

XXV SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS

ANÁLISE MULTIVARIADA DE DADOS HIDROGEOQUÍMICOS PARA A IDENTIFICAÇÃO DE UNIDADES AQUÍFERAS

Débora Maria Diniz Barbosa¹; Alice Zupo Guimarães²; Rodrigo Sérgio de Paula³ & Leila Nunes Menegasse Velasquez⁴

Resumo: Para uma adequada gestão das águas subterrâneas, é relevante ter-se amplo conhecimento hidrogeoquímico da região, o que é comumente dificultado pela ausência de dados construtivos e litológicos dos poços presentes na área. A análise multivariada de agrupamento hierárquico vem sendo utilizada no tratamento de dados hidroquímicos. Os agrupamentos são realizados de acordo com as distâncias de dissimilaridade entre as concentrações obtidas para cada parâmetro de análise em cada ponto amostrado. A partir da análise por agrupamentos hierárquicos, é possível separar as águas analisadas por unidades aquíferas a que estão relacionadas. A análise por agrupamento hierárquico foi aplicada aos dados hidrogeoquímicos obtidos a partir da amostragem de água subterrânea em 10 pontos d'água localizados na região da APA Carste de Lagoa Santa e entornos. A partir do dendrograma gerado, que considera o Método de Ward e a distância Euclidiana, foram individualizadas 03 unidades aquíferas: águas relacionadas aos calcários da Formação Sete Lagoas, águas relacionadas aos metapelitos da Formação Serra de Santa Helena e águas relacionadas às rochas do embasamento granito gnáissico. Assim, comprovou-se a aplicabilidade do método para determinar unidades aquíferas, minimizando o problema da ausência de dados dos poços na caracterização hidrogeoquímica de uma área.

Palavras-Chave – Análise Multivariada, Hidrogeoquímica, Análise de Agrupamento Hierárquico

Abstract: For an adequate management of groundwater, it is important to have a broad hydrogeochemical knowledge of the region, which is commonly hampered by the absence of constructive and lithological data from the wells present in the area. Multivariate analysis of hierarchical clustering has been used in the treatment of hydrochemical data. The groupings are carried out according to the dissimilarity distances between the concentrations obtained for each analysis parameter at each sampled point. From the analysis by hierarchical groups, it is possible to separate the analyzed waters by aquifer units to which they are related. The analysis by hierarchical clustering was applied to the hydrogeochemical data obtained from the sampling of groundwater in 10 water points located in the APA Karst region of Lagoa Santa and surroundings. From the generated dendrogram, which considers the Ward Method and the Euclidean distance, 03 aquifer units were identified: waters related to the limestones of the Sete Lagoas Formation, waters related to the metapelites of the Serra de Santa Helena Formation and waters related to the rocks of the granite-gneiss basement. Thus, the applicability of the method to determine aquifer units was proved, minimizing the problem of lack of data from the wells in the hydrogeochemical characterization of an area.

Key-Words – Multivariate analysis, Hydrogeochemistry, Hierarchical Clustering Analysis.

1) Discente de Pós-Graduação do Departamento de Geologia do Instituto de Geociências da UFMG; Av. Pres. Antônio Carlos, 6627 - Pampulha, Belo Horizonte - MG, 31270-901; (31) 994025634; deboradiniz01@gmail.com

2) Discente de Pós-Graduação do Departamento de Geologia do Instituto de Geociências da UFMG; Av. Pres. Antônio Carlos, 6627 - Pampulha, Belo Horizonte - MG, 31270-901; (31) 996647566; alicezupo@gmail.com

3) Docente do Departamento de Geologia do Instituto de Geociências da UFMG; Av. Pres. Antônio Carlos, 6627 - Pampulha, Belo Horizonte - MG, 31270-901; (31) 991167905; rodrigo.spdm@yahoo.com.br

4) Docente do Departamento de Geologia do Instituto de Geociências da UFMG; Av. Pres. Antônio Carlos, 6627 - Pampulha, Belo Horizonte - MG, 31270-901; (31) 991234393; menegasse@yahoo.com.br

1- INTRODUÇÃO

O conhecimento hidrogeoquímico de uma área pode contribuir para o estabelecimento da adequada gestão do uso das águas subterrâneas, minimizando os riscos de ocorrência de impactos decorrentes da exploração desenfreada ou contaminação dos aquíferos. A amostragem hidroquímica de águas subterrâneas em poços e nascentes, associada à conhecimentos geológicos, é o método mais empregado para a caracterização hidrogeoquímica de uma área.

Uma dificuldade recorrente na caracterização hidrogeoquímica de uma área é a ausência de dados de perfis construtivos e litológicos dos poços de captação de água subterrânea existentes, principalmente quando se trata de instrumentos mais antigos. Nesse contexto, a análise multivariada de agrupamento hierárquico de dados hidroquímicos se apresenta como uma ferramenta bastante eficaz para auxiliar na definição de unidades aquíferas. Basta-se ter alguns poços com informações conhecidas para a separação de todos os demais em clusters que representarão as unidades aquíferas, por meio da análise de similaridade das propriedades físico-químicas das águas.

A análise de agrupamento hierárquico - ou análise de clusters - é uma técnica de estatística multivariada já amplamente utilizada em estudos relacionados à hidrogeoquímica (*e.g.* Vega *et al.*, 1998; Andrade *et al.*, 2008; Salgado *et al.*, 2011; Gomes & Cavalcante, 2017; Gomes & Nascimento, 2021; Santos, 2022.)

Na última década, diversas pesquisas foram realizadas na região da APA Carste de Lagoa Santa no âmbito da hidrogeologia (*e.g.* Vieira, 2018; De Paula, 2019; Ribeiro, 2020; Teodoro, 2020). Devido à complexa dinâmica de circulação de água na região da APA Carste de Lagoa Santa e a alta vulnerabilidade intrínseca dos aquíferos cársticos, associadas ao aumento da demanda por água decorrente do intenso desenvolvimento econômico da região nos últimos anos, torna-se cada vez mais relevante caracterizar a hidrogeologia da área. Assim, com o intuito de expandir os conhecimentos referentes às unidades aquíferas que os poços e nascentes presentes na região estão associados, aplicou-se a ferramenta de análise multivariada de agrupamento hierárquico a dados hidroquímicos de águas subterrâneas.

2- CONTEXTUALIZAÇÃO DA ÁREA DE ESTUDO

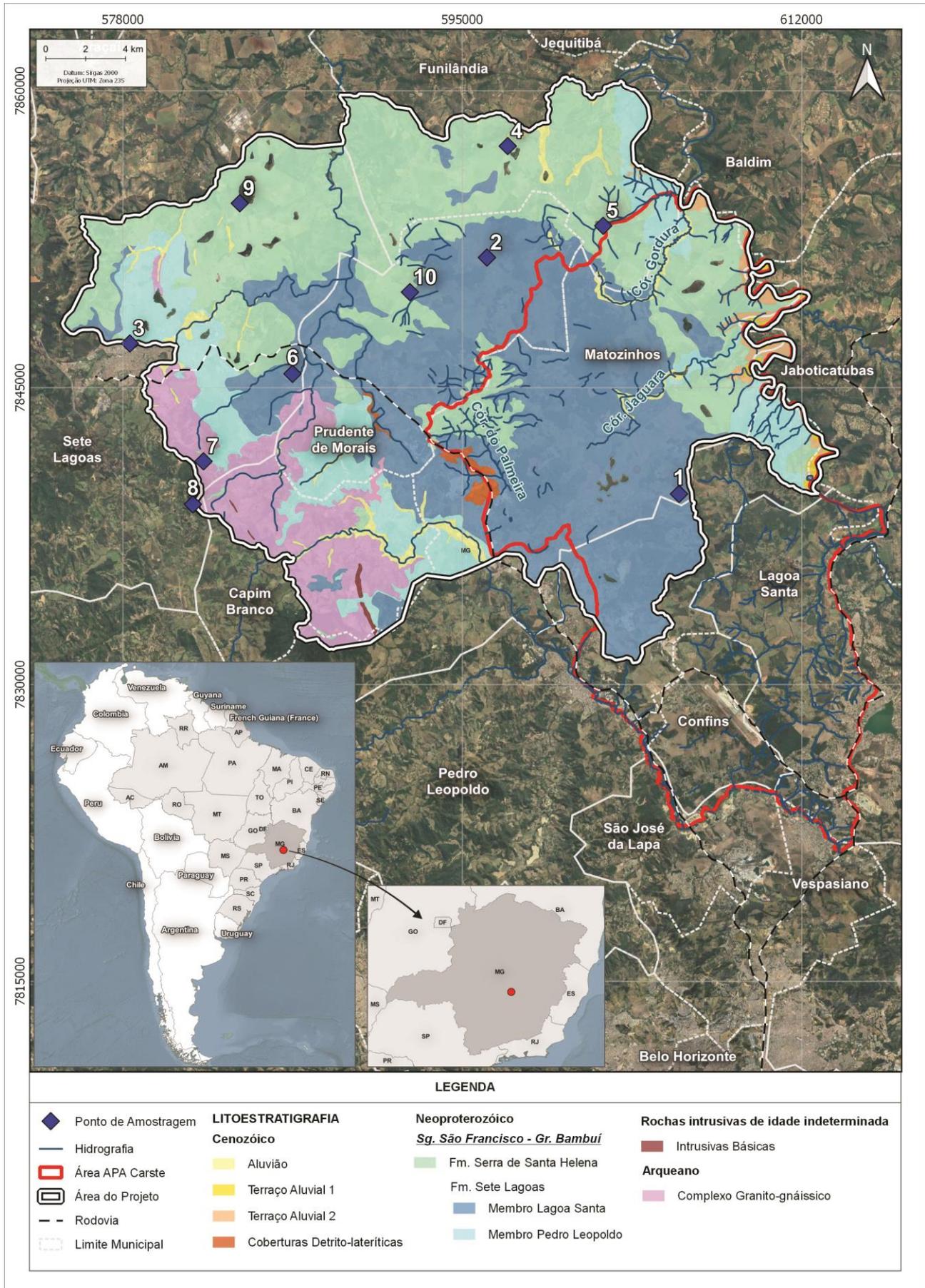
Localização

A área de estudo abrange a porção norte da APA Carste de Lagoa Santa e está localizada na região central do estado de Minas Gerais, a cerca de 60 km a norte de Belo Horizonte (Figura 1). Consiste numa área de aproximadamente 686 km², inserida na bacia do São Francisco, sub-bacia do rio das Velhas. O sistema de drenagem da região é predominantemente subterrâneo, com poucas ocorrências de cursos d'água superficiais.

Contexto Geológico e Hidrogeológico

Na área alvo, inserida na porção sul do Cráton São Francisco, o embasamento cristalino é sobreposto por sequências carbonáticas e metapelíticas do Grupo Bambuí, caracterizado na região por rochas das formações Sete Lagoas e Serra de Santa Helena. Depósitos aluvionares quaternários e terciários também ocorrem. A geologia da área pode ser observada no mapa da Figura 1.

Figura 1 – Mapa de localização, contexto geológico da área de estudo e pontos amostrados.



O embasamento cristalino é caracterizado por rochas granito-gnáissicas-migmatíticas do Complexo Belo Horizonte. O Grupo Bambuí consiste numa sequência de sedimentos plataformais neoproterozoicos depositados em ambiente marinho raso num período de intercalações de ciclos transgressivos e regressivos (Ribeiro *et al.*, 2003). A Formação Sete Lagoas subdivide-se em dois Membros: Pedro Leopoldo (inferior), composto por metacalcários com intercalações pelíticas, e Lagoa Santa (superior), constituído por metacalcários puros (Vieira *et al.*, 2018). Já a Formação Serra de Santa Helena é composta por rochas siliciclásticas metapelíticas, raramente carbonáticas (Ribeiro *et al.*, 2003).

Podem ser individualizadas quatro unidades hidrogeológicas distintas na área: Aquífero Cristalino, Aquífero Cárstico-Fissural, Aquitardo Serra de Santa Helena e Aquífero de Cobertura. O Aquífero Cristalino, constituído por rochas do embasamento, é do tipo fissural, com circulação de água condicionada às descontinuidades, bastante heterogêneo e anisotrópico. O Aquífero Cárstico-Fissural compreende as rochas calcárias da Formação Sete Lagoas, sendo predominantemente livre, semi-confinado quando sotoposto pelo Aquitardo Santa Helena, muito heterogêneo, anisotrópico e com alta produtividade (De Paula, 2019). O Aquitardo Serra de Santa Helena engloba as rochas metapelíticas da Formação Serra de Santa Helena, apresenta baixa porosidade fissural, baixa permeabilidade e alta capacidade de armazenamento. Já o Aquífero de Cobertura inclui os depósitos detrito-lateríticos e aluvionares recentes, tratando-se de um aquífero livre, heterogêneo e anisotrópico, do tipo granular, que funciona como zona de recarga dos aquíferos subjacentes (De Paula, 2019).

Amostragens e análises hidroquímicas de águas na região da APA Carste de Lagoa Santa já foram realizadas por autores como Vieira (2018) e Ribeiro (2020). De acordo com Vieira (2018), as águas relacionadas à Fm. Sete Lagoas apresentam concentrações maiores de Ca^{2+} e HCO_3^- . Águas do Membro Lagoa Santa possuem, geralmente, concentrações maiores de Ca^{2+} , HCO_3^- , SO_4^{2-} e F^- , e as do Membro Pedro Leopoldo de Mg^{2+} , Na^+ e Si^{4+} . A dificuldade de separá-las se dá devido aos diferentes tipos hidroquímicos relacionados ao Membro Pedro Leopoldo, reflexo da composição variada de suas rochas. Ainda segundo Vieira (2018), dentre todas as unidades aquíferas, as águas relativas ao embasamento granito-gnáissico são mais elevadas em Mg^{2+} , K^+ , Na^+ , Si^{4+} e Cl^- , enquanto as águas da Fm. Serra de Santa Helena exibem concentrações elevadas de K^+ e Si^{4+} , com ampla variação desses íons.

3- METODOLOGIA

Amostragem Hidroquímica

O levantamento de dados dos poços existentes na área foi realizado a partir das outorgas cadastradas no Sistema Integrado de Informação Ambiental (SIAM), de responsabilidade da Secretaria de Meio Ambiente e Desenvolvimento Sustentável (SEMAD) e considerando os poços cadastrados no Sistema de Informação de Águas Subterrâneas (SIAGAS), de responsabilidade do Serviço Geológico do Brasil (CPRM). Há aproximadamente 510 poços cadastrados na área de estudo, mas apenas cerca de 90 poços – menos de 20% do total - apresentam informações de perfil litológico e entradas d'água disponíveis.

A amostragem hidroquímica das águas subterrâneas na área de estudo foi realizada em junho de 2022, durante o período seco. Foram selecionados 10 poços de bombeamento distribuídos na região (Figura 1), que se encontravam em operação no período em que as coletas foram realizadas, sendo utilizada a própria bomba já instalada no poço para a exploração da água a ser coletada. Para a obtenção de águas representativas do aquífero, a amostragem foi realizada somente após 10 minutos de bombeamento. Os pontos de coleta considerados foram torneiras direto do poço, diretamente do cano e cano da caixa d'água, nesta ordem de prioridade. A amostragem das águas foi realizada com balde de aço inoxidável previamente ambientado por três vezes com a água a ser coletada.

Análises *in situ* de parâmetros físico-químicos (pH, Eh, condutividade elétrica, sólidos totais dissolvidos e temperatura da água e do ar) foram realizadas utilizando-se sonda multiparâmetro Ultrameter III 9PTKB Myron L Company. A sonda multiparâmetro também foi previamente ambientada com a água do ponto a ser amostrado.

As frascarias utilizadas nas coletas foram fornecidas pelo laboratório SGS Geosol, responsável pelas análises. Para as análises que requeriam, as amostras foram filtradas em membranas de 45 mesh e/ou adicionado ácido nítrico para sua preservação. Ressalta-se a utilização de luvas descartáveis durante as coletas, para evitar possíveis contaminações das amostras. Após a coleta, as amostras foram mantidas refrigeradas, em caixas contendo gelo, até serem entregues ao laboratório.

Análise de Agrupamento Hierárquico

A análise multivariada é uma importante ferramenta na manipulação de dados com muitas variáveis (Ferreira, 1996), consistindo em um conjunto de métodos estatísticos que permite a análise simultânea de medidas múltiplas para cada indivíduo ou objeto em análise. Há diversas técnicas para a análise multivariada a depender do objetivo de pesquisa específico. A metodologia para agrupamentos hierárquicos é a mais aplicada em geociências (Moreto *et al.*, 2010).

Os métodos hierárquicos particionam sucessivamente os dados, produzindo uma representação hierárquica dos agrupamentos (Everitt *et al.*, 2001). Nos métodos hierárquicos aglomerativos, inicialmente cada objeto forma o seu próprio agrupamento e, depois, os grupos vão sendo unidos até que um único agrupamento contendo todos os dados é gerado, ou seja, a definição dos agrupamentos é realizada de forma sequencial, iniciando-se com os mais simples pares de objetos até formar conjuntos maiores.

Segundo Moreto *et al.* (2010), no contexto geológico, o agrupamento dos objetos é mais bem realizado ao considerar as distâncias de dissimilaridade. Assim, a similaridade entre duas amostras será dada pela distância euclidiana, uma das técnicas mais utilizadas em análise de agrupamentos hierárquicos (Gauch, 1982), que pode ser representada pela diferença entre os valores analíticos das amostras. A distância euclidiana corresponde à medida da distância entre dois pontos baseado em suas coordenadas no plano cartesiano, sendo a distância geométrica entre os indivíduos (Hair, 2009). A distância euclidiana entre os elementos $P = (p_1, p_2, \dots, p_n)$ e $Q = (q_1, q_2, \dots, q_n)$ num espaço n -dimensional é obtida pela seguinte equação:

$$\text{Distância Euclidiana (DE)} = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (1)$$

$$DE = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

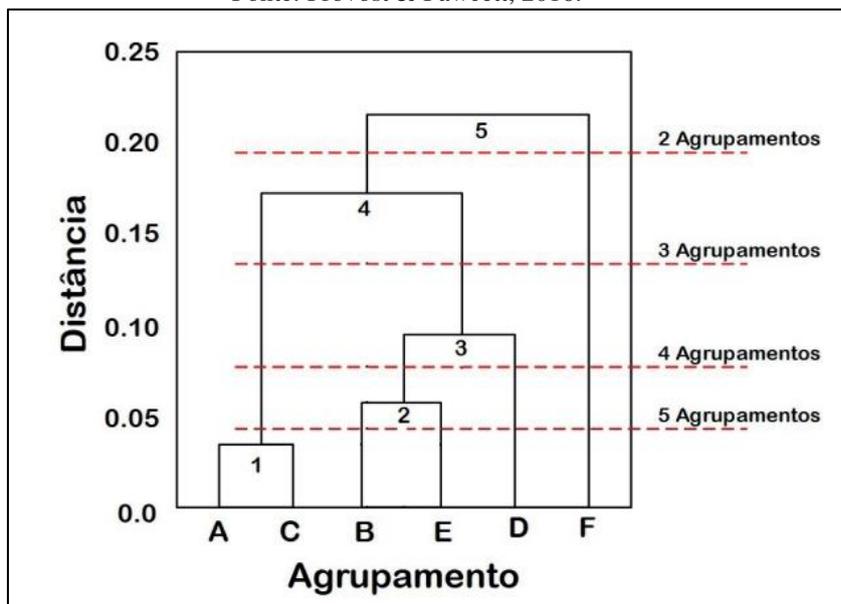
Como técnica para medir a distância entre os grupos, foi utilizado o método de Ward, também denominado método da mínima variância, que fornece os resultados mais coerentes com a realidade geológica (Monteiro *et al.*, 2010). Esse método consiste no agrupamento de objetos em grupos com a menor variância intragrupo. O método de Ward busca por partições que minimizem a perda de informação resultante de cada agrupamento (Ward, 1963). Essa perda pode ser medida pela diferença entre a soma dos erros quadráticos de cada padrão e a média da partição em que está contido. A Soma dos Quadrados dos Desvios (SQD) para cada agrupamento é dada pela equação a seguir, em que n é o número total de elementos e x_i é o i -ésimo elemento do agrupamento.

$$SQD = \sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2 \quad (2)$$

Como produto da análise multivariada de agrupamento hierárquico, será gerado um resumo gráfico da solução de cluster, denominado dendrograma. O dendrograma mostra explicitamente a hierarquia dos agrupamentos, permitindo a visualização dos dados em forma de árvore hierárquica, onde amostras semelhantes são agrupadas entre si, indicando também a ordem com que os dados foram agrupados. Quanto menor a distância entre os objetos, maior a semelhança entre eles.

Na Figura 2 é apresentado um dendrograma esquemático que exibe o agrupamento hierárquico para seis amostras: A, B, C, D, E e F (Provost & Fawcett, 2016). O eixo y representa a distância entre os agrupamentos. Na parte inferior ($y = 0$), os pontos estão separados uns dos outros. Conforme y aumenta, diferentes agrupamentos vão se formando: primeiro, A e C são agrupados, depois B e E, em seguida, o agrupamento BE é mesclado com D, e assim por diante, até que todos os agrupamentos estejam unidos no topo. As linhas de corte (linhas pontilhadas em vermelho) indicam o número de agrupamentos a ser considerado em cada caso, a depender do grau de detalhamento desejado.

Figura 2 – Dendrograma esquemático de agrupamento hierárquico para as amostras A, B, C, D, E e F.
Fonte: Provost & Fawcett, 2016.



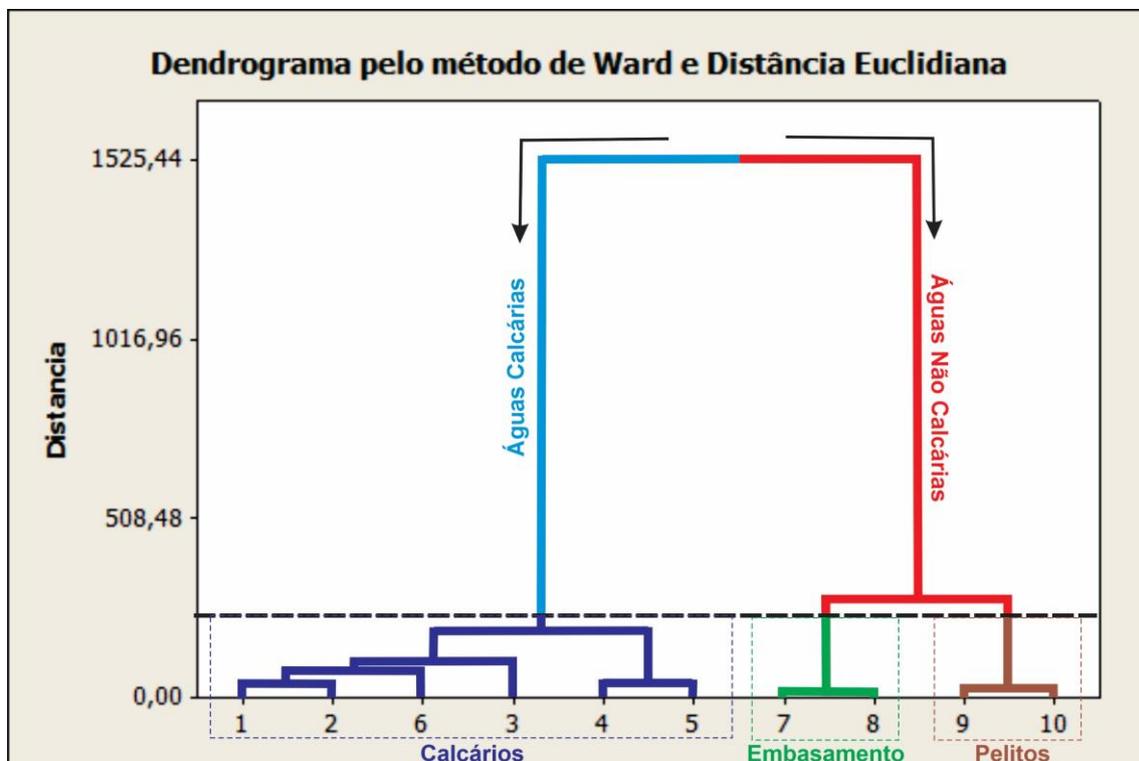
De modo sucinto, nesse estudo o agrupamento dos objetos será obtido por meio da aplicação da análise de agrupamento hierárquico, com a medida da dissimilaridade entre os indivíduos sendo dada pela distância euclidiana e a distância entre os grupos calculada a partir do método de Ward. As amostras serão agrupadas conforme a similaridade das propriedades hidrogeoquímicas das águas subterrâneas. A visualização dos clusters será expressa de forma gráfica por meio de um dendrograma elaborado a partir do *software* MINITAB versão 14, considerando o Método de Ward e utilizando a distância Euclidiana.

4- RESULTADOS E DISCUSSÃO

Os teores dos parâmetros alcalinidade de bicarbonato, cálcio dissolvido, cloreto, condutividade elétrica, magnésio dissolvido, nitrato, pH, potássio dissolvido, sílica dissolvida, sódio dissolvido, sólidos totais dissolvidos, sulfato e Eh obtidos para as 10 amostras analisadas foram utilizados para caracterizar as águas das diferentes unidades aquíferas presentes na área.

A partir da análise de cluster dos dados hidroquímicos realizada por meio do *software* MINITAB 14, foi gerado o dendrograma apresentado na Figura 3.

Figura 3 – Separação das unidades aquíferas por meio do dendrograma de dados hidroquímicos.



Ao analisar o dendrograma gerado, nota-se clara divisão entre águas calcárias, relacionadas aos calcários da Fm. Sete Lagoas, e águas não calcárias. As águas não calcárias podem ser separadas em dois grupos: águas relacionadas aos pelitos da Fm. Serra de Santa Helena e águas relacionadas ao embasamento granito-gnáissico. Dessa forma, a partir da interpretação do dendrograma foi possível individualizar três grupos de amostras, cada um correspondente a uma unidade aquífera distinta. São eles:

- i) Águas associadas à Fm. Sete Lagoas – Calcários;
- ii) Águas associadas à Fm. Serra de Santa Helena – Pelitos;
- iii) Águas associadas ao Complexo Belo Horizonte – Embasamento.

As águas calcárias se diferem das demais principalmente devido aos teores elevados de bicarbonato e cálcio. A alcalinidade de bicarbonato nas amostras que foram associadas ao calcário variou entre 216 e 159 mg/L, enquanto a concentração de cálcio dissolvido variou entre 94 e 60 mg/L. Nas amostras classificadas como não calcárias, os teores de bicarbonato e cálcio obtidos foram muito inferiores, variando entre 76 e 19 mg/L e entre 3 e 18 mg/L, respectivamente. Vale ressaltar que não foi possível separar as águas calcárias relacionadas aos membros Lagoa Santa e Pedro Leopoldo, apesar das rochas do Membro Pedro Leopoldo serem mais impuras.

As águas não calcárias, relacionadas aos pelitos da Formação Serra de Santa Helena e ao embasamento granito-gnáissico, apresentam, além de concentrações de bicarbonato e cálcio mais baixas, valores de pH, condutividade elétrica e sólidos totais dissolvidos também inferiores aos

observados nas águas calcárias. As águas associadas aos pelitos da Formação Serra de Santa Helena, em relação às águas de todas as demais unidades aquíferas, exibem os teores mais baixos de bicarbonato, cálcio, condutividade elétrica, pH, magnésio dissolvido e sólidos totais dissolvidos.

As águas relativas ao embasamento granito-gnáissico puderam ser individualizadas por apresentarem as maiores concentrações de Mg^{2+} , com variação entre 8 e 6 mg/L, e Na^+ , com intervalo de concentrações entre 5 e 6 mg/L nas amostras analisadas. Nas águas calcárias e nas águas relacionadas aos pelitos, os teores de Na^+ e Mg^{2+} não ultrapassam 3 mg/L. Os valores obtidos para sílica dissolvida também foram mais elevados, aproximadamente 50 mg/L, enquanto nas águas das demais unidades aquíferas a concentração de sílica não ultrapassa 20 mg/L. Por fim, os maiores valores de K^+ também foram observados nas águas associadas ao embasamento, sendo de cerca de 2 mg/L.

5- CONCLUSÕES

A aplicação do método de agrupamentos hierárquicos para individualizar amostras de água subterrânea de acordo com as dissimilaridades das suas propriedades físico-químicas se mostrou bastante satisfatória para a diferenciação das águas derivadas dos granitos e pelitos em relação às águas calcárias. A partir da interpretação do dendrograma, puderam ser individualizados 03 grupos, sendo: (i) águas associadas aos calcários da Formação Sete Lagoas; (ii) águas associadas aos pelitos da Fm. Serra de Santa Helena; e (iii) águas associadas às rochas do Complexo Belo Horizonte. O grupo (i) apresenta as maiores concentrações de bicarbonato e cálcio; o grupo (ii) caracteriza-se por apresentar os menores teores de bicarbonato, cálcio, condutividade elétrica, pH magnésio dissolvido e sólidos totais dissolvidos; e o grupo (iii) é caracterizado pelas maiores concentrações de Mg^{2+} , Na^+ , K^+ e sílica dissolvida.

Assim, o uso da análise multivariada, pelo emprego da técnica dos agrupamentos hierárquicos, mostrou-se bastante eficiente na interpretação de dados hidrogeoquímicos, permitindo a análise simultânea de medidas múltiplas para cada parâmetro e ponto analisado. Foi possível determinar as unidades aquíferas por meio da aplicação da análise dos agrupamentos hierárquicos, o que pode minimizar os problemas decorrentes da ausência de dados de perfis construtivos e litológicos dos poços utilizados na caracterização hidrogeoquímica de uma região.

AGRADECIMENTOS

Ao Grupo Gerdau pelo financiamento das pesquisas do Projeto de Processo FUNDEP/GERDAU/UFMG Nº 22.317, ao Centro Nacional de Pesquisa e Conservação de Cavernas (CECAV), ao Instituto de Geociências da Universidade Federal de Minas Gerais (IGC-UFMG), e ao Centro de Desenvolvimento de Tecnologia Nuclear (CDTN), ao Serviço Geológico do Brasil (CPRM) e ao Instituto Mineiro de Gestão das Águas (IGAM) pela participação e apoio durante todas as etapas envolvidas.

REFERÊNCIAS

ANDRADE, E.M., PALÁCIO, H.A., SOUZA, I.H., OLIVEIRA, L.R.A. & GUERREIRO, M.J. (2008). "Land use effects in groundwater composition of an alluvial aquifer (Trussu River, Brazil) by multivariate techniques". Environmental Research, vol. 106, no. 2, pp. 170-177. <https://doi.org/10.1016/j.envres.2007.10.008>.

DE PAULA, R.S. (2019). "Modelo conceitual de fluxo dos aquíferos pelíticos-carbonáticos da região da APA Carste de Lagoa Santa, MG". 279 f. Tese (Doutorado em Geologia) – Instituto de Geociências, Universidade Federal de Minas Gerais, Belo Horizonte.

- EVERITT, B.S., LANDAU, S.E LEESE, M. (2001). “*Cluster Analysis*”. 4 ed., Arnold.
- FERREIRA, D.F. (1996). “*Análise Multivariada*”. Lavras, Minas Gerais.
- GAUCH, H.G. (1982). “*Multivariate analysis in community ecology*”. Cambridge Univ. Press, Cambridge, U.K. 298 pp.
- GOMES, M.C.R. & CAVALCANTE, I.N. (2017). “*Aplicação da análise estatística multivariada no estudo da qualidade da água subterrânea*”. Revista Brasileira de Águas Subterrâneas, vol. 31, no. 1, pp. 134 -149. <https://doi.org/10.14295/ras.v31i1.28617>.
- GOMES, M.C.R. & NASCIMENTO, S.A.M. (2021). “*Caracterização das Águas Subterrâneas no Alto Cristalino de Salvador-Bahia com Auxílio da Estatística Multivariada*”. Anuário do Instituto de Geociências, vol. 44: 40600. https://doi.org/10.11137/1982-3908_2021_44_40600.
- HAIR, J.; BLACK, W.; BABIN, B.; ANDERSON, R.; TATHAM, R. (2009). “*Análise Multivariada de Dados*”. 6. ed. Porto Alegre: Bookman.
- MONTEIRO, A. B.; CORREIA FILHO, F. L.; FREIRE, P. K. C.; DE SOUZA, N. G.; SOARES FILHO, A. R. (2010). “*Análise Estatística Multivariada de Dados Hidroquímicos do Aquífero Cabeças – Sudeste da Bacia Sedimentar do Parnaíba – PI.*” In Anais do XVI Congresso Brasileiro de Águas Subterrâneas e XVII Encontro Nacional de Perfuradores de Poços. São Luís, MA.
- PROVOST, F.; FAWCETT, T. (2016). “*Agrupamento hierárquico*”. In: PROVOST, F.; FAWCETT, T. (Ed.) Data Science para Negócios. Rio de Janeiro, RJ: Alta Books. p.165-181.
- RIBEIRO, J. H.; TULLER, M. P.; FILHO, A.D.; PADILHA, A.V.; CÓRDOBA, C.V. (2003). “*Projeto VIDA: mapeamento geológico, região de Sete Lagoas, Pedro Leopoldo, Matozinhos, Lagoa Santa, Vespasiano, Capim Branco, Prudente de Moraes, Confins e Funilândia, Minas Gerais - Relatório final, escala 1:50.000*”. 2ª ed.; Belo Horizonte: CPRM, 54p., Mapas e anexos (Série Programa Informações Básicas para Gestão Territorial - GATE, versão digital).
- RIBEIRO, C.G. (2020). “*Compartimentação dos fluxos do sistema hidrogeológico cárstico do Grupo Bambuí a partir dos isótopos $2H$, $18O$ e $3H$ na região de Lagoa Santa, Minas Gerais*”. Dissertação de Mestrado, Instituto de Geociências da Universidade Federal de Minas Gerais, Belo Horizonte, 267 p.
- SALGADO, E.V., ANDRADE, E.M. DE, FONTENELE, S. DE B. & MEIRELES, A.C.M. (2011). “*Similaridade das variáveis hidroquímicas com o uso da análise multivariada, na Bacia do Salgado, Ceará*”. Revista Caatinga, vol. 24, no. 3, pp. 158-166.
- SANTOS, R.J.S. (2022). “*Análise multivariada e modelagem espacial de parâmetros hidrogeológicos dos sistemas aquíferos médio e inferior da bacia sedimentar do Cariri Cearense*”. 2022. 217 f. Tese (Doutorado em Engenharia Civil - Recursos Hídricos) - Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza.
- TEODORO, M.I.P. (2020). “*Hidrodinâmica das Zonas Cárstica e Cárstica-fissural do Sistema Aquífero Bambuí por meio de Traçadores Corantes e da Análise das Estruturas Rúpteis na região de Lagoa Santa, Minas Gerais*”. Dissertação de Mestrado, Instituto de Geociências da Universidade Federal de Minas Gerais, Belo Horizonte, 153 p.
- VIEIRA, L.C.M. (2018). “*Hidrogeoquímica dos Aquíferos da Região da APA Carste de Lagoa Santa, MG*”. Dissertação de Mestrado, Instituto de Geociências da Universidade Federal de Minas Gerais, Belo Horizonte, 111 p.
- VEGA, M., PARDO, R., BARRADO, E. & DEBÁN, L. (1998). “*Assesment of seasonal and polluting effects on the quality of river water by exploratory data analysis*”. Water Research, vol. 32, no. 12, pp. 3581-3592. [https://doi.org/10.1016/S0043-1354\(98\)00138-9](https://doi.org/10.1016/S0043-1354(98)00138-9).

WARD, J. (1963). *“Hierarchical grouping to optimize an objective function”*. Journal of American Statistical Association, v.58, p.236-244.