

# Desafios de clusterização de dados geoespaciais para a formação de redes industriais de CCUS

Thiago Rodrigues Cruz Justino<sup>1</sup>, Beatriz Rodrigues de Almeida<sup>1,2</sup>, Carlos Alexandre Silva dos Santos<sup>1</sup>, Jansen Cruz de Souza<sup>3</sup>, Gustavo Charles Peixoto de Oliveira<sup>1,2,4</sup>

<sup>1</sup>Laboratório TRIL, Departamento de Computação Científica, UFPB, João Pessoa - PB

<sup>2</sup>Programa de Pós-Graduação em Engenharia Mecânica, UFPB, João Pessoa - PB

<sup>3</sup>Superintendência de Tecnologia da Informação, UFPB, João Pessoa - PB

<sup>4</sup>Programa de Pós-Graduação em Informática, UFPB, João Pessoa - PB

carlos.alexandre@academico.ufpb.br, thiago.rodrigues@academico.ufpb.br, biarodriguesdsa@gmail.com, jansen.souza@academico.ufpb.br, gustavo.oliveira@ci.ufpb.br

**Resumo.** Este artigo discute alguns desafios decorrentes do processo de triagem e identificação de clusters e hubs de captura e armazenamento de CO<sub>2</sub> no Brasil, com foco em áreas estratégicas onshore e offshore. Utilizando técnicas de análise de dados geoespaciais, clusterização e algoritmos de roteamento, o estudo mapeou estruturas nos seguintes estados: Rio de Janeiro, São Paulo e Rio Grande do Norte. Apresentamos como a dispersão geográfica dos dados e a dependência de requisições por API influenciam a escalabilidade e o desempenho computacional dos algoritmos.

## 1. Introdução

Com o aumento dos esforços globais para conter as mudanças climáticas, setores econômicos buscam reduzir suas emissões de gases de efeito estufa. No Brasil, processos industriais de difícil abatimento de CO<sub>2</sub> devem responder por cerca de 20% da meta nacional de emissões (BRASIL, 2023; SEEG, 2022). Tecnologias de captura, utilização e armazenamento de carbono (CCUS) são essenciais para alcançar o Net Zero até 2050.

Projetos-piloto de CCUS em regiões costeiras, especialmente nas bacias sedimentares offshore de Campos, Santos e Potiguar, têm grande potencial para armazenamento geológico de carbono, com abatimento estimado em até 10,8 MtCO<sub>2</sub>e/ano (CIOTTA et al., 2021; OGCI, 2023). Clusters e hubs de CCUS, formados por concentrações geográficas de emissores e pontos de coleta ou armazenamento, estruturam redes que reduzem custos e facilitam a expansão em larga escala (GCCSI, 2016; SUN, 2021; OGCI, 2024).

Este estudo aborda desafios de computação de alto desempenho para clusterização geoespacial no contexto do CCUS brasileiro, visando otimizar o agrupamento de fontes e sumidouros diante da esparsidade das emissões e da infraestrutura, com critérios técnicos, econômicos e geográficos.

## 2. Metodologia

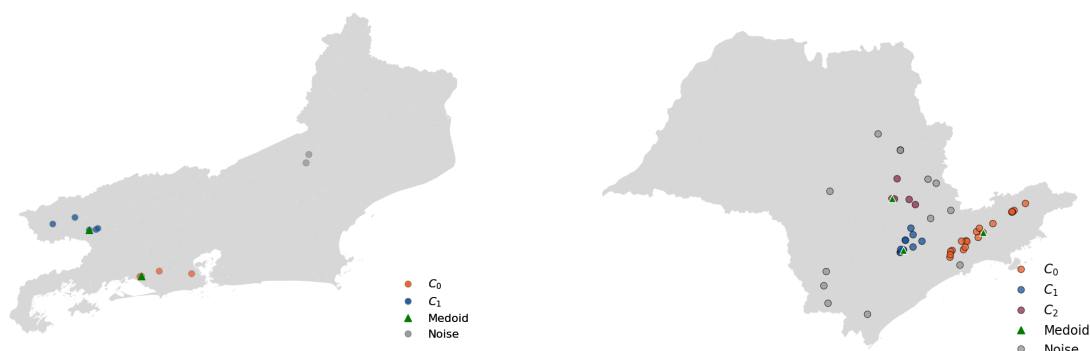
A análise de dados geoespaciais para a triagem e estruturação de arranjos industriais locais de CCUS priorizou os estados do Rio de Janeiro e de São Paulo com

base em critérios técnicos. Inicialmente, foram coletados, tratados e integrados dados georreferenciados das fontes emissoras *onshore* (indústrias) e dos potenciais locais de armazenamento *offshore* (poços).

Em seguida, as distâncias reais de deslocamento entre os pontos foram utilizadas como critério para o agrupamento, com foco na aplicação do algoritmo DBSCAN. A identificação dos medóides (indústrias com menor soma de distâncias dentro de cada cluster) permitiu a definição dos hubs de captura, por ocuparem posições centrais do ponto de vista logístico. As rotas entre os pontos foram inicialmente obtidas por meio da API do Open Source Routing Machine (OSRM), que fornece trajetos viários realistas visando o escrutínio do transporte de CO<sub>2</sub> a curtas distâncias por modais rodoviários.

### 3. Resultados e Considerações Finais

A clusterização nos estados do RJ, SP e RN revela que o uso do OSRM, comparado à métrica de Haversine, pode aumentar o tempo de processamento em mais de 14.000 vezes, como observado na Tabela 1. Esse custo computacional elevado cresce proporcionalmente ao número de pontos analisados. A matriz de distâncias é construída com dois loops aninhados, considerando apenas pares únicos. Isso reduz cálculos, mas mantém a complexidade quadrática. O problema se acentua com o uso do OSRM, pois exige chamadas externas à API. Essas chamadas são muito mais custosas que os cálculos geodésicos, porém são importantes para a precisão do estudo.



**Figura 1. Clusters de indústrias selecionadas: RJ (esq.) e SP (dir.)**

**Tabela 1. Desempenho computacional do cálculo da matriz de distâncias com relação às indústrias de cada estado**

Estado	Nº Indústrias	OSRM	Haversine
Rio de Janeiro	11	1 min e 6 seg	0.01 seg
São Paulo	48	19 min 35 seg	0.08 seg

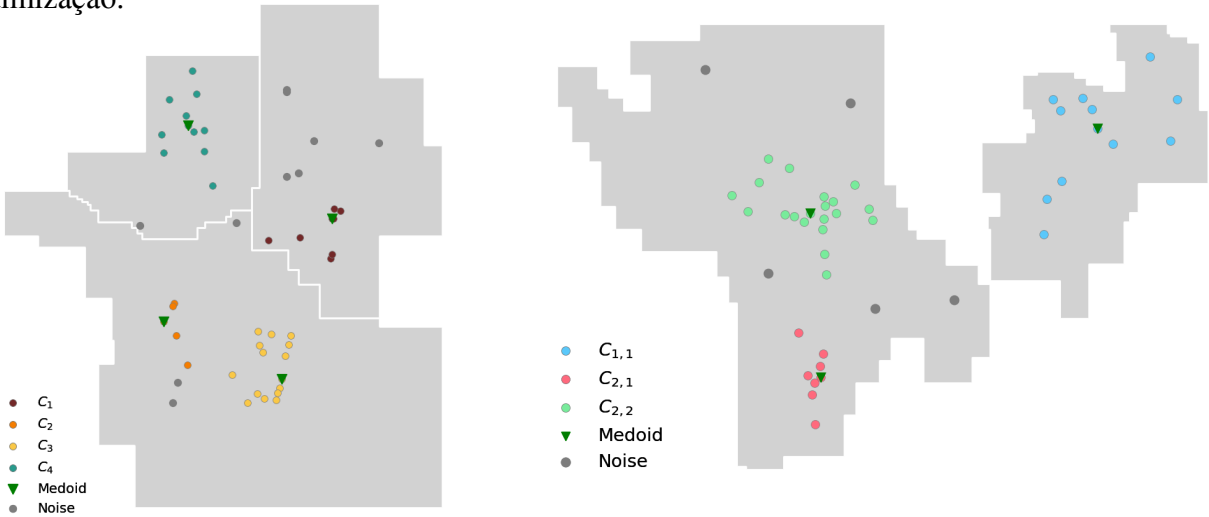
Testes de desempenho para a clusterização offshore foram realizados por meio de uma abordagem de grid search, com variação do parâmetro eps, que define a distância máxima entre dois pontos para pertencerem ao mesmo cluster no DBSCAN. A escolha do eps utilizou a técnica do cotovelo, baseada na análise gráfica das distâncias

do k-ésimo vizinho mais próximo. A Tabela 2 apresenta os tempos médios totais e específicos para o cálculo de eps, destacando o impacto dessa etapa no custo computacional. Observa-se que o número de poços, o índice  $O_w$  e a distribuição espacial (quanto mais próximos, menor o tempo de cálculo) influenciam o tempo de processamento, sugerindo que novas abordagens podem torná-lo mais eficiente.

**Tabela 2. Desempenho computacional da clusterização por bacias.**

$O_w$	Escopo	Nº Poços	Processamento (ms)		
			clustering	eps	total
0.3	Geral	387	444.97	76.87	521.84
	Santos	173	302.26	31.04	333.30
	Campos	214	172.45	21.77	194.23
0.0	Geral	494	555.41	72.83	628.24
	Santos	188	348.03	32.15	380.18
	Campos	306	269.20	22.25	291.45

A Figura 2 ilustra o resultado da clusterização de poços offshore nas bacias de Campos (esq.) e Santos (dir.), após a calibração do algoritmo DBSCAN pelo método do cotovelo. Os agrupamentos evidenciam a capacidade do modelo de segmentar áreas com alta densidade de poços. As diferentes cores representam clusters distintos, e o desempenho computacional observado reforça a aplicabilidade do método no planejamento de hubs de armazenamento de CO<sub>2</sub>, ainda que existam oportunidades para otimização.



**Figura 2. Clusters de poços selecionados da bacia de Campos: Marlim, Marlim Leste e Marlim Sul (esq.) e da bacia de Santos: Búzios, Mero (dir.)**

A aplicação de algoritmos de clusterização no contexto do CCUS revelou desafios significativos, como escassez de dados qualificados, dispersão geográfica das

fontes emissores e elevados custos computacionais. Esses fatores impactam diretamente a eficiência e a escalabilidade dos modelos propostos. Melo et al. (2024), em seu estudo sobre a plataforma GIS CCUS Brazil, também destacam a importância de abordagens robustas e adaptáveis para garantir a viabilidade de sistemas escaláveis e intersetoriais. Importante mencionar que, apesar do potencial da bacia Potiguar, esta ainda não é contemplada nas análises por ainda não possuir poços offshore selecionáveis para a clusterização.

Perspectivas futuras incluem a adoção de roteamento local para eliminar a dependência de APIs externas, redução do tempo de execução e aumento da estabilidade do processo. Além disso, pretende-se automatizar a definição de parâmetros como  $\epsilon$ , validar a automação de clusters geoespaciais e desenvolver indicadores analíticos mais sofisticados. O objetivo é criar uma solução completa, eficiente e independente para implementação em larga escala no contexto brasileiro de CCUS.

## Referências

- Ciotta, M., et al. (2021). CO<sub>2</sub> storage potential of offshore oil and gas fields in Brazil. *International Journal of Greenhouse Gas Control*, 112, Article 103492. <https://doi.org/10.1016/j.ijggc.2021.103492>
- Global CCS Institute. (2016). *Special report: Understanding industrial CCS hubs and clusters*.
- Melo, C., Cardoso, D. S., Guimarães, T. T., Felipe, L., Zielinski, J. P., Andres, F., Santos, C., Genro, L. V., Gomes, P., & Martins, M. (2024). GIS CCUS Brazil Platform: A New Tool to Promote Carbon Capture, Utilization, and Storage in Brazil's Territory. *Sustainability*, 17(7), 2856. <https://doi.org/10.3390/su17072856>
- Ministério do Meio Ambiente e Mudança do Clima. (2023). *Contribuição Nacionalmente Determinada – NDC do Brasil: Atualização 2023*. Governo Federal do Brasil. <https://www.gov.br/mma/pt-br/sustentos/clima/ndc>
- Oil and Gas Climate Initiative (OGCI). (2024). *CCUS hubs in Brazil: Making the case, breaking the barriers*. OGCI.
- Oil and Gas Climate Initiative (OGCI). (2023). *The CCUS hub: A playbook for regulators, industrial emitters and hub developers*. OGCI.
- Seeg – Sistema de Estimativas de Emissões de Gases de Efeito Estufa. (2022). *Emissões totais por setor no Brasil (1970–2021)*. Observatório do Clima. <https://seeg.eco.br>
- Sun, X., et al. (2024). Hubs and clusters approach to unlock the development of carbon capture and storage – Case study in Spain. *Applied Energy*, 300, 117418. <https://doi.org/10.1016/j.apenergy.2021.117418>
- Wang, R. (2024). Status and perspectives on CCUS clusters and hubs. *Unconventional Resources*, 4, 100065. <https://doi.org/10.1016/j.uncres.2023.100065>