



QUIMIOMETRIA MULTIVARIADA APLICADA À AVALIAÇÃO DA MOBILIDADE DE ELEMENTOS TRAÇO EM PROCESSOS GEOQUÍMICOS

Raquel R. Reinehr¹, Renan Marim¹, Vitória P. Drügg¹, João B. Vargas², Tiago A. Siqueira², Andréa R. Jelinek², Rommulo V. Conceição² e Morgana B. Dessuy¹.

¹Universidade Federal do Rio Grande do Sul, Instituto de Química, Porto Alegre, RS, Brasil, 91501-970

²Universidade Federal do Rio Grande do Sul, Instituto de Geologia, Porto Alegre, RS, Brasil, 91501-970

e-mail: rahreinehr@gmail.com

Palavras-Chave: PCA, HCA, ICP-MS.

Introdução

A análise de matrizes multidimensionais complexas é um grande desafio na química analítica, especialmente em sistemas analíticos com dependência temporal e número de variáveis muito superior ao número de amostras, o que pode dificultar a extração de informações relevantes. O processamento de dados geoquímicos é dificultado por interações complexas entre as variáveis ao longo do tempo (Abdelmalek et al., 2024). Nesse contexto, o emprego de abordagens quimiométricas surge como alternativa atrativa para a interpretação de dados geoquímicos. Técnicas como a Análise de Agrupamentos Hierárquicos (HCA) e Análise de Componentes Principais (PCA), quando aliadas a um pré-processamento adequado, otimizam a extração e o processamento dessas informações. A integração de HCA e PCA estabelece um arranjo exploratório robusto, amplamente utilizado na literatura. Enquanto HCA agrupa amostras ou variáveis por similaridade, criando um dendrograma de proximidade, a PCA reduz a dimensionalidade da matriz, projetando a variância em componentes principais ortogonais (Rencher; Christensen, 2012).

Este estudo propõe a criação de uma metodologia para tratamento quimiométrico de dados analíticos de concentração de elementos traço em produtos de reações rocha-fluido-CO₂, com diferentes tempos reacionais (0, 5, 10, 20 e 30 dias). Os elementos traço presentes no fluido da reação foram quantificados por espectrometria de massa com plasma indutivamente acoplado (ICP-MS). As ferramentas PCA e HCA foram empregadas para avaliar as similaridades e diferenças no perfil de concentração de elementos traço ao longo do tempo de reação, permitindo identificar tendências à precipitação (remoção do fluido) ou à dissolução (enriquecimento na solução). O objetivo desse estudo é desenvolver um modelo para avaliar a mobilidade de elementos traço durante reações de armazenamento de CO₂ em rochas, bem como identificar a viabilidade do uso do perfil destes elementos para monitorar o avanço dos processos geoquímicos e compreender se a reação de mineralização de CO₂ atingiu o estágio de equilíbrio.

Material e Métodos

Realizaram-se reações rocha-fluido-CO₂ em condições controladas de temperatura e pressão, por 5, 10, 20 e 30 dias. Após a abertura do reator, separou-se a fase líquida da sólida por filtração a pressão reduzida. A fase fluida foi submetida a análise por ICP-MS equipado com uma célula de colisão com gás hélio e gás argônio como gás de nebulização, gás auxiliar e gás de *make up*. Foram quantificados os elementos: Al, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Y, Zr, Nb, Mo, Cd, Sn, Cs, Ba, Eu, Hf, Ta, W, Hg, Tl, Pb e U. As curvas de



calibração foram construídas utilizando uma solução padrão multielementar, em faixas de concentração de 0,015 a 3 $\mu\text{g L}^{-1}$ para o Hg e de 1,00 a 200 $\mu\text{g L}^{-1}$ para os demais analitos, com Rh como padrão interno.

Para a aplicação de ferramentas quimiométricas é necessário, em um primeiro momento, realizar uma filtragem e pré-processamento adequado da matriz, realizando uma triagem manual dos dados para remoção de elementos com concentração abaixo do limite de quantificação (LOQ) ou com comportamento de mobilidade que não diz respeito à interação rocha-fluido- CO_2 , além de pré-processamento para eliminar diferenças de magnitude e equalizar o peso estatístico das variáveis (Ivosev; Burton; Bonner, 2008).

A quantificação dos elementos nos fluidos de 5, 10, 20 e 30 dias de reação, além do fluido inicial, gerou uma matriz analítica com colunas de variáveis (elementos determinados) e 5 linhas de amostras (fluidos de diferentes tempos reacionais). A matriz de dados foi submetida a filtragem, excluindo valores que se encontravam abaixo do limite de quantificação (LOQ) do equipamento para cada elemento. Além disso, foram excluídos os elementos i) que não apresentaram variação na concentração ao longo do período reacional e ii) presentes na composição do reator. Como as concentrações variam muito entre elementos, realizou-se o pré-processamento de coluna por autoescalamento (*autoscaling*) para tornar os dados comparáveis, processo que consiste em subtrair a média da variável de cada valor individual e o dividir pelo desvio padrão. Utilizou-se o software MatLab para o processamento quimiométrico e visualização de dados, além da inteligência artificial Gemini para estruturação e otimização dos *scripts* de programação. Por fim, realizou-se modelagem exploratória com Análise de Agrupamento Hierárquico, HCA, por distância euclidiana, pré-processamento de dados por *autoscaling* e ligação pela distância média, centroide, completa e única, além de Análise de Componentes Principais, PCA, com 4 e 3 componentes principais, e validação cruzada do tipo *Venetian Blinds* com 5 grupos, utilizando pré-processamento por *autoscaling*.

Resultados e Discussão

Determinou-se os teores dos elementos traço presentes nos fluidos resultantes das reações de 0, 5, 10, 20 e 30 dias por ICP-MS. Os elementos As, Se, Sn, Eu, Hf e Ta apresentaram concentrações baixo do LOQ, para todos os tempos de reação, e, portanto, foram retirados da matriz de dados, sendo considerados elementos imóveis nas reações. Também foram retirados Ni, Mo, Cr, Fe, W, Co e Mn da matriz de dados, uma vez que estes elementos constituem a liga Hastelloy que compõe o reator e por isso podem sofrer influência que não corresponde à mobilidade elementar rocha-fluido. Assim, após triagem manual, têm-se a matriz composta por 5 linhas (5 amostras com diferentes tempos reacionais) e 18 colunas, referentes aos elementos Al, Ti, V, Cu, Zn, Rb, Sr, Y, Zr, Nb, Cd, Sb, Cs, Ba, Hg, Tl, Pb e U. Por questão de padronização, concentrações abaixo do LOQ, para esses elementos, em algum tempo reacional, foram consideradas como zero.

A Análise de Agrupamento Hierárquico (HCA) foi realizada aplicando distância euclidiana, pré-processamento de dados por *autoscaling* e ligação pela distância média entre todos os pares de pontos dos grupos (Rencher; Christensen, 2012). O resultado encontra-se na Figura 1, onde é possível verificar que as amostras inicial e de 5 dias apresentam maior similaridade, assim como as amostras de 20 e 30 dias, mas esses dois grupos apresentam grande diferença entre si. O mesmo comportamento foi observado quando utilizada ligação centroide, completa ou única. Assim, é possível perceber que as maiores modificações na concentração dos elementos traço ocorrem no intervalo entre 5 e 20 dias, uma vez que há uma maior

similaridade entre as concentrações inicial e de 5 dias, possivelmente por uma cinética lenta. Apesar de, visualmente, haver similaridade entre 20 e 30 dias, a divergência entre eles ainda é suficiente para levantar a hipótese de que não se atingiu o equilíbrio, uma vez que há variação composicional entre estes tempos reacionais, sugerindo necessidade de estender-se o tempo experimental de reação para mapear a estabilização termodinâmica do sistema.

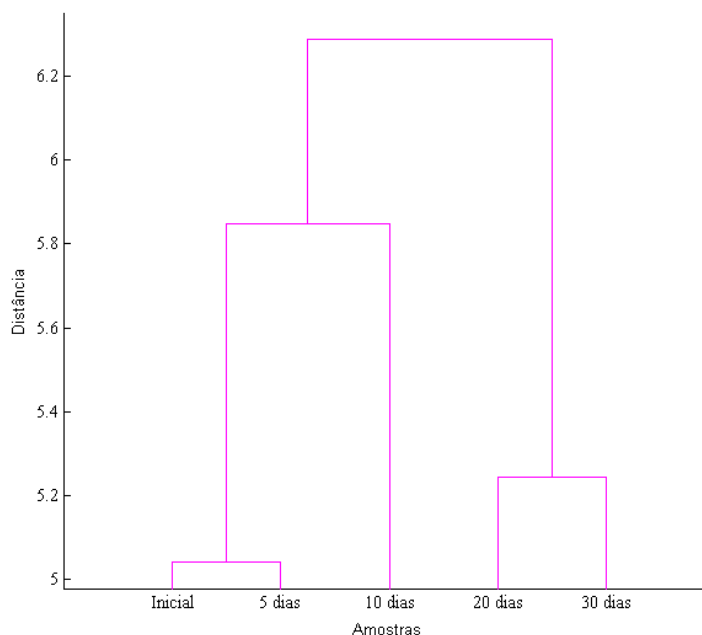


Figura 1: Dendrograma (HCA) das amostras de fluido inicial, com 5, 10, 20 e 30 dias de reação.

A seleção de número ótimo de componentes principais (PCs) no modelo PCA foi realizada por validação cruzada do tipo *Venetian Blinds* com 5 grupos, utilizando pré-processamento por coluna centralizando cada variável na média e a dividindo pelo seu desvio padrão. Quatro componentes principais explicam 100% da variância, enquanto três componentes principais explicam 85,32% da variância.

A PC 1 explica 39,36% da variação e separa muito bem as amostras inicial e 5 dias, no lado negativo, das amostras de 20 e 30 dias, no lado positivo (Figura 2), refletindo o resultado obtido na HCA. Os *loadings* (pesos), ou seja, os elementos que mais contribuem positivamente à PC1 são vários, com grande influência de Tl, Ba, Sb, Nb, Sr, Cu. Ao lado negativo da PC 1 tem-se fortemente Cs além de Pb, Zn e Ti. Já a PC 2 explica 28,45% e separa a amostra de 10 dias fortemente ao lado negativo e a amostra de 30 dias suavemente. Ao lado positivo têm-se fortemente 20 dias e a amostra inicial, com *loadings* Hg, Sb, Cd e Cu, ao passo que os *loadings* negativos são Zr, Y, Ti e Al. Já a terceira PC consegue separar muito bem as amostras inicial, 10 e 20 dias (Hg, Cu, Ti e Al) das amostras de 5 e 30 dias (Pb, Rb, Zn, U, Sr e V), explicando 17,51% da variação. A PC 4 explica 14,67% da variância, mas não traz muita informação nova, apenas separa melhor as amostras de 20, 5 e 10 dias ao lado negativo da PC 4 (Cd, Pb, Hg, Al) e positivamente as amostras inicial e de 30 dias (fortemente U e V). Assim, apesar de explicar uma porcentagem expressiva da variância observada, a quarta componente principal foi desconsiderada em prol da parcimônia do modelo. A Figura 2 contém a distribuição espacial

dos *scores* das 3 PCs, ou seja, as coordenadas de cada amostra no espaço projetado pelas componentes principais, enquanto a Figura 3 apresenta os *loadings* das três componentes principais, o grau de contribuição de cada variável original nas componentes principais.

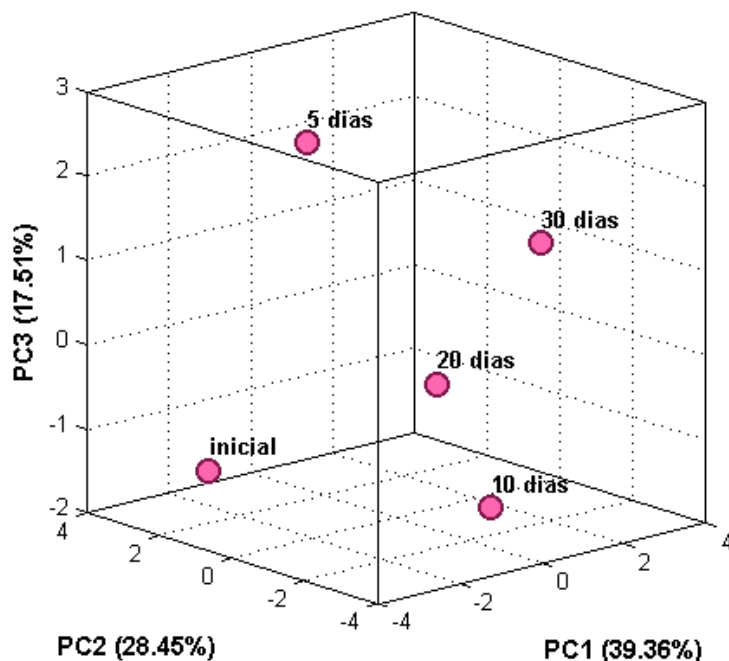


Figura 2: Projeção tridimensional dos *scores* (PC 1 x PC 2 x PC 3) da PCA.

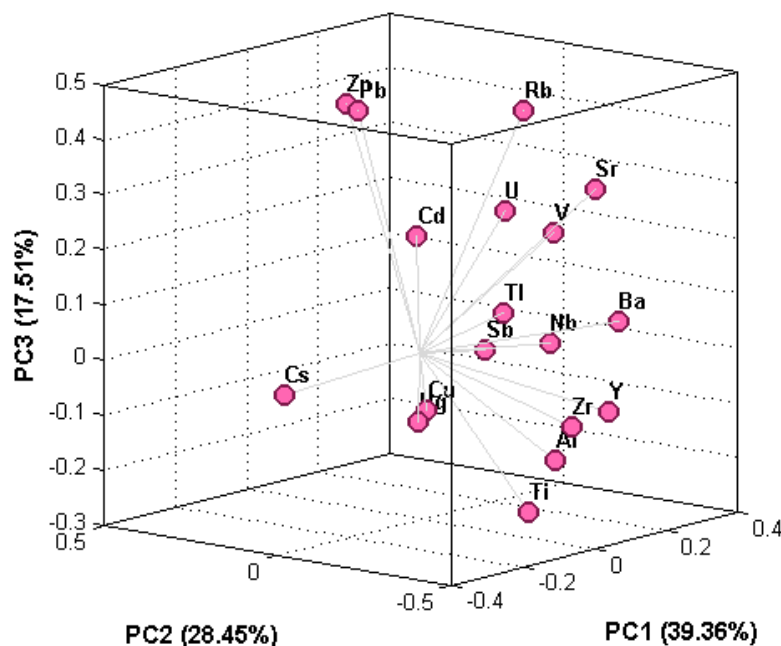


Figura 3: Projeção tridimensional dos *loadings* (PC 1 x PC 2 x PC 3) da PCA.

Com base nos resultados da análise das componentes principais dos fluidos percebe-se uma tendência em estágios iniciais de dissolução de elementos como Zn, Pb, Rb e Sr, que fortemente influenciam a PC 3, ou seja, tem alta concentração no fluido de 5 dias de reação. Já



elementos como Cu e Hg apresentam dissolução tardia, havendo um forte aumento de sua concentração aos 20 dias de reação, observados no quadrante positivo da PC 1, e elementos como V e U que apresentam mobilização máxima aos 30 dias de reação, com um aumento de sua concentração no fluido em tempos de reação maiores.

O decréscimo da concentração dos elementos no fluido, com o avanço da reação, indica que esses estão precipitando, ou seja, são incorporados e imobilizados na fase sólida. Pb e Zn apresentam elevadas concentrações aos 5 dias de reação, após esse tempo sofrem uma queda brusca, o que indica que são solubilizados em um primeiro contato com o fluido e, em tempos maiores, são imobilizados. Já o Hg e Cu apresentam teores mais altos aos 20 dias, mas são quase que totalmente removidos da fase líquida aos 30 dias de reação, dinâmica refletida na PCA pela perda de influência desses elementos nos *loadings* das PC 1 e PC 2, conforme o sistema avança em direção ao equilíbrio. Cs apresentou perfil de remoção contínua ao longo do experimento, chegando a valores muito baixos aos 30 dias de reação.

Assim, apesar de explicar 85,32% da variância total dos dados com as 3 primeiras componentes principais, a PCA permite correlacionar muito bem as amostras com diferentes tempos de reação, além de verificar importantes dinâmicas de dissolução e precipitação de elementos traço, tornando-se um modelo fundamental para elucidar a evolução do sistema.

Conclusões

A metodologia quimiométrica desenvolvida cumpre o objetivo de transformar dados brutos de ICP-MS em um modelo exploratório que permita a interpretação química das interações rocha-fluido-CO₂ ao longo de diferentes tempos reacionais. A etapa prévia de tratamento dos dados permitiu a remoção de variáveis que viriam a interferir na aplicação das ferramentas quimiométricas, seja pela concentração constante abaixo do limite de detecção ou pela inserção de comportamento que não diz respeito à mobilidade rocha-fluido como a contribuição de elementos que compõem o reator. O pré-processamento também teve grande importância, uma vez que permite que variações discretas na concentração de alguns elementos com baixa concentração tenham a mesma relevância estatística que elementos com maior presença.

Os resultados obtidos com a HCA e PCA permitiram identificar padrões de similaridade temporal no fluido e verificar esta evolução com base na concentração dos elementos traço. Apesar das três componentes principais explicarem cerca de 85% da variância total, a quarta componente foi desconsiderada em prol da parcimônia do modelo, mantendo a capacidade explicativa sem adicionar complexidade desnecessária. A primeira componente principal descreveu a maior variância observada, cerca de 39%, e foi de encontro ao observado no dendrograma da análise de agrupamentos hierárquicos, e, ao adicionar outras duas componentes, a interpretação dos gráficos de *scores* e *loadings* de três componentes principais revelaram comportamentos de dissolução e precipitação claros ao longo das amostras, diferenciando satisfatoriamente todos os tempos reacionais.

Portanto, a capacidade de integrar as variações na concentração de elementos traço ao longo dos 30 dias de reação, com parcimônia e relevância estatística, valida a abordagem quimiométrica adotada para tratamento de dados analíticos de ICP-MS, auxiliando na elucidação da evolução geoquímica do sistema.

Agradecimentos

Agradecimentos à CAPES e PPGQ-UFRGS.



Referências

ABDELMALEK, Drouiche *et al.* Groundwater quality assessment using revised classical diagrams and compositional data analysis (CoDa): Case study of Wadi Ranyah, Saudi Arabia. **Journal of King Saud University - Science**, v. 36, n. 10, p. 103463, nov. 2024.

IVOSEV, Gordana; BURTON, Lyle; BONNER, Ron. Dimensionality Reduction and Visualization in Principal Component Analysis. **Analytical Chemistry**, v. 80, n. 13, p. 4933–4944, 1 jul. 2008.

RENCER, Alvin C.; CHRISTENSEN, William F. **Methods of multivariate analysis**. 3. ed ed. Hoboken, New Jersey: Wiley, 2012.