

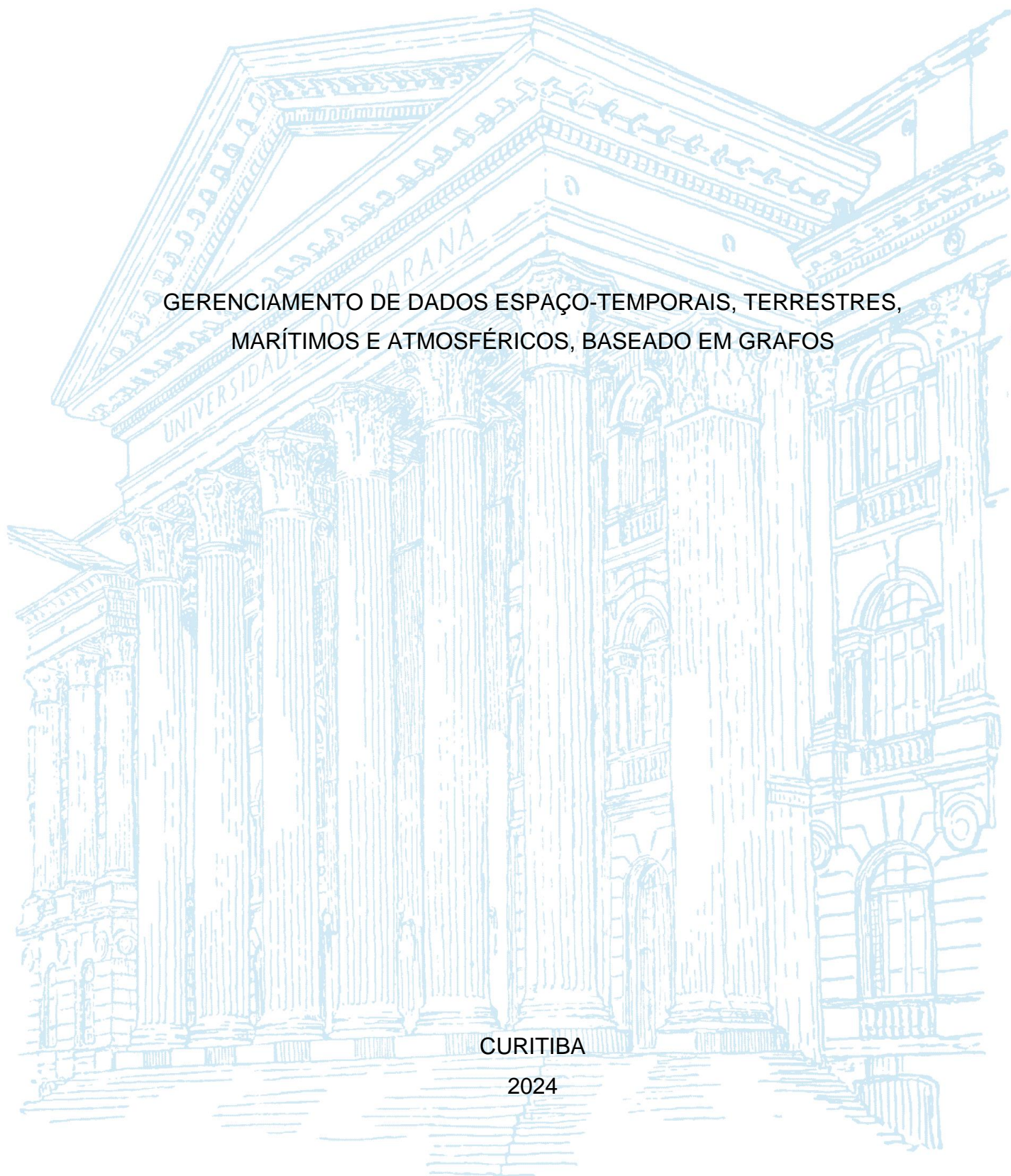
UNIVERSIDADE FEDERAL DO PARANÁ

ANA EMÍLIA DE SOUZA SILVA

GERENCIAMENTO DE DADOS ESPAÇO-TEMPORAIS, TERRESTRES,  
MARÍTIMOS E ATMOSFÉRICOS, BASEADO EM GRAFOS

CURITIBA

2024



ANA EMÍLIA DE SOUZA SILVA

GERENCIAMENTO DE DADOS ESPAÇO-TEMPORAIS, TERRESTRES,  
MARÍTIMOS E ATMOSFÉRICOS, BASEADO EM GRAFOS

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Geodésicas, Setor de Ciências da Terra, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Ciências Geodésicas.

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Luciene Stamato Delazari

CURITIBA

2024

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
UNIVERSIDADE FEDERAL DO PARANÁ  
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Silva, Ana Emília de Souza

Gerenciamento de dados espaço-temporais, terrestres, marítimos e atmosféricos, baseado em grafos / Ana Emília de Souza Silva / - Curitiba, 2024.

1 recurso on-line: PDF.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Ciências da Terra, Programa de Pós-Graduação em Ciências Geodésicas.

Orientadora: Profa. Dra. Luciene Stamato Delazari

1. Sistemas de informação geográfica. 2. Georreferenciamento.  
3. Grafos, representação dos. I. Delazari, Luciene Stamato. II. Universidade Federal do Paraná. III. Programa de Pós-Graduação em Ciências Geodésicas. IV. Título.



## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação CIÊNCIAS GEODÉSICAS da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **ANA EMÍLIA DE SOUZA SILVA** intitulada: **GERENCIAMENTO DE DADOS ESPAÇO-TEMPORAIS, TERRESTRES, MARÍTIMOS E ATMOSFÉRICOS, BASEADO EM GRAFOS**, sob orientação da Profa. Dra. LUCIENE STAMATO DELAZARI, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestra está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 29 de Fevereiro de 2024.

Assinatura Eletrônica

29/02/2024 16:41:39.0

LUCIENE STAMATO DELAZARI

Presidente da Banca Examinadora

Assinatura Eletrônica

01/03/2024 12:38:33.0

JULIA CELIA MERCEDES STRAUCH

Avaliador Externo (ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS )

Assinatura Eletrônica

29/02/2024 19:03:21.0

SILVANA PHILIPPI CAMBOIM

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

## **AGRADECIMENTOS**

A Deus, pela contínua oportunidade de desenvolvimento e evolução intelectual e moral.

Este trabalho teve como origem o Programa de Cursos de Pós-Graduação e Extraordinários em Estabelecimentos e Instituições Civis no País aprovado pela Portaria nº 91/2021 da Diretoria-Geral do Pessoal da Marinha. Após o devido processo seletivo, a autora foi designada pela Portaria nº 407/2022 da Diretoria do Pessoal Militar da Marinha a realizar o Curso de Mestrado em Ciências Geodésicas na UFPR.

Agradeço a Marinha do Brasil pela oportunidade singular de capacitação profissional e conseqüentemente o meu particular enriquecimento intelectual, social e humano.

A minha orientadora Prof.<sup>a</sup> Dr.<sup>a</sup> Luciene Stamato Delazari pela oportunidade, irrestrita atenção e suporte no desenvolvimento desta dissertação de mestrado.

Ao apoio incondicional e incansável do meu eterno professor e amigo Ivan de Araujo Medina, o qual foi parte essencial na construção deste trabalho com seus ensinamentos, sugestões e esclarecimentos.

## RESUMO

O gerenciamento de dados espaço-temporais caracterizados por multidomínios terrestres, marítimos e atmosféricos em uma estrutura de armazenamento organizada, acessível, e consistente — que suporte crescimento, flexibilidade e absorção de dados heterogêneos — onde a informação possa ser extraída com toda a sua integridade, se torna essencial em aplicações sensíveis tais como: cenários decisórios, planejamento territorial, e até na compreensão do Sistema-Terra. Apesar da sua importância, a gestão de dados espaço-temporais, relacionados ou não, dinâmicos e de amplo espectro, permanece como um desafio, devido à necessidade de construção de um modelo de dados único envolver complexidades como a multidimensionalidade e a heterogeneidade dos dados. Para otimizar o armazenamento, recuperação, gestão, e integração desses dados este trabalho propõe uma abordagem baseada em grafos. A abordagem responde aos principais desafios de integração dos dados espaço-temporais, projeta um modelo conceitual e lógico em um banco de dados orientado por grafos centrado no tempo e no espaço e os implementa, por meio de um projeto piloto com dados localizados nos Estados do Rio de Janeiro e em São Paulo. Consultas de recuperação sob a ótica operacional foram conduzidas para demonstrar a sua aplicabilidade. Os resultados alcançados demonstram que o gerenciamento de dados baseado em grafos possui as características da alta escalabilidade, alta flexibilidade, baixa manutenibilidade e pouca exigência quanto a recursos computacionais. A abordagem fornece uma nova referência para desenvolvedores na integração de dados espaço-temporais multidomínios, quaisquer que sejam eles.

Palavras-chave: Integração 1. Multidimensionais 2. Geoinformática 3. Heterogeneidade 4.

## **ABSTRACT**

The management of spatiotemporal data characterized by terrestrial, maritime, and atmospheric multidomains in an organized, accessible, and consistent storage structure — that supports growth, flexibility, and absorption of heterogeneous data — where information can be extracted with all its integrity, becomes essential in sensitive applications such as: decision-making scenarios, territorial planning, and even in understanding the Earth System. Despite its importance, the management of spatio-temporal data, whether related or not, dynamic and broad spectrum, remains a challenge, due to the need to build a unique data structure involving complexities such as multidimensionality and data heterogeneity. To optimize the storage, retrieval, management, and integration of this data, this work proposes a graph-based approach. The approach responds to the main challenges of integrating spatio-temporal data, designs a conceptual and logical model in a graph-oriented database centered on time and space and implements it, through a pilot project with data located in the States of Rio de Janeiro and São Paulo. Recovery consultations from an operational perspective were conducted to demonstrate its applicability. The results achieved demonstrate that graph-based data management has the characteristics of high scalability, high flexibility, low maintainability and little requirement for computational resources. The approach provides a new reference for developers in the integration of multi-domain spatio-temporal data, whatever they may be.

Keywords: Integration 1. Multidimensional 2. Geoinformatics 3. Heterogeneity 4.

## LISTA DE FIGURAS

FIGURA 1 – DADOS ESPAÇO-TEMPORAIS .....	21
FIGURA 2 – CLASSIFICAÇÃO DAS CLASSES DE RECURSOS NATURAIS .....	24
FIGURA 3 – CAMADAS ESTAÇO-TEMPORAIS COMO PRINCIPAL REPRESENTAÇÃO USADA EM GIS .....	25
FIGURA 4 – UM GRAFO ESPAÇO-TEMPORAL REPRESENTANDO QUATRO INSTANTÂNEOS TEMPORAIS. ....	26
FIGURA 5 – EXEMPLOS DE DADOS ESPAÇO-TEMPORAIS .....	27
FIGURA 6 – SEQUÊNCIA DE QUATRO CAPTURAS DE TELA MOSTRANDO A TEMPERATURA MÉDIA DE CINCO ANOS DA UNIÃO EUROPEIA E DA TURQUIA DE 1952 A 2008.....	28
FIGURA 7 – GRAFO ESPAÇO-TEMPORAL DA EVOLUÇÃO DAS PROVÍNCIAS DO CHILE .....	29
FIGURA 8 – ARMAZENAMENTO DO ESPAÇO E TEMPO EM GRAFOS REPRESENTADO EM UML.....	31
FIGURA 9 – EXEMPLO DE BANCO DE DADOS ORIENTADO POR GRAFOS NO SGBD ARANGODB .....	32
FIGURA 10 – EXEMPLO DE NÓS ARMAZENADOS EM UM DOCUMENTO JSON	32
FIGURA 11 – EXEMPLO DE ARESTAS ARMAZENADAS EM UM DOCUMENTO JSON .....	33
FIGURA 12 – LIMITES DO RECORTE TERRITORIAL .....	35
FIGURA 13 – RESUMO ESQUEMÁTICO DO FLUXO DE TRABALHO .....	39
FIGURA 14 – REPRESENTAÇÃO CONCEITUAL DA CONSTRUÇÃO DE UM GRAFO A PARTIR DE DADOS ESPAÇO-TEMPORAIS HETEROGÊNEOS.....	40
FIGURA 15 – PARTE INICIAL DO <i>SCRIPT</i> EM PYTHON.....	42
FIGURA 16 – VISÃO GERAL DO MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA .....	43
FIGURA 17 – MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA EXPANDIDO .....	45
FIGURA 18 – MODELO LÓGICO DE REFERÊNCIA DO MODELO CONCEITUAL.	46



FIGURA 19 – EXEMPLOS DE TERMOS UTILIZADOS COMO ATRIBUTOS PERTENCENTES AOS NÓS TIME E SPACE E TIPOS DE ARMAZENAMENTO UTILIZADO NO NÓ SPACE .....	47
FIGURA 20 – EXEMPLO DE INSTÂNCIA NO NÓ TIME .....	48
FIGURA 21 – DIAGRAMA DE CLASSES EM UML PARA O GERENCIAMENTO DE DADOS ESPAÇO-TEMPORAIS TERRESTRES, MARÍTIMOS E ATMOSFÉRICOS .....	49
FIGURA 22 – EXEMPLO DA ARESTA UPDATE .....	51
FIGURA 23 – APRESENTAÇÃO DE PARTE DA VERSÃO LÓGICA EM GRAFOS .....	52
FIGURA 24 – VISUALIZAÇÃO DO DIAGRAMA DE OBJETOS EM UML COM DADOS ESPAÇO-TEMPORAIS CONCRETOS .....	54
FIGURA 25 – AMOSTRA DO CONTEÚDO DO ARQUIVO JSON DA CLASSE TIME .....	55
FIGURA 26 – AMOSTRA DO CONTEÚDO DO ARQUIVO JSON DO RELACIONAMENTO LOCATED_IN .....	55
FIGURA 27 – VISUALIZAÇÃO DE UMA AMOSTRA DO GRAFO COM O ARANGODB.....	57
FIGURA 28 – VISÃO MACRO PARCIAL DO BANCO DE DADOS ORIENTADO POR GRAFOS PARA GESTÃO DE DADOS ESPAÇO-TEMPORAIS.....	59
FIGURA 29 – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-2011 A 31-12-2020.....	62
FIGURA 30 – PARÂMETROS FÍSICOS DO AR OBSERVADOS NO DIA 28-08-2023 ÀS 00:00. ....	66

## LISTA DE GRÁFICOS

GRÁFICO 1 – VALORES DE TEMPERATURA DA ÁGUA DO MAR OBSERVADAS ENTRE OS DIAS 28-10-2019 A 29-10-2019.....	64
GRÁFICO 2 – VALORES DE TEMPERATURA DA ÁGUA DO MAR ENTRE AS PROFUNDIDADES DE 200 À 600 M EM RELAÇÃO A PROFUNDIDADE E A LONGITUDE ENTRE 28-10-2019 A 29-10- 2019.....	65

## LISTA DE QUADROS

QUADRO 1 – DESCRIÇÃO DAS FONTES DE DADOS .....	36
QUADRO 2 – PROGRAMAS COMPUTACIONAIS QUE FORAM UTILIZADOS NA EXECUÇÃO DO FLUXO DE TRABALHO.....	38
QUADRO 3 – ESPECIFICAÇÕES DA MÁQUINA UTILIZADA PARA A EXECUÇÃO DO FLUXO DE TRABALHO.....	38
QUADRO 4 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DA CLASSE UNIDADES DE CONSERVAÇÃO AMBIENTAIS .....	62
QUADRO 5 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DA CLASSE TEMPERATURA DA ÁGUA DO MAR.....	63
QUADRO 6 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DA CLASSE TEMPERATURA DA ÁGUA DO MAR ENTRE PROFUNDIDADES DEFINIDAS .....	64
QUADRO 7 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DAS CLASSES ASSOCIADAS AOS PARÂMETROS FÍSICOS DO AR ARMAZENADOS NO BD .....	66

## LISTA DE TABELAS

TABELA 1 – COBERTURA TEMPORAL E FORMATO DE ARMAZENAMENTO DOS DADOS .....	36
TABELA 2 – UNIDADES DE MEDIDAS DOS OBJETOS UTILIZADOS NESSE ESTUDO .....	37
TABELA 3 – PADRONIZAÇÃO DOS OBJETOS UTILIZADOS NESSE ESTUDO....	41
TABELA 4 – EXEMPLOS DE DADOS ESPAÇO-TEMPORAIS CONCRETOS .....	53
TABELA 5 – SUMÁRIO GERAL DOS RESULTADOS DAS INSTÂNCIAS DAS CLASSES .....	56
TABELA 6 – SUMÁRIO GERAL DOS RESULTADOS DOS RELACIONAMENTOS	56

## LISTA DE SIGLAS E ACRÔNIMOS

AIS	- <i>Automatic Identification System</i>
AQL	- <i>ArangoDB Query Language</i>
BD	- Bancos de Dados
CC	- Casa Civil
CHM	- Centro de Hidrografia da Marinha
DECEA	- Departamento de Controle do Espaço Aéreo
ET-EDGV	- Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais
ICMBIO	- Instituto Chico Mendes de Conservação da Biodiversidade
IHO	- <i>International Hydrographic Organization</i>
INEA	- Instituto Estadual do Ambiente
MD	- Ministério da Defesa
MMA	- Ministério do Meio Ambiente
PR	- Presidente da República
REGEX	- <i>Regular Expressions</i>
S-100	- <i>IHO Universal Hydrographic Data Model</i>
S-57	- <i>IHO Transfer Standard for Digital Hydrographic Data</i>
SEAS	- Secretaria de Estado do Ambiente e Sustentabilidade
SGBD	- Sistema de Gerenciamento de Banco de Dados
SQL	- <i>Structured Query Language</i>
UML	- <i>Unified Modeling Language</i>
UNESCO	- Organização das Nações Unidas para a Educação, a Ciência e a Cultura

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>16</b>
1.1 OBJETIVOS .....	17
1.2 JUSTIFICATIVA .....	18
1.3 ESTRUTURA DO TRABALHO .....	19
<b>2 FUNDAMENTAÇÃO</b> .....	<b>20</b>
2.1 A NATUREZA DINÂMICA DOS DADOS ESPAÇO-TEMPORAIS.....	20
2.2 CLASSIFICAÇÃO DOS DADOS ESPAÇO-TEMPORAIS .....	22
2.3 REPRESENTAÇÃO COMPUTACIONAL DOS DADOS ESPAÇO-TEMPORAIS	24
2.4 INTEGRAÇÃO DE DADOS ESPAÇO-TEMPORAIS .....	27
2.5 MODELO LÓGICO ORIENTADO POR GRAFOS .....	30
2.5.1 SGBD ArangoDB.....	31
<b>3 MATERIAL E MÉTODOS</b> .....	<b>35</b>
3.1 RECORTE TERRITORIAL .....	35
3.2 DADOS UTILIZADOS.....	35
3.3 RECURSOS COMPUTACIONAIS UTILIZADOS .....	37
3.4 PROCEDIMENTO .....	38
3.4.1 Modelo conceitual de representação computacional integrada.....	39
3.4.2 Modelo lógico orientado por grafos .....	39
3.4.3 Grafo integrado: terrestre, marítimo e atmosférico .....	40
3.4.4 Recuperação integrada: espaço, tempo e classes .....	42
<b>4 RESULTADOS</b> .....	<b>43</b>
4.1 MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA .....	43
4.2 MODELO LÓGICO ORIENTADO POR GRAFOS .....	46
4.3 GRAFO INTEGRADO: TERRESTRE, MARÍTIMO E ATMOSFÉRICO .....	55
4.4 RECUPERAÇÃO INTEGRADA: ESPAÇO, TEMPO E CLASSES .....	61
<b>5 DISCUSSÃO</b> .....	<b>67</b>
5.1 MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA .....	67
5.2 MODELO LÓGICO ORIENTADO POR GRAFOS .....	68
5.3 GRAFO INTEGRADO: TERRESTRE, MARÍTIMO E ATMOSFÉRICO .....	69
5.4 RECUPERAÇÃO INTEGRADA: ESPAÇO, TEMPO E CLASSES .....	70
<b>6 CONSIDERAÇÕES FINAIS</b> .....	<b>71</b>
6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS .....	72

<b>REFERÊNCIAS.....</b>	<b>73</b>
<b>APÊNDICE 1 – CONSULTAS EM AQL.....</b>	<b>79</b>
<b>APÊNDICE 2 – EVOLUÇÃO TEMPORAL DAS UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO .....</b>	<b>84</b>

## 1 INTRODUÇÃO

O gerenciamento de dados é o processo de organização, armazenamento, acesso e preservação de dados, diretamente associado à construção de uma estrutura organizada, acessível, e consistente, que suporte crescimento, flexibilidade e absorção de dados heterogêneos, onde a informação possa ser extraída, com toda a sua integridade. Desde há muito o gerenciamento de dados é considerado uma ferramenta fundamental para a tomada de decisões estratégicas e planejamento territorial espacial em abordagens multisetoriais como meio ambiente, economia e defesa (ZAUCHA; GEE, 2019; LIU; ZHOU, 2021; UN, 2016).

Dados são registrados a todo momento e em todos os pontos do planeta, desde a atmosfera, acima da superfície oceânica e terrestre, até as camadas mais profundas do subsolo. Os dados são extraídos de ocorrências, e seus efeitos geram consequências muito difíceis de serem entendidas e modeladas pelo ser humano sem antes sofrerem as devidas integrações necessárias ao entendimento simbólico dos mesmos. Observações naturais, ações e conceitos antrópicos coexistem no mesmo espaço e podem possuir variações de valores ao longo do tempo. Podemos citar como exemplo uma estação meteorológica marinha posicionada em um ponto do espaço (x, y, z), representada cartograficamente por uma boia onde são observados diferentes parâmetros físicos como a temperatura do ar, a velocidade e direção do vento, a radiação solar, a nebulosidade, etc, que possuem diferentes valores específicos para cada ponto de tempo t.

Assim os continentes, a atmosfera e os oceanos formam um sistema integrado, interconectado e balanceado (GRAVES *et al.*, 1984). A gestão de dados distribuídos por essas componentes é a base fundamental para o desenvolvimento de aplicações sensíveis como cenários decisórios (JARDIM *et al.*, 2022), planejamento territorial (XIA *et al.*, 2023) até a compreensão do Sistema-Terra (REICHSTEIN *et al.*, 2019).

Os dados terrestres, marítimos e atmosféricos são essencialmente geoespaciais e temporais, inerentemente relacionados, dinâmicos e de amplo espectro. Pode-se citar diversos domínios de aplicação como Biológicos, Meteorológicos, Oceanográficos, Sociais, Geológicos, etc (GRASSLE, 2000; KARSENTI *et al.*, 2011; BENWAY *et al.*, 2019; WOODRUFF *et al.*, 1987; JOHNSON *et al.*, 2020; DE LA TORRE-CASTRO *et al.*, 2017; TU *et al.*, 2018).



A literatura apresenta a evolução de diferentes abordagens sobre dados terrestres, marítimos e atmosféricos como integração, modelagem, análises, mineração de dados, visualização e aplicações (BAUMANN, 2021; MATSUOKA *et al.*, 2021; JAHANBAKHT *et al.*, 2021; LENSU; GOERLANDT, 2019; JEON *et al.*, 2021). Porém ressalta-se a afirmação de Yuan *et al.* (2004) que a representação computacional de fenômenos geográficos tridimensionais e dinâmicos e a integração de dados marinhos e terrestres numa base de dados global contínua permanecem como desafios de pesquisa, vide os trabalhos recentes de Ding *et al.* (2022) e Li *et al.* (2022).

Assim, a integração de dados espaço-temporais dos ambientes terrestres, marítimos e atmosféricos em um modelo de dados único envolve complexidades como a multidimensionalidade e a heterogeneidade dos dados.

O aspecto dinâmico da multidimensionalidade dos dados é determinado por comportamentos estáticos e dinâmicos, e a coexistência de diferentes entidades no espaço-tempo. Enquanto a heterogeneidade dos dados engloba aspectos como a variedade de modelos de dados, formatos e semântica que são consequência natural dos diversos domínios de aplicações na qual os dados espaço-temporais são usualmente representados de forma estanque.

Portanto, a integração e a capacidade de recuperação desses dados baseadas no tempo e espaço são os problemas fundamentais deste estudo.

Esse trabalho assume a hipótese que a abordagem mais adequada para a construção de uma estrutura de armazenamento visando à gestão de dados espaço-temporais seja a de um banco de dados orientado por grafos, pois foram desenvolvidos para lidar com dados altamente conectados e são flexíveis o suficiente para a expansão do modelo de dados. Os grafos são amplamente difundidos nas áreas de Inteligência Artificial e Ciência de Dados.

## 1.1 OBJETIVOS

Esse trabalho tem como objetivo geral a proposição de uma abordagem para a integração de dados espaço-temporais terrestres, marítimos e atmosféricos sob a forma vetorial em um banco de dados orientado por grafos.

Ressalte-se os seguintes objetivos específicos:

- a) Propor uma nova visão de modelo conceitual para integração de dados espaço-temporais de multidomínios;
- b) Desenvolver um modelo lógico sob a ótica do modelo conceitual proposto baseado em grafos; e
- c) Realizar a recuperação integrada de dados espaço-temporais segundo as necessidades operacionais.

## 1.2 JUSTIFICATIVA

Dois pontos são fundamentais para a integração de dados espaço-temporais nos ambientes terrestres, marítimos e atmosféricos em um modelo de dados único: a multidimensionalidade e a heterogeneidade dos dados.

O aspecto dinâmico da multidimensionalidade dos dados espaço-temporais é solucionado atualmente através do armazenamento de dados em modelos lógicos do tipo relacional; muitas vezes, com algum tipo de desenvolvimento adicional, como módulos, APIs e *frameworks*, e refletindo conseqüentemente nas suas manutenibilidade, integridade e escalabilidade (GEBBERT; PEBESMA, 2014; TIAN *et al.*, 2015; HULL *et al.*, 2014) e do tipo não-relacional, como a fornecida pela teoria dos grafos.

Nos últimos anos, a área de Geociências tem utilizado extensivamente a teoria dos grafos, como pode-se citar na Geomorfologia (Heckmann *et al.*, 2015), na Hidrologia (Poulter *et al.*, 2008), na Oceanografia (PETELIN *et al.*, 2019), na Meteorologia (BENTSEN *et al.*, 2023; LIU *et al.*, 2016; ZHANG, W *et al.*, 2023; WU *et al.*, 2022), transporte marítimo (DEL MONDO *et al.*, 2021; YAN *et al.*, 2020), Cartografia (DEL MONDO *et al.*, 2013), Hidrografia (THIBAUD *et al.*, 2013) etc. Inúmeras vantagens existem no armazenamento de dados espaço-temporais utilizando o modelo lógico em grafos, dentre as quais destacam-se: os relacionamentos são entidades de primeira classe, desempenho constante independentemente do tamanho do conjunto de dados e o esquema flexível (ROBINSON *et al.*, 2015; BECHBERGER; PERRYMAN, 2020).

A heterogeneidade dos dados espaço-temporais (variedade de modelos de dados, formatos e semântica) intrínseca à integração desses dados é apresentada na literatura por meio de diversos modelos de dados, com propósitos funcionais específicos que se refletem na quantidade de objetos e correspondente semântica

pertencente a cada modelo, e desenvolvida para um ambiente predominantemente marítimo, terrestre ou atmosférico. Assim, existem classificações duplicadas, com visões semânticas distintas, diversos formatos de dados e de abrangência limitada (DING *et al.*, 2022; IHO, 2023; DSG, 2023; USGS, 2019; INSPIRE, 2023; OSM, 2023).

Assim, com base nas considerações acima se verificou a necessidade de se pesquisar uma nova visão para o gerenciamento de dados espaço-temporais sem limitações de domínios de aplicações e com a flexibilidade possibilitada pelos grafos.

### 1.3 ESTRUTURA DO TRABALHO

Esta dissertação está organizada em seis capítulos. A introdução contextualizou o problema da pesquisa e apresentou os objetivos da mesma, o restante do trabalho está estruturado da seguinte forma. O capítulo 2 discute trabalhos relevantes sobre dados espaço-temporais. O capítulo 3 detalha a elaboração dos modelos conceitual e lógico, a construção do grafo e a recuperação integrada dos dados espaço-temporais. O capítulo 4 apresenta os resultados alcançados. O capítulo 5 discute os modelos e a recuperação de dados. As considerações finais estão apresentadas no capítulo 6.

## 2 FUNDAMENTAÇÃO

A integração dos dados espaço-temporais terrestres, marítimos e atmosféricos e a capacidade de recuperação integrada dos dados baseada no tempo e espaço são a motivação para esse estudo. Nesta seção, será apresentada a natureza dinâmica dos dados espaço-temporais, trabalhos relacionados com a classificação, representação e integração de dados, além de uma breve exposição sobre o modelo lógico orientado por grafos e o Sistema de Gerenciamento de Banco de Dados (SGBD) aberto selecionado para implantação, o ArangoDB, o qual será objeto de consideração posterior.

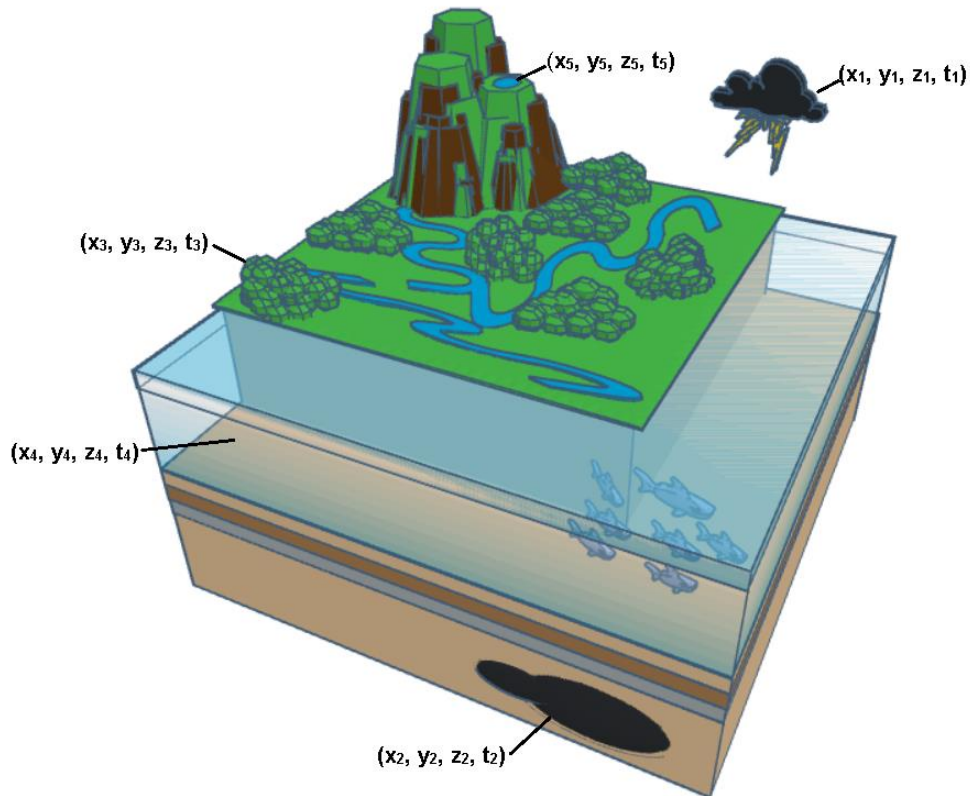
### 2.1 A NATUREZA DINÂMICA DOS DADOS ESPAÇO-TEMPORAIS

Observações naturais, ações e conceitos antrópicos ocorrem a todo momento e estão localizados em ambientes com características predominantemente terrestres, marítimas e atmosféricas do nosso planeta. Diversas entidades do mundo real coexistem no tempo e no espaço e interagem nos diversos ambientes que possuem fronteiras invisíveis entre eles.

Para os efeitos do presente trabalho, que primordialmente visa a generalização, considera-se que uma entidade qualquer se torna para as finalidades de generalização mencionadas um objeto. Uma coleção de objetos (1, ..., n) que está associado a quatro dimensões: longitude, latitude, altitude ou profundidade, e tempo, respectivamente, simbolizadas pelas letras x, y, z e t, é chamado de dado espaço-temporal.

A FIGURA 1 apresenta dados espaço-temporais, desde a atmosfera até as profundezas do solo oceânico, estando alguns objetos destacados e identificados por  $(x_i, y_i, z_i, t_i)$  variando i de 1 a 5. Observa-se que  $(x_i, y_i, z_i)$  representam posições geoespaciais diferentes inclusive no eixo z que podem estar acima,  $i=1, 3$  e  $5$ , ou abaixo na superfície de referência,  $i=2$  e  $4$ , sendo essa o nível médio dos mares. A componente temporal  $t_i$  pode assumir valores iguais para i de 1 a 5, no qual todos os objetos estão sob o mesmo domínio temporal, ou diferentes, se cada objeto foi observado em diferentes momentos  $t_i$ .

FIGURA 1 – DADOS ESPAÇO-TEMPORAIS



FONTE: Adaptada de Autodesk (2023).

Os dados espaço-temporais podem se comportar de forma dinâmica em todas as dimensões: podem envolver objetos em movimento contínuo; por exemplo, em um sistema de navegação os navios mudam de posição, mas não de forma; objetos com mudanças discretas de forma e posição, por exemplo, rios mudam de posição no sistema de informação cadastral; e a integração movimento contínuo e mudanças de forma, por exemplo, uma tempestade é medida como um fenômeno em movimento, que muda propriedades (por exemplo, intensidade) e forma ao longo do tempo (RENOLLEN, 2000; WANG *et al.*, 2000).

O tempo no qual os objetos foram observados no mundo real chama-se tempo referência. Esse é definido como momento em que o fato se torna efetivo na para fins de representação. Existem dois fatos básicos para os quais é necessário registrar o tempo: eventos e estados. Um evento ocorre em um momento exato, ou

seja, um evento que não tem duração é um ponto de tempo, exemplo, o horário de nascer do sol. Um estado é definido por um intervalo de tempo, portanto, tem uma duração, por exemplo, a precipitação teve início às 10 AM e término às 13 PM (HORNSBY; EGENHOFER, 2000; WANG; CHENG, 2001; TRYFONA; JENSEN, 1999).

A componente espacial  $(x, y)$  é representada por uma posição na superfície da Terra identificada por longitude e latitude  $(\lambda, \varphi)$  (LONGLEY *et al.*, 2005). A dimensão  $z$  pode ser suprimida, pela inexistência de semântica na representação, por exemplo, limites marítimos; pode estar ao nível do mar, profundidade e altitude iguais a 0; ou  $z$  pode absorver uma semântica de altitude ou profundidade, sendo ambas, por convenção positivas.

Também pode-se destacar a concorrência de objetos em uma mesma componente espaço e tempo; parâmetros meteorológicos e oceanográficos são exemplos desses objetos. Parâmetros físicos como velocidade e direção, temperatura, condutividade e pressão da água do mar são observados em um mesmo ponto  $(x, y, z)$  em um tempo  $t$ .

De maneira geral, um dado espaço-temporal é representado por uma tripla  $\langle s, t, a, a(s, t) \rangle$  onde  $s$  define o espaço,  $t$  o tempo,  $a$  identifica o atributo do objeto e  $a(s, t)$  define o valor específico do atributo naquele ponto (GOODCHILD *et al.*, 2007; YI; XIAO, 2015).

Portanto, o aspecto dinâmico da multidimensionalidade dos dados é fundamental para a modelagem dos dados espaço-temporais.

## 2.2 CLASSIFICAÇÃO DOS DADOS ESPAÇO-TEMPORAIS

Geralmente os dados espaço-temporais são organizados conforme as características predominantes do conjunto de dados. De modo geral, os modelos de dados são compostos por uma organização hierárquica de classes, atributos e domínios, todos com uma semântica associada (LAURINI; THOMPSON, 1992). A título ilustrativo podem-se citar classes como rodovias, trilhos, rios, lagos, limites político-administrativos federais, etc.

Devido à amplitude de objetos do mundo real diversos modelos de dados coexistem e foram construídos para um domínio de aplicação com propósitos funcionais específicos que se refletem na quantidade de objetos e correspondente

semântica pertencente a cada modelo. Laney (2001) escreveu “Até 2003/04, não existirá maior barreira à gestão eficaz de dados do que a variedade de formatos de dados incompatíveis, modelos de dados não alinhados e semântica de dados inconsistentes.” fato que se verifica até os dias atuais.

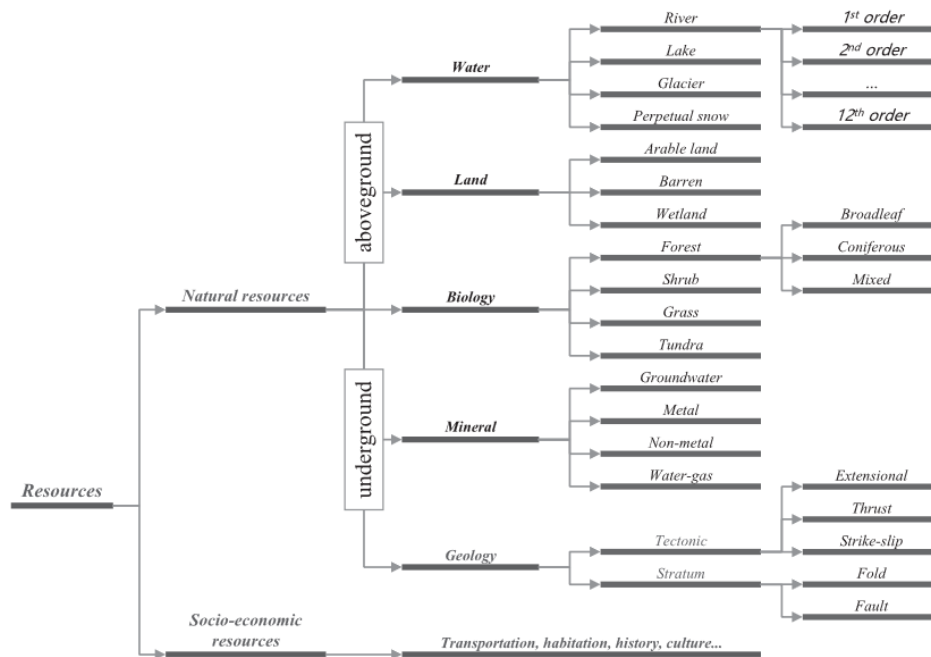
Essa heterogeneidade (ou Variedade caracterizada pelos 3Vs: Volume, Velocidade e Variedade, que são parte da definição de *Big Data*) é considerada o desafio mais importante na construção de sistemas integrados multidisciplinares e interorganizacionais (NATIVI *et al.* 2015). A heterogeneidade é potencializada pelas atuais facilidades tecnológicas como as progressivas possibilidades de acesso e redução de custos na aquisição de dados, na capacidade de processamento e armazenamento dos equipamentos eletrônicos, e conectividade que proporcionaram a expansão da aquisição de dados em quantidades e velocidades sem precedentes e, muitas das vezes, ininterruptos (DICKEY, 2003) e insubstituíveis (UNESCO, 2022).

Os domínios de aplicações desenvolvem sua própria visão sobre a organização e semântica dos objetos, como para o ambiente predominantemente marítimo tem-se o modelo de dados destinada a construção de cartas náuticas S-57 (*IHO Transfer Standard for Digital Hydrographic Data*) atualmente descontinuada e substituída pela S-100 (*IHO Universal Hydrographic Data Model*) (IHO, 2023); para o ambiente predominantemente terrestre pode-se citar a Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV) (DSG, 2023), a especificação do mapa digital topográfico do U. S. Geological Survey (USGS, 2019), a Infraestrutura de Informação Espacial na Europa (INSPIRE, 2023) e o modelo de dados geoespaciais do *OpenStreetMap* (OSM, 2023).

Devido às especializações entre os modelos de dados e as áreas de interseção entre elas, existem classificações duplicadas e com visões semânticas distintas, pode-se citar como exemplos, as áreas de Hidrografia, Oceanografia e Biologia Marinha.

Considerando isso, Ding *et al.* (2022) realizou uma generalização da classificação dos modelos de dados atuais com a incorporação de classes de cunho geral e elementar como água, terra e biologia, apresentada na FIGURA 2. Essa abordagem possibilita uma maior flexibilidade na incorporação de classes de representação de objetos, por exemplo, a classe rio está associada à água ao invés de hidrografia.

FIGURA 2 – CLASSIFICAÇÃO DAS CLASSES DE RECURSOS NATURAIS



FONTE: Ding et al. (2022)

A flexibilidade do modelo conceitual de classificação dos dados espaço-temporais é fundamental para a generalização da representação desses dados.

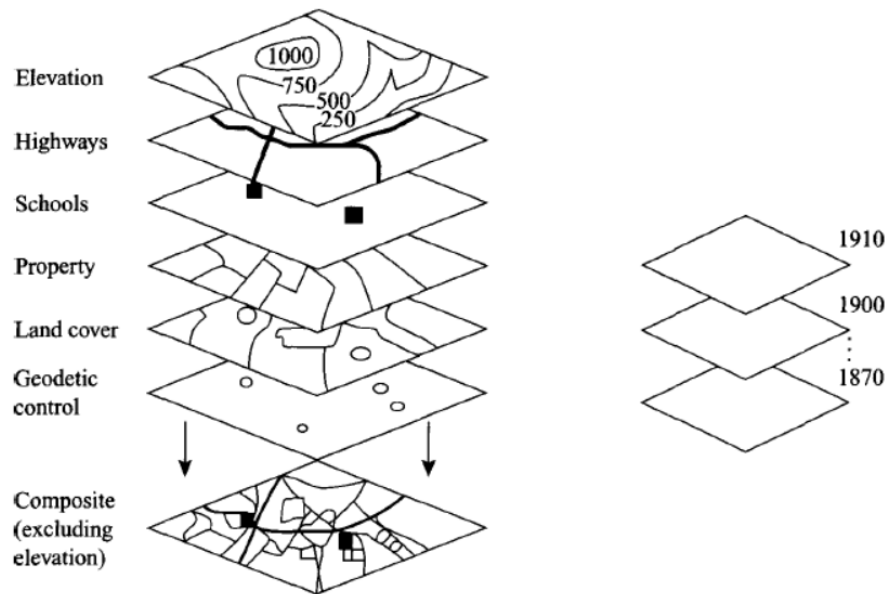
### 2.3 REPRESENTAÇÃO COMPUTACIONAL DOS DADOS ESPAÇO-TEMPORAIS

A maior parte dos estudos sobre a representação dos fenômenos no tempo e no espaço serviu para revelar que é significativamente mais complexo e mais difícil do que a sua representação apenas no espaço (PEUQUET, 1994). A abordagem cartográfica tradicional e estática fundamenta a estratégia frequentemente utilizada no contexto operacional do gerenciamento de dados espaço-temporais, implicando na supressão da dimensão temporal (PULTAR *et al.* 2009) ou um conjunto ordenado de instantâneos sequenciais, organizando o espaço ao longo do tempo, baseados em camadas independentes (ARMSTRONG, 1988) e o tratamento da dimensão z como um atributo discreto ou uma dimensão suprimida (LAURINI; THOMPSON, 1992), conforme a FIGURA 3. O problema dessa representação é a ausência do relacionamento entre as classes e a generalização do tempo para períodos. Peuquet e Duan (1995) propuseram um modelo de dados espaço-temporais centrado no tempo com valor discreto como base organizacional, implementado na linguagem C,



concebido para representar explicitamente a mudança do espaço em relação ao tempo, com o objetivo de facilitar os procedimentos de consultas relacionadas às relações temporais e as tarefas analíticas de comparação entre as diferentes sequências de alterações do espaço.

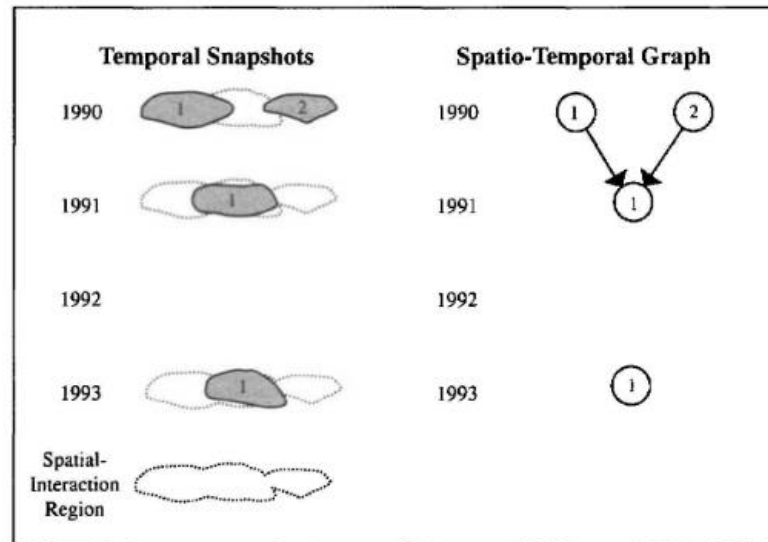
FIGURA 3 – CAMADAS ESTAÇÃO-TEMPORAIS COMO PRINCIPAL REPRESENTAÇÃO USADA EM GIS



FONTE: LAURINI; THOMPSON (1992)

De forma a facilitar a análise temporal entre instantâneos sucessivos, Wilcox *et al.* (2000) fizeram uso da ideia da teoria dos grafos para rastrear objetos poligonais a medida que eles mudam através do tempo e do espaço, armazenando os dados em tabelas no ArcInfo (FIGURA 4). Ideia semelhante foi aplicada a parcelas de campos agrícolas por Lin-Nan *et al.* (2007a).

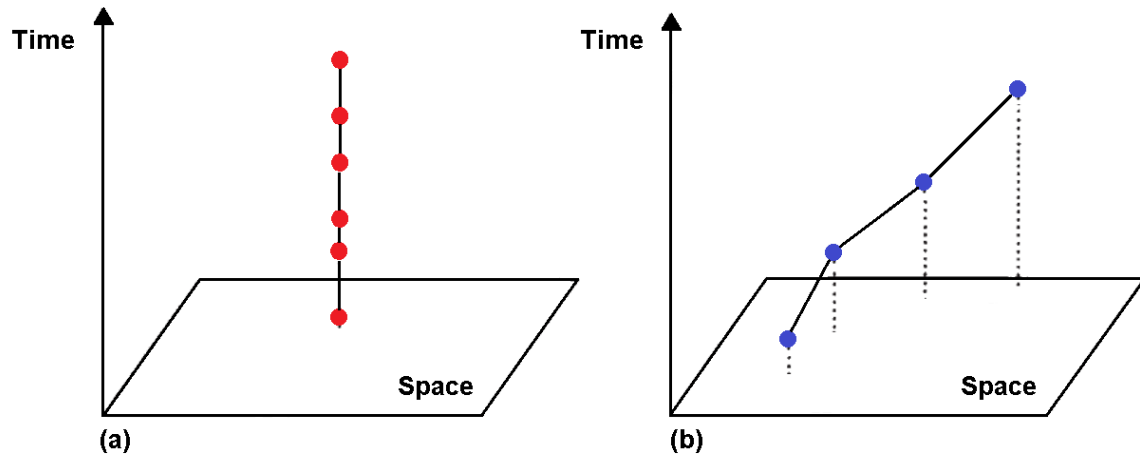
FIGURA 4 – UM GRAFO ESPAÇO-TEMPORAL REPRESENTANDO QUATRO INSTANTÂNEOS TEMPORAIS.



FONTE: WILCOX *et al.* (2000)

Em vez de armazenar dados em camadas e por períodos de tempo, as atuais abordagens de representação espaço-temporal são centradas na coexistência, conexão ou união dos elementos (WACHOWICZ, 2014). Nessa ideia, pode-se representar os dados espaço-temporais como triplas  $\langle s, t, a, a(s, t) \rangle$  e os seus correspondentes relacionamentos. A FIGURA 5 caso (a) representa o objeto estação meteorológica, onde são observados diferentes parâmetros físicos (objetos) como a temperatura do ar, a velocidade e direção do vento, a radiação solar, a nebulosidade, etc, em uma mesma posição (x, y, z) e em diferentes momentos (t). O caso (b) poderia ser ilustrado pela descrição feita por Le (2012) sobre a detecção de iodo radioativo: um nível muito baixo de iodo radioativo, I-131, foi detectado em uma amostra de ar perto de Seattle, Washington, em 20 de março de 2011; cinco dias depois, o I-131 foi detectado no leite da área de Spokane, Washington; esses eventos resultaram dos danos à usina nuclear do Japão causado pelo tsunami em 11 de março de 2011; este iodo radioativo foi liberado da usina danificada e viajou pelo ar através do Oceano Pacífico até o estado de Washington.

FIGURA 5 – EXEMPLOS DE DADOS ESPAÇO-TEMPORAIS



FONTE: O autor (2023)<sup>1</sup>.

A representação dos dados espaço-temporais como triplas  $\langle s, t, a, a(s, t) \rangle$  e os seus correspondentes relacionamentos são fundamentais para a recuperação integrada de dados espaço-temporais baseados no tempo e espaço.

## 2.4 INTEGRAÇÃO DE DADOS ESPAÇO-TEMPORAIS

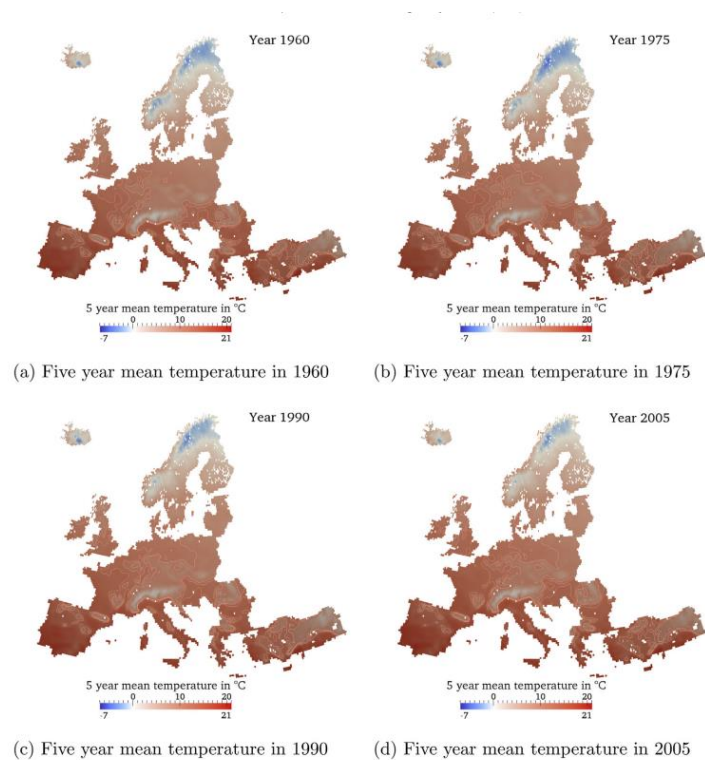
Quando se trata de integração de dados espaço-temporais a literatura apresenta amplo espectro de implementações de tecnologias como o armazenamento de dados espaço-temporais em modelos lógicos do tipo relacional e não-relacional até a ligação entre repositórios de dados através do desenvolvimento de APIs e *frameworks*. Independentemente da tecnologia utilizada o objetivo fundamental é a recuperação integrada de dados espaço-temporais baseado no tempo, espaço e em classes.

Gebbert e Pebesma (2014) propuseram uma abordagem baseada na integração de um Banco de Dados (BD) temporal e um BD relacional no GRASS através de API e módulos, escritos em Python, aplicados a dados espaço-temporais

<sup>1</sup> Ao longo do documento, algumas figuras elaboradas pela autora serão apresentadas na língua inglesa.

ambientais. Os autores apresentaram visualizações da recuperação de dados ambientais, como temperatura média do ar, baseados no tempo (FIGURA 6).

FIGURA 6 – SEQUÊNCIA DE QUATRO CAPTURAS DE TELA MOSTRANDO A TEMPERATURA MÉDIA DE CINCO ANOS DA UNIÃO EUROPEIA E DA TURQUIA DE 1952 A 2008.



Fonte: GEBBERT; PEBESMA (2014)

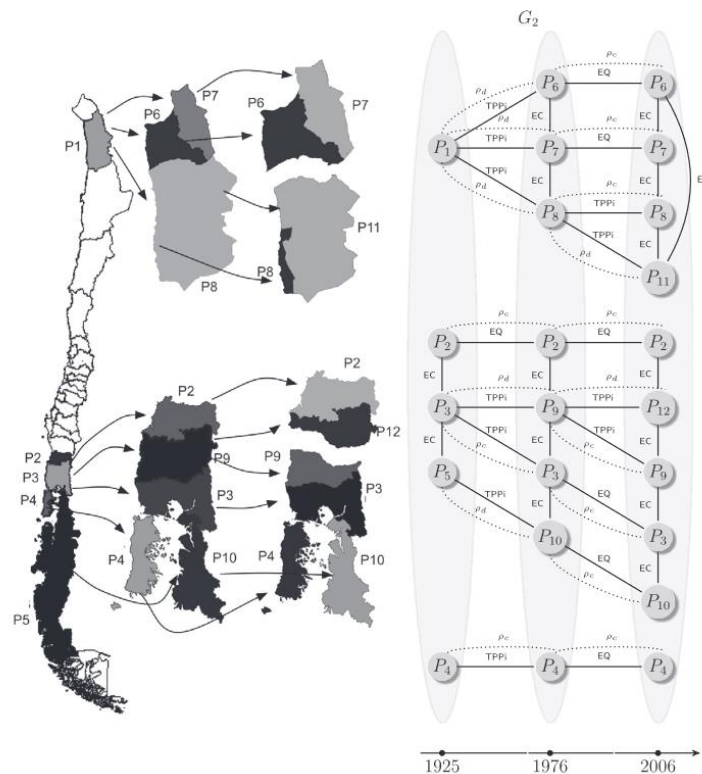
Tian *et al.* (2015) também realizaram uma recuperação integrada de classes de infraestrutura (estradas) baseada no tempo e no espaço utilizando um BD relacional. Hull *et al.* (2014) integraram dados espaço-temporais do ecossistema marinho, aquacultura, transporte marítimo, turismo, etc; através da construção de um framework. Nesse trabalho, os dados foram recuperados a partir das variáveis tempo e espaço e a visualização dos resultados foi apresentada através de gráficos e mapas.

A área de Geociências tem feito amplo uso da teoria dos grafos nos últimos anos, como pode-se citar na Geomorfologia (HECKMANN *et al.*, 2015), na Hidrologia (POULTER *et al.*, 2008), na Oceanografia (PETELIN *et al.*, 2019), na Meteorologia (BENTSEN *et al.*, 2023), transporte marítimo (DEL MONDO *et al.*, 2021), etc.

Bentsen *et al.* (2023) propuseram uma metodologia para previsão espaço-temporal da velocidade do vento baseada em grafos para aplicação em usinas eólicas. O trabalho utilizou, entre outros dados, parâmetros meteorológicos observados em estações físicas oceânicas.

Del Mondo *et al.* (2013) apresentaram uma modelagem de um grafo espaço-temporal da evolução das áreas das províncias e regiões do Chile (FIGURA 7). Del Mondo *et al.* (2021) apresentaram um estudo prospectivo do potencial de grafos espaço-temporais para a modelagem de fenômenos geográficos no contexto do transporte marítimo.

FIGURA 7 – GRAFO ESPAÇO-TEMPORAL DA EVOLUÇÃO DAS PROVÍNCIAS DO CHILE



Fonte: DEL MONDO *et al.* (2013)

Thibaud *et al.* (2013) apresentaram um modelo baseado em grafos para a representação e análise da dinâmica espaço-temporal de dunas marinhas. Yan *et al.* (2020) propuseram um método de extração de rotas marítimas baseado em dados do sistema de identificação automática (AIS) utilizando grafos. Zhang, W. *et al.* (2023) propuseram um método para monitoramento de desastres agrometeorológicos com base em um grafo espaço-temporal. Liu *et al.* (2016)

apresentaram a representação e análise de eventos de tempestade baseada em um grafo espaço-temporal direcionado. Wu *et al.* (2022) apresentaram a aplicação de grafos na área de meteorologia considerando o valor da altura (ou elevação) das estações.

Inúmeras vantagens existem no armazenamento de dados espaço-temporais utilizando o modelo lógico em grafos, dentre as quais destacam-se: os relacionamentos são entidades de primeira classe, desempenho constante independentemente do tamanho do conjunto de dados e o esquema flexível (ROBINSON *et al.*, 2015; BECHBERGER; PERRYMAN, 2020).

Pode-se concluir que todos os trabalhos citados possuem um domínio de aplicação pré-definido, conseqüentemente, autolimitado.

Wang *et al.* (2000) ressaltam que no gerenciamento de dados espaço-temporais, a simples agregação de espaço e tempo é inadequada. A simples conexão de um modelo de dados espaciais a um modelo de dados temporais resultará em um modelo de dados temporais que pode capturar dados espaciais ou em um modelo de dados espaciais que pode capturar sequências de dados espaciais referenciadas no tempo. Em vez disso, as características temporais dos objetos espaciais devem ser investigadas para produzir conceitos inerentemente espaço-temporais, como modelos de dados espaço-temporais unificadas, operadores espaço-temporais e interfaces de usuário espaço-temporais.

Com base nas considerações acima se verificou a necessidade de se pesquisar uma nova visão para o gerenciamento de dados espaço-temporais sem limitações de domínios e com a flexibilidade possibilitada pelos grafos.

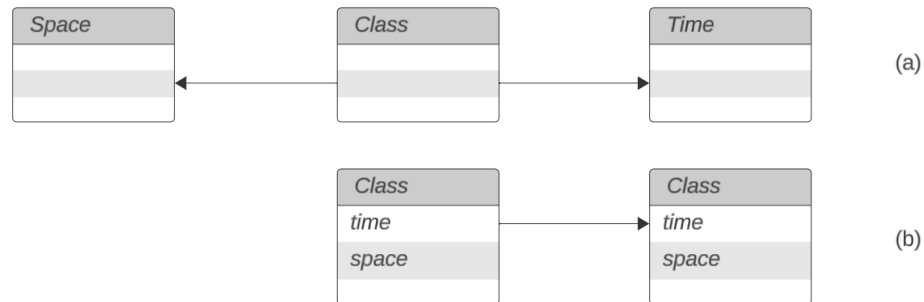
## 2.5 MODELO LÓGICO ORIENTADO POR GRAFOS

Um grafo é uma representação de nós conectados por arestas (WEST, 2000). Cada uma de suas arestas é representada por uma tupla  $\langle e^s, r, e^o \rangle$ , que indica a relação  $r$  entre entidade sujeito  $e^s$  e a entidade objeto  $e^o$  (LIU *et al.*, 2019). Os repositórios de dados que utilizam o modelo lógico baseado em grafos permitem o armazenamento de atributos associados aos nós e as arestas.

As componentes temporais e espaciais dos dados são comumente armazenadas como classes singulares (DING *et al.*, 2022) ou como atributos de

outras classes (LIU *et al.*, 2016), representados em UML (*Unified Modeling Language*) e identificados, respectivamente, pelas letras a e b da FIGURA 8.

FIGURA 8 – ARMAZENAMENTO DO ESPAÇO E TEMPO EM GRAFOS REPRESENTADO EM UML



FONTE: O autor (2023).

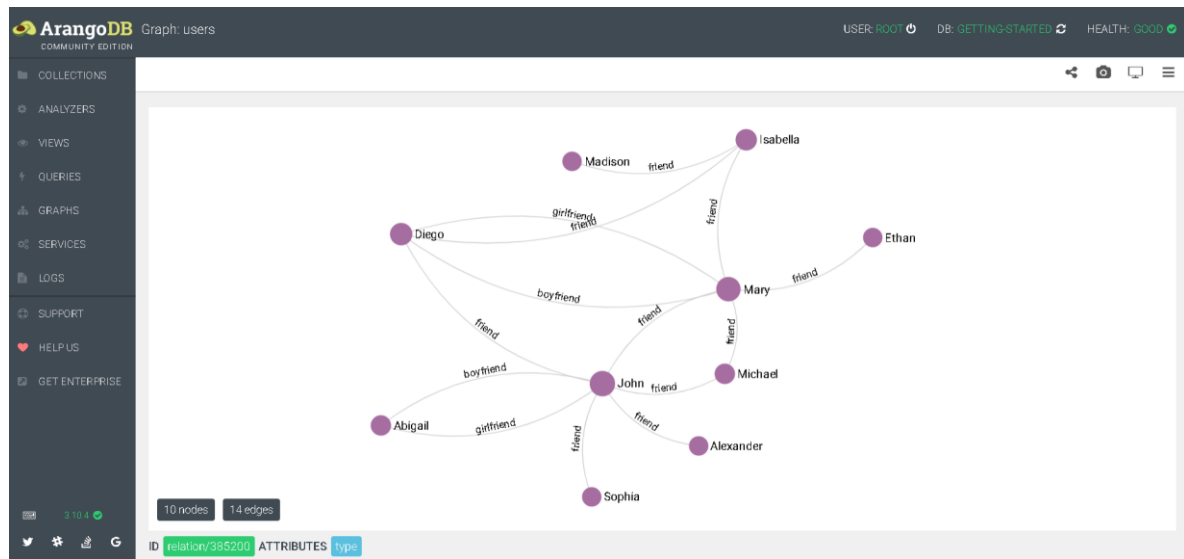
Independentemente do armazenamento da informação espacial e temporal ser realizada sob a forma de classes ou atributos, é necessário que o modelo lógico tenha classes que desempenhem a tarefa de filtros e orientem as buscas no grafo, sejam elas quais forem, evitando que as consultas o percorram por completo.

### 2.5.1 SGBD ArangoDB

ArangoDB é um SGBD não relacional multimodelo, pois suporta três modelos de dados (grafos, documentos JSON, chave-valor) (ARANGO, 2023). O ArangoDB possui duas versões: *Community Edition* e *Enterprise Edition*. A primeira é *open source* e a segunda possui recursos extras no foco da resolução de problemas em escala empresarial. O acesso aos dados, independentemente do tipo, ocorre a partir da linguagem de consulta *ArangoDB Query Language* (AQL) que possui uma estrutura similar com a *Structured Query Language* (SQL).

O ArangoDB é um SGBD que utiliza o modelo de grafo composto por classes, relacionamentos e atributos (FIGURA 9).

FIGURA 9 – EXEMPLO DE BANCO DE DADOS ORIENTADO POR GRAFOS NO SGBD ARANGODB



FONTE: O autor (2023)

Outras vantagens do ArangoDB são: a existência de *drivers* para várias linguagens de programação, incluindo Java, JavaScript e Python; possui uma interface Web em que é possível visualizar os grafos e realizar consultas; e possui uma boa documentação. No ArangoDB as instâncias são armazenadas em documentos JSON os quais armazenam os atributos na forma de pares chave-valor (FIGURA 10).

FIGURA 10 – EXEMPLO DE NÓS ARMazenADOS EM UM DOCUMENTO JSON

```

C:\Users\anaem\Videos\amazon-meta\users.json - Notepad++
Arquivo Editar Localizar Visualizar Formatar Linguagem Configurações Ferramentas Macro Executar Plugins Janela ?
users.json relation.json
1
2 { "_id": "users/100", "_key": "100", "name": "John", "age": 37, "active": true, "gender": "m" },
3 { "_id": "users/101", "_key": "101", "name": "Fred", "age": 36, "active": true, "gender": "m" },
4 { "_id": "users/102", "_key": "102", "name": "Jacob", "age": 35, "active": false, "gender": "m" },
5 { "_id": "users/103", "_key": "103", "name": "Ethan", "age": 34, "active": false, "gender": "m" },
6 { "_id": "users/104", "_key": "104", "name": "Michael", "age": 33, "active": true, "gender": "m" },
7 { "_id": "users/105", "_key": "105", "name": "Alexander", "age": 32, "active": true, "gender": "m" },
8 { "_id": "users/106", "_key": "106", "name": "Daniel", "age": 31, "active": true, "gender": "m" },
9 { "_id": "users/107", "_key": "107", "name": "Anthony", "age": 30, "active": true, "gender": "m" },
10 { "_id": "users/108", "_key": "108", "name": "Jim", "age": 29, "active": true, "gender": "m" },
11 { "_id": "users/109", "_key": "109", "name": "Diego", "age": 28, "active": true, "gender": "m" },
12 { "_id": "users/200", "_key": "200", "name": "Sophia", "age": 37, "active": true, "gender": "f" },
13 { "_id": "users/201", "_key": "201", "name": "Emma", "age": 36, "active": true, "gender": "f" },
14 { "_id": "users/202", "_key": "202", "name": "Olivia", "age": 35, "active": false, "gender": "f" },
15 { "_id": "users/203", "_key": "203", "name": "Madison", "age": 34, "active": true, "gender": "f" },
16 { "_id": "users/204", "_key": "204", "name": "Chloe", "age": 33, "active": true, "gender": "f" },
17 { "_id": "users/205", "_key": "205", "name": "Eva", "age": 32, "active": false, "gender": "f" },
18 { "_id": "users/206", "_key": "206", "name": "Abigail", "age": 31, "active": true, "gender": "f" },
19 { "_id": "users/207", "_key": "207", "name": "Isabella", "age": 30, "active": true, "gender": "f" },
20 { "_id": "users/208", "_key": "208", "name": "Mary", "age": 29, "active": true, "gender": "f" },
21 { "_id": "users/209", "_key": "209", "name": "Mariah", "age": 28, "active": true, "gender": "f" }
22
JSON file length: 2.024 lines: 22 Ln: 22 Col: 2 Pos: 2.025 Windows (CR LF) UTF-8 INS

```

FONTE: O autor (2023)



Os relacionamentos conectam as instâncias e estruturam o grafo. No ArangoDB, os relacionamentos são armazenados em documentos JSON que armazenam os atributos na forma de pares chave-valor e o relacionamento pelos atributos `_from` e `_to` que, respectivamente, indicam a direção do relacionamento, respectivamente, o nó inicial e o nó final (FIGURA 11).

FIGURA 11 – EXEMPLO DE ARESTAS ARMazenADAS EM UM DOCUMENTO JSON

```

1 [{"_from": "users/209", "_to": "users/205", "type": "friend" },
2 {"_from": "users/206", "_to": "users/108", "type": "friend" },
3 {"_from": "users/202", "_to": "users/204", "type": "friend" },
4 {"_from": "users/200", "_to": "users/100", "type": "friend" },
5 {"_from": "users/205", "_to": "users/101", "type": "friend" },
6 {"_from": "users/209", "_to": "users/203", "type": "friend" },
7 {"_from": "users/200", "_to": "users/203", "type": "friend" },
8 {"_from": "users/100", "_to": "users/208", "type": "friend" },
9 {"_from": "users/101", "_to": "users/209", "type": "friend" },
10 {"_from": "users/206", "_to": "users/102", "type": "friend" },
11 {"_from": "users/104", "_to": "users/100", "type": "friend" },
12 {"_from": "users/104", "_to": "users/108", "type": "friend" },
13 {"_from": "users/108", "_to": "users/209", "type": "friend" },
14 {"_from": "users/206", "_to": "users/106", "type": "friend" },
15 {"_from": "users/204", "_to": "users/105", "type": "friend" },
16 {"_from": "users/208", "_to": "users/207", "type": "friend" },
17 {"_from": "users/102", "_to": "users/108", "type": "friend" },
18 {"_from": "users/207", "_to": "users/203", "type": "friend" },
19 {"_from": "users/203", "_to": "users/106", "type": "friend" },
20 {"_from": "users/202", "_to": "users/108", "type": "friend" },
21 {"_from": "users/201", "_to": "users/203", "type": "friend" },
22 {"_from": "users/105", "_to": "users/100", "type": "friend" },
23 {"_from": "users/100", "_to": "users/109", "type": "friend" },
24 {"_from": "users/207", "_to": "users/109", "type": "friend" },
25 {"_from": "users/103", "_to": "users/203", "type": "friend" },
26 {"_from": "users/208", "_to": "users/104", "type": "friend" },
27 {"_from": "users/105", "_to": "users/104", "type": "friend" },
28 {"_from": "users/103", "_to": "users/208", "type": "friend" },
29 {"_from": "users/203", "_to": "users/107", "type": "boyfriend" },
30 {"_from": "users/107", "_to": "users/203", "type": "girlfriend" },
31 {"_from": "users/208", "_to": "users/109", "type": "boyfriend" },
32 {"_from": "users/109", "_to": "users/208", "type": "girlfriend" },
33 {"_from": "users/106", "_to": "users/205", "type": "girlfriend" },
34 {"_from": "users/205", "_to": "users/106", "type": "boyfriend" },
35 {"_from": "users/103", "_to": "users/209", "type": "girlfriend" },
36 {"_from": "users/209", "_to": "users/103", "type": "boyfriend" },
37 {"_from": "users/201", "_to": "users/102", "type": "boyfriend" },
38 {"_from": "users/102", "_to": "users/201", "type": "girlfriend" },
39 {"_from": "users/206", "_to": "users/100", "type": "boyfriend" },
40 {"_from": "users/100", "_to": "users/206", "type": "girlfriend" }
41 ]
42 ]

```

FONTE: O autor (2023)

O ARANGODB armazena geometrias como *LineString*, *Point*, *Polygon*, *MultiPolygon*, etc. O AQL suporta funções para consultas geoespaciais como por exemplo *is\_in\_polygon()* que seleciona feições que estão contidas dentro do polígono. Os pares de coordenadas geoespaciais, latitude e longitude, são

armazenados no sistema de referência de coordenadas geográficas, *World Geodetic System 1984* (WGS-84) em unidades de graus decimais.

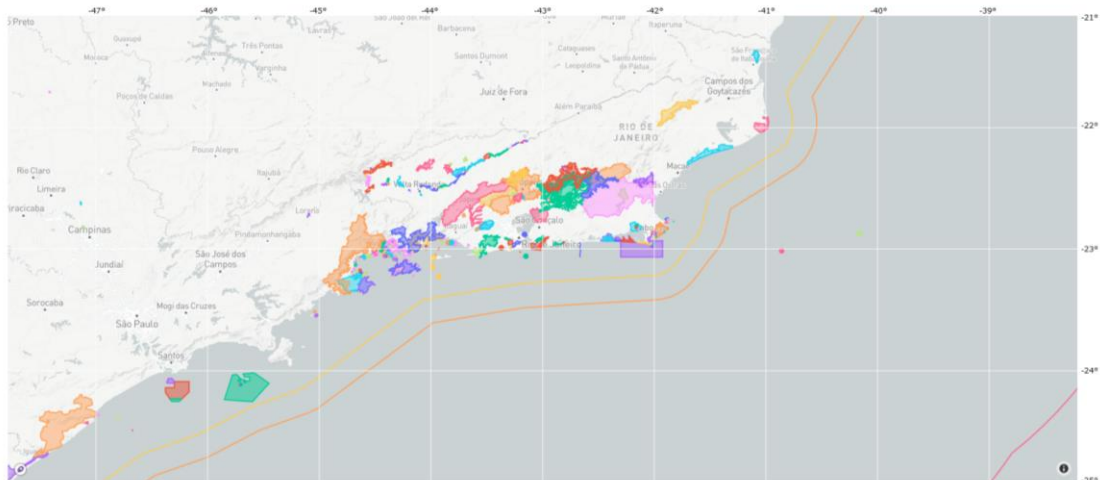
### 3 MATERIAL E MÉTODOS

Essa seção apresenta a área de cobertura dos dados espaço-temporais, a caracterização desses dados, as etapas pertencentes ao fluxo de trabalho da metodologia proposta, e os recursos computacionais que serão utilizados para a sua execução.

#### 3.1 RECORTE TERRITORIAL

Os dados espaço-temporais são compostos por objetos do mundo real localizados nos ambientes marítimos, terrestres e atmosféricos, entre os paralelos  $-21^\circ$  e  $-25^\circ$  e os meridianos  $-38.5^\circ$  e  $-47.5^\circ$ , localizados entre o Estado do Rio de Janeiro e a cidade de Santos-SP, conforme apresentado na FIGURA 12.

FIGURA 12 – LIMITES DO RECORTE TERRITORIAL



FONTE: O autor (2023).

#### 3.2 DADOS UTILIZADOS

Os dados espaço-temporais foram obtidos através de diversas fontes de dados como estações meteorológicas, leis e decretos, boias meteoceanográficas e equipamentos oceanográficos, QUADRO 1. Todos os dados foram obtidos através do portal da Infraestrutura Nacional de Dados Espaciais (INDE, 2023), exceto os dados de temperatura da água do mar que foram obtidos através do projeto Coriolis

(CORIOLIS, 2023) mantido por várias entidades francesas. Ambos os portais promovem o compartilhamento de dados geográficos tornando-os facilmente acessíveis ao público. A aplicação direta desses dados na pesquisa é o seu armazenamento e recuperação através de um BD orientado por grafos.

QUADRO 1 – DESCRIÇÃO DAS FONTES DE DADOS

Dados	Descrição
Estações meteorológicas	Estações físicas de observação de parâmetros meteorológicos.
Limites marítimos	Divisão física de áreas marítimas da Terra que usam critérios geopolíticos e fisiográficos. Lei n.8617 de 4 de janeiro de 1994 e Convenção das Nações Unidas pelo Direito do Mar.
Boias meteoceanográficas	Estações físicas flutuantes de observação de parâmetros meteorológicos e oceanográficos.
Unidades de conservação ambiental	São áreas territoriais, incluindo seus recursos ambientais químicos, biológicos e físicos, com características naturais relevantes, criadas por leis e decretos e protegidas pelo Poder Público com objetivos de conservação.
Temperatura da água do mar	Parâmetro físico observado através de equipamentos oceanográficos.

FONTE: O autor (2023).

A TABELA 1 apresenta os dados espaço-temporais que foram reunidos para a execução dessa pesquisa, indicando para cada tipo de dado a cobertura temporal e o formato de armazenamento.

TABELA 1 – COBERTURA TEMPORAL E FORMATO DE ARMAZENAMENTO DOS DADOS

Tipo de Dado	Identificador	Período	Formato
Estação meteorológica	83054 – Jacarepaguá	01/12/2022 – 04/11/2023	csv
Estação meteorológica	83117 – Ilha rasa	04/01/2023 – 04/11/2023	csv
Limites marítimos	12 NM, 24 NM e 200 NM	04/01/1994	shp
Boia meteoceanográfica	AX23 – Bacia de Campos	25/07/2023 – 30/08/2023	csv
Boia meteoceanográfica	BM02 – Bacia de Campos 2	02/02/2022 – 30/08/2023	csv
Boia meteoceanográfica	BM01- Bacia de Santos	10/12/2020 – 30/08/2023	csv
Boia meteoceanográfica	Alcatrazes	05/04/2022 – 30/08/2023	csv
Unidades de Conservação Ambientais Federais	UC	14/06/1937 – 05/06/2023	shp
Unidades de Conservação Ambientais do Estado do RJ	UC	22/05/1969 – 01/07/2023	shp
Temp. da água do mar	196312(49-67)	28/10/2019 – 29/10/2019	csv
Temp. da água do mar	168514(13-18)	31/05/2015	csv
Temp. da água do mar	16850072	31/03/2015	csv

FONTE: O autor (2023)

A TABELA 2 apresenta os objetos e as unidades de medida presentes nos dados espaço-temporais e armazenados no BD.

TABELA 2 – UNIDADES DE MEDIDAS DOS OBJETOS UTILIZADOS NESSE ESTUDO

Objeto	Símbolo	Unidade de medida	Referência
Temperatura do ar	°C	graus celsius	air_temperature
Temperatura do ar maxima	°C/dia	graus celsius por dia	max_air_temperature
Temperatura do ar minima	°C /dia	graus celsius por dia	min_air_temperature
Umidade do ar	%	porcentagem	relative_humidity
Pressão do ar	HPa	hectopascal	air_pressure
Velocidade do vento	m/s	Metros por segundo	wind_speed
Direção do vento	°	graus	wind_direction
Nebulosidade	Décimos	decimos	nebulosity
Insolação	H	ângulo horário	solar_radiation
Chuva	mm/dia	Milímetros por dia	daily_rain
Sea Pressure	Dbar	decibar	sea_water_pressure
Profundidade	M	metros	depth
Elevação	M	metros	elevation
Temperatura do mar	°C	graus celsius	sea_water_temperature
Limites marítimos	NM	Milhas náuticas	maritime_zones

FONTE: O autor (2023)

### 3.3 RECURSOS COMPUTACIONAIS UTILIZADOS

No QUADRO 2 podem ser observados os programas computacionais que foram utilizados durante a execução do fluxo de trabalho.

Geralmente os SGBD *open source* orientados por grafos possuem limites de funcionalidades, por exemplo, o tamanho do grafo permitido. O SGBD ArangoDB foi selecionado para implementar o BD orientado por grafos pois é *open source* e possui disponível ampla documentação ao mesmo tempo em que não foi detectado, para os objetivos do presente projeto, nenhum limite de funcionalidade, e ele aceita a importação e exportação de dados a partir de arquivos JSON, o que traz uma maior flexibilidade para o futuro intercâmbio de dados. Nenhuma análise de desempenho foi realizada para a seleção do SGBD, pois não interfere nos objetivos propostos nesse trabalho.

QUADRO 2 – PROGRAMAS COMPUTACIONAIS QUE FORAM UTILIZADOS NA EXECUÇÃO DO FLUXO DE TRABALHO

Programa	Descrição	Fonte	Objetivo/Aplicação
ArangoDB 3.x	SGBD <i>open source</i> , multimodelo nativo para grafos, documento e chave-valor.	ARANGO (2023)	Implementação do SGBD.
Scilab	É um software científico para computação numérica que fornece um ambiente computacional aberto para aplicações científicas	SCILAB (2023)	Processamento dos dados espaço-temporais.
Notepad++	É um editor de texto e de código fonte de código aberto.	NOTEPA D++ (2023)	Visualização dos arquivos JSON.
Python	É uma linguagem de programação de alto nível, interpretada de <i>script</i> , imperativa, orientada a objetos, funcional, de tipagem dinâmica e forte.	PYTHON (2023)	Consulta ao banco de dados
QGIS 3.24.2	É um software livre com código-fonte aberto, multiplataforma de sistema de informação geográfica que permite a visualização, edição e análise de dados georreferenciados	QGIS (2023)	Conversão dos dados para o WGS-84 e exportação para o formato JSON

FONTE: O autor (2023).

No QUADRO 3 podem ser observadas as especificações da máquina que foi utilizada durante a execução do fluxo de trabalho.

QUADRO 3 – ESPECIFICAÇÕES DA MÁQUINA UTILIZADA PARA A EXECUÇÃO DO FLUXO DE TRABALHO

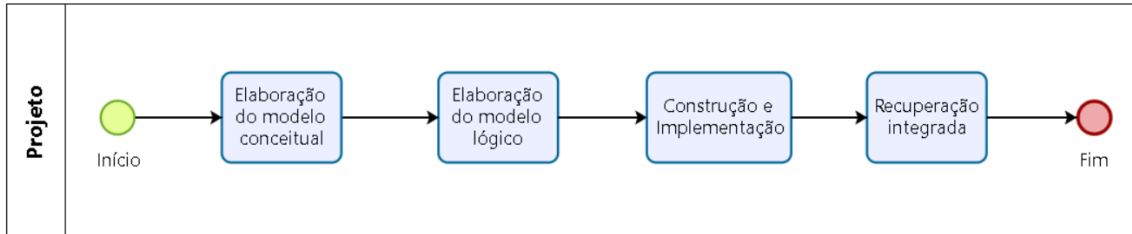
Componente	Descrição
Processador	Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.00GHz
Memória RAM	20 GB
Sistema operacional	Windows 10 Home (64 bits)
GPU	Dedicada NVIDIA GeForce MX110 2GB

FONTE: O autor (2023).

### 3.4 PROCEDIMENTO

Durante o cumprimento do objetivo da pesquisa, o mesmo apresenta duas fases bem definidas: a integração dos dados espaço-temporais e a sua recuperação através de consultas baseadas no tempo, no espaço e em classes. Essas duas fases serão executadas em um fluxo de trabalho contínuo, FIGURA 13.

FIGURA 13 – RESUMO ESQUEMÁTICO DO FLUXO DE TRABALHO



FONTE: O autor (2023)

A primeira fase engloba as atividades de elaboração do modelo conceitual e lógico, e a construção e implementação do grafo. A segunda fase é caracterizada pela etapa “recuperação integrada”. A seguir apresentam-se as descrições detalhadas dos processos integrantes do fluxo de trabalho.

#### 3.4.1 Modelo conceitual de representação computacional integrada

Essa etapa tem por objetivo principal a elaboração do modelo conceitual de representação computacional integrada de dados espaço-temporais localizados nos ambientes marítimo, terrestre e atmosférico. A ideia principal do modelo conceitual reside no fato que existem objetos do mundo real que compartilham um único local no espaço e no tempo. Além disso, as classes de representação dos objetos já existentes em outros modelos de dados podem ser reorganizadas considerando uma hierarquia de classes que possui semânticas mais abrangentes.

#### 3.4.2 Modelo lógico orientado por grafos

Essa etapa tem por objetivo principal a elaboração do modelo lógico orientado por grafos, tendo como base o modelo conceitual. Nessa etapa foram definidos os nós, as relações e principais atributos utilizando um diagrama de classes em UML.

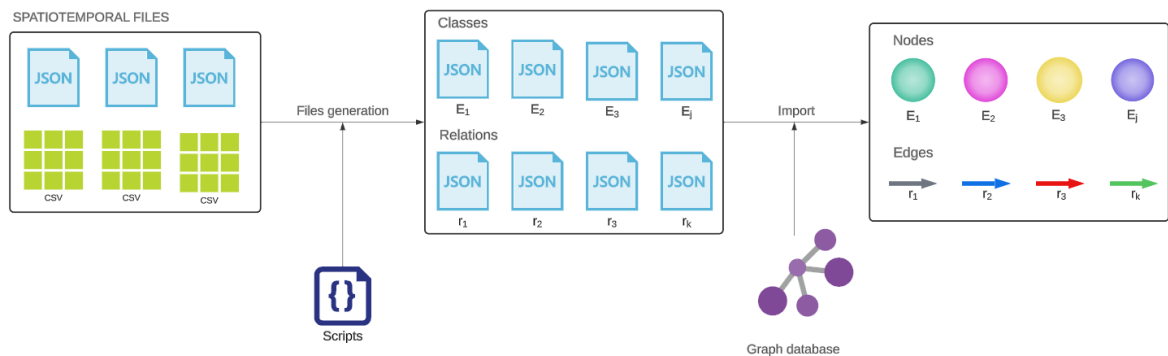
A partir de um protótipo do modelo lógico, e com a seleção de alguns dados espaço-temporais obtidos através das várias fontes de dados, o modelo inicial foi

implementado no banco de dados orientado por grafos, a partir do qual estabeleceu-se um processo cíclico de melhoria contínua.

### 3.4.3 Grafo integrado: terrestre, marítimo e atmosférico

A FIGURA 14 apresenta o processo de construção e implementação do grafo, desde os arquivos de armazenamento de dados espaço-temporais até as correspondentes representações no grafo - os nós e as arestas. O principal objetivo desse processo foi realizar o processamento dos arquivos de entrada com a extração e formatação das instâncias e a geração de arquivos JSON de classes e relacionamentos previstos no modelo lógico em UML. Posteriormente, esses arquivos foram importados para o BD orientado por grafos onde os arquivos JSON de classes e relacionamentos foram convertidos, respectivamente, para nós e arestas.

FIGURA 14 – REPRESENTAÇÃO CONCEITUAL DA CONSTRUÇÃO DE UM GRAFO A PARTIR DE DADOS ESPAÇO-TEMPORAIS HETEROGÊNEOS



FONTE: O autor (2023).

Os dados espaço-temporais estão armazenados em dois formatos de arquivos shapefile e csv. A quantidade de objetos e a organização dos dados são variáveis dependendo da fonte fornecedora dos dados. A superfície de referência do BD será a do sistema de referência WGS-84, portanto todos os dados espaço-temporais de entrada estão ou foram convertidos utilizando o programa QGIS, o qual permite essa conversão com facilidade. Os arquivos do formato shapefile foram exportados para o formato JSON, também utilizando o programa QGIS.



Geração de Arquivos: O processamento dos arquivos de entrada foi realizado por grupos reunidos de acordo com a sua semelhança de organização. Para cada grupo foi construído um *script* que lê o arquivo no formato csv ou JSON, extrai e formata as instâncias; e, finalmente, exporta sob a forma de arquivos individuais de classes e relacionamentos no formato JSON. Os *scripts* de processamento foram implementados e executados no programa Scilab. A extração e formatação das instâncias das classes e relacionamentos estão condicionadas ao modelo lógico construído na etapa “Modelo lógico orientado por grafos” do fluxo de trabalho.

Os *scripts* utilizaram, na formatação dos valores dos objetos, os critérios apresentados na TABELA 3 e ferramentas de formatação como a concatenação de *strings*, a filtragem por *Regular Expressions* (REGEX) e *replace*. Além disso, foram consideradas a limpeza de dados espúrios com critérios estabelecidos pelas fontes de dados; a padronização de ponto decimal; e a reorganização das coordenadas geográficas, latitude e longitude, para o sentido anti-horário nas geometrias do tipo *Polygon* e *MultiPolygon* (ArangoDB, 2023).

TABELA 3 – PADRONIZAÇÃO DOS OBJETOS UTILIZADOS NESSE ESTUDO

<b>Objetos</b>	<b>Padrão</b>
tempo referência	YYYY-MM-DDTHH:MM:SSZ (ISO 8601)
profundidade	Positiva para baixo, contada a partir do nível médio dos mares
elevação	Positiva para cima, contada a partir do nível médio dos mares

FONTE: O autor (2023)

Importação: Cada arquivo de classes corresponde a um tipo de nó, que registra o conteúdo não redundante de cada tipo. Cada arquivo de relacionamento significa um tipo de aresta. Assim, o passo seguinte é importar os arquivos para o BD orientado por grafos para armazenar os dados. Na prática, os arquivos de classes são primeiro importados completamente e lidos linha por linha, pelo BD, que irá converter para os correspondentes nós e, posteriormente, os arquivos de relacionamento são importados para conectar os nós e criar arestas. Nas arestas os nós de início (*\_from*) e destino (*\_to*) são usados repetidamente no processo sem recriá-los. Devido aos diversos grupos de arquivos, após a execução dos *scripts* são exportados diversos arquivos individuais de classes e relacionamentos, que são

importados para o BD. De acordo com o processo automático de identificação e associação, diferentes tipos de nós e arestas são finalmente combinados em um grafo de gerenciamento de dados completo contendo todas as informações.

#### 3.4.4 Recuperação integrada: espaço, tempo e classes

Essa etapa tem por objetivo executar consultas para recuperação integrada de dados baseado no tempo, no espaço e em classes através da linguagem AQL. Primeiramente, foram desenvolvidas consultas baseadas na tríade *WHERE*, *WHEN* e *WHO* proposta por Peuquet (1994); e consultas de caráter geral de acesso e administração de dados.

Em seguida realiza-se a recuperação integrada das classes (*WHO*) tendo como dados de entrada (*WHEN*), data específica ou um intervalo de tempo, e (*WHERE*), uma área delimitada. Os resultados foram apresentados através de visualizações 2D, 3D e na forma de gráficos, fruto dos resultados das consultas. *Scripts* em Python realizam a conexão com o ArangoDB, executam as consultas e exportam os resultados visuais.

A FIGURA 15 apresenta a parte inicial do *script* em Python que realiza a sua conexão com o ArangoDB, transcreve a consulta a ser realizada no banco de dados e executa a recuperação desses dados. Em seguida, é possível dar continuidade à codificação, incluindo as saídas de visualizações geoespaciais conforme a necessidade.

FIGURA 15 – PARTE INICIAL DO *SCRIPT* EM PYTHON

```
from pyArango.connection import *

# Conexão ao ArangoDB
conn = Connection(username="root", password="", arangoURL="http://localhost:8529")
db = conn["projeto"] # a palavra projeto refere-se ao nome do banco de dados
# Entrada da consulta AQL de recuperação de dados
aql = " "
# Execução da consulta em AQL
data = db.AQLQuery(aql, rawResults=True)
# Continuação do script
...
```

FONTE: O autor (2023).

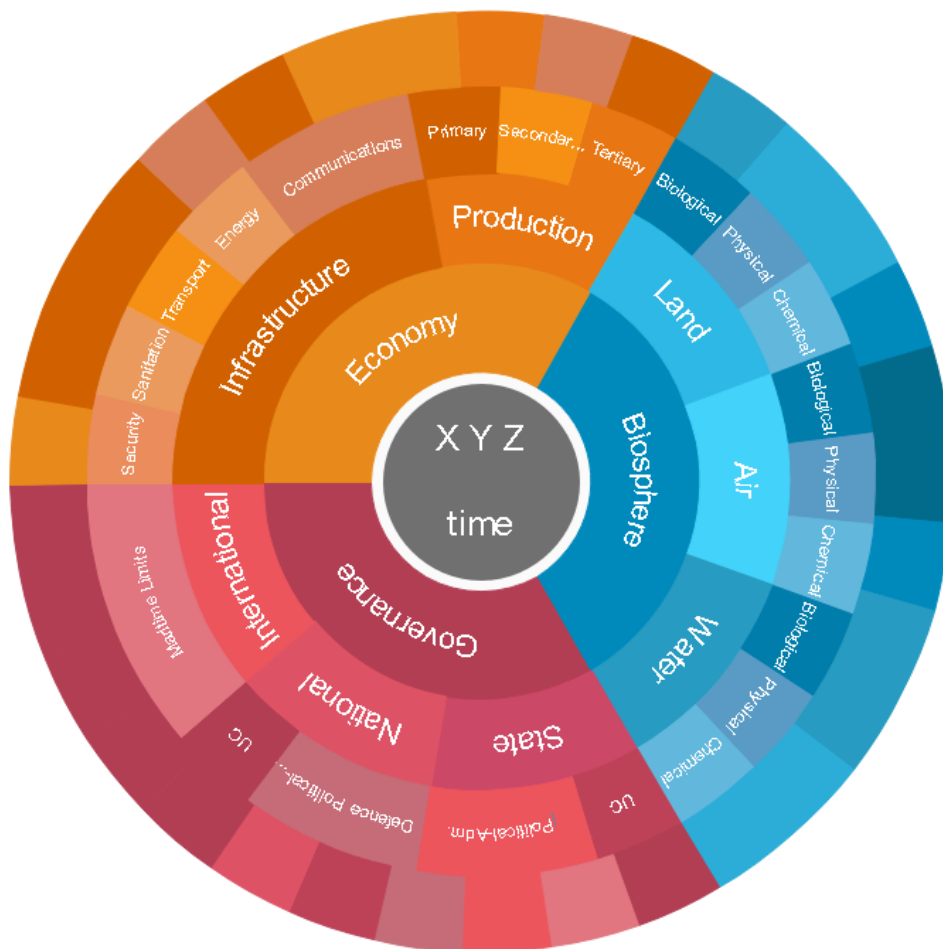
## 4 RESULTADOS

Esta seção apresenta os resultados da construção do modelo conceitual de representação integrada, do modelo lógico orientado por grafos, do grafo integrado com dados espaço-temporais terrestres, marítimos e atmosféricos; e da recuperação integrada de dados baseada no espaço, tempo e classes.

### 4.1 MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA

O objetivo do modelo conceitual de representação integrada é organizar de forma unificada os dados espaço-temporais localizados nos ambientes marítimo, terrestre e atmosférico. Uma visão macro do modelo é ilustrada na forma de um diagrama circular, FIGURA 16.

FIGURA 16 – VISÃO GERAL DO MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA



FONTE: O autor (2023).

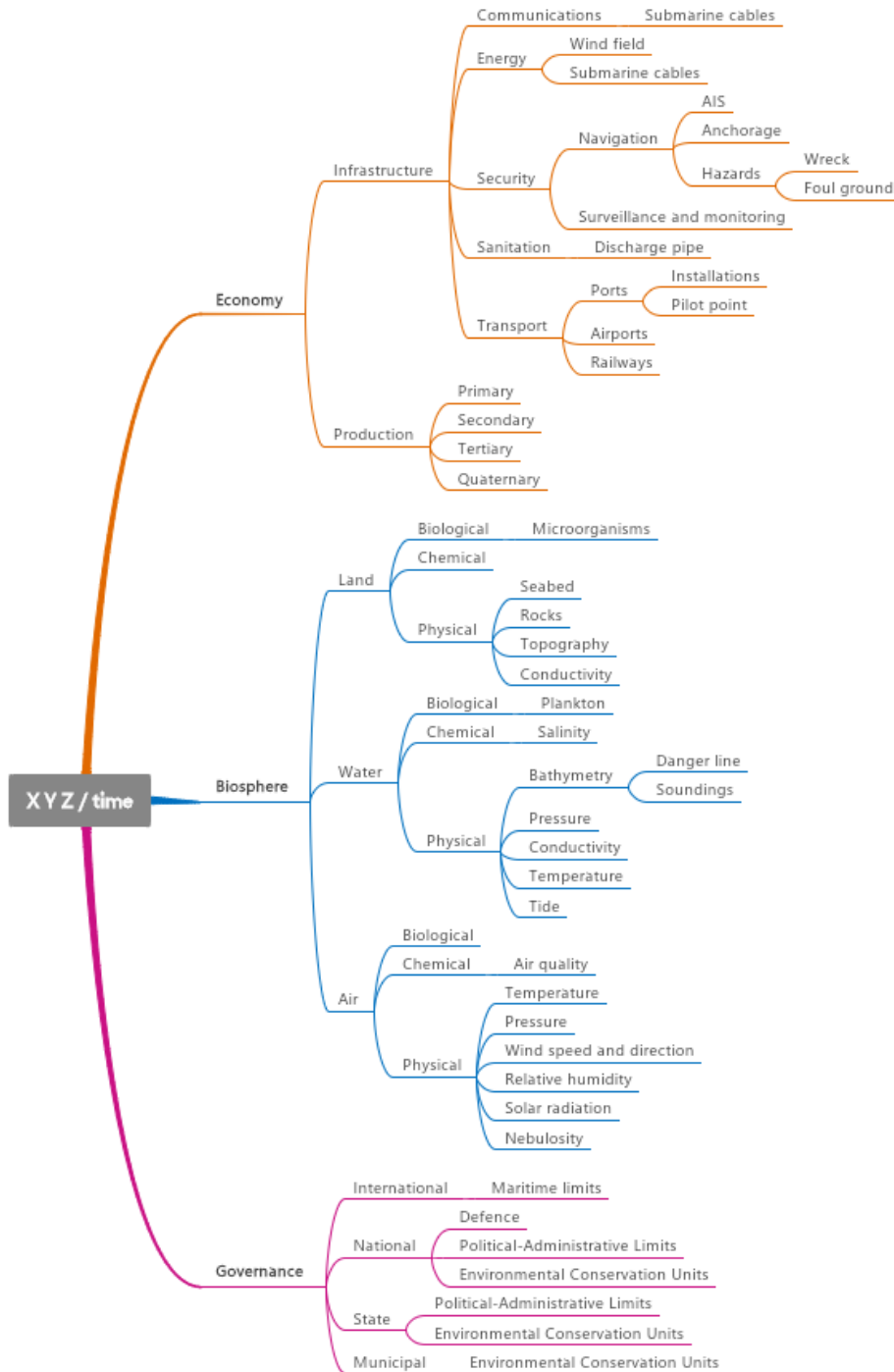
Considera-se que um valor unitário dos dados espaço-temporais são compostos por inúmeros objetos do mundo real, com seus respectivos atributos, pertencentes a diferentes classes que compartilham uma localização (x, y, z) no espaço e um tempo correspondente (t). Os dados espaço-temporais estão localizados no centro do modelo conceitual. Classificações intermediárias são dispostas em camadas concêntricas e hierarquizadas, da semântica mais abrangente para a mais específica, formando uma hierarquia de classes que culmina em classes de representação dos objetos. Considerando que os dados espaço-temporais são compostos por inúmeros objetos do mundo real, esses podem pertencer a uma, ou mais, hierarquias de classes na qual serão associados.

A primeira camada da hierarquia de classes observa três grandes divisões: a biosfera, a economia e a governança. A biosfera compreende os dados físicos, químicos e biológicos, de origem natural, localizados em uma das grandes componentes da biosfera: ar, água e terra. A economia engloba os dados relacionados com as atividades de infraestrutura e produção como: a segurança, energia, comunicações, transporte, indústrias, agropecuária, saneamento, etc. A governança compreende os dados espaço-temporais cujos objetos são entidades conceituais, de determinada jurisdição (internacional, nacional, estadual e municipal) materializadas na forma documental.

Essa visão geral do modelo conceitual de representação integrada pode ser expandida para uma visão detalhada para as classes de representação dos objetos, conforme apresentado na FIGURA 17. Nessa visão podem-se observar várias classes de representação dos objetos relacionadas a várias hierarquias de classes.

Um dos grandes diferenciais dessa representação integrada é a grande simplificação obtida na tipologia das classes bem como a extrema facilidade obtida na recuperação dos dados. Tome-se, por exemplo, a classe plâncton, que nesse modelo observa à seguinte hierarquia de classes: biosfera → água → biológico → plâncton; a posição (x, y, z) do dado espaço-temporal que contenha o objeto plâncton irá posicioná-lo em áreas oceânicas. A variável *t* e os respectivos atributos constituem uma instância que será associada a essa posição.

FIGURA 17 – MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA EXPANDIDO

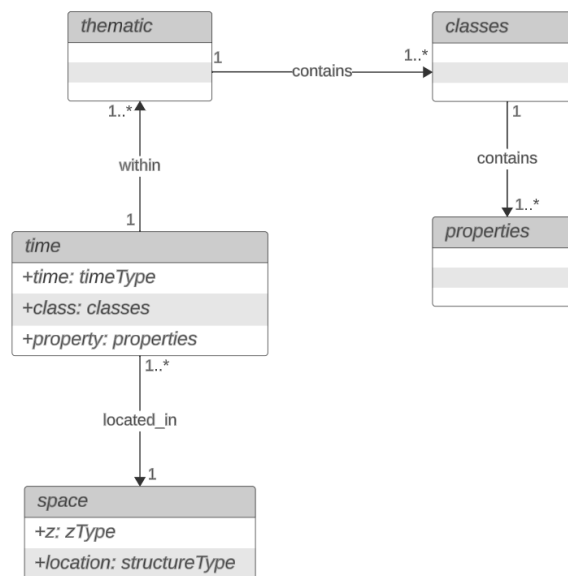


FONTE: O autor (2023).

## 4.2 MODELO LÓGICO ORIENTADO POR GRAFOS

O modelo lógico orientado por grafos é apresentado na FIGURA 18, em um diagrama de classes UML, com a quantidade mínima de classes que refletem o modelo conceitual. Cada classe representa um nó com seus próprios atributos, e cada relação representa uma aresta direcionada. Os nós, arestas e os principais atributos serão discutidos a seguir.

FIGURA 18 – MODELO LÓGICO DE REFERÊNCIA DO MODELO CONCEITUAL



FONTE: O autor (2023).

O nó *classes* armazena os identificadores das classes de representação dos objetos. O nó *properties* armazena os identificadores dos respectivos atributos pertencentes a cada classe de representação dos objetos. Esses dois nós são conectados pela aresta *contains* na relação de 1 para 1..\*, pois o modelo de uma classe pode possuir diversos atributos. O centro do modelo lógico é representado pelos nós *time* e *space*. Os atributos pertencentes ao nó *time* são definidos por uma única componente temporal, e uma, ou múltiplas, classes de representação dos objetos com seus respectivos atributos. O nó *time* utilizará para nomes dos atributos os identificadores armazenados no nó *classes* e *properties*. Os atributos pertencentes ao nó *space* são uma estrutura vetorial descrita pelos seus componentes unitários latitude e longitude; e se houver, uma única componente z

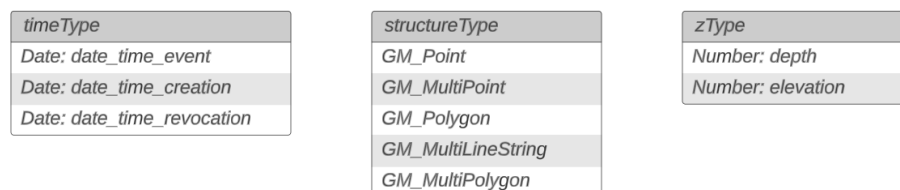
sendo altitude ou profundidade. Os nós *time* e *space* coexistem com o relacionamento, *located\_in*, de 1..\* para 1, pois cada, ou múltiplas, instâncias do nó *time* estão associadas a uma instância do nó *space*.

O nó *thematic* armazena os valores da hierarquia de classes apresentadas no modelo conceitual. A aresta *within* relaciona o nó *time* com o nó *thematic*, e permite o relacionamento de 1 para 1..\* pois as instâncias armazenadas no nó *time* podem conter objetos pertencentes a diferentes classes.

A aresta *contains* ocorre novamente entre os nós *thematic* e *classes*, pois permite o relacionamento de 1 para 1..\* , ou seja, as hierarquias de classes estão associadas a um ou a mais classes individuais.

A FIGURA 19 apresenta alguns termos identificadores utilizados como atributos pertencentes aos nós *space* e *time* e tipos de armazenamento utilizados no nó *space*. A classe *timeType* apresenta três termos identificadores utilizados para o tempo: *date\_time\_event*, *date\_time\_creation* e *date\_time\_revocation*, respectivamente, pontos de tempo registrados em eventos de observação e/ou instalação, pontos de tempo de criação e revogação de leis e decretos. A estrutura vetorial descrita pelos seus componentes unitários latitude e longitude são armazenadas conforme apresentado na classe *structureType*: ponto (*GM\_Point*), linha (*GM\_LineString*), polígono (*GM\_Polygon*), multipolígono (*GM\_MultiPolygon*), multilinhas (*GM\_MultiLineString*), entre outros. A classe *zType* apresenta os termos identificadores utilizados para a componente z, profundidade e altitude, respectivamente, *depth* e *elevation*, que são armazenados separadamente, pois ambas, por convenção, são positivas apesar de terem a mesma superfície de referência.

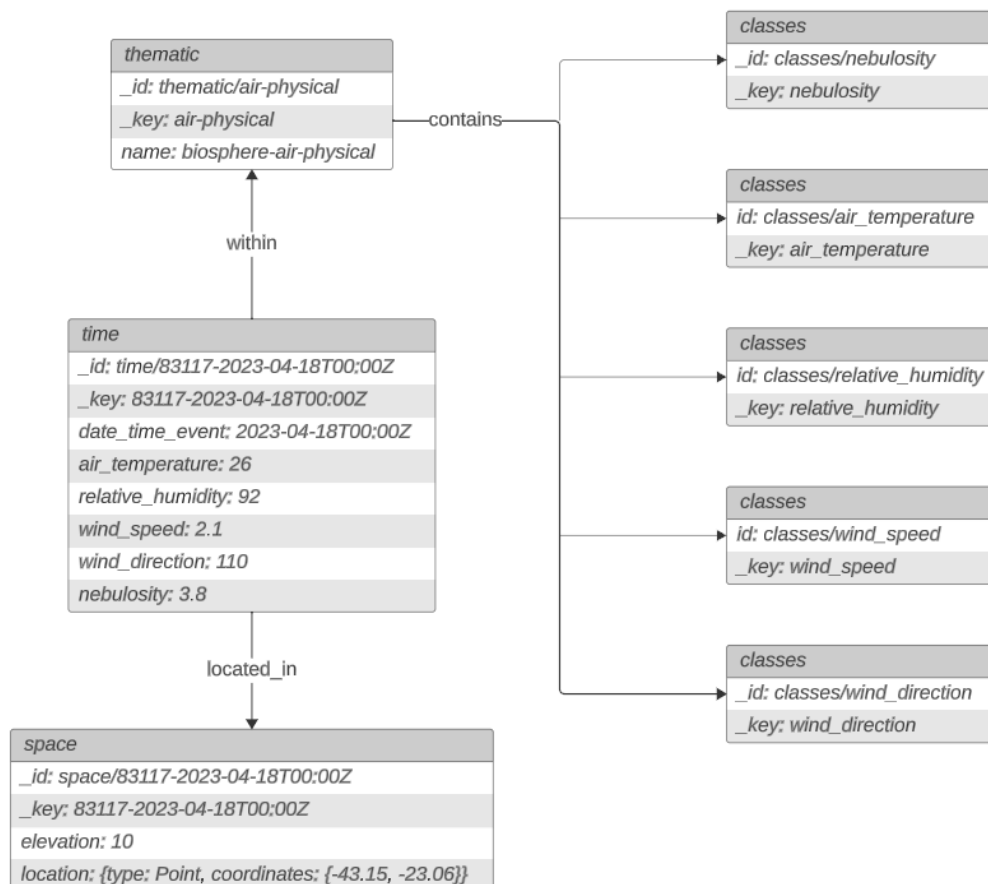
FIGURA 19 – EXEMPLOS DE TERMOS UTILIZADOS COMO ATRIBUTOS PERTENCENTES AOS NÓS TIME E SPACE E TIPOS DE ARMAZENAMENTO UTILIZADO NO NÓ SPACE



FONTE: O autor (2023).

A FIGURA 20 apresenta um exemplo de instância para o nó *time* sob a forma de um diagrama de objetos em UML. Vários objetos *air\_temperature*, *relative\_humidity*, *wind\_speed*, *wind\_direction* e *nebulosity* compartilham o ponto de tempo 2023-04-18T00:00Z e a localização (-43.15, -23.06) que pertencem a hierarquia de classes biosfera → ar → físico.

FIGURA 20 – EXEMPLO DE INSTÂNCIA NO NÓ TIME

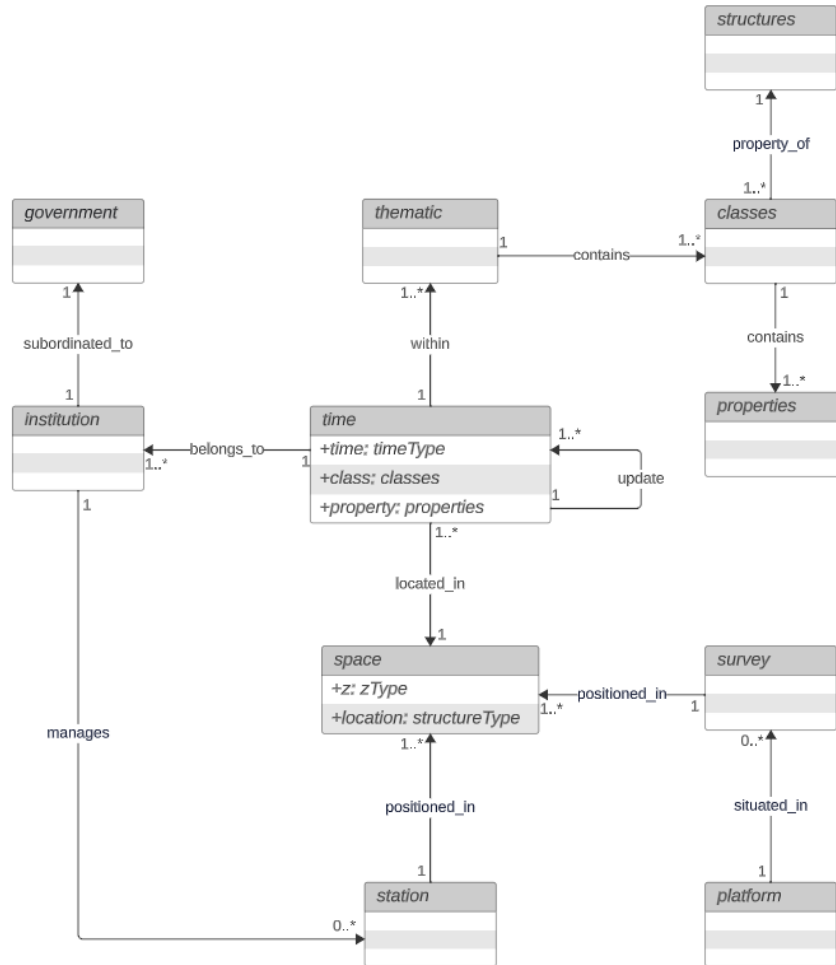


FONTE: O autor (2023).

Sob a ótica do gerenciamento de dados espaço-temporais terrestres, marítimos e atmosféricos pode-se ampliar o modelo lógico de referência apresentado na FIGURA 18 para o da FIGURA 21, correspondente a um diagrama de classes geral, o qual apresenta alguns nós e arestas adicionais necessários para o gerenciamento de dados que serão apresentados na sequência.



FIGURA 21 – DIAGRAMA DE CLASSES EM UML PARA O GERENCIAMENTO DE DADOS ESPAÇO-TEMPORAIS TERRESTRES, MARÍTIMOS E ATMOSFÉRICOS



FONTE: O autor (2023).

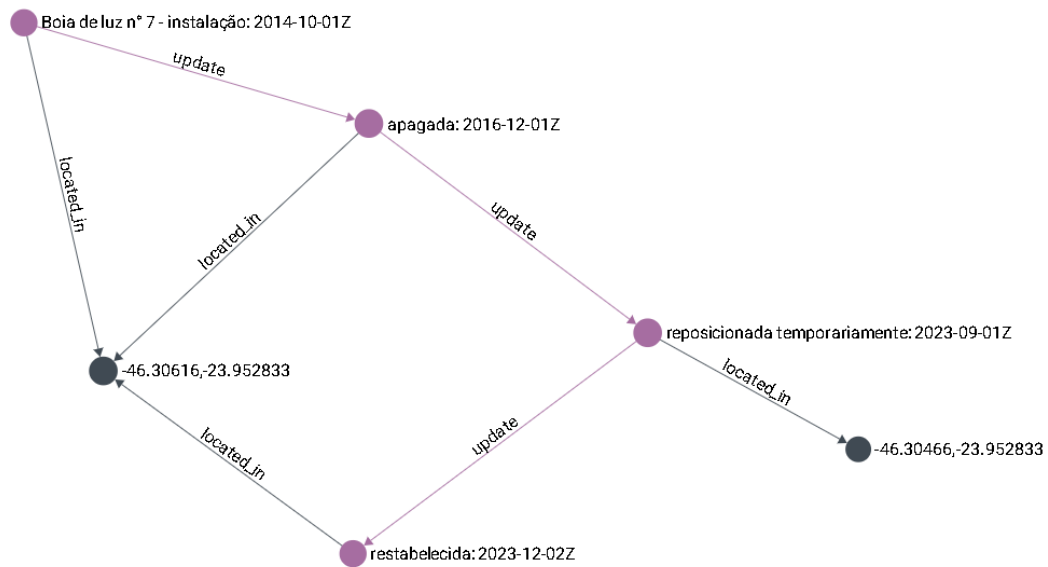
O nó *platform* armazena dados sobre os meios de transporte utilizados para realizar observações em pontos de sondagens como navios, aviões e automóveis. O nó *survey* armazena informações sobre pontos e áreas de sondagem. A aresta *situated\_in* interliga os nós *platform* e *survey* em um relacionamento de 1 para 0..\* pois um meio de transporte pode ter sido utilizado ou não. A aresta *positioned\_in* terá um relacionamento de 1 para 1..\* pois esses locais de sondagem realizam observações com variações nas componentes espaciais ao longo do tempo.

O nó *institution* representa a organização responsável pelos dados armazenados nos nós *time* e *space*. Cada organização observa subordinação específica, situada nas esferas de poder público ou privado armazenados no nó *government*. Esse relacionamento é caracterizado pela aresta *subordinated\_to* em direção ao nó *government*.

O nó *station* refere-se a estações físicas permanentes de observação de dados, podendo estar localizada em ambiente terrestre ou marítimo, fluvial e lacustre. Podem-se tomar por exemplos estações meteorológicas e boias meteoceanográficas. Essas estações possuem sempre um atributo espacial fixo, no caso de instalações terrestres, ou de projeto, no caso de instalações em ambiente aquático. Nesse último caso a aresta *positioned\_in* terá um relacionamento de 1 para 1..\* pois essas estações possuem um raio de giro e a sua manutenção periódica faz com que as componentes espaciais tenham variações de valores ao longo tempo. Essas estações são administradas por alguma organização armazenada no nó *institution*, porém nem todas as organizações possuem estações físicas, assim a aresta *manages* permite um relacionamento 0..\* em direção ao nó *station*.

A aresta *update* relaciona diferentes entradas de dados correspondentes a atualizações. Esse relacionamento pode ser de 1 para 1..\*. A FIGURA 22 apresenta um exemplo da aresta *update* com uma sequência de três atualizações: o objeto boia de luz nº 7 foi instalado no dia 01-10-2014 na localização (-46,30616, -23.952833); o mesmo objeto sofreu uma atualização, com a alteração não intencional (luz apagada) no dia 01-12-2016, mas a posição foi mantida; ocorreu uma terceira atualização, o objeto foi reposicionado temporariamente para a posição (-46.30466, -23.952833) no dia 01-09-2023; e por último, no dia 02-12-2023 as condições iniciais do objeto foram restabelecidas.

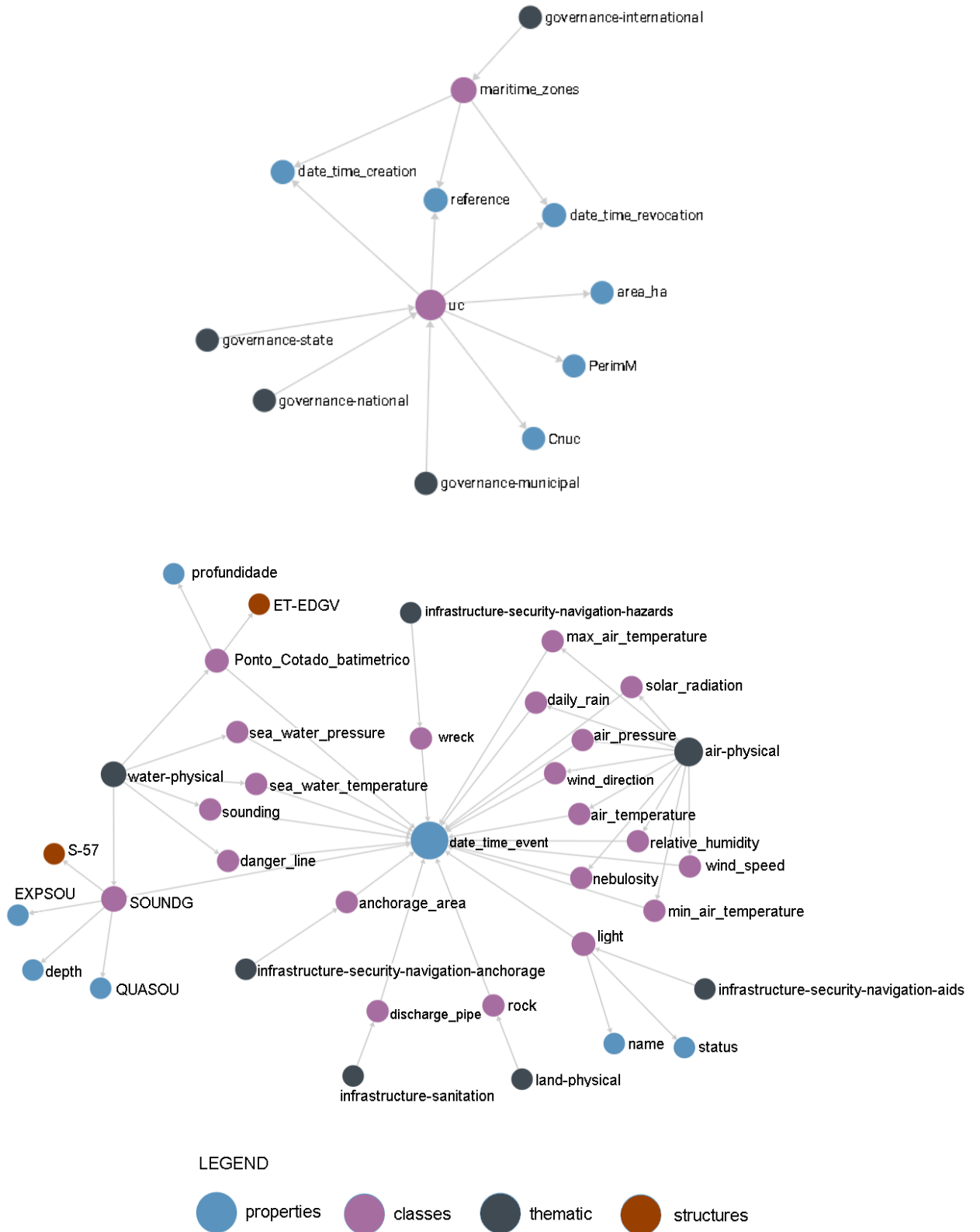
FIGURA 22 – EXEMPLO DA ARESTA UPDATE



FONTE: O autor (2023).

O nó *structures* possibilita armazenar a nomenclatura de diversos modelos de dados, para possibilitar o seu registro histórico e a sua correspondente recuperação, e a aresta *property\_of* possibilita associar as classes armazenadas no nó *classes* ao seu modelo de dados original. Essa relação é de 1 para 1..\* para cada modelo de dados. A FIGURA 23 apresenta na forma de grafos os nós *classes*, *thematic*, *properties* e *structures* e seus respectivos relacionamentos, correspondente a parte do diagrama de classes apresentado na FIGURA 18. Pode-se observar que os atributos pertencentes ao nó *properties* estão relacionados com vários atributos do nó *classes* e da mesma forma vários atributos de *classes* estão relacionados com os atributos pertencentes ao nó *thematic*. Essa figura também apresenta a coexistência de objetos com a mesma semântica, porém pertencentes a diferentes modelos de dados, nó *structure*, é o caso do objeto *Ponto\_Cotado\_batimetrico* e *SOUNDG* pertencentes aos modelos de dados a ET-EDGV e S-57, respectivamente.

FIGURA 23 – APRESENTAÇÃO DE PARTE DA VERSÃO LÓGICA EM GRAFOS



FONTE: O autor (2023)

A título ilustrativo de aplicação do modelo lógico para gerenciamento de dados, apresentado na FIGURA 21, serão utilizados dois exemplos de dados espaço-temporais concretos, conforme a TABELA 4. O resultado é apresentado na forma de um diagrama de objetos em UML, na FIGURA 24.

TABELA 4 – EXEMPLOS DE DADOS ESPAÇO-TEMPORAIS CONCRETOS

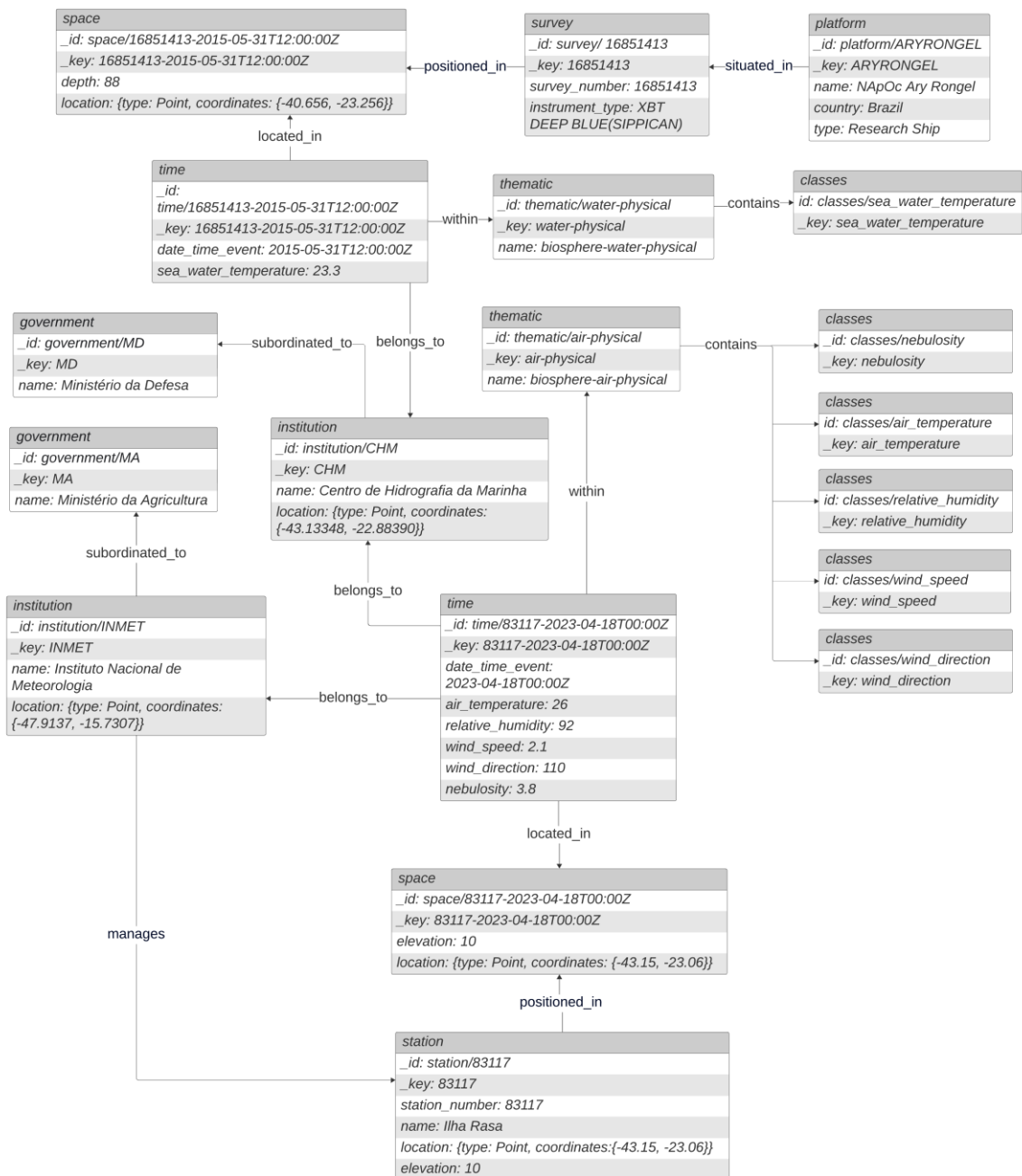
Dados	Características	Valor
1	Meio flutuante	Research Ship NApOc Ary Rongel
	Número do ponto de sondagem	16851413
	Instrumento	XBT: DEEP BLUE (SIPPICAN)
	Tempo	12:00:00 UTC 31/05/2015
	Longitude e latitude	-40.656 e -23.256
	Profundidade	88 m
	Temperatura da água do mar	23.3 °C
	Fiel depositário	Centro de Hidrografia da Marinha
	Subordinado	Ministério da Defesa
2	Estação	Ilha Rasa
	Número da estação	83117
	Longitude e latitude	-43.15 e -23.06
	Altitude	10 m
	Data	00:00 UTC 18/04/2023
	Temperatura do ar	26 °C
	Umidade relativa	92
	Velocidade e direção do vento	2.1 e 110°
	Nebulosidade	3.8
	Responsável pela estação	Centro de Hidrografia da Marinha
	Fiel depositário	Centro de Hidrografia da Marinha e Instituto Nacional de Meteorologia
	Subordinado	Ministério da Defesa

FONTE: O autor (2023)

Na FIGURA 24 cada instância pertencente ao nó *time* refere-se somente a uma localização no espaço e a uma hierarquia de classes, porém não é uma regra geral. O nó *station* em questão é uma estação terrestre, portanto sua posição espacial é fixa e igual para todas as observações tomadas nesta estação. O ponto de sondagem armazenado no nó *survey* foi observado a partir do navio de pesquisa NApOc Ary Rongel com um equipamento específico em uma localização na profundidade de 88 m. Além disso, existem dois pontos notáveis no exemplo: a instância pertencente ao nó *time*, identificada por 83117-2023-04-18T00:00Z, possui duas arestas *belongs\_to* pois esses dados pertencem a duas instituições diferentes,

CHM e INMET; em segundo lugar, as instâncias pertencentes ao nó *time* possuem diferentes relacionamentos com o nó *thematic*, pois os objetos armazenados nessas instâncias pertencem a diferentes hierarquias de classes biosfera → ar → físico e o outro biosfera → água → físico.

FIGURA 24 – VISUALIZAÇÃO DO DIAGRAMA DE OBJETOS EM UML COM DADOS ESPAÇO-TEMPORAIS CONCRETOS



FONTE: O autor (2023)

### 4.3 GRAFO INTEGRADO: TERRESTRE, MARÍTIMO E ATMOSFÉRICO

A construção do grafo integrado, exigido pelo SGBD selecionado para implementação, tem como etapa inicial a geração de arquivos JSON de classes e relacionamentos previstos no modelo lógico. A FIGURA 25 e FIGURA 26 apresentam uma amostra do conteúdo do arquivo JSON, respectivamente, da classe *time* e do relacionamento *located\_in*.

FIGURA 25 – AMOSTRA DO CONTEÚDO DO ARQUIVO JSON DA CLASSE TIME

```
[
  { "_id": "time/83054-2022-12-01Z",
    "_key": "83054-2022-12-01Z",
    "max_air_temperature": 37.1,
    "min_air_temperature": 21.0,
    "daily_rain": 7.0,
    "date_time_event": "2022-12-01Z"
  },
  { "_id": "time/83054-2022-12-03Z",
    "_key": "83054-2022-12-03Z",
    "max_air_temperature": 36.0,
    "min_air_temperature": 22.5,
    "daily_rain": 0.0,
    "date_time_event": "2022-12-03Z"
  }
]
```

FONTE: O autor (2023).

FIGURA 26 – AMOSTRA DO CONTEÚDO DO ARQUIVO JSON DO RELACIONAMENTO LOCATED\_IN

```
[
  {"_from": "time/83054-2022-12-01Z", "relation": "located_in", "_to": "space/83054"},
  {"_from": "time/83054-2022-12-02Z", "relation": "located_in", "_to": "space/83054"}
]
```

FONTE: O autor (2023).

O processamento dos arquivos de entrada extraiu 54.567 dados espaço-temporais. As tabelas abaixo apresentam o sumário dos resultados encontrados para as instâncias das classes e relacionamentos, respectivamente, TABELA 5 e

TABELA 6. Assim, foram extraídos dos dados espaço-temporais processados um total de 107.647 instâncias das classes e 235.997 relacionamentos.

TABELA 5 – SUMÁRIO GERAL DOS RESULTADOS DAS INSTÂNCIAS DAS CLASSES

<b>Entidades</b>	<b>total</b>
<i>time</i>	54.567
<i>space</i>	53.011
<i>thematic</i>	10
<i>classes</i>	20
<i>survey</i>	20
<i>platform</i>	2
<i>station</i>	6
<i>institution</i>	6
<i>government</i>	5
Total	107.647

FONTE: O autor (2023)

TABELA 6 – SUMÁRIO GERAL DOS RESULTADOS DOS RELACIONAMENTOS

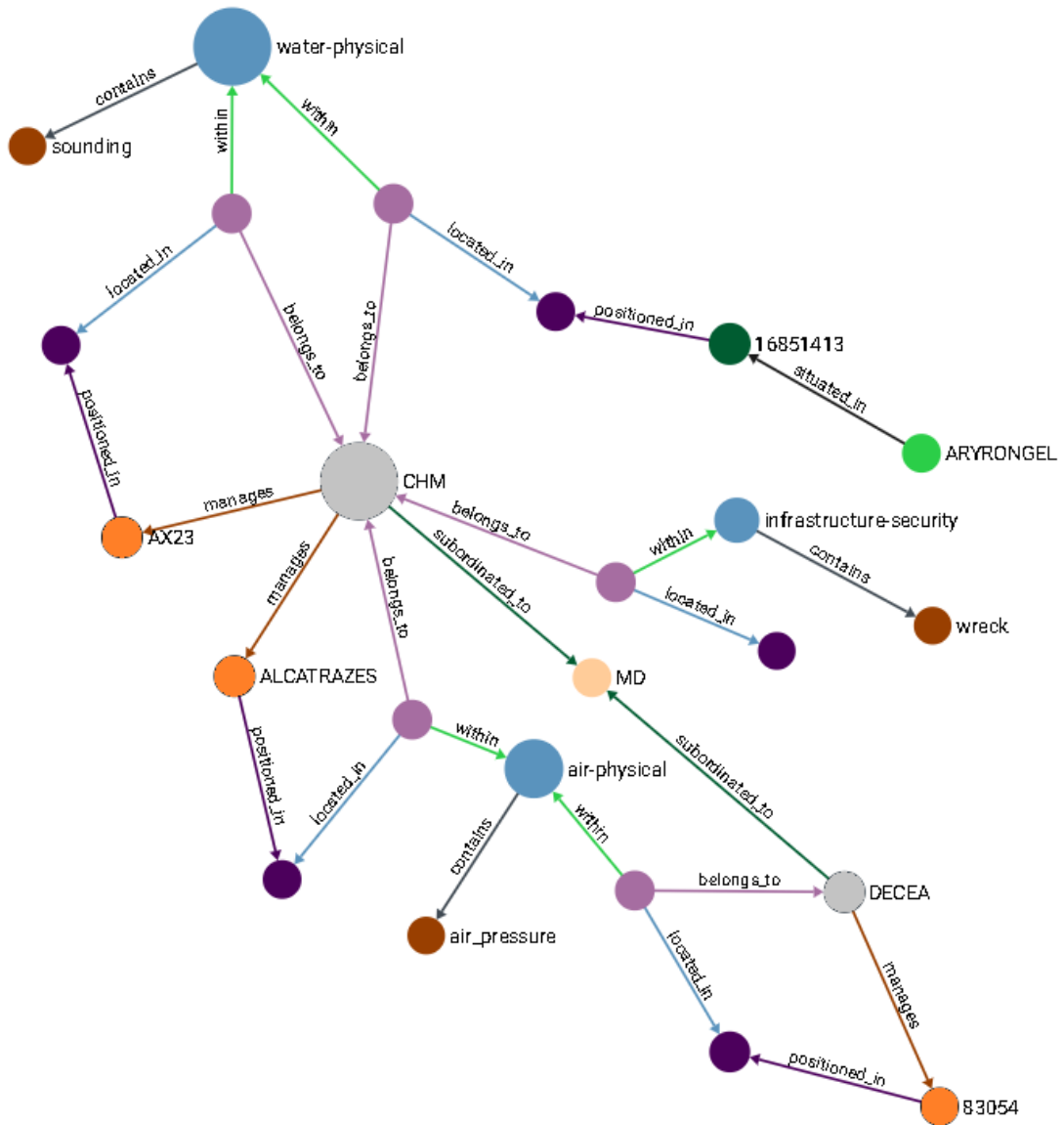
<b>Relacionamentos</b>	<b>total</b>
<i>within</i>	73.898
<i>located_in</i>	54.568
<i>contains</i>	22
<i>belongs_to</i>	54.567
<i>subordinated_to</i>	6
<i>positioned_in</i>	52.910
<i>situated_in</i>	20
<i>manages</i>	6
-	235.997

FONTE: O autor (2023)

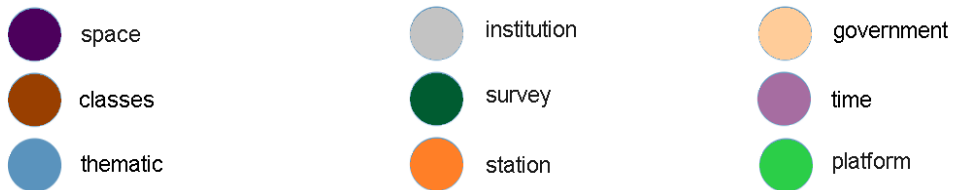
Os arquivos JSON correspondentes às classes e relacionamentos foram importados para o BD orientado por grafos e automaticamente convertidos para nós e arestas. A FIGURA 27 apresenta a disposição dos nós e arestas de uma amostra do grafo implementado no ArangoDB. Cada classe e relacionamento estão representados no grafo, respectivamente, por nós e arestas de diferentes cores. Pode-se observar os nós *time*, *space*, *thematic*, *code*, *survey*, *platform*, *station*, *institution* e *government* e as arestas *within*, *located\_in*, *contains*, *belongs\_to*, *subordinated\_to*, *positioned\_in*, *situated\_in* e *manages*.



FIGURA 27 – VISUALIZAÇÃO DE UMA AMOSTRA DO GRAFO COM O ARANGODB.



## LEGEND:

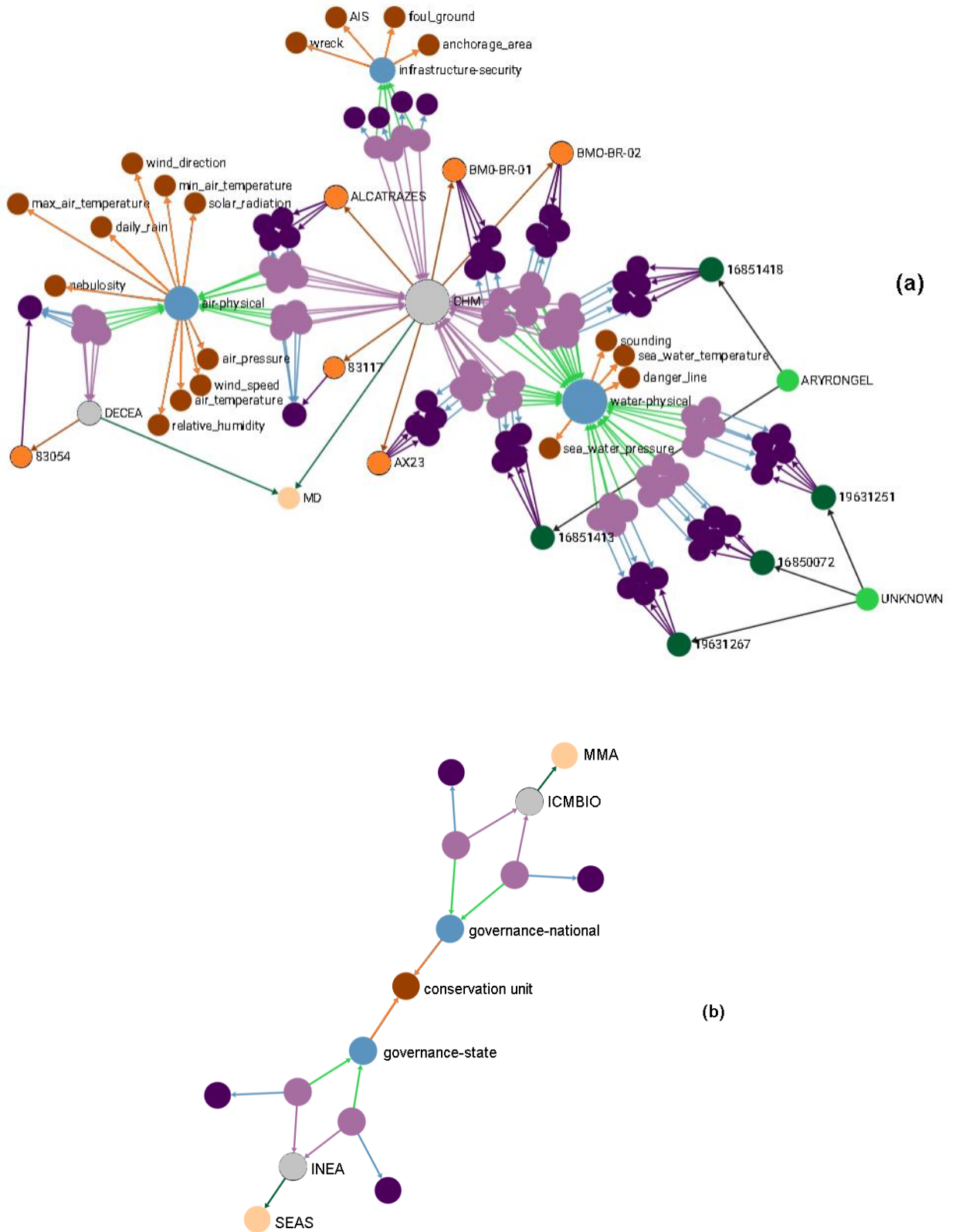


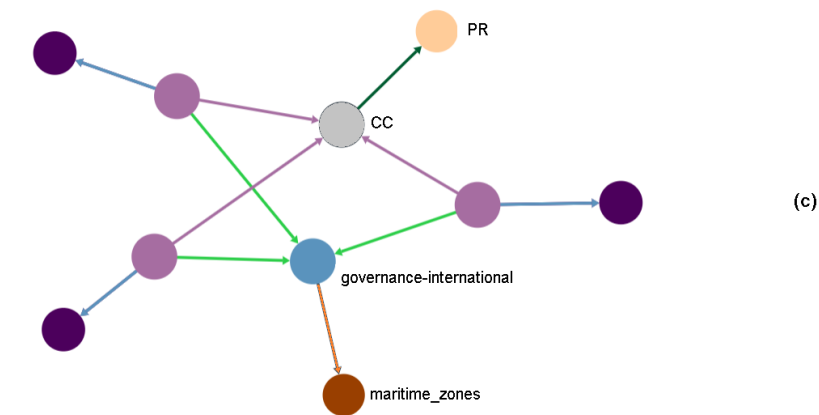
FONTE: O autor (2023)

Os grafos apresentados, FIGURA 27 e FIGURA 28, foram gerados automaticamente pelo ArangoDB. A variação do tamanho dos nós é definida automaticamente pelo SGBD, o qual destaca os nós, em relação ao tamanho, que estão envolvidos em mais relacionamentos na amostra em questão. A legenda informa as cores associadas aos nós e arestas, pois em alguns casos, os rótulos foram suprimidos do grafo para a manutenção da clareza.

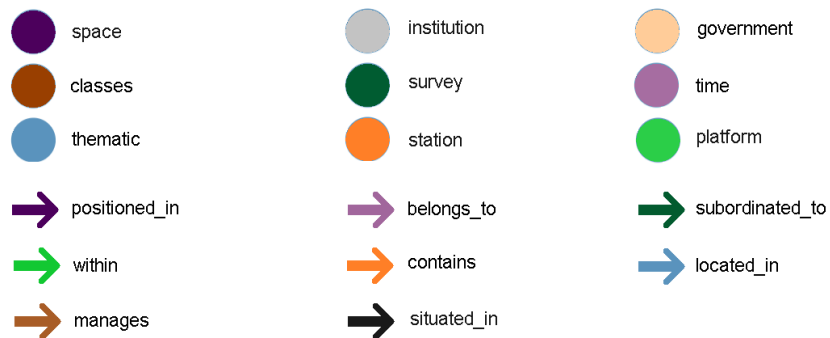
Na FIGURA 28 pode-se observar, a título ilustrativo, uma visão macro parcial do resultado da construção do banco de dados para a gestão de dados espaço-temporais. Destacam-se no grafo os nós *thematic*, em azul, *time*, em lilás e *space* em roxo. Além disso, no caso (a), podem-se observar amostras de dados espaço-temporais associadas com a biosfera no ambiente terrestre: a estação meteorológica 83054 localizada no aeroporto de Jacarepaguá/RJ e administrada pelo Departamento de Controle do Espaço Aéreo (DECEA) do Comando da Aeronáutica, a estação meteorológica 83117 localizada na Ilha Rasa/RJ e administrada pelo Centro de Hidrografia da Marinha (CHM), ambas as instituições estão subordinadas ao Ministério da Defesa (MD); dados marítimos e atmosféricos: a estação meteoceanográfica Alcatrazes, AX23, BMO-BR-02 e BMO-BR-01 administradas pelo CHM que observam parâmetros físicos do ar e da água; e dados marítimos: os pontos de sondagem 19631251, 16850072 e 19631267 observados a partir de uma plataforma desconhecida e os pontos de sondagem 16851418 e 16851413 observados a partir do navio de pesquisa Ary Rongel. Podem-se observar também amostras de dados espaço-temporais associadas com a hierarquia de classes economia → infraestrutura → segurança no ambiente marítimo: os objetos casco soçobrado, área de fundeio, fundo sujo e tráfego marítimo.

FIGURA 28 – VISÃO MACRO PARCIAL DO BANCO DE DADOS ORIENTADO POR GRAFOS PARA GESTÃO DE DADOS ESPAÇO-TEMPORAIS





## LEGEND:



FONTE: O autor (2023)

Na FIGURA 28 também se pode observar amostras de dados espaço-temporais associadas à governança, identificados pelas letras b e c da FIGURA 28. No caso (b) pode-se observar amostras de dados espaço-temporais associadas com a governança no ambiente terrestre e marítimo que transpõem os limites atmosféricos: duas unidades de conservação administradas pelo Instituto Estadual do Ambiente (INEA) do Estado do Rio de Janeiro e duas administradas pelo Instituto Chico Mendes de Conservação da Biodiversidade (ICMBIO) subordinadas, respectivamente, pela Secretaria de Estado do Ambiente e Sustentabilidade (SEAS) do Estado do Rio de Janeiro e pelo Ministério do Meio Ambiente (MMA); e no caso (c), pode-se observar amostras de dados espaço-temporais associadas ao ambiente marítimo que transpõem os limites atmosféricos: três limites marítimos definidos pela Convenção das Nações Unidas pelo Direito do Mar da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) e firmado pelo Brasil em

10 de dezembro de 1982 e promulgada pelo Presidente da República (PR) a lei n.8617 de 4 de janeiro de 1994 através da Casa Civil (CC).

Observa-se pela amostra do grafo apresentada na FIGURA 28 que o nó *thematic* não possui relacionamento entre eles, ou seja, não existe ambiguidade nas classes de representação dos objetos armazenados no nó *classe*. A hierarquia de classes *air-physical*, *water-physical* e *infrastructure-security* está indiretamente interconectada, pois a mesma instituição, CHM, é a fiel depositária desses dados, o que não ocorre com a parte inferior do grafo com as instituições INEA, ICMBIO e CC.

#### 4.4 RECUPERAÇÃO INTEGRADA: ESPAÇO, TEMPO E CLASSES

O APÊNDICE 1 apresenta as consultas na linguagem AQL desenvolvidas para a recuperação integrada de dados espaço-temporais baseadas na tríade *WHERE*, *WHEN* e *WHO* e consultas de caráter geral de acesso e administração de dados.

Em seguida apresenta-se uma sequência de resultados sob a forma de visualizações geoespaciais em 2D, 3D e gráficos que refletem os resultados de consultas de recuperação integrada de classes baseadas em dados de entrada limitados pelo tempo e espaço. As consultas e suas respectivas visualizações geoespaciais foram geradas através de *scripts* na linguagem Python.

O QUADRO 4 apresenta uma consulta em AQL de recuperação integrada da classe unidades de conservação ambientais Federais e do Estado do Rio de Janeiro para uma janela temporal de 10 anos em uma área definida pelo polígono ([[[-38.5,-21],[-47.5,-21],[-47.5,-25],[-38.5,-25],[-38.5,-21]]]).

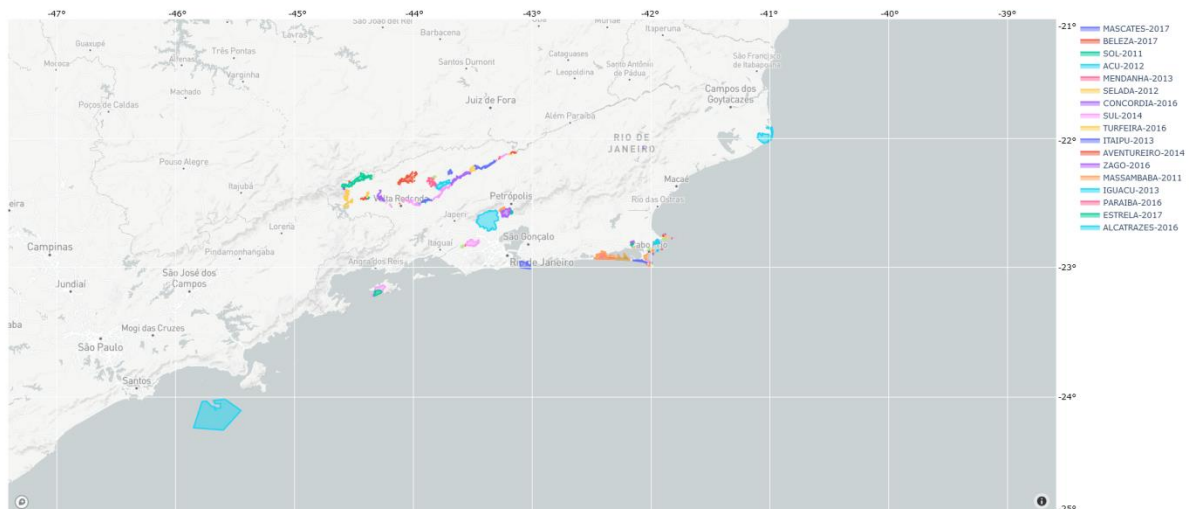
O APÊNDICE 2 apresenta uma sequência de resultados sob a forma de visualizações geoespaciais em 2D, longitude e latitude, da evolução temporal de criação de unidades de conservação ambientais federais e do Estado do Rio de Janeiro apresentadas a cada 10 anos, datadas entre 1961 e 2023. A título ilustrativo, a FIGURA 29 apresenta as 17 unidades de conservação federais e do Estado do Rio de Janeiro estabelecidas entre 01-01-2011 e 31-12-2020.

QUADRO 4 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DA CLASSE UNIDADES DE CONSERVAÇÃO AMBIENTAIS

Consulta
Quais as unidades de conservação ambientais encontradas no polígono ([[ -38.5,-21 ],[-47.5,-21 ],[-47.5,-25],[ -38.5,-25],[ -38.5,-21]]) entre um determinado periodo t1 e t2?
<pre> LET lista = 'uc'  FOR u IN code   FILTER u._key == lista   FOR v1, e1, p1 IN 1..1 INBOUND u._id contains     FOR v2, e2, p2 IN 1..1 INBOUND v1 within       OPTIONS { uniqueVertices: "path" }       FILTER v2.date_time_creation &gt; "1936-01-01" &amp;&amp; v2.date_time_creation &lt; "2023-01-03"  LET area = GEO_POLYGON([[ -38.5,-21 ],[-47.5,-21 ],[-47.5,-25],[ -38.5,-25],[ -38.5,-21]])  FOR v IN OUTBOUND v2 located_in   FILTER GEO_CONTAINS(area, v.location)   LET attributes = (     LET h = (MERGE(v2, v))     FOR name IN ATTRIBUTES(h)       FILTER REGEX_TEST(name, '[a-z]')       RETURN {         name: name,         value: h[name]       }   ) RETURN ZIP(attributes[*].name, attributes[*].value) </pre>

FONTE: O autor (2023).

FIGURA 29 – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-2011 A 31-12-2020.



FONTE: O autor (2023).

O QUADRO 5 apresenta uma consulta em AQL de recuperação integrada da classe temperatura da água do mar observada entre os dias 28 e 29 de outubro de 2019 na área definida pelo polígono ([[ -38.5,-21],[-47.5,-21],[-47.5,-27],[-38.5,-27],[-38.5,-21]]).

QUADRO 5 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DA CLASSE TEMPERATURA DA ÁGUA DO MAR

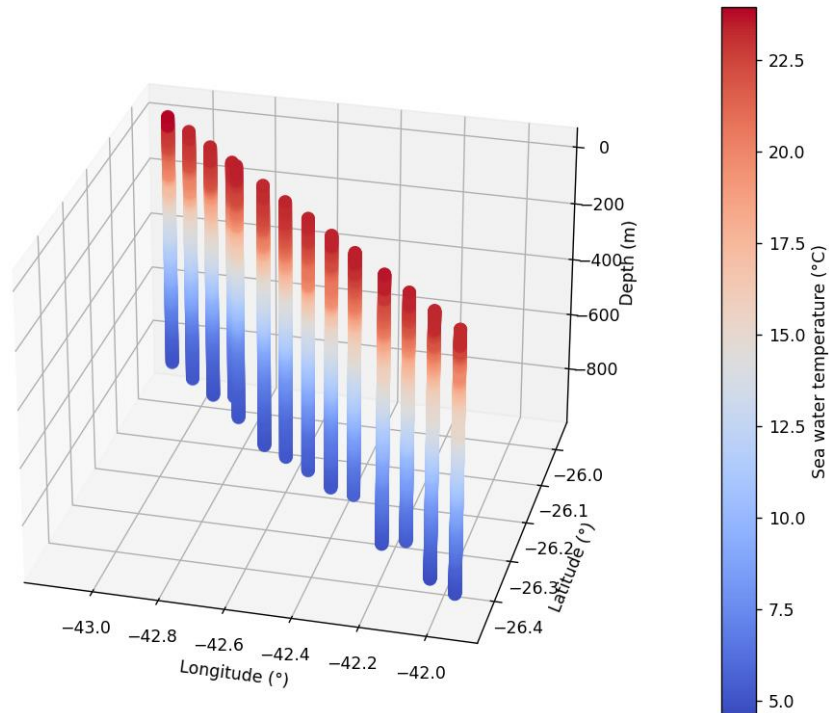
<b>Consulta</b>
Quais as temperaturas da água do mar encontradas no polígono ([[ -38.5,-21],[-47.5,-21],[-47.5,-27],[-38.5,-27],[-38.5,-21]]) entre um determinado período t1 e t2?
<pre> FOR u IN parameters   FILTER u.date_time_inst &gt;= "2019-10-28" &amp;&amp; u.date_time_inst &lt;= "2019-10-29"   LET area = GEO_POLYGON([[ -38.5,-21 ],[-47.5,-21 ],[-47.5,-27],[-38.5,-27],[-38.5,-21]])   FOR v IN OUTBOUND u located_in     FILTER GEO_CONTAINS(area, v.location)     FILTER u.sea_water_temperature != null     RETURN {temperature:u.sea_water_temperature, location: v.location, depth: v.depth} </pre>

FONTE: O autor (2023).

O GRÁFICO 1 apresenta o resultado da consulta sob a forma de visualização geoespacial 3D (longitude, latitude e profundidade) onde os valores da temperatura da água do mar aparecem em cores hipsométricas e foram observadas entre os dias 28 e 29 de outubro de 2019. Cada ponto de sondagem, representado pelas barras verticais, é definido por um conjunto de dados espaço-temporais com latitude, longitude e tempo, iniciais e constantes, e variações dos valores da temperatura da água do mar em relação a profundidade.

O QUADRO 6 apresenta uma consulta em AQL de recuperação integrada da classe temperatura da água do mar observada entre os dias 28 e 29 de outubro de 2019 na área definida pelo polígono ([[ -38.5,-21],[-47.5,-21],[-47.5,-27],[-38.5,-27],[-38.5,-21]]) e entre as profundidades de 200 e 600 metros.

GRÁFICO 1 – VALORES DE TEMPERATURA DA ÁGUA DO MAR OBSERVADAS ENTRE OS DIAS 28-10-2019 A 29-10-2019.



FONTE: O autor (2023).

QUADRO 6 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DA CLASSE TEMPERATURA DA ÁGUA DO MAR ENTRE PROFUNDIDADES DEFINIDAS

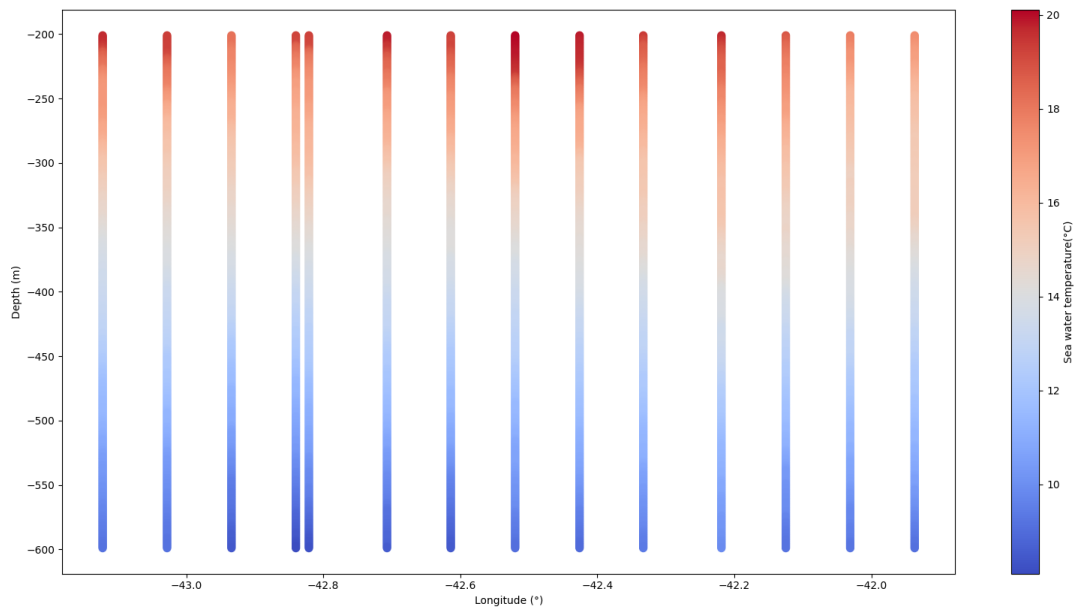
Consulta
Quais as temperaturas da água do mar encontradas no polígono ([[ -38.5,-21 ],[-47.5,-21 ],[-47.5,-27],[-38.5,-27],[-38.5,-21]]) entre um determinado período t1 e t2 e entre as profundidades d1 e d2?
<pre> FOR u IN parameters   FILTER u.date_time_inst &gt;= "2019-10-28" &amp;&amp; u.date_time_inst &lt;= "2019-10-29"   LET area = GEO_POLYGON([[ -38.5,-21 ],[-47.5,-21 ],[-47.5,-27],[-38.5,-27],[-38.5,-21]])   FOR v IN OUTBOUND u located_in     FILTER GEO_CONTAINS(area, v.location)     FILTER v.depth &gt; 200 &amp;&amp; v.depth &lt; 600     FILTER u.sea_water_temperature != null   RETURN {temperature:u.sea_water_temperature, location: v.location, depth: v.depth} </pre>

FONTE: O autor (2023).

O GRÁFICO 2 apresenta o resultado da consulta sob a forma de visualização gráfica (longitude e profundidade) onde os valores da temperatura da água do mar aparecem em cores hipsométricas, foram observadas entre os dias 28 e 29 de outubro de 2019 e entre as profundidades de 200 e 600 metros.



GRÁFICO 2 – VALORES DE TEMPERATURA DA ÁGUA DO MAR ENTRE AS PROFUNDIDADES DE 200 À 600 M EM RELAÇÃO A PROFUNDIDADE E A LONGITUDE ENTRE 28-10-2019 A 29-10-2019.



FONTE: O autor (2023).

O QUADRO 7 apresenta uma consulta em AQL de recuperação integrada das classes associadas aos parâmetros físicos do ar que estão armazenados no BD no dia 28/08/2023 às 00:00 na área definida pelo polígono ([[ -38.5,-21],[ -47.5,-21],[ -47.5,-25],[ -38.5,-25],[ -38.5,-21]]).

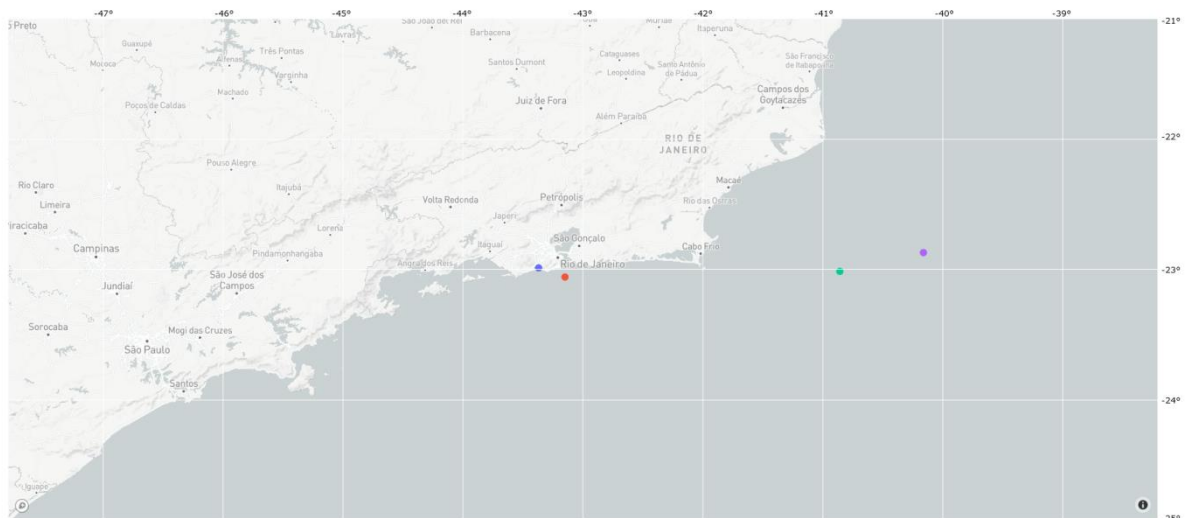
A FIGURA 30 apresenta o resultado da consulta na forma de visualização geospacial 2D, latitude e longitude. Podem ser identificadas no mapa a posição geográfica de quatro estações com cores diferentes. A legenda apresenta, para cada estação, os parâmetros físicos do ar, e respectivos valores, datados do dia 28/08/2023 às 00:00. Pode-se observar na legenda que a quantidade de parâmetros varia entre as estações e que no mesmo instante de tempo as estações oceânicas apresentaram velocidades do vento com valores bem próximos, porém diferentes das estações terrestres.

QUADRO 7 – CONSULTA EM AQL DE RECUPERAÇÃO INTEGRADA DAS CLASSES ASSOCIADAS AOS PARÂMETROS FÍSICOS DO AR ARMAZENADOS NO BD

Consulta
<p>Quais dos parâmetros selecionados são encontrados no polígono ([[ -38.5,-21],[ -47.5,-21],[ -47.5,-25],[ -38.5,-25],[ -38.5,-21]]) no dia 28-08-2023 às 00:00:00?</p> <pre> LET lista = 'wind_speed wind_direction sea_water_pressure air_pressure sea_water_temperature air_temperatur e relative_humidity nebulosity solar_radiation discharge_pipe anchorage_area danger_line maritime_z ones max_air_temperature min_air_temperature daily_rain' FOR u IN parameters   FILTER u.date_time_inst &lt;= '2023-08-28T00:00Z' &amp;&amp; u.date_time_inst &gt;= '2023-08-28T00:00:00Z'    u.date_time_creation &gt; "0" &amp;&amp; u.date_time_creation &lt; "2023-08-28T00:00Z"   LET area = GEO_POLYGON([[ -38.5,-21 ],[ -47.5,-21 ],[ -47.5,-25],[ -38.5,-25],[ -38.5,-21]])   FOR v IN OUTBOUND u located_in   FILTER GEO_INTERSECTS(area, v.location)   LET attributes = (   LET h = (MERGE(u, v))   FOR name IN ATTRIBUTES(h)   FILTER REGEX_TEST(name, lista)   RETURN {     name: name,     value: h[name],     location: h.location   }   )   FILTER attributes[0].location != null   RETURN { _key: CONCAT(ZIP(attributes[*].name, attributes[*].value)), location: attributes[0].location} </pre>

FONTE: O autor (2023).

FIGURA 30 – PARÂMETROS FÍSICOS DO AR OBSERVADOS NO DIA 28-08-2023 ÀS 00:00.



- {"relative\_humidity":100,"nebulosity":10,"air\_pressure":1021.5,"air\_temperature":16.7,"wind\_direction":350,"wind\_speed":0.5}
- {"wind\_speed":3.1,"relative\_humidity":90,"air\_temperature":16,"wind\_direction":210}
- {"wind\_speed":8.43,"air\_temperature":18.5,"sea\_water\_pressure":1022.3,"wind\_direction":147}
- {"wind\_speed":8.5,"air\_temperature":19.9,"sea\_water\_pressure":1020.5,"wind\_direction":45}

FONTE: O autor (2023).

## 5 DISCUSSÃO

Este trabalho propõe uma abordagem baseada em grafos para armazenar dados espaço-temporais marítimos, terrestres e atmosféricos bem como para a recuperação integrada de classes baseadas no espaço e tempo. Esta seção discute os principais resultados alcançados pelo trabalho, a saber: o modelo conceitual de representação integrada, o modelo lógico baseado em grafos, a construção do grafo integrado e a recuperação integrada baseada no espaço, no tempo e em classes.

### 5.1 MODELO CONCEITUAL DE REPRESENTAÇÃO INTEGRADA

O modelo conceitual de representação integrada dos dados espaço-temporais localizados nos ambientes marítimos, terrestres e atmosféricos interliga objetos do mundo real, que compartilham o mesmo tempo e espaço, independentemente das classes de representação dos objetos e dos respectivos atributos. O relacionamento entre os dados espaço-temporais e as classes de representação de objetos é realizado através de uma hierarquia de classes orientada do geral para o específico.

A hierarquia de classes do modelo é o único esquema de caráter permanente, as classes de representação de objetos serão adicionadas a qualquer tempo.

O presente trabalho teve como potencial de ativação o trabalho de DING *et al.* (2022) do qual foi absorvida a ideia de palavras-chave selecionadas para a generalização do modelo como terra, água, biologia, geologia; e associação das classes de representação dos objetos. Diferentemente dos autores, na organização hierárquica do modelo não foi utilizada uma linha de referência que separasse classes acima e abaixo da superfície terrestre; no presente modelo, essa informação será obtida pela posição (x, y, z) do objeto. O presente modelo conceitual ampliou a ideia de generalização de DING *et al.* (2022) para as várias componentes da biosfera com subdivisões intermediárias como química, física e biológica, de forma a aumentar o espectro das classes para o ambiente marítimo e atmosférico.

A ideia de representação espaço-temporal centradas na coexistência, conexão ou união dos elementos apresentadas pelas literaturas (WACHOWICZ,

2014; PEUQUET; DUAN, 1995; GOODCHILD *et al.*, 2007) também foi imperativa nesse trabalho e está representada no núcleo do modelo conceitual.

A vantagem desse modelo conceitual é a sua expansibilidade e flexibilidade. O modelo é capaz de absorver qualquer classe pertencente a modelos de dados, pretéritos ou futuros, na medida em que ocorram, sem necessidade de alterar todo um esquema de banco de dados.

Não existem classes organizadas por domínios de aplicações, pela inexistência, neste modelo, de interseção de atuação entre as diversas áreas como, por exemplo, hidrografia, oceanografia, meteorologia e biologia marinha. Além disso, a sua concepção foi centrada na ideia de que para qualquer objeto existe uma posição (x, y, z) no espaço e um momento no tempo.

## 5.2 MODELO LÓGICO ORIENTADO POR GRAFOS

Wang *et al.* (2000) ressalta que a tecnologia tradicional de banco de dados relacional não é adequada para gerenciar dados espaço-temporais, que são multidimensionais com modelos e comportamentos complexos. Além disso, Peuquet (1994) disse que os padrões de descrição de relações espaço-temporais são complexos e inter-relacionados e capturar esta complexidade requer um método descritivo poderoso e muito flexível. Ressalta-se também os comentários de Zhang, C. *et al.* (2023) relativamente à escolha do BD orientado por grafos: em um banco de dados relacional, a maioria dos esforços de desenvolvimento reside em um projeto de esquema deliberativo porque é muito mais difícil modificar o esquema uma vez predefinido. Em contraste, um grafo permite uma mudança flexível do modelo lógico. Mesmo adicionando uma propriedade de um nó ou criando novas arestas, dispensa a necessidade de utilizar todo o registro, apenas os nós correspondentes.

Além das observações realizadas pelos autores citados não se acredita ser possível a construção de um modelo lógico que reflita a essência do modelo conceitual fora da solução baseada em grafos.

O modelo lógico adaptou a ideia de generalização apresentada no modelo conceitual, observando que as componentes espaço, tempo e classes não são opcionais, porém, a flexibilidade do BD orientado por grafos permite que para uma mesma classe se tenha atributos diferentes, sendo esse o caso da classe *time*, como também não possuam valores nulos. Além disso, observa-se que todas as

classes e relacionamentos do modelo lógico atuam como possíveis filtros de recuperação dos dados espaço-temporais.

Com a observação das classes e relacionamentos componentes do modelo lógico observa-se que o tempo de recuperação dos dados espaço-temporais baseados no tempo, espaço e classes é mínimo devido à pré-filtragem realizada pelas classes *thematic* e *classes*. Assim as consultas são direcionadas para as instâncias que possuam os atributos buscados.

Os modelos conceitual e lógico propostos possibilitam a integração de modelos de dados progressos que podem coexistir com outros, mantendo as suas nomenclaturas originais. Ressalte-se, porém, que é necessária a manutenção da semântica das classes componentes desses diversos modelos de dados coexistentes para a recuperação integrada de dados.

Os dados espaço-temporais podem sofrer alterações ao longo do tempo, o modelo lógico orientado por grafos possui a capacidade de obter a recuperação através de consultas baseadas no rastreamento do fluxo histórico dos dados com a adição de uma aresta, no nosso caso a aresta *update*.

### 5.3 GRAFO INTEGRADO: TERRESTRE, MARÍTIMO E ATMOSFÉRICO

O conhecimento prévio sobre o funcionamento do SGBD escolhido como, criar, ler, atualizar, eliminar dados e a realização de consultas, para implementação é fundamental na fase de pré-processamento, pois existem características na formatação dos dados exigidas pelo SGBD como orientação das coordenadas geográficas e o formato de entrada das primitivas gráficas.

Os arquivos nos formatos *shapefile* e *comma separated values (csv)* são facilmente trabalhados no pré-processamento.

Verificou-se nesse trabalho a mesma facilidade de implementação do BD orientado por grafos, SGBD Neo4j, relatado por Zhang, C. *et al.* (2023) devido a não necessidade de construir tabelas e ao importar os dados diretamente dos arquivos.

A escolha do ArangoDB mostrou-se adequada para as finalidades de avaliação do modelo conceitual e lógico colimadas. Constatou-se que a visualização geoespacial fornecida pelo ArangoDB está limitada a 1000 resultados. Essa limitação foi introduzida no sentido de se evitar travamento dos computadores de menor capacidade eventualmente utilizados. Porém os resultados podem ser

exportados ou acessados em sua completude absoluta a partir de *scripts* na linguagem Python.

#### 5.4 RECUPERAÇÃO INTEGRADA: ESPAÇO, TEMPO E CLASSES

Um cenário decisório pode ser estabelecido de forma investigativa, quando se busca explicar a consequência de um evento pretérito (AHMADUN *et al.*, 2020; SOARES *et al.*, 2020); preditivo, quando se busca a determinação da probabilidade de ocorrência do evento no futuro, exemplo, Li *et al.* (2022), e corretivo quando se constrói um cenário do momento presente. Independentemente do tipo de cenário decisório a busca de respostas será através da análise e síntese de dados integrados fidedignos, organizados e consistentes.

A recuperação integrada de dados espaço-temporais dos ambientes terrestre, marítimo e atmosférico possibilita a seleção de dados de acordo com o cenário decisório. Os resultados de recuperação integrada apresentados por esse trabalho demonstraram a possibilidade de resgatar um conjunto de dados baseados numa janela temporal específica e para um volume definido geograficamente. A definição do tempo e espaço são dois parâmetros básicos para um cenário decisório.

Os resultados alcançados por esse trabalho são análogos ao Gebbert e Pebesma (2014), Tian *et al.* (2015), e outros citados anteriormente, porém sem nenhum tipo de desenvolvimento adicional, como módulos, APIs e frameworks, e refletindo conseqüentemente nas suas manutenibilidade, integridade e escalabilidade.

A linguagem de consulta AQL do SGBD ArangoDB possui uma curva de aprendizado lenta, porém a documentação oficial do SGBD disponível é abrangente e completa, orientando o usuário no desenvolvimento das consultas para recuperação de quaisquer dados de interesse contidos no ArangoDB.

Sob o ponto de vista operacional pode-se aplicar, por exemplo, a recuperação de dados integrados com propósitos de investigação de acidentes, óleo no mar, zonas pesqueiras ou operações de empregos de meios flutuantes em uma determinada época do ano, onde parâmetros físicos como o regime de ventos e correntes podem influenciar.

## 6 CONSIDERAÇÕES FINAIS

A abordagem desenvolvida nesse trabalho oferece a capacidade armazenamento e gestão de grandes conjuntos de dados espaço-temporais e recuperação desses dados baseados no tempo, espaço e em classes. A busca de uma estrutura integrada facilitará a gestão, análise e distribuição de dados.

A proposta deste trabalho consistiu na construção de um modelo conceitual de representação integrada para dados espaço-temporais marítimos, terrestres e atmosféricos e um modelo lógico baseado em grafos para o gerenciamento de dados espaço-temporais.

Nessa abordagem, com a motivação da necessidade de integração de dados espaço-temporais pertencentes a vários domínios de aplicação, foi criado um modelo conceitual centrado no tempo e espaço. Foi realizada a implementação do modelo lógico em um BD orientado por grafos e foram conduzidas consultas de recuperação de dados sob a ótica operacional para demonstrar a sua aplicabilidade.

No processo confirmou-se a hipótese inicial de que a abordagem mais adequada para a construção de uma estrutura de armazenamento visando à gestão de dados espaço-temporais seja a de um BD orientado por grafos, devido à sua flexibilidade intrínseca e ao alto relacionamento entre os dados espaço-temporais. Foram constatadas as premissas iniciais de baixa manutenibilidade, alta integridade e alta escalabilidade ao mesmo tempo em que muito reduz a necessidade de recursos computacionais.

Os resultados observados demonstram que o gerenciamento de dados baseado em grafos pode suportar de forma eficaz a absorção de modelos de dados pretéritos e futuros. A abordagem fornece uma nova visão para a integração de dados espaço-temporais. O projeto apresentou-se como um laboratório de implantação do modelo conceitual e lógico proposto nesse trabalho.

Os modelos conceitual e lógico são flexíveis e podem ser atualizados com a incorporação de outros dados espaço-temporais na medida em que são processados e incorporados ao BD sendo esse objetivo somente alcançável sob o paradigma de orientação por grafos.

## 6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Recomenda-se a aplicação do modelo conceitual e lógico para além das classes de representação dos objetos exploradas no projeto piloto. Integrações de dados espaço-temporais em grande escala podem ser a base para aplicações como planeamento territorial, criação de *digital twins*, sistemas de apoio a decisões, etc.



## REFERÊNCIAS

- AHMADUN, F.-R.; WONG, M. M. R.; MAT SAID, A. Consequences of the 2004 Indian Ocean Tsunami in Malaysia. **Safety Science**, v. 121, p. 619–631, 2020.
- ARANGO. ArangoDB. Disponível em: <<https://arangodb.com/>>. Acesso em: 20/11/2023.
- ARANGODB. ArangoDB Geo-Spatial Indexes. Disponível em: <<https://docs.arangodb.com/3.11/index-and-search/indexing/working-with-indexes/geo-spatial-indexes/>>. Acesso em: 20/11/2023.
- ARMSTRONG, M. Temporality in spatial database. **Proceedings of the Proceedings: GIS/LIS, San Antonio**, v. 2, 1988.
- AUTODESK. Tinkercad. Disponível em: <<https://www.tinkercad.com/dashboard>>. Acesso em: 29/11/2023.
- BAUMANN, P. A general conceptual framework for multi-dimensional spatio-temporal data sets. **Environmental Modelling & Software**, v. 143, p. 105096, 2021.
- BECHBERGER, D.; PERRYMAN, J. **Graph Databases in Action**. 1st edition ed. Shelter Island, New York: Manning, 2020.
- BENTSEN, L. Ø.; WARAKAGODA, N. D.; STENBRO, R.; ENGELSTAD, P. A Unified Graph Formulation for Spatio-Temporal Wind Forecasting. **Energies**, v. 16, n. 20, p. 7179, 2023. Multidisciplinary Digital Publishing Institute.
- BENWAY, H. M.; LORENZONI, L.; WHITE, A. E.; et al. Ocean Time Series Observations of Changing Marine Ecosystems: An Era of Integration, Synthesis, and Societal Applications. **Frontiers in Marine Science**, v. 6, 2019. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmars.2019.00393>>. Acesso em: 19/10/2023.
- CORIOLIS. Coriolis: In situ data for operational oceanography. Disponível em: <<https://www.coriolis.eu.org/>>. Acesso em: 20/11/2023.
- DEL MONDO, G.; PENG, P.; GENSEL, J.; CLARAMUNT, C.; LU, F. Leveraging Spatio-Temporal Graphs and Knowledge Graphs: Perspectives in the Field of Maritime Transportation. **ISPRS International Journal of Geo-Information**, v. 10, n. 8, p. 541, 2021. Multidisciplinary Digital Publishing Institute.
- DEL MONDO, G.; RODRÍGUEZ, M. A.; CLARAMUNT, C.; BRAVO, L.; THIBAUD, R. Modeling consistency of spatio-temporal graphs. **Data & Knowledge Engineering**, v. 84, p. 59–80, 2013.
- DICKEY, T. D. Emerging ocean observations for interdisciplinary data assimilation systems. **Journal of Marine Systems**, The Use of Data Assimilation in Coupled Hydrodynamic, Ecological and Bio-geo-chemical Models of the Ocean. Selected papers from the 33rd International Liege Colloquium on Ocean Dynamics, held in Liege, Belgium on May 7-11th, 2001., v. 40–41, p. 5–48, 2003.

DING, Y.; XU, Z.; ZHU, Q.; et al. Integrated data-model-knowledge representation for natural resource entities. **International Journal of Digital Earth**, v. 15, n. 1, p. 653–678, 2022. Taylor & Francis.

DSG. Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais. Disponível em: <[https://bdgex.eb.mil.br/portal/index.php?option=com\\_content&view=article&id=81:et-edgv&catid=60&Itemid=353&lang=pt](https://bdgex.eb.mil.br/portal/index.php?option=com_content&view=article&id=81:et-edgv&catid=60&Itemid=353&lang=pt)>. Acesso em: 24/11/2023.

GEBBERT, S.; PEBESMA, E. A temporal GIS for field based environmental modeling. **Environmental Modelling & Software**, v. 53, p. 1–12, 2014.

GOODCHILD, M. F.; YUAN, M.; COVA, T. J. Towards a general theory of geographic representation in GIS. **International Journal of Geographical Information Science**, v. 21, n. 3, p. 239–260, 2007. Taylor & Francis.

GRASSLE, J. F. The Ocean Biogeographic Information System (OBIS): an on-line, worldwide atlas for accessing, modeling and mapping marine biological data in a multidimensional geographic context. **Oceanography**, v. 13, n. 3, p. 5–7, 2000. Oceanography Society.

GRAVES, N. J.; DUNLOP, O. J.; TORNEY-PURTA, J. V. **Teaching for international understanding, peace and human rights**. UNESCO, 1984.

HECKMANN, T.; SCHWANGHART, W.; PHILLIPS, J. D. Graph theory—Recent developments of its application in geomorphology. **Geomorphology**, v. 243, p. 130–146, 2015.

HORNSBY, K.; EGENHOFER, M. J. Identity-based change: a foundation for spatio-temporal knowledge representation. **International Journal of Geographical Information Science**, v. 14, n. 3, p. 207–224, 2000. Taylor & Francis.

HULL, S.; DICKIE, I.; TINCH, R.; SAUNDERS, J. Issues and challenges in spatio-temporal application of an ecosystem services framework to UK seas. **Marine Policy**, v. 45, p. 359–367, 2014.

IHO. International Hydrographic Organization. Disponível em: <<https://iho.int/en/iho-publications>>. Acesso em: 24/11/2023.

INDE. Infraestrutura de Dados Espaciais. Disponível em: <<https://inde.gov.br/>>. Acesso em: 20/11/2023.

INSPIRE. Infrastructure for spatial information in Europe. Disponível em: <<https://inspire.ec.europa.eu/>>. Acesso em: 22/11/2023.

JAHANBAKHT, M.; XIANG, W.; HANZO, L.; RAHIMI AZGHADI, M. Internet of Underwater Things and Big Marine Data Analytics—A Comprehensive Survey. **IEEE Communications Surveys & Tutorials**, v. 23, n. 2, p. 904–956, 2021.

JARDIM, R.; DOS SANTOS, M.; NETO, E.; et al. Design of a framework of military defense system for governance of geoinformation. **Procedia Computer Science**, The 8th International Conference on Information Technology and Quantitative

Management (ITQM 2020 & 2021): Developing Global Digital Economy after COVID-19., v. 199, p. 174–181, 2022.

JEON, M.; NOH, Y.; JEON, K.; LEE, S.; LEE, I. Data gap analysis of ship and maritime data using meta learning. **Applied Soft Computing**, v. 101, p. 107048, 2021.

JOHNSON, A. E.; MCCLINTOCK, W. J.; BURTON, O.; et al. Marine spatial planning in Barbuda: A social, ecological, geographic, and legal case study. **Marine Policy**, v. 113, p. 103793, 2020.

KARSENTI, E.; ACINAS, S. G.; BORK, P.; et al. A Holistic Approach to Marine Eco-Systems Biology. **PLoS Biology**, v. 9, n. 10, p. e1001177, 2011.

LANEY, D. 3D Data Management: Controlling Data Volume, Velocity, and Variety. **Gartner**, 2001. Disponível em: <<https://studylib.net/doc/8647594/3d-data-management--controlling-data-volume--velocity--an...>>. Acesso em: 24/10/2023.

LAURINI, R.; THOMPSON, D. **Fundamentals of Spatial Information Systems**. 1ª edição ed. London: Elsevier Science Publishing Co Inc, 1992.

LE, Y. Challenges in Data Integration for Spatiotemporal Analysis. **Journal of Map & Geography Libraries**, v. 8, n. 1, p. 58–67, 2012. Routledge.

LENSU, M.; GOERLANDT, F. Big maritime data for the Baltic Sea with a focus on the winter navigation system. **Marine Policy**, v. 104, p. 53–65, 2019.

LI, H.; ZHU, Q.; ZHANG, L.; et al. Integrated representation of geospatial data, model, and knowledge for digital twin railway. **International Journal of Digital Earth**, v. 15, n. 1, p. 1657–1675, 2022. Taylor & Francis.

LIN-NAN, Y.; LIN, P.; LI-MIN, Z.; LI-LIAN, Z. Application of a spatial-temporal data model based on graph theory for the field parcel updating geographical information system (FPUGIS). **New Zealand Journal of Agricultural Research**, v. 50, n. 5, p. 895–902, 2007. Taylor & Francis.

LIU, J.; ZHANG, Q.; FU, L.; WANG, X.; LU, S. Evolving Knowledge Graphs. IEEE INFOCOM 2019 - IEEE Conference on Computer Communications. **Anais...** . p.2260–2268, 2019. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/8737547>>. Acesso em: 24/11/2023.

LIU, W.; LI, X.; RAHN, D. A. Storm event representation and analysis based on a directed spatiotemporal graph model. **International Journal of Geographical Information Science**, v. 30, n. 5, p. 948–969, 2016. Taylor & Francis.

LIU, Y.; ZHOU, Y. Territory spatial planning and national governance system in China. **Land Use Policy**, v. 102, p. 105288, 2021.

LONGLEY, P. A.; GOODCHILD, M. F.; MAGUIRE, D. J.; RHIND, D. W. **Geographic Information Systems and Science**. 2nd edição ed. Chichester: John Wiley & Sons, 2005.

MATSUOKA, K.; SKOGLUND, A.; ROTH, G.; et al. Quantarctica, an integrated mapping environment for Antarctica, the Southern Ocean, and sub-Antarctic islands. **Environmental Modelling & Software**, v. 140, p. 105015, 2021.

NATIVI, S.; MAZZETTI, P.; SANTORO, M.; et al. Big Data challenges in building the Global Earth Observation System of Systems. **Environmental Modelling & Software**, v. 68, p. 1–26, 2015.

NOTEPAD++. Notepad++. Disponível em: <<https://notepad.plus/>>. Acesso em: 20/11/2023.

OSM. OpenStreetMap. Disponível em: <<https://www.openstreetmap.org/>>. Acesso em: 24/11/2023.

PETELIN, B.; KONONENKO, I.; MALAČIČ, V.; KUKAR, M. Frequent subgraph mining in oceanographic multi-level directed graphs. **International Journal of Geographical Information Science**, v. 33, n. 10, p. 1936–1959, 2019. Taylor & Francis.

PEUQUET, D. J. It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. **Annals of the Association of American Geographers**, v. 84, n. 3, p. 441–461, 1994. Routledge.

PEUQUET, D. J.; DUAN, N. An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. **International journal of geographical information systems**, v. 9, n. 1, p. 7–24, 1995. Taylor & Francis.

POULTER, B.; GOODALL, J. L.; HALPIN, P. N. Applications of network analysis for adaptive management of artificial drainage systems in landscapes vulnerable to sea level rise. **Journal of Hydrology**, v. 357, n. 3, p. 207–217, 2008.

PULTAR, E.; RAUBAL, M.; COVA, T. J.; GOODCHILD, M. F. Dynamic GIS Case Studies: Wildfire Evacuation and Volunteered Geographic Information. **Transactions in GIS**, v. 13, n. s1, p. 85–104, 2009.

PYTHON. Python. Disponível em: <<https://www.python.org/>>. Acesso em: 20/11/2023.

QGIS. QGIS. Disponível em: <[https://qgis.org/pt\\_BR/site/](https://qgis.org/pt_BR/site/)>. Acesso em: 20/11/2023.

REICHSTEIN, M.; CAMPS-VALLS, G.; STEVENS, B.; et al. Deep learning and process understanding for data-driven Earth system science. **Nature**, v. 566, n. 7743, p. 195–204, 2019. Nature Publishing Group.

RENOLLEN, A. Modelling the Real World: Conceptual Modelling in Spatiotemporal Information System Design. **Transactions in GIS**, v. 4, n. 1, p. 23–42, 2000.

ROBINSON, I.; WEBBER, J.; EIFREM, E. **Graph Databases: New Opportunities for Connected Data**. 2nd edition ed. Beijing Boston Farnham: O'Reilly Media, 2015.

SCILAB. Scilab. Disponível em: <<https://www.scilab.org/>>. Acesso em: 20/11/2023.

SOARES, M. DE O.; TEIXEIRA, C. E. P.; BEZERRA, L. E. A.; et al. Oil spill in South Atlantic (Brazil): Environmental and governmental disaster. **Marine Policy**, v. 115, p. 103879, 2020.

THIBAUD, R.; DEL MONDO, G.; GARLAN, T.; MASCRET, A.; CARPENTIER, C. A Spatio-Temporal Graph Model for Marine Dune Dynamics Analysis and Representation. **Transactions in GIS**, v. 17, n. 5, p. 742–762, 2013.

TIAN, Y.; JI, Y.; SCHOLER, J. A Prototype Spatio-temporal Database Built on Top of Relational Database. 2015 12th International Conference on Information Technology - New Generations. **Anais...** . p.14–19, 2015. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7113442>>. Acesso em: 22/11/2023.

DE LA TORRE-CASTRO, M.; FRÖCKLIN, S.; BÖRJESSON, S.; OKUPNIK, J.; JIDDAWI, N. S. Gender analysis for better coastal management – Increasing our understanding of social-ecological seascapes. **Marine Policy**, v. 83, p. 62–74, 2017.

TRYFONA, N.; JENSEN, C. S. Conceptual Data Modeling for Spatiotemporal Applications. **Geoinformatica**, v. 3, n. 3, p. 245–268, 1999.

TU, E.; ZHANG, G.; RACHMAWATI, L.; RAJABALLY, E.; HUANG, G.-B. Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey From Data to Methodology. **IEEE Transactions on Intelligent Transportation Systems**, v. 19, n. 5, p. 1559–1582, 2018.

UN. FIRST GLOBAL INTEGRATED MARINE ASSESSMENT (FIRST WORLD OCEAN ASSESSMENT). Disponível em: <[https://www.un.org/Depts/los/global\\_reporting/WOA\\_RegProcess.htm](https://www.un.org/Depts/los/global_reporting/WOA_RegProcess.htm)>. Acesso em: 19/10/2023.

UNESCO. Guide for establishing an IODE National Oceanographic Data Centre. , 2022. Disponível em: <<https://unesdoc.unesco.org/ark:/48223/pf0000382882>>. Acesso em: 20/10/2023.

USGS. **US Topo Product Standart**. Techniques and Methods, , 2019.

WACHOWICZ, M. **Object-Oriented Design for Temporal GIS**. London: CRC Press, 2014.

WANG, D.; CHENG, T. A spatio-temporal data model for activity-based transport demand modelling. **International Journal of Geographical Information Science**, v. 15, n. 6, p. 561–585, 2001. Taylor & Francis.

WANG, X.; ZHOU, X.; LU, S. Spatiotemporal data modelling and management: a survey. Proceedings 36th International Conference on Technology of Object-Oriented Languages and Systems. TOOLS-Asia 2000. **Anais...** . p.202–211, 2000. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/885919>>. Acesso em: 24/11/2023.

WEST, D. B. **Introduction to Graph Theory**. Subsequent edition ed. Upper Saddle River, N.J: Pearson College Div, 2000.

WILCOX, D. J.; HARWELL, M. C.; ORTH, R. J. Modeling Dynamic Polygon Objects in Space and Time: A New Graph-based Technique. **Cartography and Geographic Information Science**, v. 27, n. 2, p. 153–164, 2000. Taylor & Francis.

WOODRUFF, S. D.; SLUTZ, R. J.; JENNE, R. L.; STEURER, P. M. A Comprehensive Ocean-Atmosphere Data Set. **Bulletin of the American Meteorological Society**, v. 68, n. 10, p. 1239–1250, 1987. American Meteorological Society.

WU, J.; ORLANDI, F.; O’SULLIVAN, D.; DEV, S. LinkClimate: An interoperable knowledge graph platform for climate data. **Computers & Geosciences**, v. 169, p. 105215, 2022.

XIA, H.; YUAN, S.; PRISHCHEPOV, A. V. Spatial-temporal heterogeneity of ecosystem service interactions and their social-ecological drivers: Implications for spatial planning and management. **Resources, Conservation and Recycling**, v. 189, p. 106767, 2023.

YAN, Z.; XIAO, Y.; CHENG, L.; et al. Exploring AIS data for intelligent maritime routes extraction. **Applied Ocean Research**, v. 101, p. 102271, 2020.

YI, S.; XIAO, Y. Spatio-temporal properties (STP) graph model for multiple-entity variables estimation. 2015 23rd International Conference on Geoinformatics. **Anais...** p.1–5, 2015. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7378563>>. Acesso em: 24/11/2023.

YUAN, M.; MARK, D. M.; EGENHOFER, M. J.; PEUQUET, D. J. Extensions to Geographic Representations. **A Research Agenda for Geographic Information Science**, 2004. CRC Press.

ZAUCHA, J.; GEE, K. (ORGS.). **Maritime Spatial Planning: past, present, future**. Cham: Springer International Publishing, 2019.

ZHANG, C.; HE, B.; GUO, R.; MA, D. A graph-based approach for representing addresses in geocoding. **Computers, Environment and Urban Systems**, v. 100, p. 101937, 2023.

ZHANG, W.; PENG, L.; GE, X.; et al. Spatio-Temporal Knowledge Graph-Based Research on Agro-Meteorological Disaster Monitoring. **Remote Sensing**, v. 15, n. 18, p. 4403, 2023. Multidisciplinary Digital Publishing Institute.

## APÊNDICE 1 – CONSULTAS EM AQL

### 1) Consultas de caráter geral

QUADRO – CONSULTAS DE CARÁTER GERAL

Id	Consulta	AQL
1	Quais os parâmetros e correspondentes valores foram observados no polígono pol1?	<pre>FOR u IN parameters LET area = GEO_POLYGON([[ -42.42865,-18.79761 ],[ -43.41192,-18.90138 ],[ -43.56024,-23.71109],[ - 42.15673,-23.63562],[ - 42.42865,-18.79761]]) FOR v IN OUTBOUND u located_in FILTER GEO_CONTAINS(area, v.location) FILTER u.air_temperature != null RETURN u</pre>
2	Quais os parâmetros e correspondentes valores encontrados dentro do polígono pol1 a uma profundidade entre p1 e p2?	<pre>FOR u IN parameters LET area = GEO_POLYGON([[ -35.77058,-19.18352 ],[ -49.14916,-19.25754 ],[ -47.94718,-24.74226],[ - 34.76458,-27.32460],[ -35.77058,-19.18352]]) FOR v IN OUTBOUND u located_in FILTER GEO_CONTAINS(area, v.location) FILTER v.depth &gt; 12 &amp;&amp; v.depth &lt; 16  LET attributes = ( LET h = (MERGE(u, v)) FOR name IN ATTRIBUTES(h) FILTER REGEX_TEST(name, '^[a-z]') RETURN { name: name, value: h[name] } ) RETURN ZIP(attributes[*].name, attributes[*].value)</pre>
3	Quais os parâmetros e correspondentes valores encontrados entre os tempos t1 e t2 e entre os parâmetros p1 e p2?	<pre>FOR u IN parameters FILTER u.air_temperature &gt; 20 &amp;&amp; u.air_temperature &lt; 25 &amp;&amp; u.date_time_inst &gt; "2023-01-01" &amp;&amp; u.date_time_inst &lt; "2023-01-03" FOR v IN OUTBOUND u located_in RETURN u</pre>
4	Onde está localizado o parâmetro cujo valor encontra-se entre v1 e v2?	<pre>FOR u IN parameters FILTER u.air_temperature &gt; 20 &amp;&amp; u.air_temperature &lt; 25 FOR v IN OUTBOUND u located_in RETURN v.location</pre>
5	Quando e em que lugar ocorreu o parâmetro p1 com valores v1 e v2?	<pre>FOR u IN parameters FILTER u.air_temperature &gt; 20 &amp;&amp; u.air_temperature &lt; 25 FOR v IN OUTBOUND u located_in RETURN {temperatura: u.air_temperature, data:u.date_time_inst, local: v.location }</pre>
6	Quais são os parâmetros presentes no BD no período t1 e t2?	<pre>LET result = SORTED_UNIQUE( FOR u IN parameters FILTER u.date_time_inst &gt; "2023-01-01" &amp;&amp; u.date_time_inst &lt; "2023-01-03" FOR name IN ATTRIBUTES(u)</pre>

Id	Consulta	AQL
		<pre> FILTER REGEX_TEST(name, '^[a-z]') RETURN name ) RETURN result </pre>
7	Qual é o número total de registros de um parâmetro específico?	<pre> FOR v2 IN time   FILTER v2.sea_water_salinity_avg != null   COLLECT WITH COUNT INTO length RETURN length </pre>
8	A partir de um parâmetro quais registros com o valor do parâmetro específico e o tempo estão constantes no BD?	<pre> LET lista = 'sea_water_pressure' FOR u IN code   FILTER u._key == lista   FOR v1, e1, p1 IN 1..1 INBOUND u._id contains     FOR v2, e2, p2 IN 1..1 INBOUND v1 within       OPTIONS { uniqueVertices: "path" }       FOR name IN ATTRIBUTES(v2)         FILTER REGEX_TEST(name, lista)         RETURN {[name]: v2.air_temperature, date_time_inst:v2.date_time_inst } </pre>
9	A partir de um parâmetro quais registros com o valor do parâmetro específico estão constantes no BD?	<pre> LET lista = 'air_temperature' FOR u IN code   FILTER u._key == lista   FOR v1, e1, p1 IN 1..1 INBOUND u._id contains     FOR v2, e2, p2 IN 1..1 INBOUND v1 within       OPTIONS { uniqueVertices: "path" }    LET attributes = (     FOR name IN ATTRIBUTES(v2)      FILTER REGEX_TEST(name, lista)     RETURN {        name: name,       value: v2[name]     }   )   FILTER attributes != []   RETURN ZIP(attributes[*].name, attributes[*].value) </pre>
10	A partir de um parâmetro quais registros com todos os parâmetros adjacentes estão constantes no BD?	<pre> LET lista = 'sea_water_pressure'  FOR u IN code   FILTER u._key == lista   FOR v1, e1, p1 IN 1..1 INBOUND u._id contains     FOR v2, e2, p2 IN 1..1 INBOUND v1 within       OPTIONS { uniqueVertices: "path" }       FOR name IN ATTRIBUTES(v2)         FILTER REGEX_TEST(name, lista)         RETURN v2 </pre>
11	Quais os parâmetros, respectivos valores e posição geospacial foram observados entre o intervalo de tempo t1 e t2?	<pre> FOR u IN parameters   FILTER u.date_time_inst &gt; "2023-01-01" &amp;&amp; u.date_time_inst &lt; "2023-01-03"   FOR v IN OUTBOUND u located_in   LET attributes = (     LET h = (MERGE(u, v))     FOR name IN ATTRIBUTES(h)     FILTER REGEX_TEST(name, '^[a-z]')     RETURN {       name: name,       value: h[name]     }   ) </pre>



Id	Consulta	AQL
		<pre> } ) RETURN ZIP(attributes[*].name, attributes[*].value) </pre>
12	Quais os parâmetros, respectivos valores observados no polígono pol1 a uma profundidade d1?	<pre> FOR u IN parameters LET area = GEO_POLYGON([[[-35.77058,-19.18352 ],[-49.14916,-19.25754      ],[-47.94718,-24.74226],[- 34.76458,-27.32460],[-35.77058,-19.18352]]) FOR v IN OUTBOUND u located_in FILTER GEO_CONTAINS(area, v.location) FILTER v.depth == 100  LET attributes = ( LET h = (MERGE(u, v)) FOR name IN ATTRIBUTES(h) FILTER REGEX_TEST(name, '^[a-z]') RETURN {   name: name,   value: h[name] } ) RETURN ZIP(attributes[*].name, attributes[*].value) </pre>

FONTE: O autor (2023).

2) quando + onde → o quê: Descreve os objetos ou conjunto de objetos (o quê) que estão presentes em um determinado local ou conjunto de locais (onde) em um determinado momento ou conjunto de horários (quando).

#### QUADRO – CONSULTAS TEMPO + ESPAÇO

Id	Consulta	AQL
1	Quais os parâmetros e correspondentes valores foram observados no polígono pol1 no período t1 e t2?	<pre> FOR u IN parameters FILTER u.date_time_inst &gt; "2023-01-01" &amp;&amp; u.date_time_inst &lt; "2023-01-03" LET area = GEO_POLYGON([[[-42.42865,-18.79761 ],[-43.41192,-18.90138      ],[-43.56024,-23.71109],[- 42.15673,-23.63562],[- 42.42865,-18.79761]]) FOR v IN OUTBOUND u located_in FILTER GEO_CONTAINS(area, v.location) FILTER u.air_temperature != null RETURN u </pre>
2	Quais os valores de um parâmetro específico foram observados no polígono pol1 no período t1 e t2?	<pre> FOR u IN parameters FILTER u.date_time_inst &gt; "2023-01-01" &amp;&amp; u.date_time_inst &lt; "2023-01-03" LET area = GEO_POLYGON([[[-42.42865,-18.79761 ],[-43.41192,-18.90138      ],[-43.56024,-23.71109],[- 42.15673,-23.63562],[- 42.42865,-18.79761]]) FOR v IN OUTBOUND u located_in FILTER GEO_CONTAINS(area, v.location) FILTER u.air_temperature != null RETURN u.air_temperature </pre>
3	Quais os parâmetros e correspondentes valores, e posição espacial encontrados no polígono p1 no período t1 e t2?	<pre> FOR u IN parameters FILTER u.date_time_inst &gt; "2011-01-05"    u.date_time_creation &gt; "0" </pre>

Id	Consulta	AQL
		<pre> LET area = GEO_POLYGON([[[-35.77058,-19.18352 ],[-49.14916,-19.25754      ],[-47.94718,-24.74226],[- 34.76458,-27.32460],[-35.77058,-19.18352]]) FOR v IN OUTBOUND u located_in FILTER GEO_CONTAINS(area, v.location) LET attributes = (   LET h = (MERGE(u, v))   FOR name IN ATTRIBUTES(h)   FILTER REGEX_TEST(name, '^[a-z]')   RETURN {     name: name,     value: h[name]   } ) RETURN ZIP(attributes[*].name, attributes[*].value) </pre>

FONTE: O autor (2023).

3) quando + o quê → onde: Descreve o local ou conjunto de locais (onde) ocupados por um determinado objeto ou conjunto de objetos (o quê) em um determinado momento ou conjunto de momentos (quando).

QUADRO – CONSULTAS TEMPO + PARÂMETROS

Id	Consulta	AQL
1	Quais os respectivos valores e posição geospacial foram observados para o parâmetro p1 entre o intervalo de tempo t1 e t2?	<pre> LET parâmetro = 'air_temperature' FOR u IN parameters   FILTER u.date_time_inst &gt; "2023-01-01" &amp;&amp; u.date_time_inst &lt; "2023-01-03"   FOR v IN OUTBOUND u located_in   LET attributes = (     LET h = (MERGE(u, v))     FOR name IN ATTRIBUTES(h)     FILTER REGEX_TEST(name, CONCAT(parâmetro,' location'))     RETURN {       name: name,       value: h[name]     }   ) RETURN ZIP(attributes[*].name, attributes[*].value) </pre>
2	Quais os respectivos valores e posição geospacial foram observados para o parâmetro p1 com valores entre v1 e v2 entre o intervalo de tempo t1 e t2?	<pre> LET parâmetro = 'air_temperature' FOR u IN parameters   FILTER u.air_temperature &gt; 20 &amp;&amp; u.air_temperature &lt; 25 &amp;&amp; u.date_time_inst &gt; "2023- 01-01" &amp;&amp; u.date_time_inst &lt; "2023-01-03"   FOR v IN OUTBOUND u located_in   LET attributes = (     LET h = (MERGE(u, v))     FOR name IN ATTRIBUTES(h)     FILTER REGEX_TEST(name, </pre>

Id	Consulta	AQL
		<pre> CONCAT(parâmetro, 'location')) RETURN {   name: name,   value: h[name] } ) RETURN ZIP(attributes[*].name, attributes[*].value) </pre>

FONTE: O autor (2023).

4) onde + o quê → quando: Descreve os tempos ou conjunto de tempos (quando) que um determinado objeto ou conjunto de objetos (o quê) ocupou um determinado local ou conjunto de locais (onde).

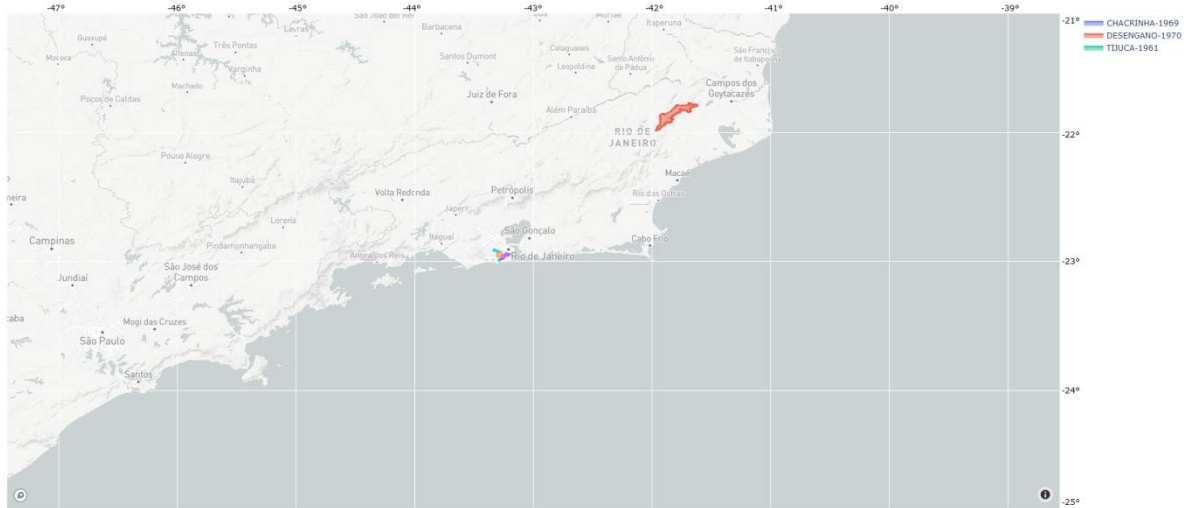
#### QUADRO – CONSULTAS ESPAÇO + PARÂMETROS

Id	Consulta	AQL
1	Quando ocorreu um específico parâmetro p1 observados no intervalo de v1 e v2, no polígono pol1 a uma profundidade d1?	<pre> FOR u IN parameters   FILTER u.sea_water_temperature &gt; 0 &amp;&amp;   u.sea_water_temperature &lt; 25   LET area = GEO_POLYGON([[ -35.77058, -19.18352   ], [-49.14916, -19.25754 ], [-47.94718, -24.74226], [-   34.76458, -27.32460], [-35.77058, -19.18352]])   FOR v IN OUTBOUND u located_in   FILTER GEO_CONTAINS(area, v.location)   FILTER v.depth == 100    LET attributes = (   LET h = (MERGE(u, v))   FOR name IN ATTRIBUTES(h)   FILTER REGEX_TEST(name, '[a-z]')   RETURN {     name: name,     value: h[name]   }   ) RETURN ZIP(attributes[*].name, attributes[*].value) </pre>
2	Quando ocorreu um específico parâmetro, com os seus respectivos valores entre v1 e v2 e localização observados no polígono pol1?	<pre> FOR u IN parameters   FILTER u.sea_water_temperature &gt; 1 &amp;&amp;   u.sea_water_temperature &lt; 25   LET area = GEO_POLYGON([[ -35.77058, -19.18352   ], [-49.14916, -19.25754 ], [-47.94718, -24.74226], [-   34.76458, -27.32460], [-35.77058, -19.18352]])   FOR v IN OUTBOUND u located_in   FILTER GEO_CONTAINS(area, v.location)   RETURN {temp: u.sea_water_temperature,   location:v.location, data:u.date_time_inst } </pre>

FONTE: O autor (2023).

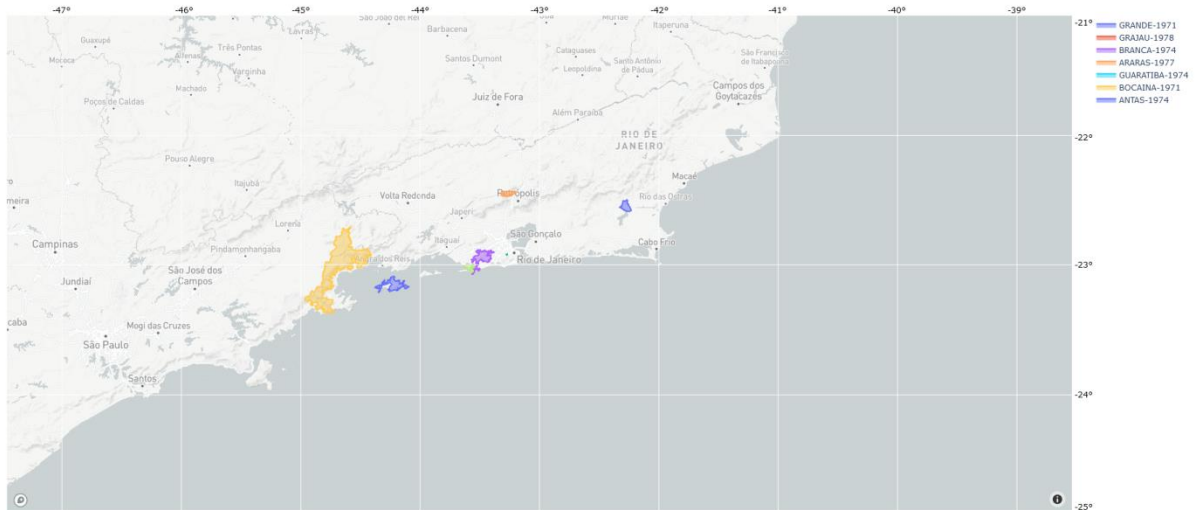
## APÊNDICE 2 – EVOLUÇÃO TEMPORAL DAS UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO

FIGURA – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-1961 A 31-12-1970.



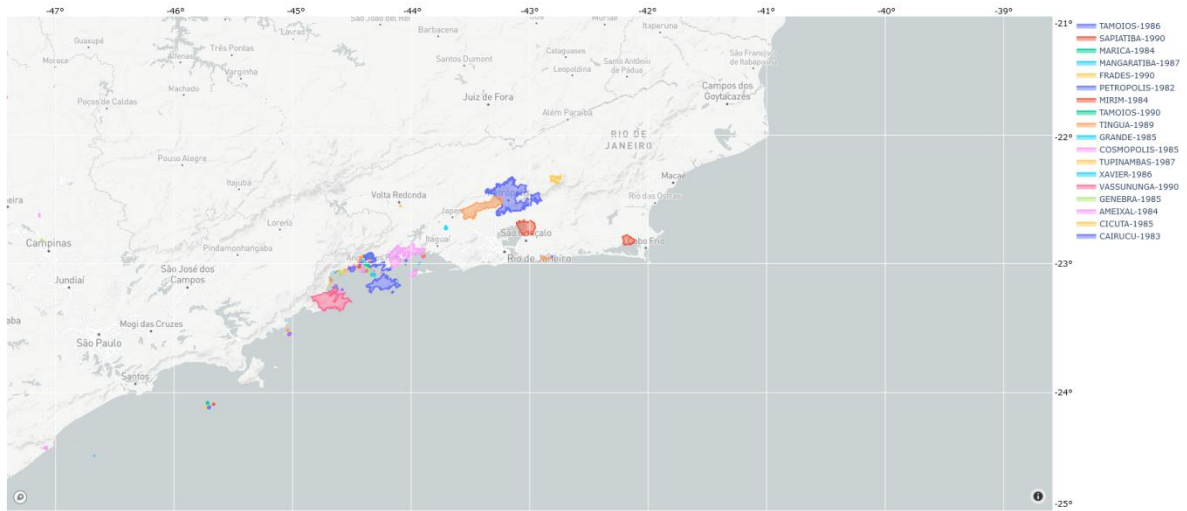
FONTE: O autor (2023).

FIGURA – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-1971 A 31-12-1980



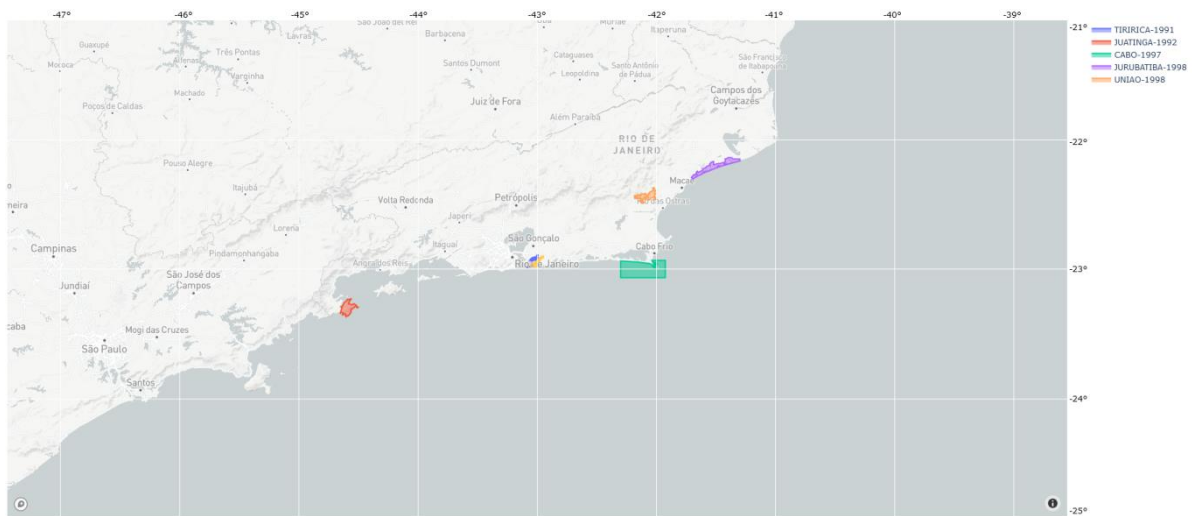
FONTE: O autor (2023).

FIGURA – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-1981 A 31-12-1990.



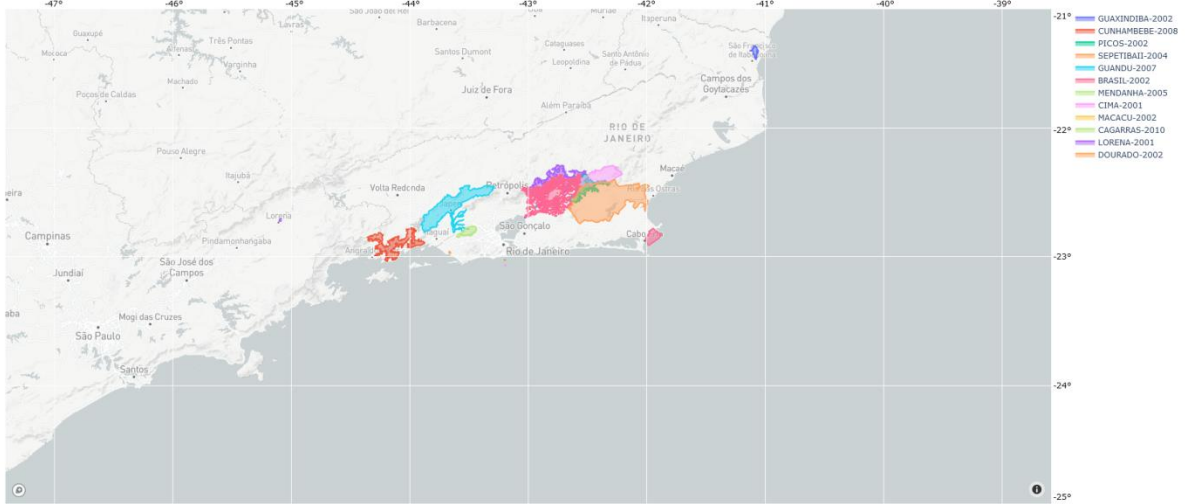
FONTE: O autor (2023).

FIGURA – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-1991 A 31-12-2000.



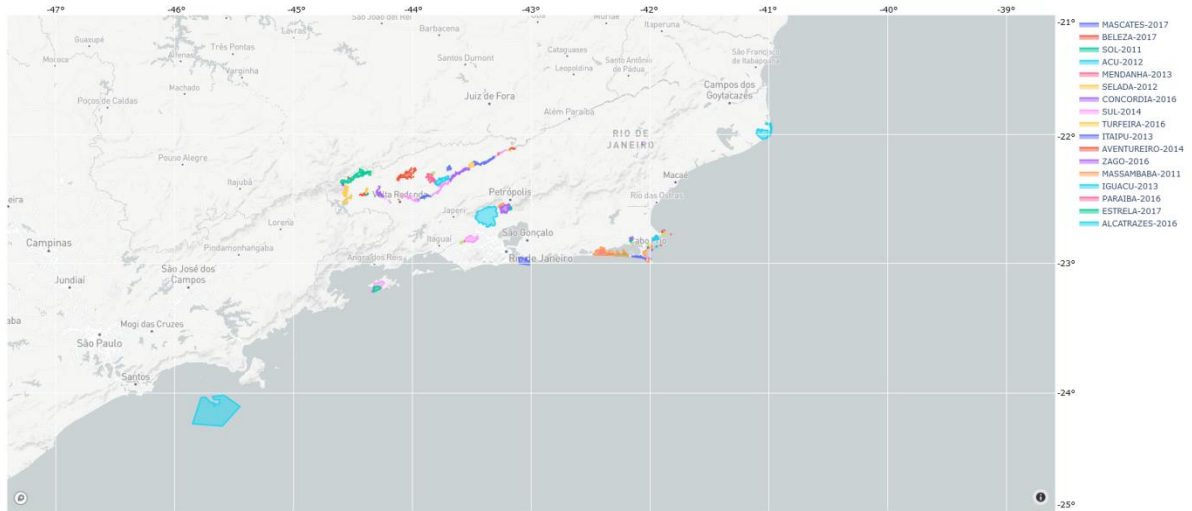
FONTE: O autor (2023).

FIGURA – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-2001 A 31-12-2010.



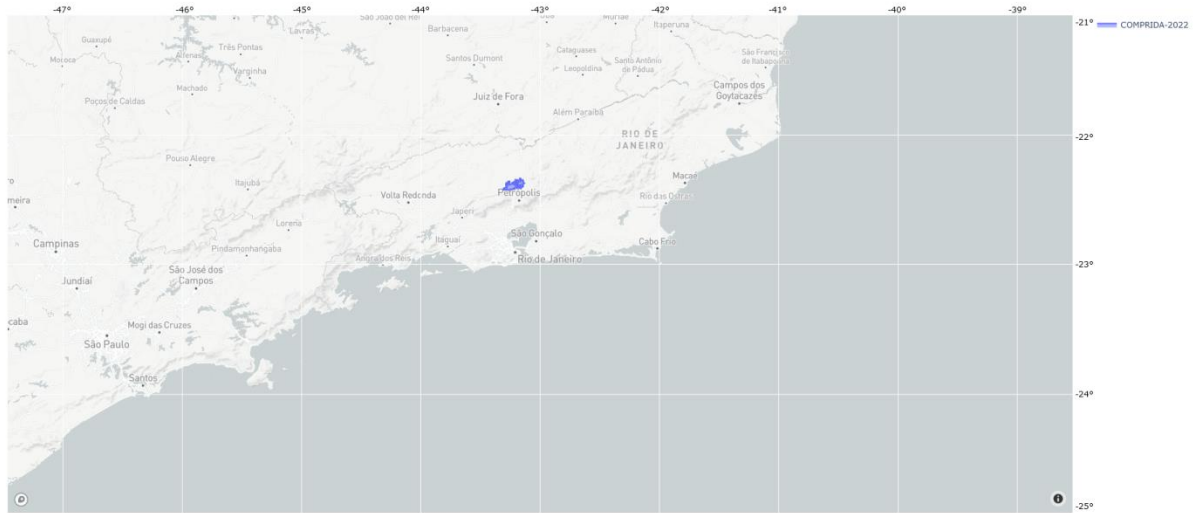
FONTE: O autor (2023).

FIGURA – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-2011 A 31-12-2020.



FONTE: O autor (2023).

FIGURA – UNIDADES DE CONSERVAÇÃO FEDERAIS E DO ESTADO DO RIO DE JANEIRO ESTABELECIDAS ENTRE 01-01-2021 A 31-12-2023.



FONTE: O autor (2023).