



**REENCONTROS
NOVOS ESPAÇOS
OPORTUNIDADES**

XXXIV SIC Salão Iniciação Científica

**26 - 30
SETEMBRO
CAMPUS CENTRO**

Evento	Salão UFRGS 2022: SIC - XXXIV SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2022
Local	Campus Centro - UFRGS
Título	Cluster Tree para definições de domínios estacionários
Autor	JOAO PEDRO GROSS LAGUE
Orientador	DIEGO MACHADO MARQUES

Cluster Tree para definição de domínios estacionários

João Pedro Gross Lague¹, Diego Machado Marques¹

¹ Universidade Federal do Rio Grande do Sul, e-mail: joaopedrogrosslague@gmail.com;
diego.marques@ufrgs.br;

A modelagem de recursos naturais compreende diversas etapas durante seu desenvolvimento e um dos primeiros passos desse fluxo de trabalho é a definição de domínios estacionários. A definição de domínios estacionários pode não ser uma tarefa trivial, e depender de parâmetros físicos e/ou químicos que não estão devidamente amostrados. Sendo assim, a definição de domínios estacionários não é algo intrínseco à definição geológica, mas pode estar ligada com a homogeneidade estatística dos domínios modelados, ou estacionariedade das populações estipuladas, sendo cada população um domínio estacionário, observando o ponto de vista geoestatístico. Ou seja, há necessidade de conexão espacial nestas populações estatísticas ditas estacionárias. Em um cenário hipotético, espera-se que domínios estacionários reflitam características do depósito geológico a partir de suas similitudes químico/físicas decorrentes do fenômeno geológico envolvido na sua gênese. O entendimento das características estatísticas do fenômeno associado ao arcabouço geológico envolvido leva o particionamento do depósito em domínios estacionários, conectados espacialmente. Visando a diminuição da ambiguidade geológica envolvida no processo de modelagem, alguns estudos vêm propondo a utilização de algoritmos de aprendizagem não supervisionada, popularmente conhecidos como técnicas de agrupamento de dados. Tais técnicas utilizam algoritmos onde o conjunto de dados de entrada é composto por amostras não categorizadas, sendo o aprendizado puramente baseado na similitude multivariada das amostras. A partir do conjunto de dados de entrada são extraídos padrões e comportamentos que permitem reunir amostras em grupos denominados de “clusters”. O trabalho proposto visa uma abordagem não-convencional de técnicas de agrupamento de dados, a partir do uso de três algoritmos de aprendizagem distintos tanto por seus métodos aplicados quanto pela relação multivariada e espacial dos clusters definidos para fins de validação de bases de dados históricas. A abordagem é dita não convencional pois compreende a utilização dos algoritmos de maneira hierárquica, seguindo uma árvore de decisão geoquímica aplicados a um conjunto de dados de um depósito de ferro. A extração dos clusters, representados pelos domínios de estimativa, foi feita de maneira sequencial e binária para que o interpretador pudesse chegar aos resultados esperados. Dessa maneira, estruturou-se uma árvore de clusters a fim de determinar cada classe individualmente. A estruturação ocorre de maneira a aninhar o encadeamento dos algoritmos de acordo com a classificação previamente estabelecida. Foram propostos alguns métodos típicos de validação dos resultados obtidos através de fluxos de classificatórios para três algoritmos implementados confrontados com as classes originais. Para melhor entendimento dos agrupamentos fornecidos em cada nível da árvore de clusters foram propostas métricas de performance dos agrupamentos, a fim de avaliar a semelhança de cada agrupamento calculado em relação às suas amostras.

Palavras-chave: Árvore de Clusters, Domínios Estacionários, Modelagem Geológica.