

DIFERENCIAÇÃO DE HORIZONTES DE SOLOS DE UMA ENCOSTA DA SERRA DO MAR POR ALGORITMO DE AGRUPAMENTO (k-MEANS) UTILIZANDO PARÂMETROS QUÍMICOS.

Alexandre Matheus Bliska ¹; Carolina Verbicaro Perdomo ²; Estéfano Seneme Gobbi ³; Francisco Sérgio Bernardes Ladeira ⁴; Miriam Gonçalves Miguel ⁵;

Resumo – A caracterização e classificação de solos em áreas sujeitas a movimentos gravitacionais de massa é de extrema importância social, econômica e ambiental. As caracterizações de solos realizadas nessas áreas, porém, passam pela subjetividade do profissional que realiza o estudo. Visando mitigar as subjetividades inerentes a esses estudos, foi utilizado o algoritmo de agrupamento k-Means para avaliar as similaridades entre os horizontes de solo de 10 trincheiras dos morros da Ilha de São Vicente, constituinte da Serra do Mar, em Santos, SP. Os elementos químicos Al e Na foram descartados dos testes por não apresentarem bons resultados na Análise de Silhueta. O algoritmo foi capaz de distinguir consistentemente grupos através dos parâmetros P, C, CTC, K, Mg e Ca das amostras de solo estudadas. Os resultados indicam que existem 2 ou 3 grupos diferentes dentre essas amostras.

Abstract – The characterization and classification of soils in areas subject to mass gravitational movements is of extreme social, economic and environmental importance. However, the characterization of soils in these areas goes through the subjectivity of the professional who carries out the study. Aiming to mitigate the subjectivities inherent to these studies, the k-Means clustering algorithm was used to evaluate the similarities between the soil horizons of 10 trenches of the hills of São Vicente Island, constituent of Serra do Mar, in Santos, Southeastern Brazil. The chemical elements Al and Na were removed from the tests because they did not show good results in Silhouette Analysis. The algorithm was able to distinguish consistently groups through the P, C, CTC, K, Mg and Ca parameters of the soil samples studied. The results indicate that there are 2 or 3 different groups among these samples.

Palavras-Chave – Movimento de massa; Algoritmo de agrupamento; Instabilidade de encosta; Solos tropicais; Horizontes de solo.

¹ Mestrando, Geól., Universidade Estadual de Campinas, Campinas-SP, (19) 99722-6277, alebliska@gmail.com

² Doutoranda, Geóg., Universidade Estadual de Campinas, Campinas-SP, (19) 3521-2352, caveper@hotmail.com

³ Professor Doutor, Pontifícia Universidade Católica de Campinas, Campinas-SP, (19) 3343-7202, estefano.gobbi@puc-campinas.edu.br

⁴ Professor Doutor, Universidade Estadual de Campinas, Campinas-SP, (19) 3521-4579, fsbladeira@ige.unicamp.br

⁵ Professora Associada, Universidade Estadual de Campinas, Campinas-SP, (19) 3521-2393, mgmiguel@fec.unicamp.br

1. INTRODUÇÃO

Os movimentos gravitacionais de massa são eventos geológicos sob a influência da gravidade que envolvem solo, rocha e/ou vegetação ao longo da vertente (Tominaga, 2007). Esses eventos acontecem em ambientes terrestres, costeiros e marítimos, e diversos movimentos de solo podem ocorrer. Dentre os movimentos mais comuns estão as quedas de blocos de rochas, escorregamentos de solo/rocha, rastejo e corridas de lama/detritos (Augusto Filho, 1992).

A Serra do Mar, em Santos-SP, é uma das regiões do Brasil que sofre com movimentos de massa, muitas vezes levando às perdas humanas, impactos ambientais e danos econômicos e materiais. Segundo Nery (2014), foram mais de 4 mil mortes relacionadas a esses eventos. A detecção e prevenção de eventos como esse se faz, portanto, essencial para o planejamento e manejo urbano. Assim, para que sejam realizados estudos que auxiliem no conhecimento e prevenção desses eventos, é preciso que haja a caracterização dos solos presentes nas áreas de risco.

Nos estudos em áreas de risco, é a caracterização de solos que requer maior atenção, muitas vezes pouco dada. Diversos parâmetros são utilizados para realizar essa classificação através da avaliação de dados morfológicos, físicos, químicos e mineralógicos. Existem ainda diferentes tipos de classificação, baseadas nesses parâmetros, como o Sistema Brasileiro de Classificação de Solos (SiBCS), revisto pela última vez pela Embrapa (2013) e Sistema Unificado de Classificação de Solos (SUCS), baseado na classificação de Casagrande (1948). Todos esses fatores são somados ao conhecimento, habilidade, experiência e ao julgamento subjetivo do profissional que realiza a classificação.

É visando evitar as subjetividades inerentes à classificação dos solos por profissionais e operadores humanos que esse trabalho se faz importante. Métodos computacionais podem tornar prática e inequívoca a classificação dos solos em estudos de áreas de risco aos movimentos gravitacionais. Este estudo utiliza um algoritmo de agrupamento para avaliar a similaridades entre os horizontes de solos de 10 trincheiras abertas na vertente de uma encosta da Serra do Mar, fazendo uso de parâmetros de caracterização química.

2. CARACTERIZAÇÃO DA ÁREA ESTUDADA

2.1. Serra do Mar

A Serra do Mar, localizada na costa brasileira, se estende por 1500 km pela região sul e sudeste. É afetada por movimentos gravitacionais frequentes, principalmente em condições de intensa pluviosidade, durante o verão, entre dezembro e março. Sua origem é associada ao tectonismo Cenozóico e é constituída por uma variedade de litologias metamórficas, como migmatitos, e complexos ígneos (Almeida & Carneiro, 1998). Sua altitude varia entre 800 e 1200 m.

Devido à alta pluviosidade da região, chegando a 4.500mm/ano, as encostas com inclinação elevadas e solos com descontinuidades hidrogeológicas, a Serra do Mar é uma região extremamente suscetível a eventos de deslocamento de massa. Desde 1928 foram contabilizadas mais de 4000 mortes como resultados de eventos como esse. Somente entre 2010 e 2011 dois grandes eventos levaram a perda de 1200 vidas (Nery, 2014).

Augusto Filho e Virgili (1998) apontam que os movimentos de massa ocorrem em uma cadeia de eventos, que parte da formação da rocha e inclui sua formação geológica e também geomorfológica. Já Shelby (1993) diz que diversos condicionamentos geomorfológicos da encosta podem influenciar na ocorrência desses eventos, como posição e formado da encosta, cobertura vegetal, vibração (p.ex. estradas e estradas de ferro), água e intemperismo. Já Pachauri e Pant (1992) e Larsen e Torres-Sánchez. (1998) apontam para características intrínsecas à geologia e origem do solo, como litologia e estrutura geológica, presença de minerais orgânicos, solos residuais ou de transporte e presença de fragmentos de rochas.

Autores e estudos clássicos (Projeto RadamBrasil, 1983; Rodrigues, 1965) determinam que a região formada pelas escarpas da Serra do Mar possui orientação geral SW-NE ao longo do litoral paulista, com morfologia resultante de tectonismos regionais seguido de processos erosivos intensos. Esses processos levaram à uma topografia de vales alongados, com drenagens retilíneas e escarpas íngremes, com grandes desníveis altimétricos.

Já Ab'Sáber (1955 e 1965) expõe as vertentes da Serra do Mar como de extrema complexidade, resultante de eventos tectônicos, eustáticos e erosivos (Pós-Cretáceos). Assim, o arranjo da drenagem veio a se formar como última feição morfológica na região.

2.2. Área de Estudo

A área de estudos alvo deste artigo se encontra na cidade de Santos/SP (Figura 1).

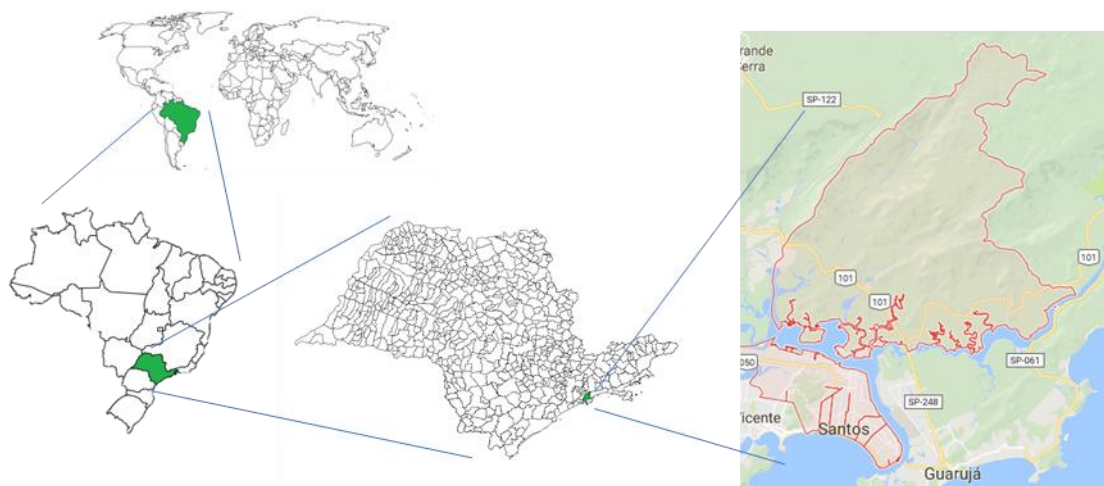


Figura 1 - Localização da área de estudos, Santos-SP (fonte: criado pelo autor).

A área de estudo está localizada em uma vertente na face norte dos morros situados na Ilha de São Vicente (Figura 2), Santos/SP, que possuem, segundo IPT (1979), cota máxima de 200 metros acima do nível do mar, distribuídos em uma área de 7,92Km². O maciço que compõe os morros pertence ao mesmo conjunto definido pela Serra do Mar e, portanto, apresenta os escorregamentos como principal agente evolutivo de relevo.

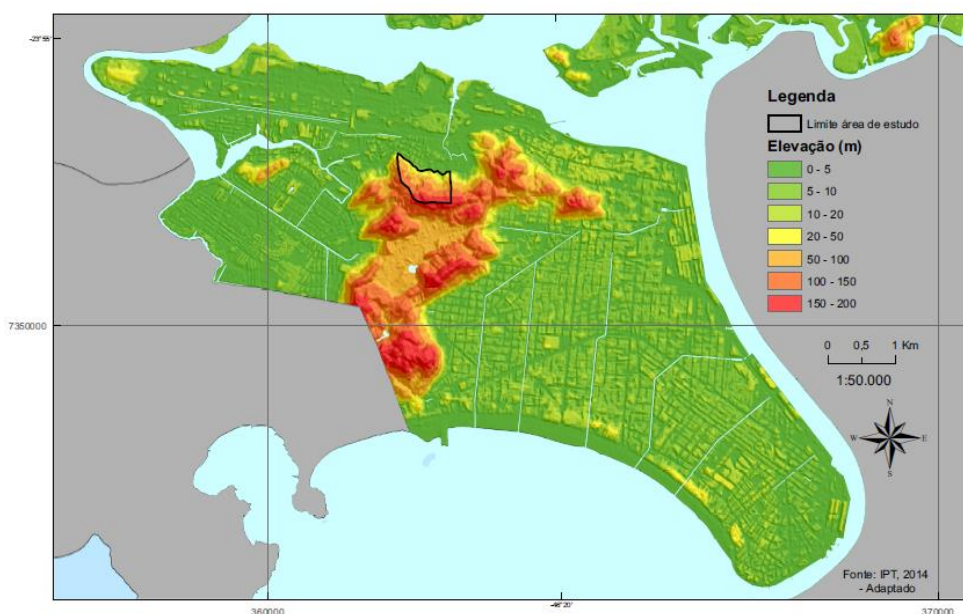


Figura 2 - Localização da área de estudos na Ilha de São Vicente, Santos/SP (adaptado de Gobbi, 2017).

3. ANÁLISE DE GRUPOS – k-Means

Segundo Estivill-Castro (2002), em seu estudo sobre as ferramentas de agrupamento Cluster Analysis, (Análise de Grupos), é um estudo estatístico que procura agrupar amostras de um conjunto de dados baseados apenas na informação que caracteriza esses objetos. Fundamentalmente os objetos contidos num grupo devem ser o mais similar entre si e o mais diferente de objetos de outros grupos. Quanto maior a similaridade, ou homogeneidade, entre as amostras contidas em cada grupo, e quanto maior a diferença entre diferentes grupos, melhor e mais distinto é o grupamento dos dados.

MacQueen (1967) coloca o k-Means como um dos algoritmos não supervisionados mais simples para se resolver problemas de grupamento de dados. Este algoritmo agrupa o conjunto de dados fornecidos em um número (k) de grupos determinado a priori pelo usuário. O algoritmo, então, define k centros, ou centroides, um para cada grupo. É importante que tais centroides sejam colocados de maneira inteligente, já que diferentes localizações podem gerar resultados diferentes. Assim, inicialmente, a melhor disposição é colocá-los o mais espaçado possível no conjunto de dados afim de garantir os melhores resultados. O passo seguinte realizado pelo algoritmo é associar cada ponto ao centroide mais próximo, até que todos os pontos estejam associados a um centroide, e se conclua o primeiro grupamento. Em seguida são definidos novos centroides no centro de cada grupo criado no passo anterior, e, de posse dos k novos centroides, cada ponto do conjunto de dados é novamente associado ao centroide mais próximo. Dessa forma, um ciclo é gerado entre a definição dos centroides e o conjunto de dados, e a cada interação os centroides mudam de posição. O ciclo se repete até que não haja mudança na posição de centroides, indicando que a soma das distâncias de cada ponto ao seu centroide é mínima. Uma descrição básica do algoritmo seria: 1: Selecionar K pontos como centroides iniciais, 2: Formar K grupos associando cada ponto ao centroide mais próximo, 3: Redefinir o centroide de cada grupo, 4: Repetir passos 1 a 3 até não haja mudanças de centroide.

3.1. Definição do número de Clusters - Método da Análise por Silhueta

O método de Análise por Silhueta (Mathworks, 2018) é utilizado para validar a consistência dos resultados obtidos pela Análise de Grupos. Esse método avalia a distância entre os grupos que foram definidos pelo algoritmo k-Means. O gráfico de Silhueta gerado pela análise demonstra quão perto cada ponto de um grupo está dos grupos adjacentes, permitindo determinar qual o melhor número de grupos de maneira visual. O exemplo da Figura 3 foi criado utilizando dados em um universo de apenas duas dimensões, ou seja, apenas duas características determinam cada amostra. Pode ser observada as divisões de grupos para valores de $k = 2, 4$ e 6 . Assim, para pequenos conjuntos de dados, formados por poucas variáveis, é fácil observar os melhores valores de k a serem utilizados. Porém, para conjuntos de dados com milhares de pontos e variáveis, essa decisão precisa ser feita de forma computacionalmente assistida.

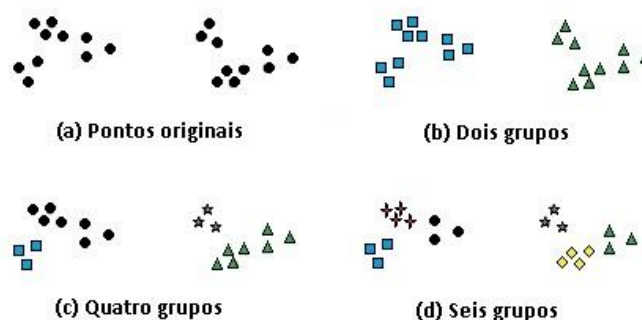


Figura 3 - Exemplo da divisão de dados em universo bidimensional.

Na Análise por Silhueta os coeficientes de Silhueta variam em um raio [-1,1], onde valores próximos a +1 indicam que a amostra está distante dos grupos adjacentes. Já valores próximos a 0 indicam que a amostra está na fronteira, ou muito próximo à, de dois grupos vizinhos. Já valores negativos indicam que a amostra está designada, possivelmente, ao grupo errado.

4. AMOSTRAS DE SOLO

As amostras de solo estudadas foram coletadas por Gobbi (2017) em 10 trincheiras (numeradas 1, 2, 3, 4, 5, 6, 7, 8, 9 e 10) dispostas na vertente da encosta estudada, conforme a Figura 4. Foram coletadas de 1 a 3 amostras de solo em todos os horizontes A, B e C identificados.

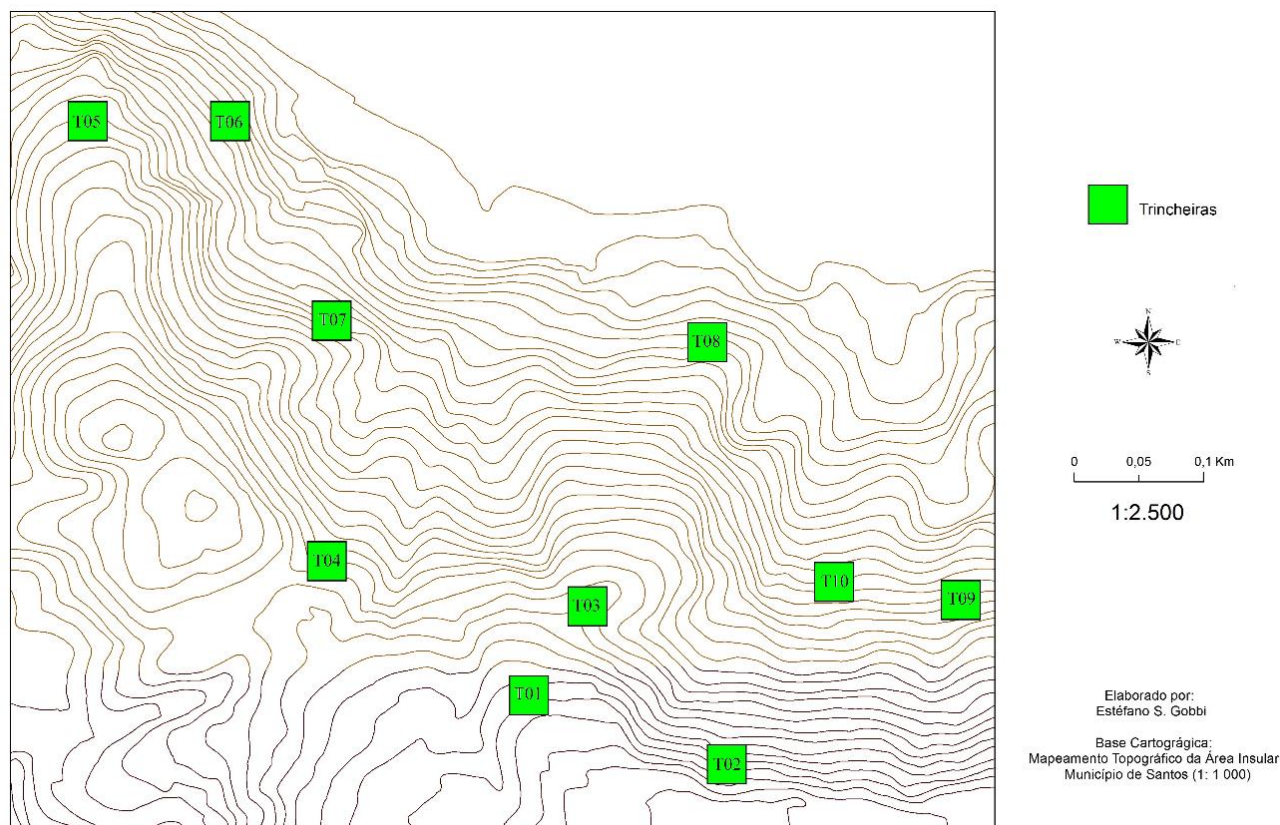


Figura 4 - Disposição das trincheiras na área de estudos (adaptado de Gobbi, 2017).

A composição química de cada amostra foi determinada segundo Camargo et al (2009) e apresentada em Gobbi (2017). Os parâmetros de Carbono (C), Capacidade de Troca Catiônica (CTC), Fósforo (P), Alumínio (Al), Magnésio (Mg), Cálcio (Ca), Sódio (Na) e Potássio (K), apresentados na Tabela 1, foram utilizados na realização deste estudo.

Tabela 1. Composição química de cada amostra de solo estudada (modificada de Gobbi, 2017)

Identificação	Al	Ca	Mg	K	Na	P ppm	C %	CTC cmol _c /kg
	cmol _c /kg							
TR 1 – A	0,1	8,0	1,5	0,56	0,20	43,0	3,5	12,66
TR 1 - Bt	1,0	0,7	0,4	0,04	0,13	2,0	0,6	6,67
TR 1 - Bt1	1,3	0,6	0,3	0,07	0,08	1,0	1,3	7,25
TR 1 - Bt2	1,0	0,7	0,4	0,05	0,12	1,0	0,8	6,87
TR 1 – Bt/C	1,2	0,5	0,3	0,04	0,10	7,0	0,4	5,74

TR 1 – C	0,4	1,0	0,5	0,04	0,21	2,0	0,5	4,75
TR 2 - A	0,3	5,2	1,3	0,38	0,23	69,0	3,1	13,11
TR 2 - Bt1	1,7	1,2	0,6	0,11	0,15	4,0	1,8	10,26
TR 2 - Bt2	1,6	1,0	0,4	0,07	0,12	2,0	1,2	9,19
TR 2 - C1	1,3	0,5	0,3	0,06	0,17	1,0	0,5	6,03
TR 2 - C2	2,0	0,6	0,4	0,07	0,19	1,0	0,6	6,86
TR 2 - C3	1,5	0,7	0,4	0,08	0,18	1,0	0,5	6,76
TR 3 – A	1,4	1,1	0,7	0,20	0,22	4,0	3,7	11,62
TR 3 - AB	1,8	0,4	0,2	0,07	0,13	1,0	2,1	8,60
TR 3 - Bt1	0,8	0,6	0,4	0,04	0,13	3,0	0,7	5,77
TR 3 - Bt2	0,6	0,4	0,3	0,04	0,17	4,0	0,4	4,91
TR 3 – C	0,6	0,7	0,4	0,03	0,26	3,0	0,5	5,19
TR 4 – A	1,1	1,1	0,6	0,22	0,12	17,0	3,2	13,04
TR 4 – B	2,0	0,4	0,2	0,04	0,10	41,0	1,4	12,74
TR 4 – C	1,0	0,3	0,1	0,06	0,11	98,0	0,7	6,97
TR 5 – A	2,3	0,6	0,4	0,21	0,09	13,0	2,9	12,70
TR 5 – B	2,0	0,3	0,2	0,07	0,07	1,0	0,8	9,64
TR 5 - A/B	1,6	0,3	0,2	0,08	0,06	1,0	1,1	9,84
TR 5 - C1	1,5	0,3	0,2	0,07	0,08	1,0	0,7	7,45
TR 5 - C2	1,8	0,4	0,2	0,07	0,07	1,0	0,5	6,74
TR 5 - C3	1,5	0,2	0,2	0,06	0,06	2,0	0,4	5,32
TR 6 – A	0,6	2,5	1,1	0,56	0,08	97,0	2,1	9,84
TR 6 - Bt1	1,5	0,5	0,3	0,18	0,03	22,0	0,8	8,01
TR 6 - Bt2	1,2	0,4	0,2	0,12	0,03	3,0	0,7	6,55
TR 7 – A	1,1	1,0	0,6	0,17	0,04	3,0	1,5	8,01
TR 7 - B1	1,8	0,7	0,4	0,09	0,05	1,0	0,6	6,84
TR 7 - B2	2,2	0,4	0,2	0,08	0,05	2,0	0,6	6,73
TR 7 - C1	2,8	0,6	0,4	0,06	0,07	1,0	0,5	5,53
TR 7 - C2	1,9	0,7	0,9	0,08	0,14	1,0	0,5	5,82
TR 8 – A	1,7	2,9	1,0	0,24	0,16	45,0	5,4	19,30
TR 8 - B1	1,5	0,7	0,4	0,09	0,15	2,0	1,7	8,94
TR 8 - B2	1,0	0,6	0,5	0,07	0,37	17,0	0,6	6,94
TR 8 - C1	1,0	0,4	0,3	0,04	0,52	98,0	0,6	5,66
TR 8 - C2	1,2	0,3	0,2	0,07	0,43	11,0	0,4	4,00
TR 8 - C3	1,5	0,3	0,2	0,06	0,40	3,0	0,4	3,56
TR 9 – A	1,1	1,3	0,6	0,25	0,08	2,0	2,8	12,43
TR 9 - B1	1,6	0,3	0,1	0,09	0,04	1,0	1,4	7,13
TR 9 - B2	0,9	0,6	0,4	0,08	0,07	1,0	0,6	5,15
TR 9 - B3	1,1	0,3	0,2	0,07	0,11	2,0	0,5	4,88
TR 9 – C	1,3	0,2	0,2	0,08	0,27	39,0	0,4	3,95
TR 10 - A1	1,1	2,0	1,1	0,45	0,11	19,0	2,8	12,26
TR 10 - A2	1,4	0,5	0,3	0,13	0,09	16,0	2,0	10,02
TR 10 - B1	1,2	1,0	0,4	0,15	0,07	9,0	0,7	8,62
TR 10 - B2	2,1	1,8	2,9	0,09	0,24	5,0	0,5	16,63
TR 10 - B3	2,6	1,2	2,5	0,07	0,29	12,0	0,3	13,26

5. RESULTADOS

Foram realizados testes através do algoritmo k-Means, para todos os parâmetros, e estes avaliados segundo o método de Análise por Silhueta. Inicialmente os testes foram feitos utilizando todos os parâmetros juntos, mas o algoritmo não foi capaz de identificar grupos consistentemente. Em seguida os testes foram individuais, definindo quais parâmetros eram melhor separados pelo algoritmo. Por fim foram realizados testes em duplas e trios, utilizando os parâmetros químicos que apresentaram as melhores respostas nos testes individuais. Dessa forma os resultados passaram a ser consistentes, sendo foi possível identificar os parâmetros que forneciam os melhores resultados na identificação dos perfis de solo.

Os parâmetros Al e Na não apresentaram bons resultados. A Figura 5 (a) exemplifica que a consistência dos grupos encontrados, para os testes utilizando apenas o Alumínio, é inicialmente baixa (pouco acima de 0,58), e apenas com 23 grupos, o valor se torna ótimo. Isso significa que as amostras só estão bem definidas em seus grupos quando o número de grupos utilizado é alto, o que gera grupos com poucas amostras. Na Figura 5 (b), para o Sódio, os valores médios de Silhueta ficaram em torno de 0,6, tornando-se ótimos apenas com 30 grupos. Esses valores de Silhueta ruins apresentados pelo Alumínio podem ser explicados pela sua presença na composição dos minerais microclina e oligoclásio, que constituem os granitos e migmatitos da área estudada, e que se mantém no perfil mesmo após a pedogênese, compondo aluminossilicatos, como apontado por Kerr (1952).

Na Figura 5 (c) as amostras apresentaram um valor elevado de consistência (acima de 0,75) com 2, 3, 4 e 5 grupos quando avaliado segundo seu conteúdo de Carbono. Isso significa que as amostras puderam ser definidas consistentemente em um número pequeno de grupos. O mesmo comportamento consistente na divisão de grupos foi observado nos demais parâmetros: P, CTC, K, Mg e Ca.

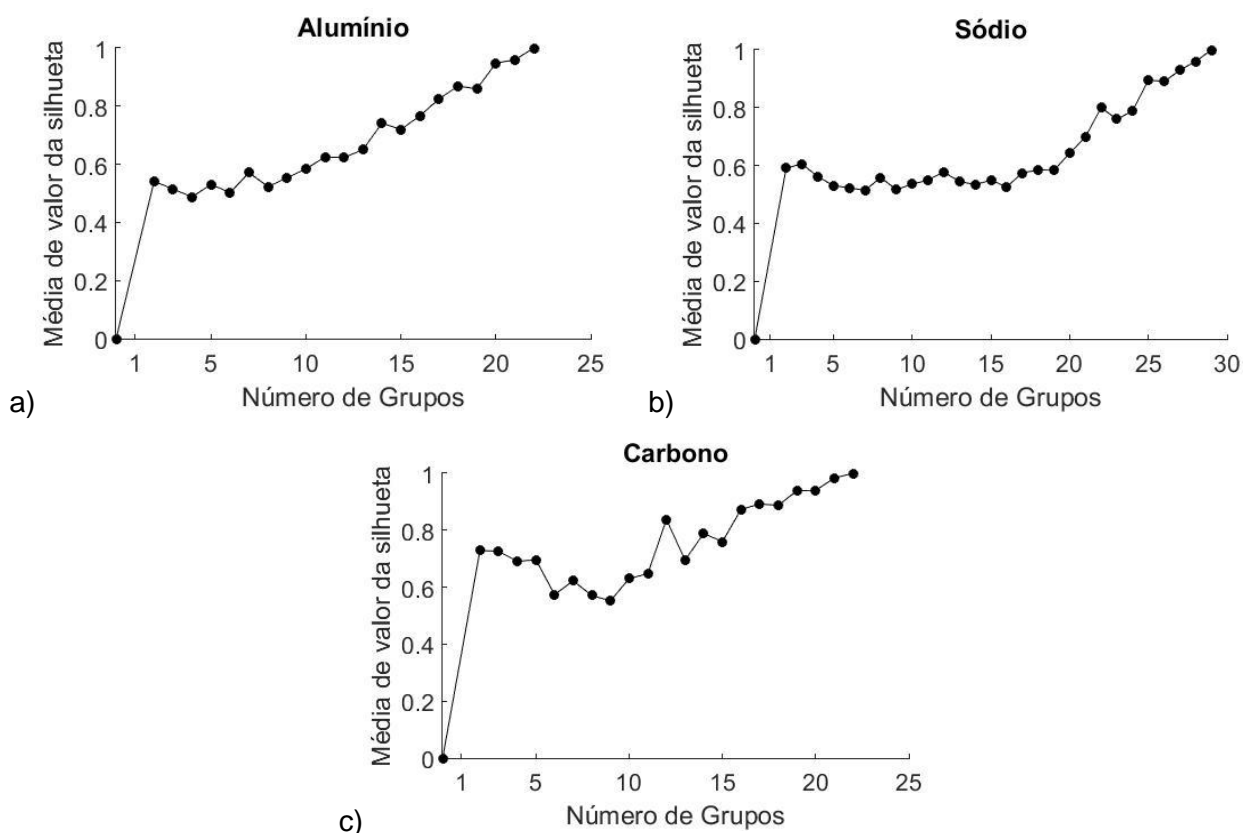


Figura 5 - Resultados dos valores de Silhueta para valores de grupos variando de 2 a 50, ao serem utilizados os parâmetros (a): Al; (b) Na e; (c) C.

Para os parâmetros de P, C e CTC, utilizando os dados de todas as amostras em conjunto, o algoritmo avaliou o valor de Silhueta para valores de grupos variando de 2 a 50. A Figura 6 (a) apresenta os resultados encontrados. Por ela é possível observar que o algoritmo foi capaz de identificar, com boa consistência, 2 e 3 grupos de horizontes de solo diferentes, com o valor de Silhueta próximo a 0,8. Valores de Silhueta próximo a 0,8 são obtidos novamente apenas com o número de clusters próximo a 35, o que indica que muitas amostras foram colocadas em grupos contendo apenas ela mesma.

O mesmo processo foi realizado em conjunto para os parâmetros de Ca, Mg e K, sendo o algoritmo capaz de encontrar com boa consistência 2 e 3 grupos de horizontes diferentes, com valores de Silhueta novamente próximos a 0,8. A Figura 6 (b) ilustra os resultados obtidos. É importante observar que os bons resultados obtidos para P, K e Mg levam a bons resultados também em CTC.

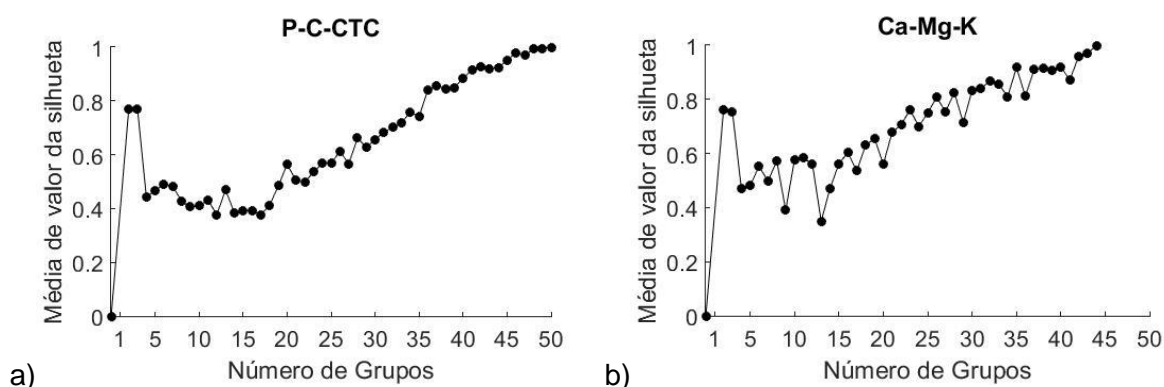


Figura 6 – Resultados dos valores de Silhueta para valores de grupos variando de 2 a 50, ao serem utilizados em conjunto os parâmetros (a): P, C e CTC; (b) Ca, Mg e K.

Em seguida foi realizada uma avaliação de todos os 6 parâmetros utilizados anteriormente (P, C, CTC, Ca, Mg e K) em conjunto, novamente variando o número de grupos de 2 a 50. O resultado apresentado na Figura 7 indica que para 2 e 3 grupos os valores foram próximos de 0,8.

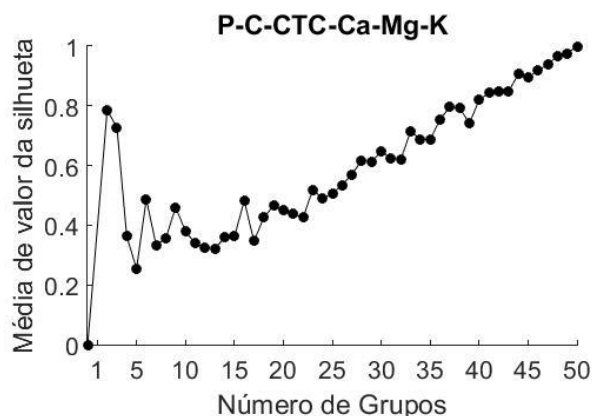


Figura 7 - Resultados dos valores de Silhueta para valores de grupos variando de 2 a 50, ao serem utilizados os parâmetros P, C, CTC, Ca, Mg e K em conjunto.

Gobbi (2017) explica que a diferenciação de C e Ca no horizonte A ser explicada pela presença de matéria orgânica, devido à cobertura vegetal existente na área estudada, e pela prática de calagem, realizada por moradores da região durante o plantio de mandioca. O algoritmo foi capaz de perceber essa variação entre os horizontes e diferenciá-los.

6. CONCLUSÕES

Nem todos os parâmetros químicos disponíveis puderam ser utilizados para a identificação de horizontes de solo. Foi necessário selecionar os parâmetros, já que alguns deles não apresentavam suficiente diferenciação entre as amostras de solo estudadas. Os elementos químicos Al e Na foram descartados dos testes por não apresentarem bons resultados na Análise de Silhueta. Já os parâmetros P, C, CTC, K, Mg e Ca puderam ser utilizados.

Determinados os parâmetros a serem utilizados, o algoritmo foi capaz de diferenciar consistentemente grupos dentre as amostras. Os resultados indicam que, por terem sido obtidos valores de Silhueta próximos a 0,8 apenas quando os dados foram divididos em 2 e 3 grupos, algumas amostras não estão bem diferenciadas entre si mesmo quando todos os 6 parâmetros são utilizados. Ainda assim, os resultados indicam que existem 2 ou 3 grupos diferentes dentre as amostras estudadas.

As amostras foram obtidas a partir de horizontes A, B e C de solo, assim, é esperado que o algoritmo seja capaz de diferenciar pelo menos 3 horizontes. O fato de, em alguns testes, o algoritmo indicar a possibilidade de haver 4 ou mais grupos dentre as amostras indica que, possivelmente, haja particularidades entre os horizontes de solos de trincheiras diferentes.

REFERÊNCIAS

- AB'SÁBER, A. N. (1955) *Contribuição à Geomorfologia do litoral Paulista*. In: Revista Brasileira de Geografia, ano XVII, nº1. Rio de Janeiro, p. 3-48.
- AB'SÁBER, A. N. (1965) *A evolução geomorfológica*. In: A Baixada Santista – Aspectos geográficos. Vol.1: p.49-66. Ed. da Universidade de São Paulo. São Paulo.
- ALMEIDA, F. F. M. de; CARNEIRO, C. D. R. (1998) *Origem e evolução da Serra do Mar*. In: Revista Brasileira de Geociências. 28 (2) São Paulo.p. 135-150.
- AUGUSTO FILHO, O. (1992) *Caracterização geológico-geogenética voltada à estabilização de encostas: uma proposta metodológica*. In Conferência Brasileira sobre Estabilidade de Encostas, 1, Rio de Janeiro. Anais... Rio de Janeiro: ABMS/ABGE/PUCRJ, 1992, v.2, p.721-733.
- AUGUSTO FILHO, O.; VIRGILLI, J. C. (1998) *Estabilidade de Taludes*. In: BRITO, S. N. A., Oliveira, A. M. S. (Ed.). Geologia de Engenharia. 1 ed. São Paulo: ABGE. Cap. 15., p.243-269.
- CASAGRANDE, A., (1948) "Classification and Identification of Soils," *Transactions*, ASCE, p. 901.
- CAMARGO, O.A.; MONIZ, A.C.; JORGE, J.A.; VALADARES, J.M.A.S. (2009) *Métodos de Análise Química, Mineralógica e Física de Solos do Instituto Agrônomo de Campinas*. Campinas, Instituto Agrônomo, 2009. 77 p. Boletim técnico, 106, Edição revista e atualizada.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). (2013) *Sistema Brasileiro de Classificação de Solos*. Rio de Janeiro. 3ª ed.
- ESTIVILL-CASTRO, V (2002). *Why so many clustering algorithms — A Position Paper*. ACM SIGKDD Explorations Newsletter. 4 (1). p. 65–75.
- GOBBI, E. S. Variação do manto intempérico nas encostas dos morros residuais em Santos-SP. Tese de Doutorado. IGe/UNICAMP – Campinas, SP: 2017. 414p.
- INSTITUTO DE PESQUISAS TECNOLÓGICAS DE SÃO PAULO (IPT) (1979) *Levantamento das condicionantes do meio físico e estabelecimento de critérios normativos para a ocupação urbana dos Morros de Santos e São Vicente (Carta Geotécnica)*. São Paulo.
- Kerr, P.F. (1952) *Formation and Occurrence of Clay Minerals*. Clays and Clay Minerals. 1 (1): 19–32.

- LARSEN, M. C. e A. J. TORRES-SÁNCHEZ (1998) *The Frequency and distribuion of recent landslides in three montane tropical regions of Puerto Rico*. *Geomorpholgy* 24: 309-331.
- MACQUEEN, J.B. (1967) *Some methods for classification and analysis of multivariate observations*. *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*. Vol. 1, p. 281–297. Berkeley: University of California Press.
- MATHWORKS, (2018) *k-Means Clustering*. Disponível em:
<<https://www.mathworks.com/help/stats/k-means-clustering.html>> Acesso em: 10 ago. 2017.
- NERY, T. D., (2014) *Susceptibility to Shallow Landslides in a Drainage Basin in the Serra do Mar, São Paulo, Brazil, Predicted Using the SINMAP Mathematical Model*. In: *Bulletin of Engineering Geology and the Environment*. 74(2).
- PACHAURI, A. K.; PANT, M. (1992) *Landslide Hazard Mapping Based on Geological Attributes*. *Engineering Geology*, 32, p. 81-100.
- SELBY, M. J. (1993) *Mass Wasting of Soils*. In: *Hillslope Materials and Processes*. p. 249-355. Oxford University Press.
- TOMINAGA, L. K. (2007) *Avaliação de Metodologias de Análise de Risco a Escorregamentos: Aplicação de um Sensaio em Ubatuba, SP*. Dpto. Geografia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo. Tese de Doutorado.