), frequentemente se concentram em características relevantes para solos de clima temperado, que não apresentam os intensos processos de intemperismo comuns em regiões tropicais.

O USCS, por exemplo, baseia-se no tamanho dos grãos e na plasticidade para a classificação. Embora eficaz para solos temperados (FOOKES, 1994), essas propriedades podem não captar adequadamente o comportamento dos solos tropicais, que frequentemente são compostos por saprolitos ou lateritas. Esses solos residuais herdam características da rocha-mãe, incluindo texturas residuais e cimentação entre os grãos. Essas particularidades podem influenciar significativamente o comportamento geotécnico, e seu impacto não pode ser totalmente explicado apenas pela plasticidade e pelo tamanho dos grãos.

Pinto (2006) enfatiza a importância de sistemas de classificação regionais que considerem essas especificidades. O sistema MCT (Miniatura, Compactado, Tropical), proposto por Nogami e Villibor (1981), exemplifica essa abordagem, sendo especificamente desenvolvido para solos tropicais. Além disso, sistemas informais locais, como a classificação "argila vermelha porosa" utilizada em São Paulo, Brasil, podem ser valiosos para regiões específicas devido ao foco em uma área limitada com menor variabilidade dos tipos de solo.

Diante das limitações dos sistemas convencionais, este estudo propõe um sistema de classificação de solos local, especificamente adaptado aos solos da província mineral do Quadrilátero Ferrífero, em Minas Gerais, Brasil. Esse sistema visa fornecer uma compreensão mais precisa do comportamento do solo em comparação com os métodos existentes. Para isso, foi compilado um conjunto de dados contendo 101 amostras de solos naturais e compactados da região, com variáveis como ângulo de atrito efetivo, coesão, índice de plasticidade, gravidade específica das partículas e teor de finos.

A Análise de Componentes Principais (PCA) e o agrupamento *k-means* foram empregados para analisar o conjunto de dados e desenvolver o sistema de classificação proposto. A PCA ajudou a identificar relações entre variáveis, reduzir a dimensionalidade dos dados e representar visualmente as informações em duas dimensões. O agrupamento *k-means* revelou três grupos distintos de solos com comportamentos geotécnicos similares. Com base nessa análise, foi desenvolvido um gráfico de classificação para categorizar os solos nesses três grupos distintos de comportamento. Essa abordagem é semelhante à utilizada por Carvalho e Ribeiro (2020) para classificar solos parcialmente saturados por meio de ensaios de penetração com cone, destacando ainda mais a eficácia das estatísticas multivariadas na classificação de solos.

As seções restantes deste artigo estão estruturadas da seguinte forma: a Seção 2 detalha o conjunto de dados, a metodologia utilizada para o desenvolvimento do sistema e revisa os métodos PCA e *k-means*. A Seção 3 apresenta informações de base sobre essas técnicas. A Seção 4 apresenta os resultados e a discussão, e a Seção 5 conclui a pesquisa.

## 2. Materiais e métodos

### 2.1. Organização dos dados

O conjunto de dados utilizado para a elaboração do sistema de classificação de solos proposto é composto por 101 amostras, nas quais foram medidos o teor de finos, a coesão, a densidade dos sólidos, o índice de plasticidade e o ângulo de atrito. Ensaios laboratoriais foram realizados nessas amostras de solos naturais e compactados provenientes de diferentes minas localizadas na província do Quadrilátero Ferrífero. Os ensaios realizados incluíram caracterização granulométrica, ensaios triaxiais de compressão CD e CIU, determinação da densidade dos sólidos (Gs) e limites de Atterberg. Esses ensaios foram conduzidos de acordo com as recomendações das normas (ABNT, 2016a; ABNT, 2016b; ABNT, 2016c; ABNT, 2016d; ISO, 2018).

As amostras utilizadas na composição do conjunto de dados apresentam grande variedade em termos de características granulométricas, resistência, densidade e plasticidade, abrangendo solos residuais de filitos, quartzitos e itabiritos, solos aluvionares e coluvionares, além de aterros compactados construídos com os mesmos materiais.

De acordo com Das e Sobhan (2013), minerais que contêm ferro em sua composição, principalmente óxidos, tendem a apresentar valores mais elevados de Gs quando comparados a minerais silicáticos, como caulinita e quartzo. Stefanou e Papazafeiriou (2013) também discutem a influência significativa do teor de finos e do Gs na resistência de cinco tipos de

solos, onde foi estabelecida uma correlação positiva entre essas duas variáveis e a resistência à penetração.

Considerando uma avaliação geotécnica preliminar de um determinado solo, sabe-se que a generalização de que solos com maior conteúdo granular (areias e cascalhos) tendem a ser mais resistentes do que solos com maior plasticidade e teor de finos (argilas e siltes) não é recomendada, conforme afirmado por Das e Sobhan (2013). Dessa forma, essa informação não deve ser utilizada como prerrogativa em métodos de projeto, como ocorre em alguns métodos empíricos de dimensionamento de rodovias.

Portanto, o conjunto de dados foi construído com base em cinco variáveis, ou seja, coesão ( $c'$ ) e ângulo de atrito ( $\phi'$ ), responsáveis por representar as características de resistência dos materiais analisados, e plasticidade (PI), teor de finos e gravidade específica das partículas (Gs), responsáveis por representar as características físicas dos materiais analisados. As dez primeiras amostras do conjunto de dados são apresentadas na Tabela 1. Para a implementação do sistema de classificação de solos proposto, os dados foram padronizados (Tabela 2).

*Tabela 1 – As dez primeiras amostras do conjunto de dados.*

ID	Teor de finos (%)	$c'$ (kPa)	Gs	PI (%)	$\phi'$ (°)
1	50.10	5.00	2.69	16.20	18.20
2	65.00	27.00	3.24	23.00	27.10
3	76.00	16.00	3.00	17.70	29.00
4	68.14	13.00	2.96	16.00	30.00
5	42.67	16.60	2.86	9.00	25.53
6	74.00	10.60	2.91	13.00	31.15
7	67.00	0.00	2.69	9.00	35.00
8	63.00	14.29	2.97	17.00	29.00
9	58.00	20.00	3.19	21.00	27.94
10	42.00	8.00	2.82	12.00	32.00

*Fonte: Autores (2024).*

*Tabela 2 – As dez primeiras amostras do conjunto de dados padronizado.*

ID	Teor de finos (%)	$c'$ (kPa)	Gs	PI (%)	$\phi'$ (°)
1	-0,91	-0,75	-0,84	0,30	-2,39
2	-0,07	1,59	0,98	1,27	-0,54
3	0,56	0,42	0,20	0,51	-0,15
4	0,11	0,10	0,04	0,27	0,06
5	-1,34	0,48	-0,27	-0,73	-0,87
6	0,44	-0,16	-0,10	-0,16	0,30
7	0,05	-1,29	-0,84	-0,73	1,10
8	-0,18	0,23	0,09	0,41	-0,15
9	-0,46	0,84	0,81	0,99	-0,37
10	-1,37	-0,44	-0,40	-0,30	0,48

*Fonte: Autores (2024).*

## 2.2. Metodologia

A metodologia para o desenvolvimento do sistema de classificação de solos proposto foi estruturada em quatro etapas (ver Figura 1). Na primeira etapa, os dados foram organizados e submetidos a uma análise estatística descritiva para compreender suas principais características. Além disso, *outliers* multivariados foram removidos utilizando a distância de Mahalanobis (MAHALANOBIS, 1928), uma métrica que considera a correlação entre variáveis e permite a identificação de *outliers* em espaços multidimensionais. Essa abordagem, baseada na média e na matriz de covariância das variáveis,

possibilita uma avaliação mais precisa da distância de um ponto de dado em relação à distribuição geral.

Na segunda etapa, foi aplicada a Análise de Componentes Principais (PCA) para reduzir a dimensionalidade do conjunto de dados, facilitando a interpretação e otimizando os processos subsequentes.

A terceira etapa do estudo consistiu na aplicação da análise de agrupamento *k-means*, um método amplamente utilizado para segmentação de dados com base em suas semelhanças (AHMED, SERAJ e ISLAM, 2020). O algoritmo *k-means* tem como objetivo particionar um conjunto de dados em *k clusters*, associando cada ponto de dado ao *cluster* cujo centroíde está mais próximo. No entanto, conforme destacado por Ahmed, Seraj e Islam (2020), a definição do número ideal de *clusters* (*k*) é um dos principais desafios dessa abordagem, pois um valor inadequado pode comprometer a qualidade da classificação.

Para determinar esse número ideal, foi adotada uma regra da maioria, baseada na avaliação de 30 diferentes índices estatísticos recomendados na literatura (CHARRAD et al., 2015). Esses índices são métricas quantitativas que auxiliam na análise da compacidade e separação dos *clusters*, contribuindo para a seleção do valor mais adequado de *k*. Dentre os 30 índices utilizados, destaca-se o Índice de Silhueta, que mede o quanto bem um ponto de dado se ajusta ao seu *cluster* e o quanto distinto ele é dos demais, atribuindo um valor entre -1 e 1, onde valores mais altos indicam melhor qualidade de agrupamento (DINH, FUJINAMI e HUYNH, 2019).

Adicionalmente, foi empregado o Critério de Calinski-Harabasz para calcular a razão entre dispersão inter-cluster e dispersão intra-cluster, indicando *clusters* mais bem definidos à medida que seu valor aumenta. O Índice de Davies-Bouldin também foi considerado, avaliando a compacidade e separação dos *clusters*, onde valores mais baixos indicam *clusters* mais bem definidos. Por fim, foi utilizado o método WSS (Within-cluster Sum of Squares), que mede a variabilidade interna dos *clusters*. Esse método é frequentemente combinado com o Método do Cotovelo (Elbow Method), uma técnica empregada para identificar o ponto em que o aumento no número de *clusters* deixa de melhorar significativamente a segmentação.

A regra da maioria foi aplicada identificando o valor de *k* que foi mais frequentemente indicado como ideal entre os diferentes índices. Assim, o número final de *clusters* foi determinado de maneira mais robusta, minimizando a subjetividade e garantindo que o processo de agrupamento representasse com precisão a estrutura dos dados.

Por fim, na quarta etapa, os *clusters* gerados foram analisados, os limites de classificação do sistema de classificação de solos proposto foram estabelecidos, e os resultados foram apresentados e discutidos.

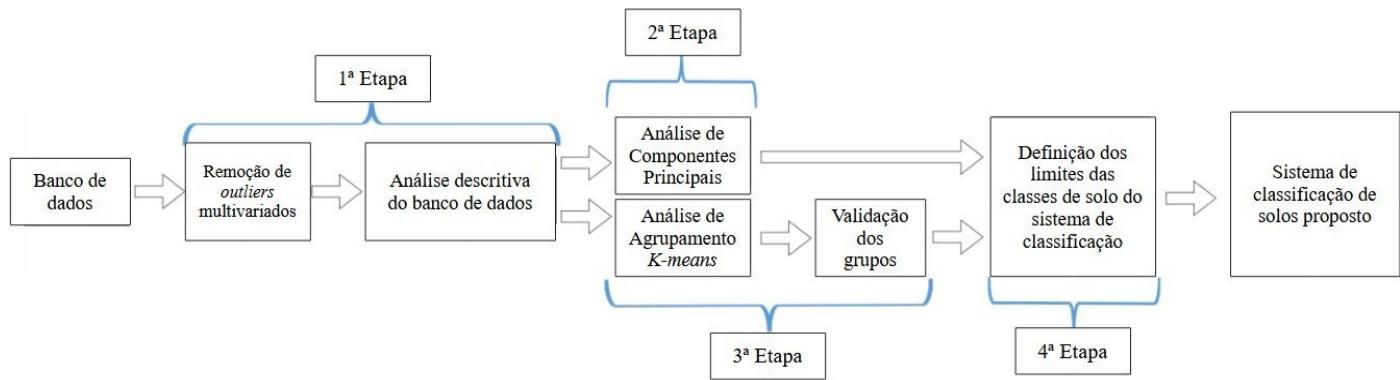


Figura 1 – Metodologia.  
Fonte: Autores (2024).

### 3. Resultados e discussão

A remoção de *outliers* multivariados foi realizada por meio da análise da distância de Mahalanobis dos pontos de dados em relação ao centroíde do conjunto de dados, com o objetivo de obter um conjunto de dados realista, excluindo informações não confiáveis. Para esse fim, foi realizada uma análise de *outliers* multivariados com o intuito de determinar quais amostras deveriam ser removidas do conjunto de dados. Foram identificados 22 *outliers*, dos quais 12 apresentaram comportamentos incoerentes para solos naturais e compactados. Valores de  $\varphi'$  acima de  $40^\circ$  foram descartados para solos predominantemente finos, e valores de  $c'$  acima de 35 kPa foram descartados para solos arenosos, buscando aderência aos valores presentes na literatura (DAS e SOBAN, 2013; MAIOLINO, 1985).

A extração dos *outliers* resultou em um conjunto de dados com teor de finos entre 26% e 94,37%, valores de coesão

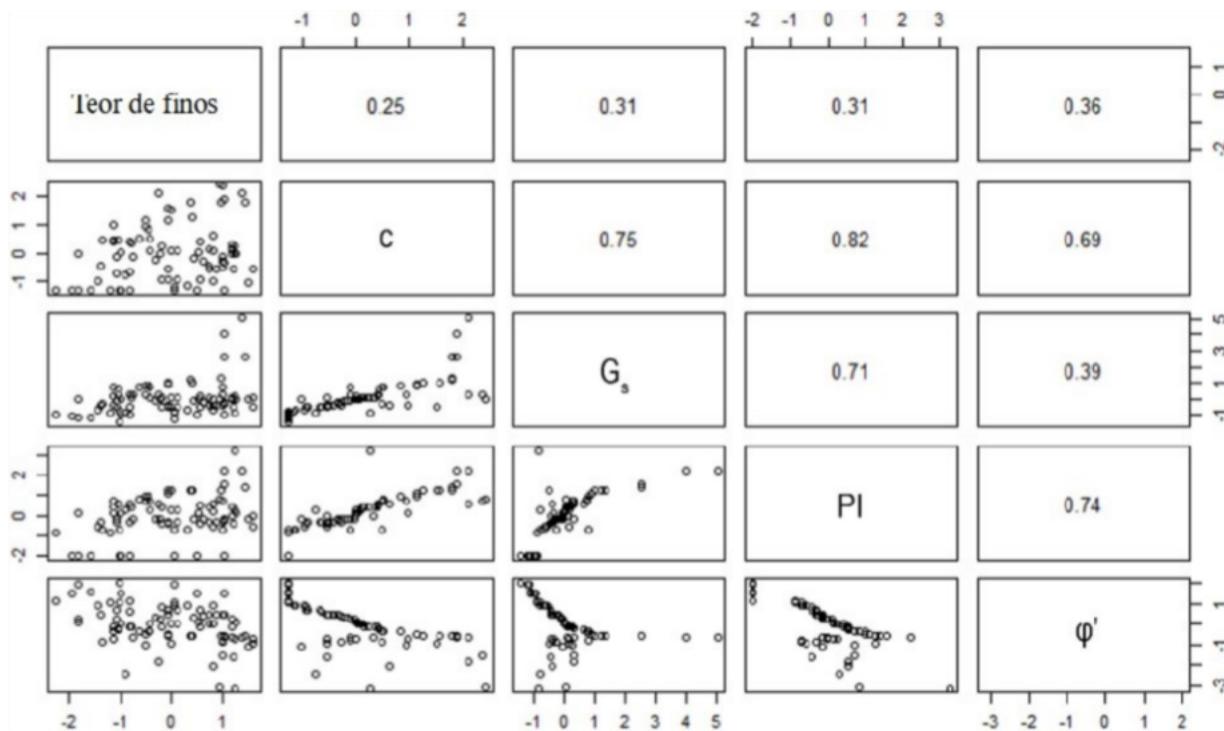
entre 0 e 35 kPa, ângulo de atrito variando entre 14,60° e 39,40°, valores de PI entre 0 e 37% e valores de  $G_s$  entre 2,50 e 4,46. Com a remoção dos dados inconsistentes do conjunto de dados, 89 amostras foram consideradas para análise posterior, realizada com o objetivo de propor um sistema de classificação de solos. A Tabela 3 apresenta as estatísticas descritivas do conjunto de dados das variáveis.

*Tabela 3 – Resumo estatístico descritivo do conjunto de dados analisado.*

	Teor de finos (%)	$c'$ (kPa)	$G_s$	PI (%)	$\phi'$ (°)
Média	66.18	12.09	2.94	14.10	29.70
Desvio padrão	17.60	9.40	0.30	7.00	4.82
1º quartil	51.52	4.61	2.80	11.0	26.97
3º quartil	83.12	16.39	3.00	18.00	32.84
Mediana	66.50	11.96	2.92	14.00	30.00
Mínimo	26.00	0.00	2.50	0.00	14.60
Máximo	94.37	35.00	4.46	37.00	39.40

*Fonte: Autores (2024).*

Foi realizada a análise das correlações entre as variáveis adotadas no presente estudo (Figura 2). Observou-se correlação significativa entre as variáveis na matriz de correlação, com valores alcançando 0,82, classificados como uma correlação forte (DOWDY, WEARDON e CHILKO, 2004). O teste de Bartlett (HAIR et al., 2009) foi realizado e foi possível concluir que os dados apresentam correlação suficiente para aplicar técnicas estatísticas multivariadas, com um valor de  $p$  igual a  $9,25 \times 10^{-53}$ . A PCA via matriz de correlação foi realizada com o objetivo de reduzir a dimensionalidade dos dados. A Tabela 4 apresenta os coeficientes (coeficientes das combinações lineares) dos cinco componentes principais gerados.



*Figura 2 – Matriz de dispersão das variáveis do estudo.*

*Fonte: Autores (2024).*

Tabela 4 – Coeficientes das componentes principais.

PC	Teor de finos	$c'$	Gs	PI	$\phi'$	Porcentagem da variância explicada	Porcentagem acumulada da variância explicada
PC <sub>1</sub>	0.266	0.509	0.451	0.516	-0.448	64,578	64,578
PC <sub>2</sub>	0.938	-0.254	-0.166	-0.150	-0.070	17,098	81,676
PC <sub>3</sub>	0.165	0.016	0.677	-0.073	0.713	12,097	93,773
PC <sub>4</sub>	0.056	0.679	-0.035	-0.728	-0.070	3,545	97,317
PC <sub>5</sub>	0.134	0.464	-0.557	0.419	0.530	2,683	100

Fonte: Autores (2024).

O critério de Kaiser (KAISER, 1970) e a análise do scree-plot (CATTELL, 1966) foram utilizados para determinar o número de componentes principais que deveriam ser mantidos na análise (ver Figura 3). Eles sugeriram a retenção de apenas um componente principal. Embora haja uma indicação de retenção de apenas uma variável, o sistema de classificação de solos proposto também considera o segundo componente principal em sua concepção, uma vez que o segundo componente principal tem um alto caráter explicativo do comportamento geotécnico dos dados.

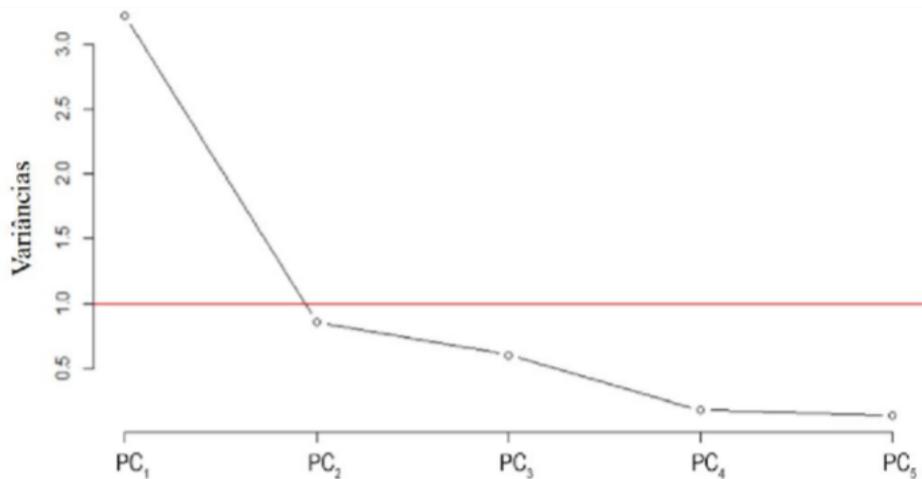


Figura 3 – Scree plot para determinar o número de componentes principais.

Fonte: Autores (2024).

A primeira componente principal apresentou correlação positiva com o teor de finos, coesão, GS e PI, e correlação negativa com o ângulo de fricção. Também é importante verificar que as correlações apresentam, em módulo, valores semelhantes, exceto para o teor de finos, que é cerca de metade dos outros. A segunda componente principal teve uma correlação muito alta com o teor de finos, sendo o principal fator que controla seu comportamento. As outras variáveis apresentaram menor importância para a definição dessa componente principal. A Figura 4 resume essas informações em um gráfico *biplot*. O uso dessas duas componentes abrange a importância de todas as variáveis, sendo capaz de explicar 81,68% da variabilidade dos dados originais.

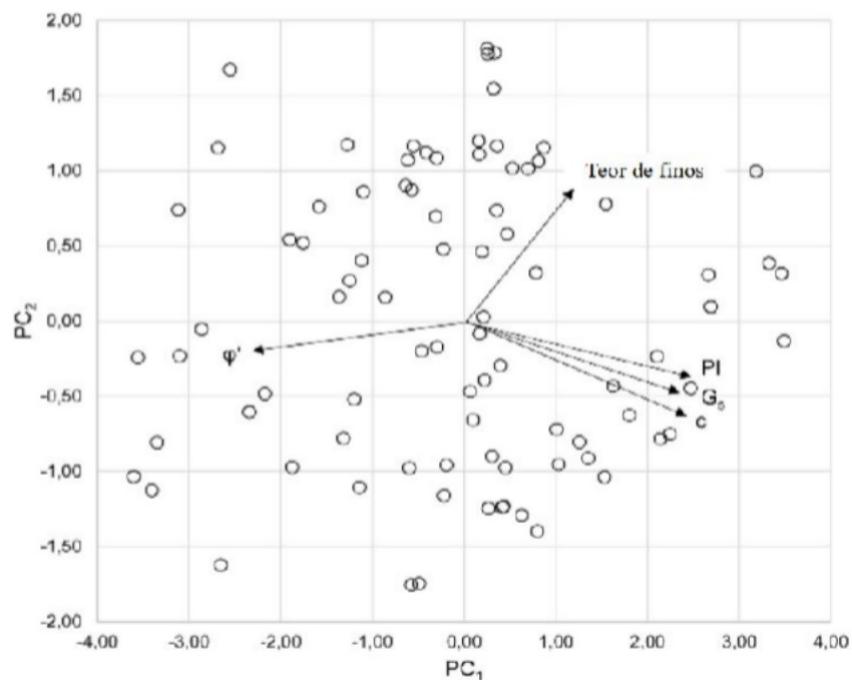


Figura 4 – Gráfico biplot.

Fonte: Autores (2024).

A clusterização dos dados foi realizada por meio do método k-means (Figura 5). A distância utilizada para o cálculo da dissimilaridade das amostras foi a distância Euclidiana e a distância de Mahalanobis. A primeira apresentou melhor desempenho e ajuste, sendo utilizada para propor o novo sistema de classificação de solos. O número de clusters adotado foi igual a três, conforme recomendado pela regra da maioria de Charrad et al. (2015), com base em 30 índices utilizados para determinar o número ótimo de clusters (Figura 6). A validação da análise de cluster foi realizada por meio da inspeção visual da Figura 5. Os grupos estão bem definidos, com uma região de sobreposição muito pequena.

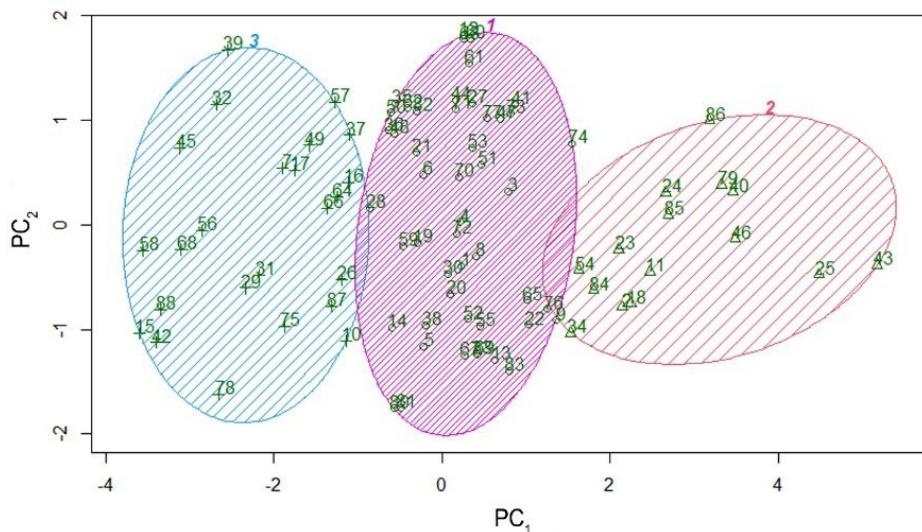


Figura 5 – Resultado da análise de agrupamentos.

Fonte: Autores (2024).

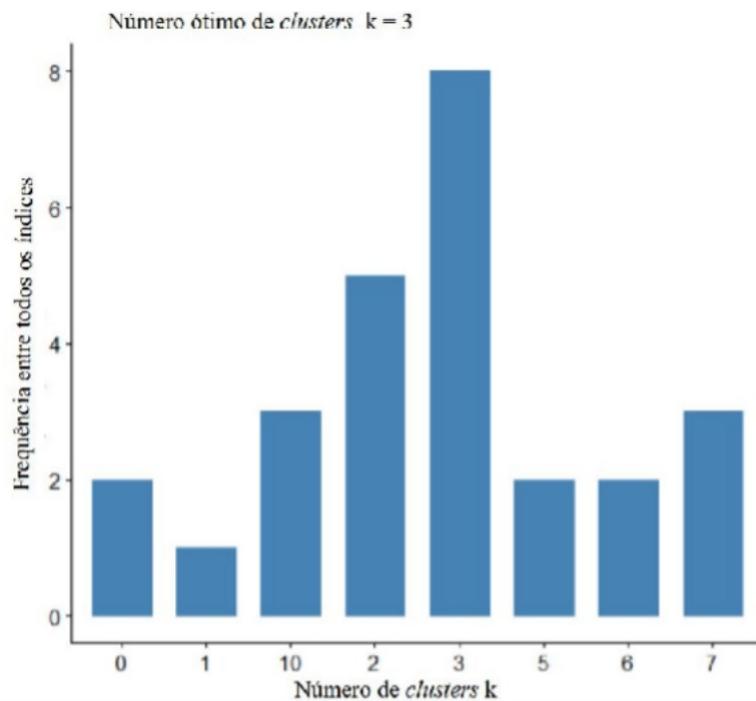


Figura 6 – Aplicação da regra da maioria para determinar o número ótimo de clusters para os dados.  
Fonte: Autores (2024).

As características das amostras que compõem cada *cluster* foram analisadas para rotulá-los, definindo três regiões de classificação de materiais no espaço cartesiano dos componentes principais 1 e 2. Contornos lineares,  $\alpha$  e  $\beta$ , foram propostos para a definição dos limites entre as classes de solo. Os contornos propostos consistem em limites lineares que interceptam os pontos de interseção dos *clusters*. As classes de solos foram rotuladas como classes A, B e C. O gráfico de classificação proposto é mostrado na Figura 7.

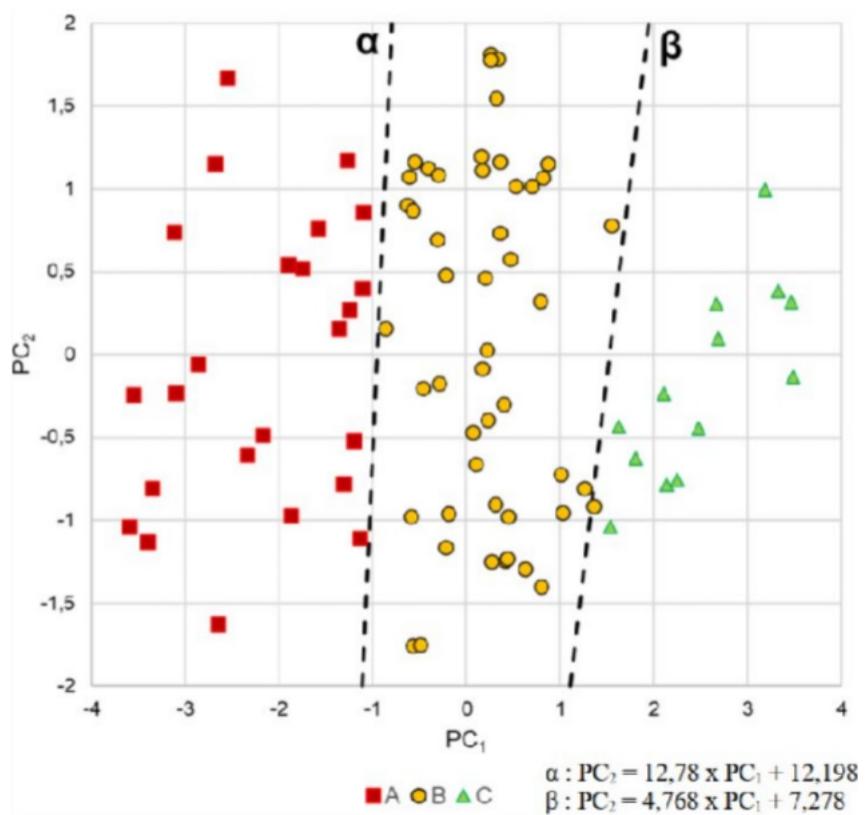


Figura 7 – Sistema proposto de classificação de solos.

Fonte: Autores (2024).

A descrição das classes de solo propostas para o sistema de classificação é apresentada abaixo.

Classe A inclui solos com os menores valores de PI e Gs. Esses solos apresentam valores de Gs inferiores a 2,82, indicando a predominância de minerais com menor peso unitário, como quartzo (Gs = 2,65) e caulinita (Gs = 2,60). De acordo com Das e Sobhan (2013), esses solos têm baixos valores de atividade e, consequentemente, baixa plasticidade. Além disso, a predominância de minerais de baixo peso unitário na Classe A indica baixo teor de ferro e, consequentemente, baixa cimentação. Essas informações são apoiadas por Stefanou e Papazafeirou (2013) e Cruz, Rodrigues e Fonseca (2013), que discutem a influência positiva da cimentação na coesão do solo. Além disso, esse tipo de solo apresenta teor de finos variando de 26% a 85%, com 75% das amostras apresentando teor de finos inferior a 67,25%. O teor de finos dessas amostras e a presença de minerais com Gs inferior a 2,82 indicam características predominantemente arenosas desses materiais. Essas características levam a maiores valores observados de φ' (DAS e SOBAN, 2013). Portanto, é possível inferir que os solos da Classe A correspondem a materiais com uma matriz fina arenosa e inativa, com concentração de minerais de baixo peso unitário e baixa cimentação.

Classe B inclui solos com comportamento transicional entre as Classes A e C. Para esses materiais, valores intermediários de Gs e PI indicam a presença de um maior percentual de minerais argilosos ativos e maior densidade em comparação com os valores observáveis para a Classe A. Coesão mais pronunciada é observada para esses materiais, indicativa de cimentação entre os grãos. No entanto, a ocorrência de ângulos de fricção moderados corrobora a observação de frações granulares de pelo menos 16% em 75% das observações. Assim, os solos da Classe B podem ser compreendidos como materiais transitórios entre as Classes A e C, com comportamentos intermediários.

Classe C inclui materiais que exibem comportamento oposto ao dos materiais da Classe A para as cinco variáveis analisadas. Pode-se observar, portanto, que a Classe C tende a abranger materiais predominantemente finos, com características argilosas e ativas, para os quais as preocupações de projeto frequentemente inerentes a materiais do tipo CH (argilas de alta plasticidade), distribuídas entre as Classes A, B e C, são aplicáveis. Os maiores valores de coesão observados para os materiais da Classe C podem estar associados a fenômenos de pré-consolidação e cimentação (CRUZ, RODRI- GUES e FONECA, 2013), sendo este último derivado da presença de minerais de alta densidade, como

hematita e goethita, justificado pela observação de valores de  $G_s$  superiores a 3,05 em 75% das observações analisadas. No entanto, os altos valores de  $c'$  observados para a Classe C são contrastados por materiais de baixa competência geotécnica, indicando a possibilidade de não processamento dos fenômenos mencionados

As amostras do conjunto de dados foram classificadas de acordo com o Sistema Unificado de Classificação de Solos (USCS) e comparadas com o sistema de classificação de solos proposto (Figura 8).

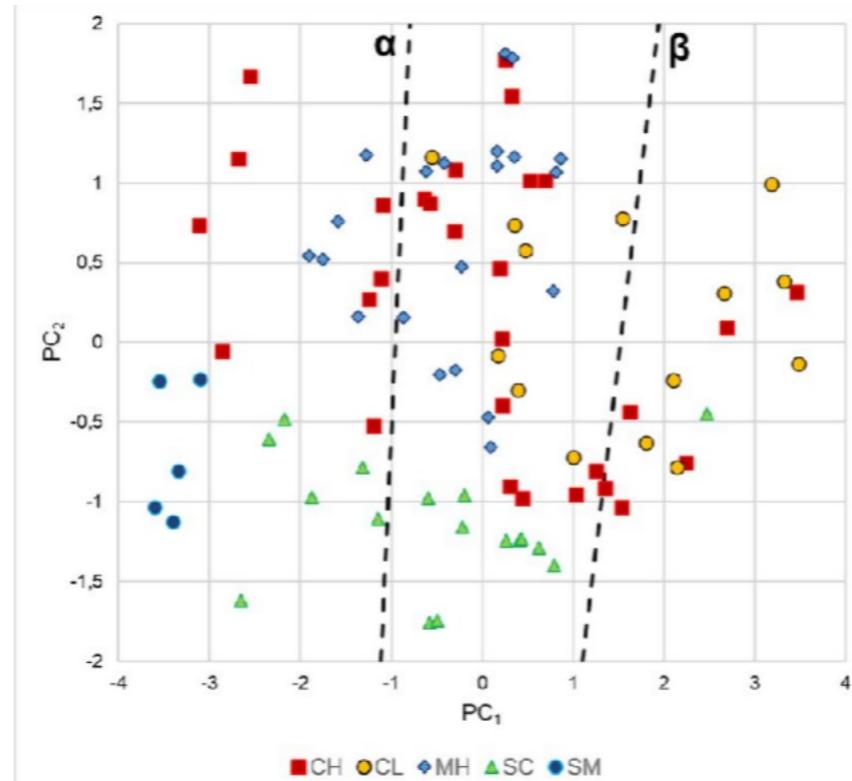


Figura 8 – USCS versus sistema proposto de classificação de solos das amostras do conjunto de dados.

Fonte: Autores (2024).

Os solos classificados pelo USCS como areias siltosas (SM) foram alocados na região correspondente à classe A do sistema de classificação proposto. Solos com argilas de alta e baixa plasticidade (CH e CL), siltes de alta plasticidade (MH) e areias argilosas (SC), considerando o USCS, apresentaram uma dispersão significativa no sistema de classificação proposto. Cada classe do USCS foi alocada em pelo menos duas classes do gráfico proposto (Figura 8). Os solos CH e CL foram alocados na região correspondente à classe C do sistema de classificação proposto. Portanto, as observações mencionadas apontam que o USCS não é eficiente na diferenciação do comportamento geotécnico dos solos analisados, pois, por exemplo, argilas de alta plasticidade, siltes de alta plasticidade, argilas de baixa plasticidade e areias argilosas apresentam comportamentos semelhantes na região entre as linhas  $\alpha$  e  $\beta$ .

Quanto ao conjunto de dados, foram elaborados boxplots das classes A, B e C do sistema de classificação proposto e boxplots das classes SM, SC, MH, CH e CL do sistema USCS, considerando as variáveis  $\phi'$ ,  $c'$ ,  $G_s$ , PI e teor de finos. As Figuras 9 a 13 apresentam os boxplots.

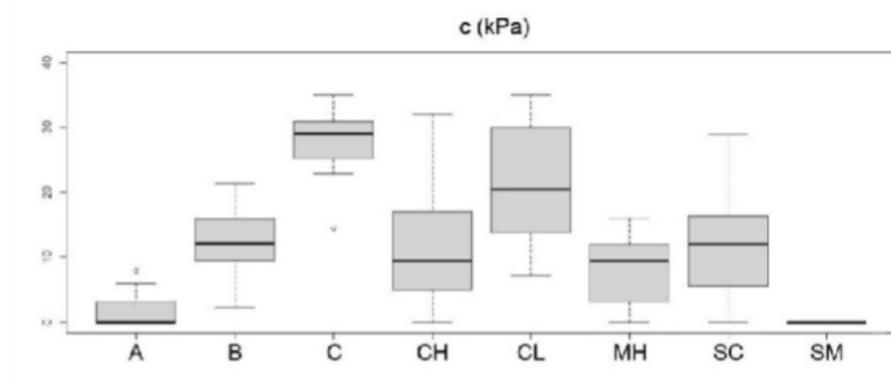


Figura 9 – Boxplots dos valores de  $c'$  para as classes analisadas.  
Fonte: Autores (2024).

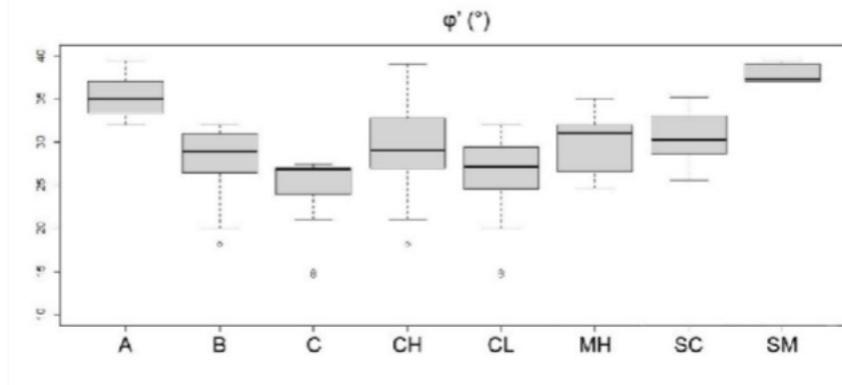


Figura 10 – Boxplots dos valores de  $\phi'$  para as classes analisadas.  
Fonte: Autores (2024).

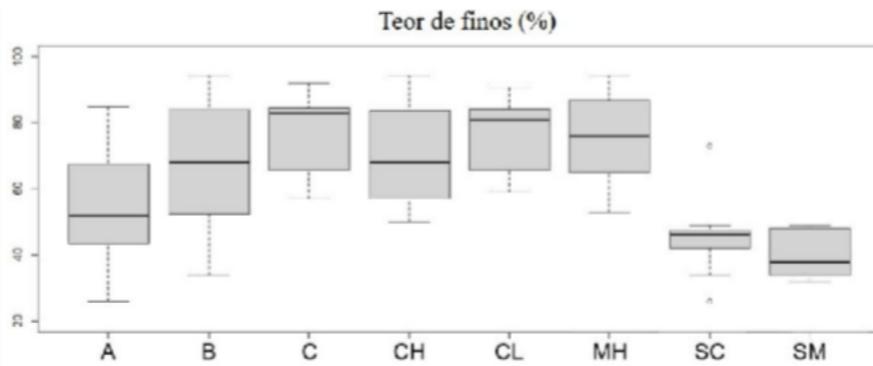


Figura 11 – Boxplots dos valores de teor de finos para as classes analisadas.  
Fonte: Autores (2024).

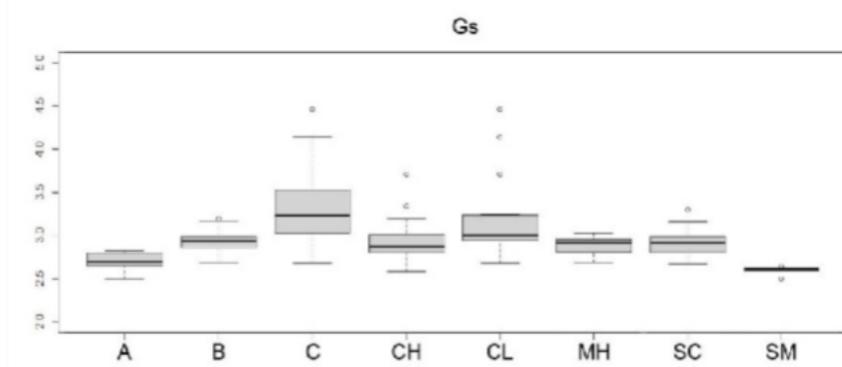


Figura 12 – Boxplots dos valores de  $G_s$  para as classes analisadas.

Fonte: Autores (2024).

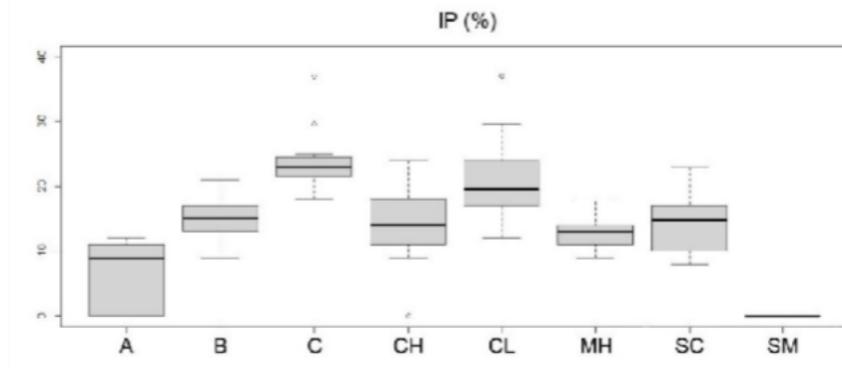


Figura 13 – Boxplots dos valores de PI para as classes analisadas.

Fonte: Autores (2024).

As classes A, B e C do sistema de classificação proposto apresentaram comportamentos geotécnicos distintos, os quais foram bem demarcados pelos diferentes intervalos de  $\phi'$ ,  $c'$ , PI,  $G_s$  e teor de finos para cada classe. No caso das classes do USCS, elas apresentam valores dispersos e aleatórios de  $\phi'$ ,  $c'$ , PI,  $G_s$  e teor de finos para cada classe, conforme mostrado nas Figuras 9 a 13. É possível observar que argilas, siltos e areias, apesar de estarem em classes diferentes no USCS, apresentam características semelhantes, do ponto de vista de resistência e comportamento, indicando a não aplicabilidade do USCS para os solos da província mineral do Quadrilátero Ferrífero. Os boxplots (Figuras 9 a 13) demonstram a capacidade do sistema de classificação proposto em discriminar os solos de acordo com seus comportamentos de resistência.

#### 4. Conclusões

Este estudo propõe um novo sistema de classificação de solos especificamente desenvolvido para a província do Quadrilátero Ferrífero, no Brasil. Técnicas estatísticas multivariadas, incluindo análise de componentes principais (PCA) e análise de agrupamento  $k$ -means, foram utilizadas para analisar um conjunto abrangente de amostras de solo. A PCA reduziu efetivamente a dimensionalidade dos dados, enquanto identificava relações-chave entre as propriedades do solo. Subsequentemente, o agrupamento  $k$ -means agrupou eficientemente os solos com base em seu comportamento geotécnico, definindo classes distintas (A, B e C). Essas classes foram ainda caracterizadas ao estabelecer os limites  $\alpha$  e  $\beta$ , que se traduzem diretamente nas variações de resistência dos solos.

Os resultados revelam limitações do amplamente utilizado sistema USCS na classificação precisa dos solos do Quadrilátero Ferrífero (FORTES, MERIGHI e NETO, 2002). Em contraste, o sistema proposto demonstrou oferecer uma eficiência superior na categorização dos solos com base em seu comportamento geotécnico, refletindo a

variabilidade única do solo dentro desta província mineral.

A adoção deste novo sistema de classificação abre caminho para o desenvolvimento de diretrizes de projeto específicas para as classes A, B e C. Essas diretrizes têm um grande potencial para aprimorar as práticas de engenharia geotécnica na região. No entanto, devido ao escopo regional do estudo e ao tamanho limitado da amostra, é fundamental ter cautela ao aplicá-lo em projetos de engenharia individuais. Recomenda-se mais pesquisas para explorar correlações entre essas classes de solo (A, B e C) e outros comportamentos geotécnicos relevantes, conduzindo, assim, a uma aplicação mais abrangente e prática deste sistema de classificação na engenharia geotécnica.

### Agradecimentos

Os autores gostariam de agradecer à Universidade Federal de Ouro Preto (UFOP), que, por meio de sua bolsa de pesquisa, financiou este estudo.

### Referências

- AHMED, M.; SERAJ, R.; ISLAM, S. M. S. 2020. The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. *Electronics*. 9(8):1295. DOI: <https://doi.org/10.3390/electronics9081295>
- Associação Brasileira de Normas Técnicas (ABNT). NBR 6458:2016: *Soil - Determination of Specific Gravity of Grains*. Rio de Janeiro: ABNT, 2016a.
- Associação Brasileira de Normas Técnicas (ABNT). NBR 6459:2016: *Soil - Determination of Liquid Limit*. Rio de Janeiro: ABNT, 2016b.
- Associação Brasileira de Normas Técnicas (ABNT). NBR 7180:2016: *Soil - Determination of Plastic Limit*. Rio de Janeiro: ABNT, 2016c.
- Associação Brasileira de Normas Técnicas (ABNT). NBR 7181:2016: *Soil - Particle Size Analysis*. Rio de Janeiro: ABNT, 2016d.
- CARVALHO, L. O.; RIBEIRO, D. B. Application of kernel k-means and kernel x-means clustering to obtain soil classes from cone penetration test data. *Soils and Rocks*, 43(4), 607-618, 2020. Disponível em: <https://doi.org/10.28927/SR.434607>. Acesso em: 12/01/2024.
- Cattell, R. B. The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2):245-276, 1966.
- CHARRAD, M.; GHAZZALI, N.; BOITEAU, V.; NIKNAFS, A. NbClust: Determining the Best Number of Clusters in a Data Set. R Package, version 3.0, 2015. Disponível em: <https://CRAN.Rproject>. Acesso em: 15/01/2024.
- Cruz, N.; Rodrigues, C.; Foneca, A. V. An Approach to Derive Strength Parameters of Residual Soils from DMT Results. *Soils and Rocks*, 37 (3): 195–209, 2013.
- Das, B. M; Sobhan, K. *Principles of geotechnical engineering*. 8<sup>a</sup>. ed. Cengage Learning, Stamford, CT, 2013.
- DINH, D. T.; FUJINAMI, T. & HUYNH, V. N. 2019. Estimating the Optimal Number of Clusters in Categorical Data Clustering by Silhouette Coefficient. In: Chen, J., Huynh, V., Nguyen, G.N., Tang, X. (eds) *Knowledge and Systems Sciences. KSS 2019. Communications in Computer and Information Science*, vol 1103. Springer, Singapore. DOI: [https://doi.org/10.1007/978-981-15-1209-4\\_1](https://doi.org/10.1007/978-981-15-1209-4_1)
- Dowdy, S.; Weardon, S.; Chilko, D. *Statistics for Research*. 3<sup>a</sup> ed. Wiley Interscience, Hoboken, NJ, 2004.
- Fookes, P. G. A review: Genesis and classification of tropical residual soils engineers. *Geotechnics in the African Environment*. 423-442, 1994.
- Fortes, R. M.; Merighi, J. V.; Neto, A. Z. Método de Teste de Azulejo para Identificação Rápida de Solos Tropicais. *Anais do 2º Congresso Português de Estradas*, Lisboa, Portugal, 18–20 novembro de 2002. Centro Rodoviário Português, 2002.

- 
- Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. *Multivariate Data Analysis*. 7<sup>a</sup>. ed. Bookman, 2009.
- International Organization for Standardization (ISO). ISO 17892-9:2018: *Geotechnical investigation and testing Laboratory testing of soil - Part 9: Triaxial compression tests for saturated cohesive soils*. Geneva: ISO, 2018.
- Kaiser, H. F. A second generation little Jiffy. *Psychometrika*, 35(4):401–415, 1970.
- MAHALANOBIS, P. C. 1928. A statistical study of the Chinese hut, Man in India, 8, p.107-122.
- MAIOLINO, A. L. *Resistência ao Cisalhamento de Solos Compactados: Uma Proposta de Tipificação*. Dissertação de Mestrado, Programa de pós-graduação em Engenharia. Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 1985.
- Nogami, J. S.; Villibor, D. F. A New Soil Classification for Road Purposes. Proceedings of the Brazilian Symposium on Tropical Soils in Engineering. ABMS, 30–41, 1981.
- Pinto, C. S. *Basic Soil Mechanics Course in 16 Lessons*. 3<sup>a</sup>. ed. Oficina de Textos, São Paulo, SP, 2006.
- Stefanou, S.; Papazafeirou, A. Effects of spatial variability on soil liquefaction: Some design recommendations. *Eurasian Journal of Soil Science*, 2 (2):122–130, 2013.