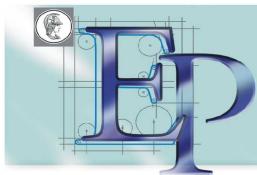


Projeto de Formatura – 2025



PCS - Departamento de Engenharia de Computação e Sistemas Digitais

Engenharia de Computação

Tema: *Framework para Imputação de Dados Faltantes em Séries Espaço-temporais de Sensores Físicos*

Motivação

A perda de dados em sensores ambientais é um problema recorrente que compromete previsões climáticas, estudos oceanográficos e análises de qualidade do ar. Falhas na transmissão ou armazenamento geram lacunas que distorcem resultados e dificultam a identificação de eventos extremos. Apesar da existência de diversos métodos de imputação, ainda não há uma ferramenta integrada que permita comparar, de forma reproduzível e visual, o desempenho das técnicas em diferentes cenários espaço-temporais.

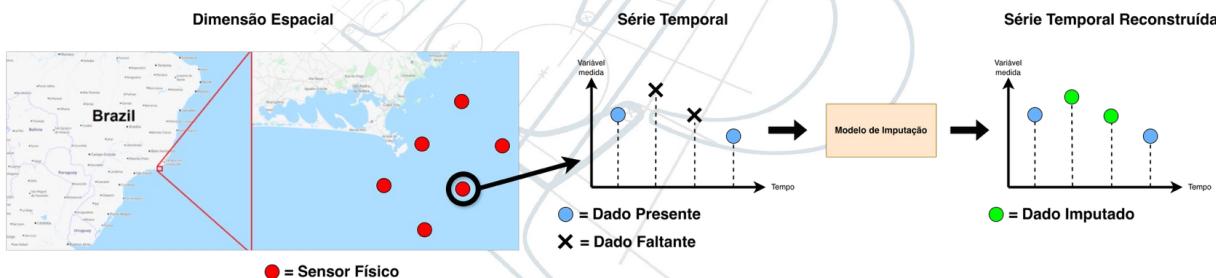
Métodos Selecionados

Os métodos contemplados foram:

SPIN – Spatiotemporal Partial Information Network: modelo baseado em *Graph Neural Networks* que utiliza atenção espaço-temporal esparsa.

PriSTI – Probabilistic Spatiotemporal Imputer: método de difusão probabilística que reconstrói séries tempo-espaciais ao modelar a propagação de informações no grid por meio de um processo de difusão invertido.

KNN – K-Nearest Neighbors Imputer: técnica estatística simples que estima valores faltantes com base nos vizinhos mais próximos, servindo como baseline eficiente para comparações.



Metodologia

A metodologia de testes consiste em partir de séries temporais completas dos *datasets* ROMS e Beijing AQI-36, aplicar máscaras artificiais de falha segundo diferentes padrões (falhas aleatórias, em blocos e espacialmente concentradas) e, em seguida, utilizar os métodos de imputação selecionados para reconstruir os valores omitidos. Para cada experimento, os valores imputados são comparados com o *ground truth* original, permitindo quantificar o erro por meio de métricas como MAE e MSE. A análise combina resultados numéricos e inspeção visual direta das séries reconstruídas, possibilitando avaliar a robustez dos métodos em diferentes intensidades e formatos de falha. Todo o processo é integrado ao framework desenvolvido, que permite selecionar *dataset*, método e padrão de falha, executar a imputação e visualizar, no mapa e nas séries temporais, o desempenho de cada técnica.

Datasets

ROMS – Cabo Frio (SSH): Modelo oceânico numérico que fornece séries temporais completas de elevação da superfície do mar, permitindo gerar cenários controlados de falhas artificiais e avaliar a resposta dos métodos em um ambiente espaço-temporal estável.

Beijing AQI-36 (PM2.5): Conjunto de medições reais de PM2.5 obtidas em 36 estações urbanas, apresentando variabilidade espacial significativa e ocorrência natural de falhas, possibilitando testar os métodos em condições reais de sensores ambientais.

Integrante: Bruno Maia de Oliveira Duarte

Professora Orientadora: Profa. Dra. Anna Helena Reali Costa

Coorientadores: Dr. Fábio José Muneratti Ortega e Dr. Marcel Rodrigues de Barros