
Use of the fuzzy algorithm in samples with imprecise descriptions to define typological domains

Uso do algoritmo fuzzy em amostras com descrições imprecisas para definição de domínios tipológicos

Received: 15-06-2024 | Accepted: 19-07-2024 | Published: 23-07-2024

Silvânia A Braga de Castro

ORCID: <https://orcid.org/0000-0002-1343-660X>

Centro Federal de Educação Tecnológica de Minas Gerais, Brasil

E-mail: silvaniabraga@cefetmg.br

Elenice Maria Schons Silva

ORCID: <https://orcid.org/0000-0003-1360-6450>

Universidade Federal de Catalão, Brasil

E-mail: eschons@ufcat.edu.br

Leandro Resende Mattioli

ORCID: <https://orcid.org/0000-0001-9626-362X>

Centro Federal de Educação Tecnológica de Minas Gerais, Brasil

E-mail: leandromattioli@cefetmg.br

André Carlos Silva

ORCID: <https://orcid.org/0000-0002-9760-0728>

Universidade Federal de Catalão, Brasil

E-mail: ancarsil@ufcat.edu.br

ABSTRACT

The identification of homogeneous spatial domains is an important step in mineral resource modeling. Inspired by the growing integration of machine learning into geostatistical modeling, this study addresses the fuzzy algorithm for automated definition of typological classes, thus improving the reliability of information. The biggest challenge in building the 3D model is optimizing the time spent grouping the domains since 53% of the samples in the database have inaccurate lithological descriptions. A combination of statistical analyzes of grades and geological characteristics was carried out to set up the fuzzy inference system based on expert knowledge. The results were verified by comparing the clustering obtained with other machine learning techniques. The algorithm proved to be effective in maintaining a better separation of typologies. Other benefits achieved using computational intelligence were the quantification of uncertainty in lithological descriptions; replicability and standardization of concepts for defining groups and automation of steps, allowing agile updating of the model when new data is inserted.

Keywords: Typological domains; Geological uncertainty; Computational intelligence.

RESUMO

A identificação de domínios espaciais homogêneos é uma etapa importante na modelagem de recursos minerais. Inspirado na crescente integração do aprendizado de máquina à modelagem geoestatística, esse estudo aborda o algoritmo fuzzy para definição automatizada das classes tipológicas, visando melhorias na confiabilidade da informação. O maior desafio na construção do modelo 3D é otimizar o tempo gasto no agrupamento dos domínios uma vez que 53% das amostras do banco de dados se encontram com descrições litológicas imprecisas. Uma combinação entre as análises estatísticas dos teores e as características geológicas foi realizada para montar o sistema de inferência fuzzy com base no conhecimento especialista. Os resultados foram verificados comparando o agrupamento obtido com outras técnicas de aprendizado de máquina. O algoritmo se mostrou eficaz ao manter uma melhor separação das tipologias. Outros benefícios atingidos utilizando a inteligência computacional foi a quantificação da incerteza das descrições litológicas; replicabilidade e padronização dos conceitos para a definição dos grupos e automação das etapas permitindo atualização ágil do modelo quando houver inserção de novos dados.

Palavras-chave: Domínios tipológicos; Incerteza geológica; Inteligência computacional

INTRODUÇÃO

Incluir informações geológicas, mineralógicas e metalúrgicas para construir um modelo de recursos é uma abordagem necessária para sucesso do planejamento de mina e do processamento mineral (WILLIAMS & RICHARDSON, 2004; KNIGHT et al., 2011; VANN et al., 2011).

A elaboração dos modelos ocorre após as etapas de amostragem e ensaios de laboratórios, e visa a predição de informação em lugares onde não se têm dados por metodologias convencionais de interpolação ou por métodos geoestatísticos (LOZANO & BENNET, 2003). As técnicas geoestatísticas mais comuns são baseadas em fortes suposições de estacionaridade de segunda ordem, ou seja, a média, a variância e a covariância dos atributos permanecem constantes em todo o domínio e não dependem da localização dos pontos, mas apenas da distância entre eles (LARRONDO & DEUTSCH, 2004).

Portanto, a definição dos domínios é uma das primeiras tarefas na modelagem (BRAGA & COSTA, 2016) e estes devem ser escolhidos como zonas estatisticamente homogêneas, geologicamente representativas (VANN et al., 2011; ROMARY et al., 2015), e com dados suficientes para permitir uma inferência viável das estatísticas necessárias (LARRONDO & DEUTSCH, 2004).

Técnicas como modelagem de seções paralelas, modelagem implícita e métodos geoestatísticos avançados são utilizados para elaboração dos modelos (NEWTON, 2011) mas, na prática, geralmente é um exercício manual que exige muito trabalho e está sujeito a erros de julgamento subjetivos e a um resultado amplamente determinístico que ignora a incerteza associada à interpretação (ABILDIN et al., 2023). Quando as litologias são bem estabelecidas e correlacionadas com a mineralização, a delimitação das unidades ocorre de forma direta a partir da descrição amostral (MOREIRA, 2020). Porém, sobreposições de estruturas durante e após o processo genético de formação das rochas associado a uma quantidade crescente de informação podem tornar a tarefa ainda mais dispendiosa (LARRONDO & DEUTSCH, 2004; DUNHAM & VANN, 2007). O uso da matemática, da estatística e da computação para tratar problemas de dimensionalidade dos bancos de dados e agrupar amostras de acordo com suas similaridades são contribuições significativas para a extração do conhecimento (NETO & MOITA, 1998). Técnicas de agrupamento são muito utilizadas para análise exploratória de dados por serem ferramentas confiáveis e rápidas na identificação de domínios geológicos (RAJABINASAB & ASGHARI, 2018). Esses métodos têm por finalidade dividir as

amostras de modo que os grupos sejam diferentes uns dos outros, mas os elementos pertencentes ao mesmo grupo sejam homogêneos e parecidos entre si (IZBICKI & SANTOS, 2020). Existem duas classes de agrupamento: agrupamento difuso e o rígido. O rígido é uma forma de agrupamento em que cada elemento pode pertencer a somente um grupo (cluster), caso contrário é o agrupamento difuso como, por exemplo, o algoritmo fuzzy (FELIX, 2018).

O k-means é um agrupamento do tipo rígido, muito aplicado devido sua facilidade e simplicidade de implementação (YASOJIMA, 2020); seu objetivo é diminuir a variabilidade intragrupos (SANCHES, 2003). O agrupamento hierárquico segue uma sequência de rodadas com duas abordagens possíveis: aglomerativo e divisivo. Assim como no K-means os objetos são comparados com base em uma medida de distância. É importante destacar que grande parte dos métodos de agrupamento tradicionais como k-means e o hierárquico não levam em consideração aspectos relacionados a continuidade espacial dos dados e a incerteza dos grupos, fator importante para banco de dados de geociências. Somente parâmetros estatísticos são utilizados nessas técnicas de agrupamento (ROSSI & DEUTSCH, 2014).

O agrupamento difuso é baseado na manipulação matemática de conjuntos nebulosos. Zadeh (1965) introduziu a teoria de objetos matemáticos que chamou de conjuntos nebulosos, cujo limites não são precisos, contrário ao conceito clássico de conjunto. Ou seja, a mudança de não-pertencimento para pertencimento a um conjunto difuso pode ser gradual em vez de abrupta. Essa mudança gradual é expressa por uma função de pertinência, que caracteriza completa e exclusivamente um determinado conjunto fuzzy ou conjunto nebuloso (KOVACIC & BOGDAN, 2006; ALMEIDA & EVSUKOFF, 2003). Segundo Almeida e Evsukoff (2003), na teoria clássica de conjuntos, o tratamento de classes de objetos e suas inter-relações em um universo são definidas. Por exemplo, pode-se definir um universo U discreto que reúne todos os números entre -10 e 10 do conjunto Z dos números inteiros. Os objetos que possuem características semelhantes são agrupados em conjuntos que consistem em uma coleção de objetos ou elementos do universo de discurso. Então, poderíamos definir um conjunto A , que reúne os elementos de U positivos, pela enumeração de todos os seus elementos:

$$A: \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \quad (1)$$

Já um conjunto fuzzy F de um universo de discurso U é caracterizado por uma função característica real, comumente chamada de função de pertinência, mapeada por $\mu_F: U \rightarrow [0,1]$ que associa a cada $x \in U$ um número real $\mu_F(x)$ no intervalo $[0,1]$

representando o grau de pertinência de x em F . A ilustração de conjuntos fuzzy depende basicamente da natureza e da dimensão do universo de discurso. Uma representação analítica é mostrada em (2) para o caso de um “conjunto dos números pequenos (P) entre - 5 e 5”, no universo de discurso U (ALMEIDA & EVSUKOFF, 2003):

$$\mu_P(x) = \{0.0/-10; 0.0/-9; 0.0/-8; 0.0/-7; 0.0/-6; 0.0/-5; 0.2/-4; 0.4/-3; 0.6/-2; 0.8/-1; 1.0/0; 0.8/1; 0.6/2; 0.4/3; 0.2/4; 0.0/5; 0.0/6; 0.0/7; 0.0/8; 0.0/9; 0.0/10\} \quad (2)$$

Neste caso, são representados todos os elementos do universo de discurso separados por ponto e vírgula. Cada elemento deste universo é representado na forma $\mu_P(x_i) / x_i$, onde o primeiro termo representa o grau de pertinência de x_i no conjunto P e o segundo termo identifica o próprio elemento x_i . No mundo real (e na maioria das aplicações na área de engenharia) existem propriedades que são vagas, incertas ou imprecisas e, desta forma, impossíveis de se caracterizar por predicados da lógica de conjuntos.

Na maioria dos depósitos minerais o pertencimento ou não das litologias a um determinado domínio também é gradual, não sendo possível determinar um limiar ou contato nítido que individualize as litologias. Na mina em estudo esse fenômeno é observado e o comportamento estrutural da jazida impõe grandes desafios para as alternativas de agrupamento. Por essa razão a teoria dos conjuntos fuzzy foi utilizada na tentativa de apresentar uma abordagem viável para lidar com a incerteza do fenômeno geológico e com as amostras com descrições ambíguas.

ÁREA DE ESTUDO, GEOLOGIA E MINERALIZAÇÃO

A Mina em estudo está posicionada no Complexo Catalão I, no estado de Goiás-GO. As concentrações supergênicas de apatita residual, pirocloro, monazita, minerais de titânio (perovskita, ilmenita e anatase) e vermiculita são resultantes da ação intempérica. Das principais concentrações minerais do complexo, somente a apatita e o pirocloro estão sendo atualmente explorados.

A jazida tem estrutura dômica com aproximadamente 6 km de diâmetro e foi formada por um conjunto de magmas neocretácios alcalinos intrusivos em metassedimentos neoproterozóicos da Faixa Brasília entre a borda NE da Bacia do Paraná e a borda SW do Craton do São Francisco (GIBSON et al., 1995). As rochas intrusivas pertencem a três séries petrogenéticas distintas: bebedourítica, foscorítica e carbonatítica (BROD et al., 2004), intimamente relacionadas entre si por uma complexa combinação de cristalização fracionada, imiscibilidade de líquidos e desgaseificação/metassomatismo

(CORDEIRO et al., 2011a; RIBEIRO, 2008). As rochas ocorrem preferencialmente na forma de diques finos, com espessura centimétrica a decimétrica, raramente métrica, que se recortam em diferentes direções, formando uma estrutura conhecida como *stockwork*. Apenas ocasionalmente podem ser encontrados pequenos corpos mais homogêneos, verticais ou com alto ângulo de mergulho, lateralmente contínuos por até algumas dezenas de metros. Neste contexto, a mina não apresenta unidades expressivas de uma única litologia, mais sim zonas de predominância de diques, veios ou *plugs* (BROD et al., 2004). Além da variabilidade lateral, que reflete a heterogeneidade dos protólitos, o perfil de alteração intempérica apresenta variações verticais intensas (RIBEIRO, 2008). Ou seja, a maioria dos processos responsáveis pela formação dos depósitos minerais são multiescalares (escala regional a microscópica) e multivariados (processos de formação das rochas e superficiais dinâmicos) (FRIMMEL & NWAILA, 2020). Isso dificulta a descrição litológica nas regiões de contato podendo levar a uma classificação tipológica distinta dos grupos predominantes, comprometendo a assertividade das representações espaciais dos corpos geológicos. A incerteza na descrição das amostras não ocorre somente pela complexidade do fenômeno mais também pelo suporte amostral restrito dos furos de sondagem.

METODOLOGIA

O banco de dados em estudo apresenta mais de 40 mil amostras de sondagem rotativa diamantada com suporte regularizado, distribuídas numa malha de 75m por 75m. Porém, somente 18.583 amostras possuem informações completas de litologia e teor. Trinta e dois tipos litológicos estão descritos e nove variáveis químicas quantificadas. Esse banco consta também com 1.548 ensaios metalúrgicos de bancada com informação isotópica restrita a poucas litologias, o que restringe o uso dessa informação para modelagem dos domínios. O conhecimento geológico e operacional da mina foi usado para dividir o depósito em oito tipologias predominantes principais considerando as características geológicas e químicas compatíveis com as características de processamento. Porém, as imprecisões nas descrições (mais de uma litologia reconhecida por amostras) estão presentes em 53% do banco de dados e dificultam o direcionamento das amostras para compor os domínios de estimativa.

Para tomar a decisão correta sobre qual grupo a amostra deve pertencer, o geomodelador precisa combinar o conhecimento geológico com as informações quantitativas da rocha, sendo essa tarefa muito exaustiva já que as descrições são

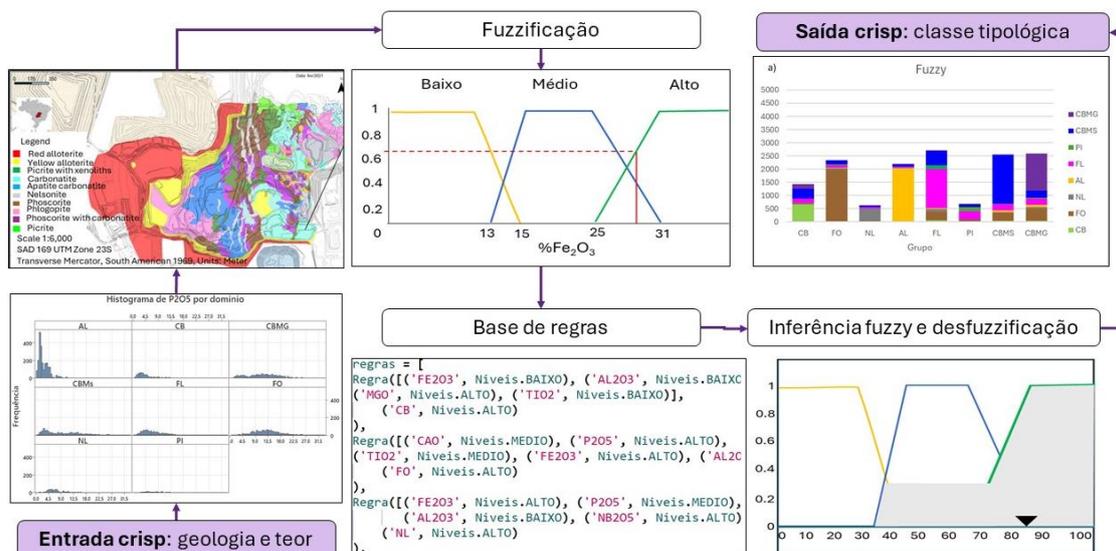
ambíguas, exigindo sólido conhecimento do geólogo de recursos ou geoestatístico. A lógica fuzzy pode ser vista como uma alternativa de lidar precisamente com a imprecisão, assim como a teoria da probabilidade pode ser vista como uma tentativa de lidar precisamente com a incerteza. (DEMICCO, 2004).

Uma adaptação do fenômeno estudado para um sistema de modelagem fuzzy foi realizada cuja classificação das unidades pode ser representada como uma função $f: R^{(n+1)} \rightarrow R$, com $y = f(x)$, onde a saída do sistema y é o domínio a qual a amostra pertence em função da entrada de x , com dimensão $n+1$, onde x_1, \dots, x_n são as expressões químicas extraídas das variáveis significativas para cada tipologia. Assim, o processo de agrupamento pode ser conhecido por uma função, mas não se tem uma expressão matemática explícita que relacione as variáveis de entrada com as variáveis de saída. É nessa etapa que o conhecimento especializado é utilizado, para dosar as relações entre os parâmetros de entrada com as informações geológicas. Na linguagem fuzzy estas relações são escritas via termos linguísticos, como no exemplo: “As rochas carbonatíticas do depósito possuem alto teor de MgO”.

O problema é como determinar o menor inteiro considerado alto (3, 5, 10, ...). Esse problema pode ser resolvido se for utilizada uma forma alternativa de descrever um conjunto. Ao invés de determinar o menor inteiro pertencente ao conjunto, podemos associar diferentes grau de pertinência para os teores daquele conjunto, obtendo valores parciais ou graduados no intervalo [0, 1], que agora se torna uma função de pertinência.

O modelo fuzzy esquemático proposto para predição das tipologias é mostrado na Figura 1.

Figura 1. Componentes do modelo de predição fuzzy.



Em zonas de transição ou sobreposição a variabilidade espacial é alta e o teor médio local oscila muito. Por essa razão um subconjunto do banco de dados original (47%) foi gerado contendo somente as amostras com uma única descrição litológica, consideradas mais uniformes ou precisas. Os teores foram analisados sobre critérios estatísticos (Figura 2) a fim de criar uma identidade química para cada domínio. A mediana e os quartis foram utilizados para definir os conjuntos de números reais a, b, c, d (Figura 3) que são os parâmetros da função de pertinência e para contornar os *outliers*. Diferentes valores para os conjuntos de números reais foram testados tendo como referências as análises estatísticas até obter o melhor ajuste dos limiares da função.

Uma expressão química para cada domínio foi definida considerando as variáveis mais significativas. O processo de fuzzificação, transformou os valores das variáveis de entrada para a linguagem usual, ou seja, os teores foram classificados qualitativamente em termos linguísticos (baixo, médio e alto), utilizando o conhecimento do especialista e, quantitativamente, atribuindo-se um grau de compatibilidade aos termos linguísticos através da função de pertinência trapezoidal.

Figura 2. Análise estatística dos teores utilizada como entrada *crisp* no modelo fuzzy.

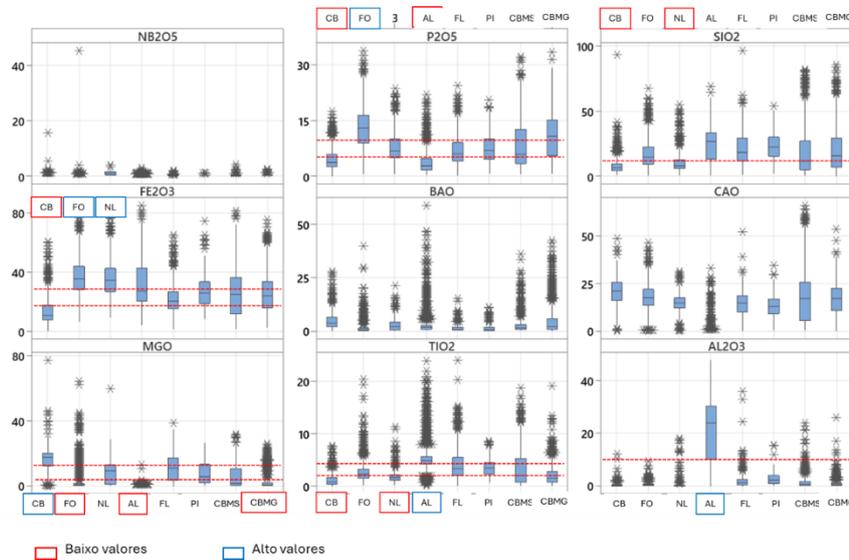
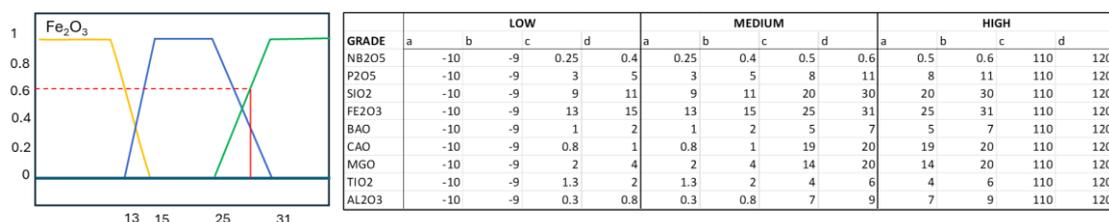


Figura 3. Função de pertinência e termos linguísticos dos conjuntos fuzzy de entrada.



O núcleo do controlador fuzzy é formado pela base de regras e por um método de inferência. A base de regras descreve as relações entre as variáveis linguísticas de entrada e de saída do sistema de fuzzificação. Todas as proposições da base de regras do modelo de predição utilizado são da forma: if <antecedente> then <consequente>. Por exemplo:

SE Fe_2O_3 é baixo, **SE** MgO é alto, **SE** Al_2O_3 é baixo, **SE** TiO_2 é baixo, **ENTÃO** CB.

A base de regras foi montada para os oito domínios de estimativa: dolomita carbonatito (CB), foscoritos (FO), flogopititos (FL), picrito brechado (PI), apatita carbonatito (CBMS) e foscorito com carbonatito (CBMG). Feito isso, o método de inferência fuzzy uniu cada uma das proposições da base de regras, por meio de operações lógicas (AND), para simular o processo de dedução humana e produzir a variável fuzzy de saída. Tipicamente, utilizam-se modelos de inferência fuzzy específicos conforme as propriedades sintáticas definidas, ou seja, o modelo de processamento definido para o sistema de conhecimento vai depender basicamente da forma de armazenamento de informações escolhida (ALMEIDA& EVSUKOFF, 2003), mas o que geralmente é mais utilizado é o método Mamdani (mínimos), adotado neste trabalho.

O passo seguinte foi a defuzzificação do conjunto fuzzy de saída [0,1], em um número real y , nesse caso o valor de saída y foi uma estimativa da classe pertencente. Os resultados das análises deste trabalho são baseados no método de defuzzificação do centroide (COG), considerando intervalos percentuais variando de 0% a 100%, e as implementações computacionais foram realizadas em linguagem python.

Por fim uma macro foi gerada para: primeiro, ordenar do maior para o menor as porcentagens das amostras com possibilidade de pertencer a mais de uma tipologia (medida de incerteza); segundo verificar a saída fuzzy com a descrição geológica dos furos de sondagem; terceiro, agrupar os domínios conforme as etapas 1 e 2. Esse processo foi repetido algumas vezes variando os parâmetros da função de pertinência a fim de se obter os melhores resultados nas simulações. O modelo final escolhido foi então verificado por comparação com o agrupamento k-means e hierárquico; por aderência com as descrições litológicas; reprodução das características estatísticas e por validação visual.

Disponibilidade de dados

O repositório de códigos para reprodução da metodologia aplicada pode ser encontrado no GitHub através do link <https://github.com/madamsilte/classfuzzy.git>.

RESULTADOS

Vinte simulações foram realizadas alterando os atributos e os valores da função pertinência conforme as análises estatísticas. A Tabela 1 traz um resumo dos resultados com maiores variações, sendo o teste de número cinco escolhido considerando a aderência com a descrição geológica e um maior número de amostras pertencente ao menos um dos oito domínios definidos no conjunto fuzzy de entrada (81,3% do banco original foi classificado).

Para a validação com as amostras de sondagem, admitiu-se que as descrições geológicas, ainda que imprecisas, estavam corretas. Portanto, se saída fuzzy fosse compatível com uma das litologias descritas nos furos de sondagem, o resultado era tido como certo, caso contrário considerado como incorreto. Ou seja, uma amostra informada como FL-FO no furo de sondagem e que apresentou alta pertinência para FO, teve a classificação fuzzy admitida, e a amostra foi migrada para o domínio FO. A porcentagem de acerto foi aproximadamente 70%, com menores índices para as tipologias com características químicas muito semelhantes (caso dos flogopititos e picrito com xenólitos).

Tabela 1. Resultados das simulações do modelo fuzzy.

Tipologias	Classe fuzzy aderente com a descrição geológica (%)				
	1	2	3	4	5*
CB	91.6	97.6	98.5	98.2	77.1
FO	71.8	64.0	61.7	64.5	69.6
NL	71.7	84.7	83.6	86.1	73.8
AL	87.6	91.5	91.5	91.7	91.6
FL	63.8	63.0	65.4	65.9	66.4
PI	28.9	53.3	53.4	46.0	61.0
CBMS	61.4	97.6	63.9	50.0	67.9
CBMG	64.0	65.2	76.9	51.6	84.2
Quantidade de amostras classificadas	15.733	9.532	10.634	14.618	15.101

A figura 4 traz uma comparação do método fuzzy, com outras duas ferramentas do aprendizado de máquinas, k-means e hierárquico. É possível observar que o algoritmo fuzzy mostrou-se mais eficiente para classificar e, conseqüentemente, separar os domínios com menos misturas de tipologias quando comparado aos demais. Nos algoritmos tradicionais a tipologia predominante em cada cluster não foi facilmente identificada. Em relação ao comportamento e o padrão das distribuições estatísticas das variáveis é possível observar que eles foram reproduzidos pelos domínios fuzzy (Figura 5).

Figura 4. Resultados dos agrupamentos. Em a) fuzzy; b) k-means e c) hierárquico.

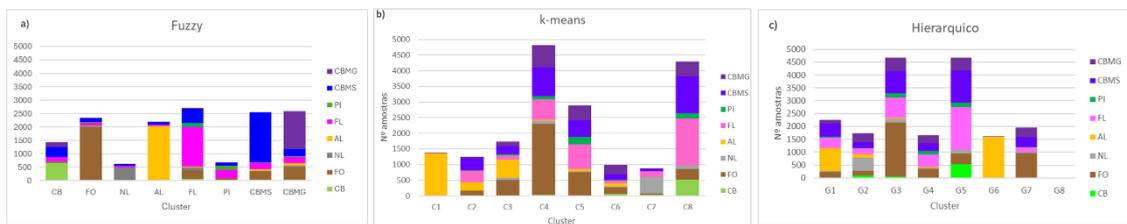
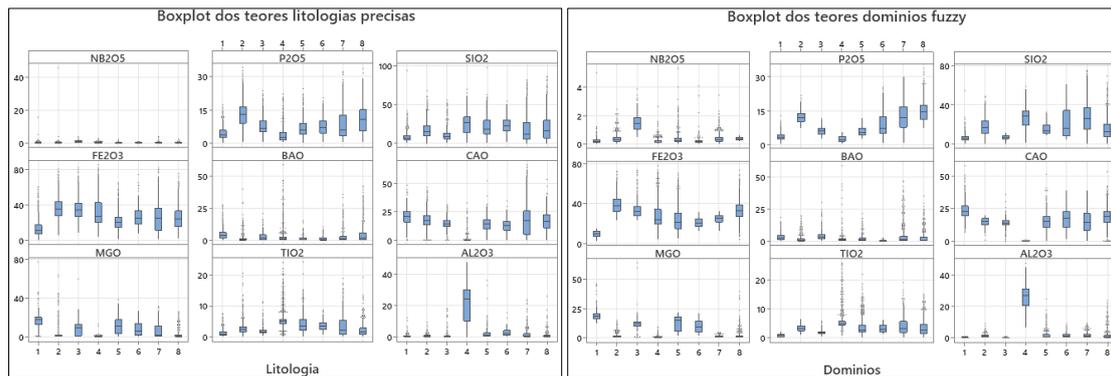
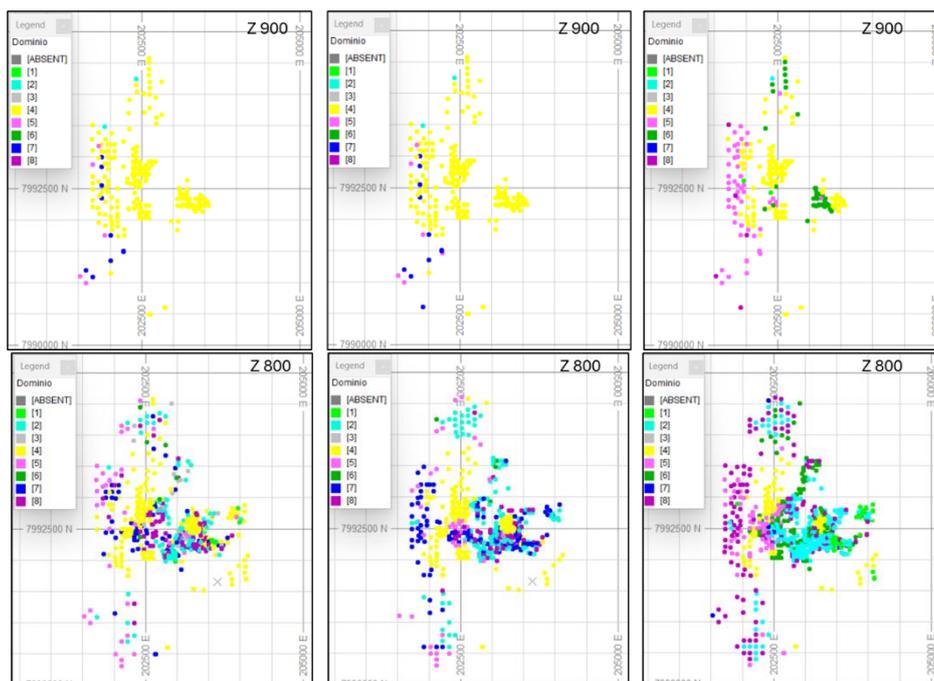


Figura 5. Análise estatística dos teores das amostras precisas e dos domínios fuzzy.



Uma inspeção visual seção a seção variando as cotas foi realizada entre os dados de sondagem (amostras precisas) e os agrupamentos fuzzy e k-means. Nota-se, na figura 6, uma similaridade maior entre as distribuições dos domínios obtidos pela lógica fuzzy com o banco de dados geológico.

Figura 6. Verificação visual: amostras sondagem, agrupamento fuzzy e k-means.



Como informação adicional à classificação dos domínios, o algoritmo difuso atribuiu um grau de associação das tipologias a todas as amostras. Esse grau de pertinência pode ser usado como uma medida de probabilidade e, portanto, pode fornecer uma maneira simples de quantificar a incerteza do agrupamento. Em uma representação nítida do conceito “foscorito”, atribuir-se-ia 1 às combinações de variáveis que se encaixam perfeitamente na definição de FO e zero para as demais. Na representação dos conjuntos nebulosos, a consideração da natureza transitória dos limites é expressa nas funções de pertinências, podendo uma determinada amostra ser simultaneamente membro de mais de um conjunto difuso, por exemplo: foscorito (FO) e foscorito com carbonatito (CBMG). A Tabela 2 apresenta o número de amostras que receberam grau de pertinência para uma ou mais classe. É possível notar que 54% das amostras receberam uma única classe e que 87% têm chances de pertencer no máximo a duas classes, com aderência aproximada de 70% com a descrição litológica dos furos. Essa informação auxilia os debates sobre onde exatamente estão as regiões de maiores incertezas nos contatos litológicos redirecionando os trabalhos de geologia.

Tabela 2. Resultados das simulações do modelo fuzzy.

Pertinência fuzzy em uma ou mais classes				Acertos fuzzy	
Quantidade de classes	Nº amostras	%	% acumulada	Nº amostras corretas	%
1	8,150	54.0	54.0	5,386	66.1
2	4,929	32.6	86.6	3,596	73.0
3	1,663	11.0	97.6	1,436	86.3
4	313	2.1	99.7	286	91.4
5	46	0.3	100.0	40	87.0
6	0	0.0	-	-	-
7	0	0.0	-	-	-
8	0	0.0	-	-	-
Soma	15.101	100	-	10.744	71.1

CONCLUSÕES

Algumas amostras do banco de dados receberam pertinência zero para os oito domínios pré-estabelecidos. Possivelmente por apresentar variações pontuais (problemas multiclases com estruturas geológicas mais intensas), já que apresentaram baixa compatibilidade com as expressões químicas utilizadas no conjunto fuzzy de entrada; ou até mesmo por viés analítico em campanhas mais antigas necessitando revisitar os furos de sondagem para averiguação. Nesses casos, fontes de informação adicionais de caracterização mineralógica, poderiam ser úteis para ajuste do modelo tornando a

predição mais assertiva. Ainda assim, o número de amostras classificadas foi suficiente para permitir a definição dos domínios com mais eficiência que os métodos de agrupamento tradicionais. O processo de separação dos clusters não foram conclusivos no k-means e nem no hierárquico, mesmo utilizando todo o banco de dados disponível.

A vantagem de combinar métodos computacionais à modelagem de recursos é a replicabilidade do procedimento, reduzindo a subjetividade na interpretação dos dados a depender do geomodelador. A capacidade de definir classes consistente usando algoritmo fuzzy não apenas padroniza o procedimento para definição dos domínios, mas também introduz uma ferramenta de análise de incerteza. Esse grau de incerteza das amostras pode direcionar os trabalhos de mapeamento geológico e as campanhas de amostragem, bem como auxiliar na escolha de amostras para caracterização e testes em pequena escala.

Pode-se concluir que o modelo fuzzy cumpriu os objetivos iniciais ao otimizar a definição dos grupos utilizando informações dos teores combinadas ao conhecimento especializado, reduzindo o tempo de análise pelo geólogo ou geoestatístico. Também é possível garantir maior agilidade quando novos dados forem gerados.

Sugere-se para trabalhos futuros, combinar outras fontes de informação como granulometria, textura, composição mineralógica para a calibração e atualização do algoritmo. Bem como visitar os furos cuja descrição difere das classes fuzzy, a fim de confirmar a suposição adotada neste estudo de que mesmo imprecisas as informações de sondagem estão corretas, já que diferentes campanhas passaram por diferentes padronizações das descrições ao longo do tempo.

AGRADECIMENTOS

Os autores agradecem à Universidade Federal de Catalão, ao Centro Federal de Educação Tecnológica de Minas Gerais e à CMOC Brasil.

REFERENCIAS

ABILDIN, Y., XU, C., DOWD, P., ADELI, A., 2023. Geometallurgical Responses on Lithological Domains Modelled by a Hybrid Domaining Framework. *Minerals*, 13, 918. <https://doi.org/10.3390/min13070918>.

ALMEIDA, P.E.M., EVSUKOFF, A.G., 2003. Sistemas Fuzzy. In: REZENDE, S.O. (coord.). *Sistemas inteligentes: Fundamentos e Aplicações*, Cap. 7, Manole, p. 169-202.

BRAGA, S. A., COSTA, J. F. C. L. 2016. Krigagem dos indicadores aplicada a modelagem das tipologias de minério fosfatados da mina F4. *HOLOS*, 1, 394-403. <https://doi.org/10.15628/holos.2016.3870>.

BROD, J.A., RIBEIRO, C.C., GASPAR, J.C., BROD, T.C., BARBOSA, E.S.R., RIFFEL, B.F., SILVA, J.F., CHABAN, N., FERRARI, A. J. D., 2004. Excursão 1. Geologia e mineralizações dos complexos alcalino-carbonatíticos da Província Ígnea do Alto Paranaíba, Congresso Brasileiro de Geologia, XLII, Araxá, MG, pp. 1 29.

CORDEIRO, P.F.O., BROD, J.A., PALMIERI, M., OLIVEIRA, C.G., BARBOSA, E.S.R., SANTOS, R.V., GASPAR, J.C., ASSIS, L.C., 2011a. The Catalão I niobium deposit, central Brazil: resources, geology and pyrochlore chemistry. *Ore Geology Reviews*. <https://doi.org/10.1016/j.oregeorev.2011.06.013>.

DEMICCO, R. V, KLIR, G. J., 2004. *Fuzzy Logic in Geology*. Academic Press. An imprint of Elsevier Science. 352P.

DUNHAM, S. AND VANN, J., 2007. Geometallurgy, geostatistics and project value - does your block model tell you what you need to know? in 'Project evaluation conference, Melbourne, Victoria', pp. 19–20.

FRIMMEL, H. E., & NWAILA, G. T., 2020. Geologic evidence of syngenetic gold in the Witwatersrand Goldfields, South Africa. In T. Sillitoe, R. Goldfarb, F. Robert, & S. Simmons (Eds.), *Geology of the major gold deposits and provinces of the world*, (Special Publication 23, pp. 645–668). Society of Economic Geologists. <https://doi.org/10.5382/SP.23.31>

GIBSON, S.A., THOMPSON R.N., DICKIN A.P., LEONARDOS O.H., 1995a. High-Ti and low-Ti mafic potassic magmas: Key to plume-lithosphere interactions and continental flood-basalt genesis. *Earth Planet Sci. Lett.* 136: 149-165.

IZBICKI, R., SANTOS, T. M., 2020. *Aprendizado de máquina: uma abordagem estatística*. [livro eletrônico], Carlos, SP. 272p.

KNIGHT, R., OLSON HOAL, K., & ABRAHAM, A. P. G., 2011. Threedimensional geometallurgical data integration for predicting concentrate quality and tailings composition in a massive sulfide deposit. In *The first AusIMM international geometallurgy conference*, Brisbane, Australia (pp. 227–232).

KOVACIC, Z.; BOGDAN, S., 2006. *Fuzzy Controller Design Theory and Application*. Press is an imprint of Taylor & Francis Group, 416p.

LARRONDO, P.F.; DEUTSCH, C.V, 2004. Methodology for Geostatistical Model of Gradational Geological Boundaries: Local Non-stationary LMC. *Cent. Comput. Geostat.*

LOZANO, C.; BENNET, C., 2003. Geometallurgical modeling applied to production forecasting, plant design and optimisation. *Universidad de Chile – Departamento Ingeniería de Minas Taller PROCEMIN*, 11p.

MOREIRA, G. C., 2020. *Análise de agrupamento aplicada à definição de domínios de estimativa para a modelagem de recursos minerais*. Dissertação (Mestre em Engenharia de Minas, Metalurgia e de Materiais), Universidade Federal do Rio Grande do Sul, Porto Alegre, 102p.

NETO, J. M. M.; MOITA, G. C., 1998. Uma introdução à análise exploratória de dados multivariados. *Química Nova*, 21(4) p.

NEWTON, M. & GRAHAM, J., 2011. Spatial modelling and optimisation of geometallurgical indices, in 'The First AusIMM International Geometallurgy Conference', Brisbane, pp. 247–261.

RAJABINASAB, B.; ASGHARI, O., 2018. Geometallurgical domaining by cluster analysis: Iron ore deposit case study. *Natural Resources Research* 28. 665-684. <https://doi.org/10.1007/s11053-018-9411-6>.

RIBEIRO, C. C., 2008. Geologia, geometalurgia, controles e gênese dos depósitos de fósforo, terras raras e titânio do complexo carbonatítico Catalão I, GO. 508 f. Tese de doutorado em Geologia-Universidade de Brasília, Brasília.

ROMARY, T., RIVOIRARD, J., DERAISME, J., QUINONES, C., FREULON, X., 2012. Domaining by clustering multivariate geostatistical data. In: *Geostatistics Oslo*, pp. 455–466. https://doi.org/10.1007/978-94-007-4153-9_37

ROSSI, M. E.; DEUTSCH, C. V., 2014. Mineral resource estimation. [S.l.]: Springer Science & Business Media.

SANCHES, M. K., 2003. Aprendizado de máquina semi-supervisionado: proposta de um algoritmo para rotular exemplos a partir de poucos exemplos rotulados. Dissertação (Mestre em Ciências da computação. Universidade de São Paulo, 142p.

VANN, J., JACKSON, J., COWARD, S., & DUNHAM, S. 2011. The Geomet curve—A model for implementation of geometallurgy. In *The first AusIMM international geometallurgy conference*, Brisbane, Australia (pp. 35–43).

WILLIAMS, S. R. E RICHARDSON, J. M., 2004. Geometallurgical Mapping: A New Approach That Reduces Technical Risk. *Proceedings 36th Annual Meeting of the Canadian Mineral Processors*, Paper 16, 28p.

YASOJIMA, C. T. K., 2020. Modelo de Krigagem Automática Baseada em Agrupamento. Tese (Doutor em Ciências da Computação. Universidade Federal do Pará, 86p.