

ANA BEATRIZ MITSUTANI

**CÁLCULO DE MÉTRICAS DE EFICIÊNCIA NO
PORTO DE SANTOS A PARTIR DO
ROTULAMENTO DE DADOS DE AIS**

São Paulo
2024

ANA BEATRIZ MITSUTANI

**CÁLCULO DE MÉTRICAS DE EFICIÊNCIA NO
PORTO DE SANTOS A PARTIR DO
ROTULAMENTO DE DADOS DE AIS**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para obten-
ção do Título de Engenheiro de Produção.

São Paulo
2024

ANA BEATRIZ MITSUTANI

**CÁLCULO DE MÉTRICAS DE EFICIÊNCIA NO
PORTO DE SANTOS A PARTIR DO
ROTULAMENTO DE DADOS DE AIS**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para obten-
ção do Título de Engenheiro de Produção.

Orientador:

Prof. Dr. Daniel de Oliveira Mota

São Paulo
2024

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Mitsutani, Ana Beatriz

Cálculo de métricas de eficiência no Porto de Santos a partir do rotulamento de dados de AIS / A. B. Mitsutani - São Paulo, 2024.

103p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Sistema de Identificação Automática. 2.Rotulamento de dados. 3.Porto de Santos. 4.Monitoramento portuário. 5.Eficiência operacional. 6.Gestão portuária. I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

“Onde quer que haja mulheres e homens, há sempre o que fazer, há sempre o que ensinar, há sempre o que aprender”

- Paulo Freire

AGRADECIMENTOS

Gostaria de agradecer à minha família, meu irmão Daniel, minhas primas, tias e avós que me apoiam e acompanham desde sempre, e ao meu namorado, Kevin, por me ensinar outras formas de enxergar as coisas e me dar mais companheirismo, amor e compreensão que eu poderia desejar.

Agradeço ao meu colega, Koji, pelo incentivo e apoio a iniciar este trabalho, e aos meus amigos e colegas de classe, Thiago, Melissa, João Pedro, Isabella e Catarina, que compartilharam comigo a jornada que foi estudar na Escola Politécnica e em quem pude confiar de olhos fechados durante minha vida acadêmica.

Ao colega de laboratório, Fray, que mesmo sem nenhuma responsabilidade sobre este projeto o abraçou como se fosse seu, e ao professor Daniel Oliveira Mota, não apenas por ter desempenhado com maestria o papel de orientador durante este ano, mas também pela paciência e por acreditar no meu potencial.

Por fim, agradeço e dedico este trabalho principalmente aos meus pais, pois tudo o que conquistei até hoje é graças a eles. Ao meu pai, Heitor, que também engenheiro me inspirou a seguir este caminho e me ensinou desde pequena a pensar de forma crítica sobre os fenômenos da natureza e dos homens, e à minha mãe, Edna, que é de longe a mulher mais forte que eu conheço e que me motiva a continuar evoluindo como ser humano todos os dias. Obrigada por sempre tentarem me entender e por todas outras as formas que vocês já me apoiaram.

RESUMO

Resumo: O presente trabalho tem como objetivo propor um método automatizado para monitorar as operações nos terminais do Porto de Santos a partir do rotulamento de dados do Sistema de Identificação Automática (AIS), um recurso de rastreamento de embarcações. A análise envolve a coleta, pré-processamento e rotulamento de dados dinâmicos, que são utilizados para definir as trajetórias das embarcações e calcular métricas de eficiência operacional, identificando os principais pontos de melhorias no Porto de Santos. Três algoritmos específicos foram implementados para otimizar o processo de rotulamento, permitindo a criação de um conjunto de dados estruturado para avaliar o desempenho dos terminais portuários. A pesquisa contribui para a modernização da gestão portuária, destacando a importância da análise de dados na melhoria de produtividade e redução de congestionamentos no Porto de Santos.

Palavras-chave: Sistema de Identificação Automática, rotulamento de dados, Porto de Santos, monitoramento portuário, eficiência operacional, gestão portuária.

ABSTRACT

Abstract: This study aims to propose an automated method for monitoring terminal operations at the Port of Santos using Automatic Identification System (AIS) data labeling, a vessel tracking resource. The analysis involves the collection, preprocessing, and labeling of dynamic data, which is used to define vessel trajectories and calculate operational efficiency metrics, identifying key improvement areas at the Port of Santos. Three specific algorithms were implemented to optimize the labeling process, enabling the creation of a structured dataset to assess terminal performance. The research contributes to the modernization of port management, highlighting the importance of data analysis in enhancing productivity and reducing congestion at the Port of Santos.

Keywords - Automatic Identification System, data labeling, Port of Santos, port monitoring, operational efficiency, port management.

LISTA DE FIGURAS

1.1	Vista aérea do canal do Porto de Santos	p. 13
1.2	Terminais do Porto de Santos	p. 14
1.3	Crescimento da tonelagem de carga movimentada no Porto de Santos entre 2013 e 2023	p. 15
2.1	Dimensões da eficiência portuária	p. 19
2.2	Dimensões da eficiência operacional: uma sequência	p. 20
2.3	Funcionamento esquemático das entidades do sistema AIS	p. 23
2.4	Níveis que compõe um porto inteligente	p. 25
2.5	Atrações por tipo de navio em Piraeus	p. 28
2.6	Rotas da Matriz OD para o Porto de Cingapura	p. 30
2.7	Índices de Moran da velocidade das áreas do Porto de Cingapura	p. 32
2.8	Índices de Moran de densidade de navios das áreas do Porto de Cingapura	p. 33
2.9	Passos para obtenção de trajetórias	p. 35
2.10	Representação matricial dos resultados de classificação com 5 categorias	p. 36
2.11	Representação matricial dos resultados de classificação binária	p. 37
2.12	Representação gráfica da acurácia	p. 38
2.13	Representação gráfica da precisão	p. 38
2.14	Representação gráfica do <i>recall</i>	p. 39
3.1	Sequencia das principais operações de tratamentos dos dados	p. 41
3.2	Área total para a qual foi coletada dados AIS	p. 43
3.3	Retângulo com a área filtrada para análise	p. 46
3.4	Distribuição do calado das embarcações	p. 47
3.5	Distribuição das velocidades não nulas dos dados em análise	p. 48
3.6	Geolocalização dos códigos utilizados para rotulação	p. 50

3.7	Rotulamento inicial dos dados de posição (terminais)	p. 52
3.8	Rotulamento inicial dos dados de posição (canal)	p. 53
3.9	Caso 1: oscilações na velocidade enquanto o navio está atracado. Navio se movendo está representado em azul e parado em vermelho	p. 54
3.10	Caso 2: Oscilações na velocidade enquanto o navio está navegando. Navio se movendo está representado em azul e parado em vermelho	p. 55
3.11	Representação da primeira parte da reconstrução de trajetórias	p. 57
3.12	Representação da segunda e terceira partes da reconstrução de trajetórias	p. 57
3.13	Perfil da velocidade original do STAR CRIMSON	p. 61
3.14	Perfil da velocidade ajustada do STAR CRIMSON	p. 62
3.15	Trajetórias geradas após a aplicação do algoritmo 1 para o STAR CRIMSON	p. 64
3.16	Trajetórias geradas após a aplicação do algoritmo 2 para o STAR CRIMSON	p. 64
3.17	Trajetórias do navio SOL DO BRASIL antes da aplicação do Algoritmo 3	p. 65
3.18	Trajetórias do navio SOL DO BRASIL após da aplicação do Algoritmo 3	p. 66
4.1	Comparação entre o número de atracações calculadas a partir do rotulamento e as reais, por terminal	p. 69
4.2	Representação gráfica da relação dos resultados	p. 70
4.3	Trajetórias do navio APL TURKEY criadas pelo Algoritmos 1	p. 72
4.4	Ampliação das trajetórias somente de parada do navio APL TURKEY criadas pelo Algoritmos 1	p. 73
4.5	Trajetórias do navio APL TURKEY criadas pelo Algoritmos 1	p. 74
4.6	Ampliação das trajetórias somente de parada do APL TURKEY criadas após os ajustes do Algoritmos 2	p. 75
4.7	Trajetória única de parada do APL TURKEY após aplicação do Algoritmo 3	p. 76
4.8	Trajetórias do navio APL TURKEY criadas pelo Algoritmo 3	p. 77
4.9	Número total de trajetórias geradas após a aplicação de cada algoritmo	p. 78

4.10 Número médio de dados registros por trajetória após a aplicação de cada algoritmo	p. 78
4.11 Número médio de trajetória por navio na base após a aplicação de cada algoritmo	p. 79
4.12 Comparação entre o número de atracações calculadas a partir do rotulamento e as reais, por terminal	p. 80
4.13 Comparação entre os erros de falso positivo antes e depois do tratamento	p. 81
4.14 Comparação entre os erros de falso negativo antes e depois do tratamento	p. 82
4.15 Gráfico de ocupação dos terminais no Porto de Santos	p. 84
4.16 Atracações por dia no Porto de Santos	p. 85
4.17 Atracações por mês no Porto de Santos	p. 86
4.18 Distribuição do tempo de permanência das embarcações no Porto de Santos	p. 87
4.19 Taxa de Ocupação dos terminais do Porto de Santos	p. 88
4.20 Relação entre TP e número de atracações dos terminais do Porto de Santos	p. 89
4.21 Relação entre TEO e TO dos terminais do Porto de Santos	p. 90
4.22 Relação entre TEO e número de operações dos terminais do Porto de Santos	p. 90
4.23 Número de operações registradas por tipo de carga	p. 91
4.24 Taxa de Ocupação (TO) por tipo de carga no Porto de Santos	p. 93
4.25 Relação entre TP e o tipo de carga movimentado para os dois métodos de cálculo	p. 94
4.26 Relação entre TEO e número de operações dos terminais do Porto de Santos	p. 95

LISTA DE TABELAS

3.1	Campos da tabela de informações dinâmicas recebidas via AIS	p. 44
3.2	Campos da tabela de informações estáticas recebidas via AIS	p. 44
3.3	Vértices da área filtrada para análise	p. 46
3.4	Estrutura das Tabelas de Trajetória	p. 58
3.5	Estrutura das Tabelas de Trajetória	p. 59
4.1	Resultados das atracações por tipo a partir do rotulamento inicial . . .	p. 70
4.2	Métricas de qualidade para o rotulamento inicial	p. 71
4.3	Resultados das atracações por tipo a partir do rotulamento após os tra- tamentos	p. 80
4.4	Métricas de qualidade para o rotulamento após os tratamentos	p. 81

1 INTRODUÇÃO

1.1 Portos

Bichou e Gray (2004) definem um porto como uma organização multifuncional que desempenha um papel relevante para a economia global ao intermediar os modais terrestres com os marítimos e fluviais, servindo como ponto de entrada e saída para rotas regionais e internacionais de comércio. As aplicações e contribuições do porto vão além de sua estrutura física; ele é um sistema dinâmico e estratégico, integrando aspectos econômicos, logísticos, sociais e ambientais, e desempenhando um papel vital nas cadeias de suprimentos globais e no desenvolvimento socioeconômico de regiões inteiras. Os portos podem ser de diferentes tipos, incluindo portos de carga, passageiros, pesqueiros e militares, sendo a maior parte um misto destes.

Atualmente, cerca de 80% do volume do comércio global de mercadorias é transportado por via marítima. Esse percentual é ainda mais alto em países em desenvolvimento, onde o transporte marítimo desempenha um papel crucial para a economia devido à dependência de importações e exportações de larga escala (UNCTAD, 2024), afirmando definitivamente a importância desse modal de transporte, e consequentemente das estruturas portuárias, para a economia mundial.

Os desafios logísticos enfrentados pelos portos são complexos e variados, refletindo a importância dessas instituições. O *scheduling* eficiente das embarcações para evitar congestionamentos e atrasos, a gestão de grandes volumes de carga e a integração de diferentes modos de transporte são alguns exemplos de tarefas que exigem coordenação precisa e uso intensivo de tecnologia. No entanto, muitos portos ainda enfrentam dificuldades significativas para se modernizarem e melhorarem sua eficiência operacional.

Neste contexto, o uso de dados tem ganhado cada vez mais relevância em diversos ramos, impulsionado pelo aumento exponencial da quantidade de dados disponíveis devido aos avanços tecnológicos. As organizações que conseguem coletar, analisar e utilizar esses dados para tomar decisões mais inteligentes e rápidas estão se destacando

no mercado, melhorando sua eficiência operacional, reduzindo custos, e aumentando a produtividade. A indústria portuária não é exceção, e o uso intensivo de dados neste ramo tem o potencial de transformar a eficiência operacional, aprimorando áreas como movimentação de cargas, gestão de tráfego e de estoques, monitoramento da manutenção e da emissão de poluentes.

Dessa forma, o presente trabalho busca utilizar técnicas de análise de dados aplicadas à informações coletadas por meio do Sistema de Identificação Automático (AIS) com o objetivo de criar recursos que possibilitem a compreensão do estado atual e a proposição de melhorias da eficiência operacional dos terminais do Porto de Santos.

1.2 Complexo portuário de Santos

1.2.1 Contexto histórico

O início das atividades no Porto de Santos remonta ao século XVI (Autoridade Portuária de Santos, 2023), quando operações portuárias rudimentares de carga e descarga foram iniciadas na região. Essas operações, em princípio conduzidas de forma independente, adquiriram relevância devido à posição estratégica do Porto para a exportação da produção cafeeira, o que culminou em sua privatização no final do século XIX. Por meio de uma concessão a um grupo posteriormente denominado Companhia Docas de Santos (CDS), o Porto foi desenvolvido e reconhecido como o primeiro porto organizado do Brasil.

Ao longo das décadas subsequentes, o Porto de Santos passou por uma série de modernizações significativas, expandindo sua infraestrutura para incluir terminais especializados em contêineres, grãos, líquidos, papel, celulose e outros produtos. Foram estabelecidas também diversas malhas rodoviárias, ferroviárias e dutoviárias para facilitar o transporte de cargas entre o porto e os principais pontos de produção e distribuição em todo o país.

Na década de 1980, após o término da concessão à Companhia Docas de Santos, a administração do porto foi novamente transferida para o setor público, sob a gestão da antiga Companhia Docas do Estado de São Paulo (Codesp). Esta entidade manteve o monopólio das operações até a promulgação da Lei 8.630/93, quando o porto começou a operar sob o regime conhecido como Landlord Port. Atualmente, a Autoridade Portuária de Santos S.A. (APS), anteriormente conhecida como Codesp, atua como Autoridade Portuária, atuando apenas no campo das funções administrativas, enquanto as operações são realizadas por terminais portuários privados.

Figura 1.1: Vista aérea do canal do Porto de Santos



Fonte: Autoridade Portuária de Santos (2024)

No presente momento, o Porto de Santos é considerado o maior porto da América Latina, tendo movimentado cerca de 173 milhões de toneladas de carga no ano de 2022, e ainda passa por diversas melhorias em sua infraestrutura e operações, que se refletem no volume crescente de movimentação de carga. Dessa forma, é considerado um elemento de grande relevância para a economia brasileira como ponto de comércio nacional e internacional e de desenvolvimento regional.

1.2.2 Organização Administrativa

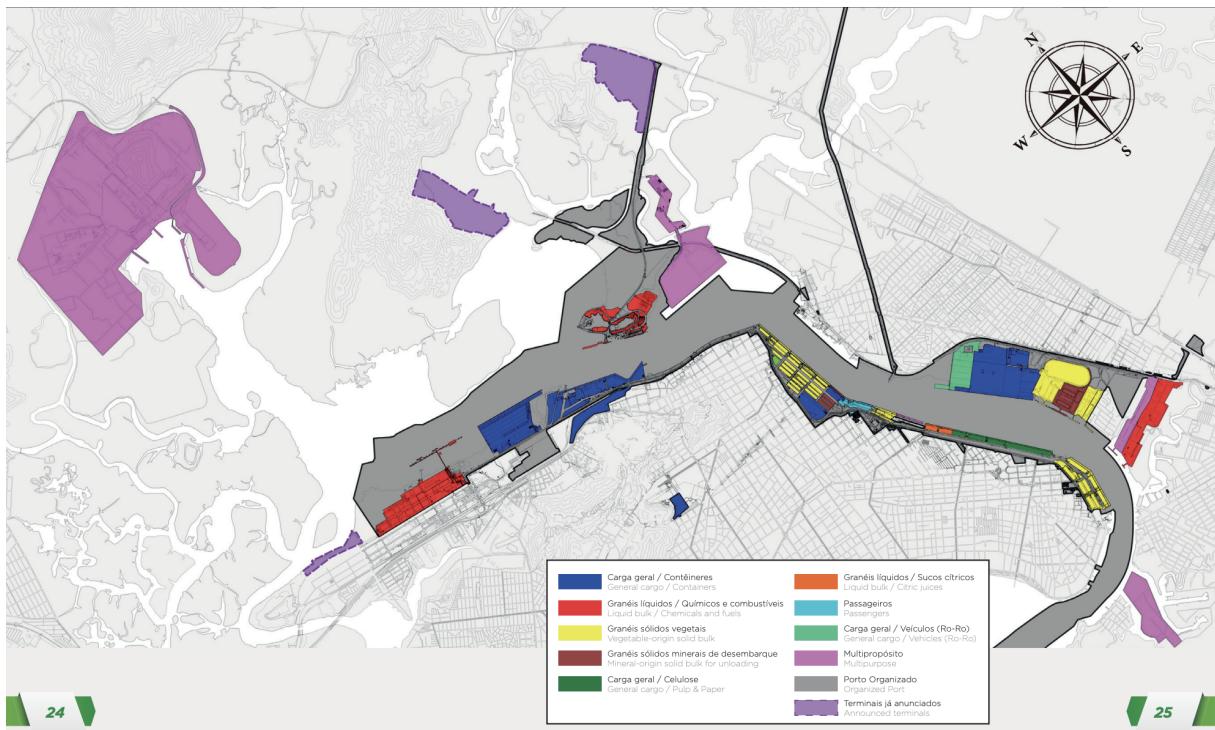
No complexo portuário de Santos, a exploração segue regras baseadas na localização, dividida em duas áreas distintas: dentro ou fora do Porto Organizado. Dentro do Porto Organizado, os terminais estão sob a jurisdição da Autoridade Portuária de Santos (APS), e o poder público realiza contratos com prazo de ocupação para conceder terminais ao setor privado. Já fora desta área, o método de exploração da atividade portuária é por meio de autorizações de uso, sendo estes terminais nomeados Terminais de Uso Privado (TUPs).

Há ainda uma terceira categoria de terminais, os Terminais Retropórtuários, cuja principal função é realizar o desembaraço aduaneiro, operação que permite a finalização fiscal das operações de importação dos demais terminais. Além disso, estas entidades também executam serviços operacionais, administrativos e aduaneiros para garantir o

funcionamento logístico do complexo.

O Porto atualmente conta com um total de 53 terminais operando, sendo 39 terminais dentro do porto organizado, 6 TUPs e 8 Terminais Retroportuários. A figura 1.2 apresenta a relação entre os tipos de cargas movimentadas pelo Porto de Santos e seus respectivos terminais do porto organizado.

Figura 1.2: Terminais do Porto de Santos



Fonte: Autoridade Portuária de Santos (2023)

1.2.3 Relevância Econômica

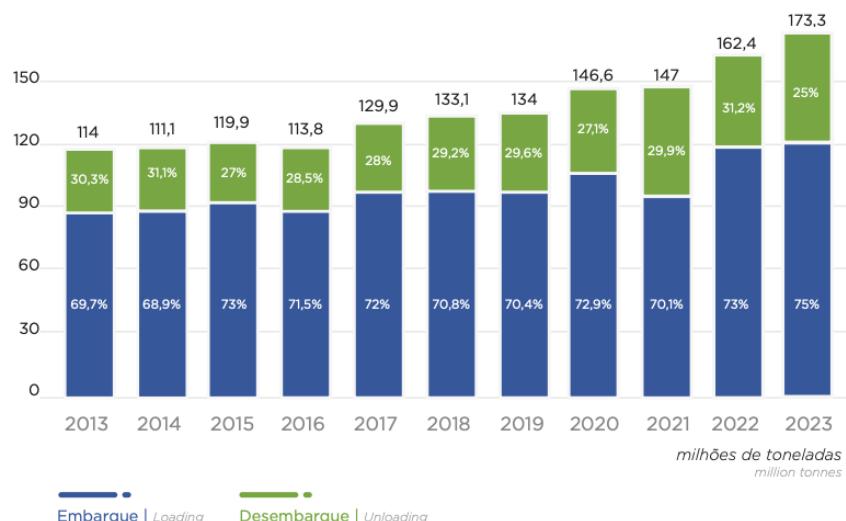
Segundo os relatórios da Autoridade Portuária para o ano de 2023, o Porto de Santos foi responsável por 28% da corrente comercial comercial do Brasil, em um valor estimado de 607 Bilhões de dólares. Isto se deve, além da importância histórica da instituição, à sua localização estratégica e às extensas malhas rodoviária e ferroviária, que permitem que o Porto tenha polos econômicos com fácil acesso às suas instalações, como o estado de São Paulo, o maior polo econômico, e do Mato Grosso, Mato Grosso do Sul e Goiás, grandes produtores agrícolas.

Com movimentação anual de contêineres de 4,9 milhões de TEU (*Twenty-Foot Equivalent Unit*), o Porto de Santos é o porto mais relevante do Brasil nesse quesito, o segundo maior da América Latina e o 41º mundialmente. É também o porto que mais exporta

granéis sólidos no país, com mais de 70 milhões de toneladas de grãos e 11 milhões de toneladas de minerais transportados anualmente, impactando diretamente a balança comercial do país e promovendo o crescimento de setores agrícolas e industriais.

A posição estratégica do Porto de Santos também facilita o acesso de produtos importados ao mercado brasileiro, com destaque para itens industriais como peças automotivas, compostos agrícolas, componentes eletrônicos e medicamentos. Dessa forma, o porto não apenas fortalece as exportações brasileiras, mas também desempenha um papel fundamental na distribuição de produtos essenciais para a economia nacional.

Figura 1.3: Crescimento da tonelagem de carga movimentada no Porto de Santos entre 2013 e 2023



Fonte: Autoridade Portuária de Santos (2023)

Além da movimentação de cargas, o Porto de Santos também tem uma importância social na geração de empregos locais. A prefeitura municipal de Santos estima que suas atividades geram mais de 50 mil empregos associados, em áreas como transporte, logística e comércio exterior, sendo assim de grande relevância para a economia da cidade e seus arredores.

1.3 Motivação

Apesar da importância do Porto de Santos para a economia brasileira, este complexo sistema ainda é carente de tecnologias já difundidas no mercado e amplamente empregadas por organizações semelhantes. Uma tecnologia, que foi explorada neste estudo, consiste no uso de dados para monitorar as operações e analisar possíveis melhorias. Muitos dados que

poderiam contribuir para uma gestão mais eficiente não são coletados, são descartados, ou, embora coletados, são extraídos e registrados por meio de processos manuais, aumentando a propensão a erros humanos. Embora demande grandes esforços pontuais, a implantação do uso de dados de forma automatizada permitiria que as operações do Porto atinjam um nível de produtividade de excelência quando comparado com outros portos das mesmas dimensões.

Em um cenário global cada vez mais impulsionado por tecnologias como *IoT*, ciência de dados e inteligência artificial, a coleta e análise de dados como operações rotineiras faz-se imprescindível na indústria portuária e tem se tornado cada vez mais acessível. A falta de integração de dados coloca o Porto de Santos em desvantagem competitiva devido a operações ineficientes em um momento em que outros portos já estão avançando no caminho da digitalização e automação.

1.4 Objetivo

Este trabalho tem como objetivo propor um método de automação do monitoramento das operações marítimas do Porto de Santos baseando-se na coleta e tratamento de dados gerados pelo sistema AIS (*Automatic Identification System*), que são disponibilizados publicamente. Embora o foco deste estudo seja a utilização dos dados para extração de trajetórias e cálculo de métricas de eficiência e utilização dos terminais, as aplicações dos dados do sistema AIS são mais amplas, como a prevenção de acidentes marítimos, rastreamento de cargas em tempo real, e o sequenciamento de navios em filas de espera. A exploração dessas informações, é um passo fundamental para modernizar as operações e garantir a competitividade do Porto de Santos no cenário global.

1.5 Estrutura do Trabalho

Para atingir o objetivo proposto, este trabalho está organizado da seguinte forma:

Inicialmente, no capítulo 1, é feita uma contextualização para introduzir o leitor ao tema, oferecendo uma visão geral sobre o Porto de Santos, incluindo sua história e relevância para a economia brasileira. Ainda neste capítulo, são apresentadas a problemática que será abordada ao longo do estudo, as motivações que justificam a pesquisa e os objetivos que orientam a condução dos trabalhos.

No capítulo 2, é feito um levantamento dos estudos já realizados que abordam tópicos

relevantes para a pesquisa que servem de embasamento teórico e fundamenta a metodologia desenvolvida para o tratamento dos dados. São discutidas as abordagens que sustentam os métodos aplicados, garantindo uma base sólida para as etapas posteriores.

No capítulo 3, a metodologia adotada é descrita em detalhe, desde a coleta e o pré-processamento dos dados até as etapas de transformação dos dados brutos em informações estruturadas e confiáveis para análise. A metodologia é dividida em duas seções principais: a primeira apresenta o método de rotulação, que categoriza os dados coletados com rótulos de posição a partir de pontos de referência dentro do Porto e a segunda seção detalha o método de reconstrução de trajetórias das embarcações, que visa aprimorar o processo de rotulagem e permitir a identificação precisa de eventos portuários. Ambos os métodos consistem em um conjunto de operações fundamentadas em premissas bem definidas.

No capítulo 4, são apresentados e analisados os resultados obtidos com a aplicação desses métodos. Inicialmente, mostra-se o rotulamento inicial e uma avaliação de sua qualidade, destacando os principais problemas e limitações encontrados. Em seguida, discute-se a qualidade dos rótulos após as melhorias proporcionadas pelo método de reconstrução de trajetórias, comparando esses resultados com os parciais obtidos anteriormente. Além disso, são exibidos os resultados específicos do processo de reconstrução de trajetórias.

Por fim, no capítulo 5, é levantada a relevância dos resultados obtidos e suas possíveis aplicações no contexto do Porto de Santos, e apresentam-se os próximos passos que poderiam aprimorar e aprofundar os achados deste trabalho.

2 REVISÃO BIBLIOGRÁFICA

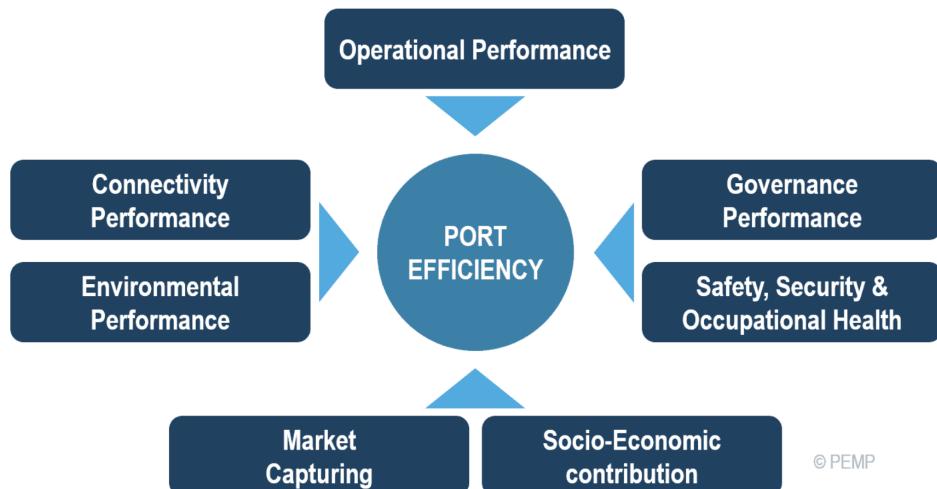
2.1 Eficiência de portos

Avaliar a eficiência de um porto é um fator crítico para o bom funcionamento destas instituições. Portos eficientes são capazes de maximizar o fluxo de bens e passageiros por suas instalações, além de diminuir os custos logísticos de funcionamento (DALANDI et al., 2024). Por meio deste acompanhamento é possível assegurar a conformidade com regulamentações nacionais e internacionais, identificar ineficiências e implantar medidas de melhoria, garantindo a competitividade do porto frente ao cada vez mais competitivo cenário global.

De acordo com Chung (2013), duas medidas primárias para medir a eficiência de um porto são o tempo de circulação de navio (*ship turn around time*), que é o tempo médio entre a entrada de um navio no porto e sua saída, e a carga total operada por dia por navio no porto. Utilizar somente estas duas métricas porém, ignora outras dimensões das operações portuárias que devem ser consideradas, que incluem o número de embarcações, os tipos de embarcação, as instalações disponíveis, a capacidade dos navios que operaram, as dimensões do porto, entre outros, sendo necessárias medidas mais complexas para medir a real eficiência do porto.

A lista de indicadores utilizados para avaliar a eficiência dos portos e terminais está em constante expansão, tornando a mensuração do desempenho portuário cada vez mais complexo (NOTTEBOOM; PALLIS; RODRIGUE, 2022). Além das métricas tradicionais, a avaliação de eficiência agora abrange fatores financeiros, operacionais, produtivos e de marketing, comparando o desempenho atual com os anos anteriores e com o de correntes. O desempenho financeiro das entidades portuárias, incluindo operadores de terminais e órgãos gestores, juntamente com o modelo de governança adotado, são aspectos fundamentais analisados em todos os portos. Com o crescente papel dos portos nas cadeias de suprimentos e os desafios relacionados à sustentabilidade ambiental, novos indicadores foram incorporados, como a conectividade (avaliando capacidade, custos e

Figura 2.1: Dimensões da eficiência portuária



Fonte: Notteboom, Pallis e Rodrigue (2022)

congestionamento), o desempenho ambiental, além de métricas de segurança e saúde no trabalho. Adicionalmente, o sucesso do porto em termos de captação de mercado e sua contribuição socioeconômica são aspectos cruciais que ressaltam a importância estratégica de sua operação.

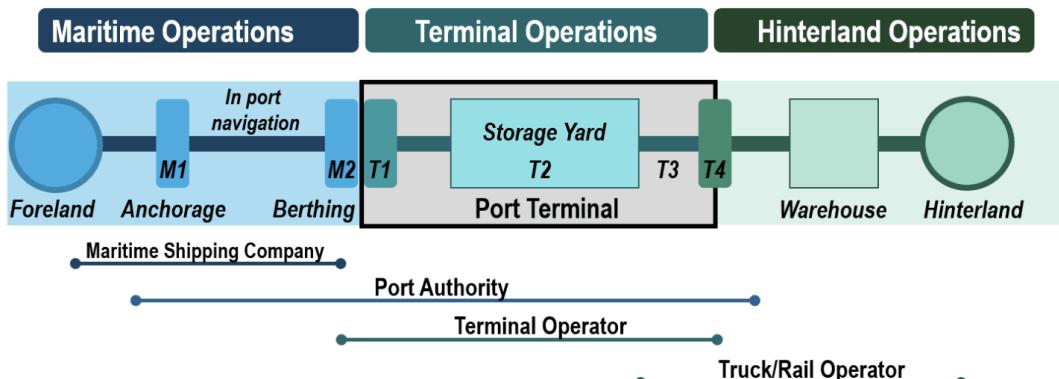
2.1.1 Eficiência operacional

Tratando-se mais especificamente da eficiência operacional, Notteboom, Pallis e Rodrigue (2022) dividem os processos do porto em uma sequência lógica composta por três principais dimensões: operações marítimas, operações no terminal e operações internas.

As operações marítimas são aquelas que ocorrem até a atracação, como a navegação na orla, a ancoragem e atracação no terminal em si, e são de responsabilidade primordialmente da companhia marítima da embarcação, mas também da autoridade portuária e da operadora do terminal. As operações no terminal envolvem principalmente as operações de carga e descarga, sendo as principais partes a autoridade portuária e a operadora do terminal. Por fim, as operações internas compreendem as atividades realizadas já completamente em terra, como a estocagem das cargas e a distribuição dos bens transportados, e são de responsabilidade das empresas ferro e rodoviárias.

Dentre as medidas relacionadas à eficiência operacional das operações nos terminais, Chung (2013) divide o tempo de circulação das embarcações no porto em dois momentos, aquele em que o navio esteve atracado em um terminal e o tempo fora de terminais

Figura 2.2: Dimensões da eficiência operacional: uma sequência



Fonte: Notteboom, Pallis e Rodrigue (2022)

(navegando), e destaca a proporção entre o tempo esperando para atracar e o tempo gasto no terminal como uma importante medida de congestionamento.

No estudo presente, foram exploradas três métricas com mais profundidade para medir a eficiência operacional dos terminais do Porto de Santos, o Tempo de Permanência Médio (TP), a Taxa de Ocupação (TO), e o Tempo Médio entre as Operações (TEO). Cada medida está detalhada a seguir, juntamente à justificativa de sua relevância.

2.1.2 Tempo de Permanência Médio (TP)

O **Tempo de Permanência Médio** ou *Average vessel time at berth* é uma métrica que avalia o tempo médio que uma unidade de carga ou um navio permanece no terminal. Essa métrica é crucial, pois tempos de permanência elevados podem indicar ineficiências operacionais, congestionamento e custos adicionais para os operadores portuários e clientes.

Diversos fatores influenciam no tempo de permanência de uma embarcação nos terminais, como a quantidade de carga, as características do navio, e os recursos e infraestrutura disponíveis no terminal. Apesar disso, de forma geral, um tempo de permanência reduzido é um reflexo de processos aduaneiros eficientes, manuseio rápido de carga e logística interna eficaz. A redução do tempo de permanência pode aumentar a capacidade do porto de lidar com volumes maiores de carga, além de melhorar a satisfação dos clientes devido à redução de custos e tempos de espera. (CHUNG, 2013)

Matematicamente, o TP é definido por:

$$TP_i = \frac{\sum_{j=1}^n O_{ij}}{n} \quad (2.1)$$

Onde:

- O_{ij} é o tempo de operação da j -ésima operação no terminal i ;
- n é o número total de operações realizadas no período analisado.

Segundo Chung (2013), o tempo total de uma embarcação em um porto (*turnaround time*) é composto pela soma do tempo de espera (TE), o tempo de atracação e desatração (TA), e o tempo de permanência no terminal (equação 2.2)

$$Turnaround = TP + TA + TP \quad (2.2)$$

Enquanto os dois primeiros componentes representam de forma geral uma pequena parte do tempo total, o tempo de permanência no terminal é o principal fator que pode reduzir o *turnaround* (CHUNG, 2013), sendo assim de grande relevância para mensurar o aumento da eficiência.

2.1.3 Tempo Médio Entre as Operações (TEO)

O **Tempo Médio Entre as Operações** mede o tempo médio em que um terminal esteve ocioso entre o término de uma operação e o início da próxima. Essa métrica é complementar ao tempo médio de permanência, podendo ser um indicador para a eficiência do agendamento e da coordenação das operações portuárias.

A fórmula para o TEO é:

$$TEO_i = \frac{t_{\text{análise}} - \sum_{j=1}^n O_{ij}}{n} \quad (2.3)$$

Onde:

- $t_{\text{análise}}$ é o tempo total disponível no período analisado;
- $\sum_{j=1}^n O_{ij}$ é o tempo total de operação no terminal i no período;
- n é o número total de operações realizadas.

Uma medida similar ao TEO é o Tempo médio fora (*Average vessel time outside*), que mede o tempo médio que as embarcações estiveram no porto fora de terminais (navegando). Caso o tempo médio fora seja alto, há um indicativo de gargalos e esperas, que podem ser especialmente problemáticos pois os navios fora de terminais congestionam a circulação no canal (CHUNG, 2013).

2.1.4 Taxa de Ocupação (TO)

A **Taxa de Ocupação** é uma métrica que indica a proporção de tempo em que as instalações portuárias estiveram efetivamente em uso, com uma embarcação realizando operações de carga e descarga, em relação ao tempo total disponível. Essa métrica é fundamental para avaliar a utilização dos recursos portuários e identificar subutilização ou gargalos, sendo de grande importância na decisão de investir em melhorias nos terminais existentes ou de expandir com novos terminais.

A TO é calculada por:

$$TO = \frac{\sum_{j=1}^n O_{ij}}{t_{\text{análise}}} \times 100\% \quad (2.4)$$

Onde:

- $\sum_{j=1}^n O_{ij}$ é o tempo total de operação no terminal i no período;
- $t_{\text{análise}}$ é o tempo total disponível no período analisado.

Embora taxas de ocupação baixas (<50%) possam sugerir subutilização dos recursos e alta ociosidade, taxas muito altas (>70%) podem indicar riscos de congestionamento e tempos de espera prolongados nas áreas de fundeio (MWASENGA, 2012).

O Ministério da Índia ainda sugere que a Taxa de Ocupação dos berços utilize horas como medida de tempo, ao invés de dias. Ao utilizar dias nos quais houveram atracções sobre número total de dias para calcular a TO, mesmo dias de baixa operação com somente uma ou duas atracções curtas, são contabilizados, fazendo que a métrica fique inflacionada sem que necessariamente o terminal esteja sendo utilizado em grande parte do tempo. No caso do porto de Cochin, na Índia, a taxa de ocupação caiu de 100% para somente 16,5% em um dos terminais de navios cargueiros ao mudar a base de cálculo de dias para horas (Comptroller and Auditor General of India, 2009).

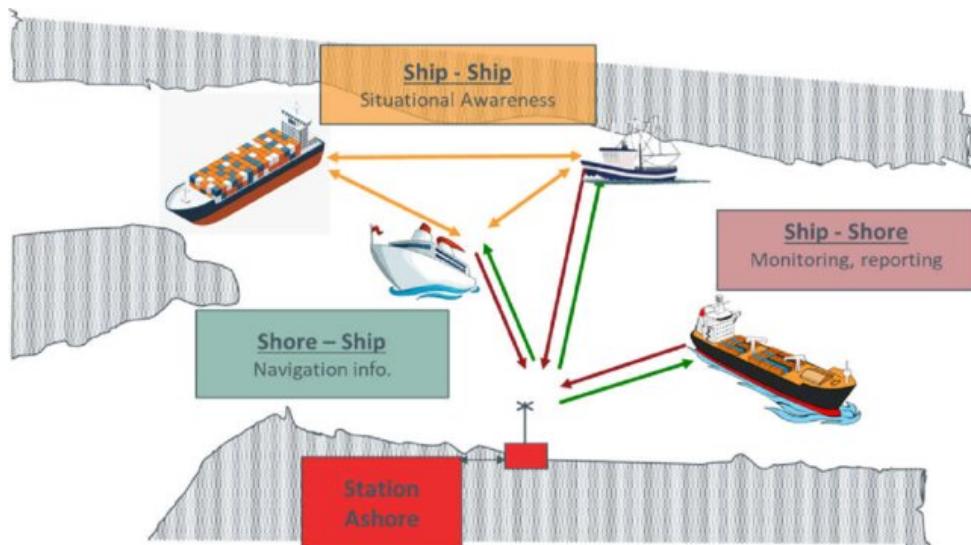
A utilização conjunta das métricas TP, TO e TEO permite uma análise abrangente

do desempenho operacional portuário. Enquanto o TP foca na eficiência individual das operações, a TO avalia a utilização dos recursos e o TEO identifica potenciais ineficiências entre as operações.

2.2 Sistema de Identificação Automática (AIS)

O Sistema de Identificação Automática é a principal ferramenta de rastreamento utilizado globalmente para coletar e enviar dados sobre uma embarcação e sua trajetória, como identificação do navio, porto de destino, tipo de carga, velocidade e coordenadas geográficas (YANG et al., 2019). Nesse sistema, as embarcações funcionam como entidades transmissoras, enviando dados em uma determinada frequência de forma recorrente, que são captadas por uma rede de receptores diversos, composta por outras embarcações, satélites ou receptores terrestres.

Figura 2.3: Funcionamento esquemático das entidades do sistema AIS



Fonte: Johnson et al. (2021)

Os transmissores que enviam dados para o sistema de AIS podem ser divididos em duas categorias: Classe A e Classe B. Os equipamentos de Classe A transmitem informações num intervalo que varia entre 2 e 10 segundos, de acordo com a velocidade da embarcação no momento de envio, ou a cada três minutos, caso o navio esteja ancorado, e incluem 3 grupos de dados: informações estáticas do navio, informações dinâmicas e informações relacionadas à viagem. Os transmissores de Classe B funcionam com intervalos de tempo maiores, diminuindo a frequência de envio, e são mais limitados quanto

aos campos enviados, não incluindo o IMO, o indicativo de chamada (*callsign*), o calado e as informações dinâmicas (YANG et al., 2019).

No momento de sua primeira implementação, há mais de duas décadas, a função original do sistema era permitir que navios visualizassem o tráfego em áreas próximas, evitando acidentes (YANG et al., 2019). Entretanto, atualmente a Organização Internacional Marítima (IMO) impõe que todos os navios com mais de 300 toneladas, e todos os navios de passageiros tenham equipamentos de transmissão de Classe A, tornando o sistema de AIS em um fornecedor de posicionamento em alta frequência e em tempo real para a maior parte da frota comercial do mundo. Tal dimensão possibilitou que seu potencial tenha sido explorado em diversas outras áreas da navegação, como:

- Estimar atrasos e horários de chegada de embarcações;
- Aplicações de segurança, identificando atividades suspeitas;
- Rastreio de cargas e encomendas;
- Análise da eficiência operacional em portos e terminais;
- Relatórios de dados gerais de navegação e extração de estatísticas;
- Investigação post-mortem de acidentes, uma vez que as informações da AIS são mais precisas que de radares;
- Auxílio da navegação, atracação, manobragem e outras operações portuárias;
- Operações de busca e resgate marítimas;

Atualmente existem diversos serviços que realizam a captura das informações das embarcações e distribuem por meio de softwares, bases de dados privadas e APIs.

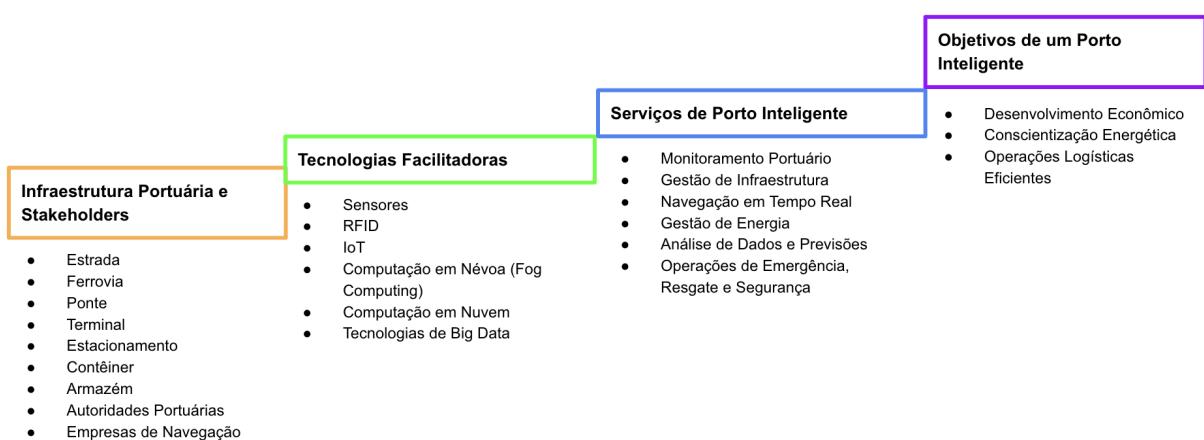
Dentre as aplicações dos dados de AIS, seu uso para gestão portuária tem se tornado um tema relevante sob o ponto de vista de pesquisa. Rajabi et al. (2018) propõe uma arquitetura de porto inteligente utilizando dados de AIS para estudar o tráfego marítimo e diminuir impactos ambientais. A arquitetura é dividida em quatro grandes camadas, que estruturam os blocos fundamentais para um porto inteligente, sendo os níveis mais baixos necessários para alcançar os níveis mais altos. São as camadas propostas:

1. **Infraestrutura portuária e stakeholders:** inclui toda a infraestrutura disponível no porto, como terminais, contêineres, ruas, estacionamentos e galpões de estoque,

e as partes responsáveis por elas, incluindo empresas de navegação e a autoridade portuária;

2. **Tecnologias capacitadoras:** é constituído pelas tecnologias que devem ser aplicadas à camada de infraestrutura para que o porto possa ser considerado inteligente, sendo o uso de *IoT* um recurso chave, com o uso de sensores conectados capazes de coletar, processar e analisar dados em tempo real, bem como comunicar as partes interessadas instantaneamente;
3. **Serviços de Portos inteligentes:** inclui os serviços fornecidos pela aplicação das tecnologias capacitadoras na infraestrutura portuária de forma adequada. Incluem o monitoramento do porto com a centralização de informações de posições de embarcações, ocupação de terminais, nível da água do mar, etc, tornando tais dados facilmente acessíveis aos *stakeholders*. Outros serviços destacados são: o gerenciamento da infraestrutura com o monitoramento do estado de estradas e equipamentos, da temperatura e do vento, navegação em tempo real, garantindo a eficiência do tráfego marítimo e disponibilizando tais informações, o gerenciamento do impacto ambiental através do monitoramento da eficiência energética, o uso de análise de dados e previsões utilizando *big data* para prever trajetórias e acidentes, e, por fim, as operações de emergência, resgate e segurança otimizadas pelo uso de dados em tempo real;
4. **Metas dos Portos inteligentes:** desenvolvimento econômico, uso responsável da energia e eficiência nas operações logísticas. Só são possíveis através da concretização e integração dos níveis anteriores.

Figura 2.4: Níveis que compõe um porto inteligente



Fonte: Rajabi et al. (2018)

Os principais desafios enfrentados por um porto tradicional na implementação da arquitetura proposta de portos inteligentes incluem o alto investimento econômico em equipamentos e treinamento, a necessidade de gestores e especialistas altamente qualificados para coordenar a ampla gama de tecnologias envolvidas em um porto inteligente, a exigência por uma infraestrutura robusta de dados para receber, processar e analisar um grande volume de informações de forma segura, e a necessidade de redesenhar processos já existentes por conta da implementação das novas tecnologias.

No caso do porto de Le Havre, um dos mais relevantes no comércio e transporte de passageiros na França e na Europa, Rajabi et al. (2018) utiliza os dados de AIS coletados e processados automaticamente para categorizar as embarcações que transitaram por Le Havre de duas formas.

Os dados foram coletados ao longo do ano de 2016 por três fontes: uma antena local na Universidade de Le Havre, dados da comunidade do *AISHub*, uma iniciativa aberta que disponibiliza informações globais, e por fim da GPMH (*Grand Port Maritime du Havre*), instituição pública de comércio e indústria que é a responsável pela administração do Porto de Le Havre.

Os dados das três fontes são unidos e então processados utilizando uma biblioteca *open source* capaz de decodificar as mensagens e AIS e é realizada a limpeza dos dados, eliminando erros em mensagens e corrigindo erros humanos. As informações são armazenadas no banco de dados, permitindo o monitoramento das embarcações e das operações portuárias realizadas, e a disponibilização dos dados é feita por meio de uma interface web, que fornece relatórios e análises para as partes interessadas.

A primeira categorização proposta decorrente da coleta de dados de AIS é a classificação baseada na capacidade de contêineres, medida em TEUs, onde foram propostas 5 categorias. O tamanho dos navios porta-container (navios maiores têm maior TEUs) tem impacto direto na administração do porto, pois uma mudança nesse perfil pode levar a atrasos.

A segunda categorização das embarcações baseia-se nos terminais de chegada. Foram considerados quatro terminais, sendo três deles (*France, TNMSC e Ocean*) partes do Porto 2000, e um deles não (*Atlantic*) e analisou-se a distribuição dos navios que atracaram em relação à capacidade, conforme a primeira classificação. Identificou-se que o Porto 2000 recebe a maior quantidade de navios de forma geral. Isto se deve a alguns acordos entre o porto e as companhias que administraram os terminais, que acabam por favorecer alguns berços em detrimento de outros, que têm suas capacidades subutilizadas.

2.3 Rotulamento de status de navegação

Martincic et al. (2020) apresenta três métodos para rotular dados históricos de AIS do porto de Piraeus (Grécia) com um status de navegação. Os rótulos de status possíveis são: navegando, ancorado ou atracado, e os dados foram coletados durante mais de um ano pela plataforma *AISHub*. A eficácia dos métodos é validada com os dados coletados manualmente pela autoridade do Porto e os dados validados são utilizados para fornecer um processo de extração de viagens de embarcações e tempo de gasto em uma posição. A partir das viagens identificadas, foram avaliados as atrações realizadas no porto por tipo de embarcação.

O primeiro método se baseia na localização e velocidade do navio para identificar seu status de navegação. Navios que estejam parados dentro de um determinado polígono definido manualmente recebem o status de ancorado ou atracado, e, por exclusão, os que se encontram fora desta área ou com velocidade não nula são rotulados como navegando. O problema desta classificação é que no caso em que a embarcação ancora ou atraca próximo à borda do polígono, o status de navegação ficaria oscilando, mostrando múltiplas ancoragens que na realidade não ocorreram.

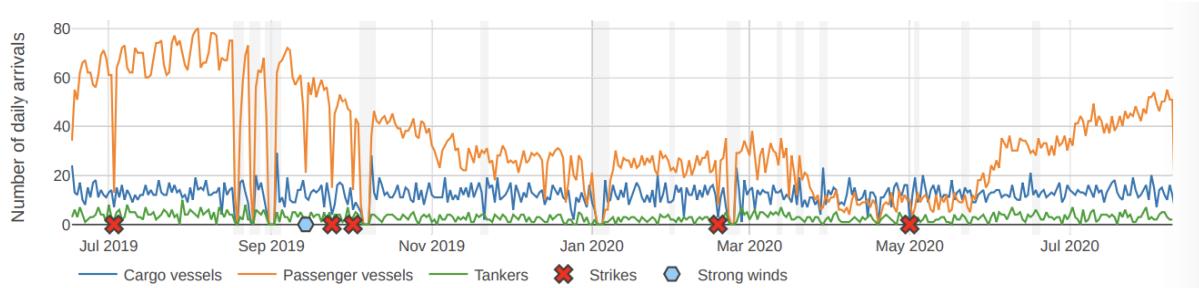
O segundo método utiliza as velocidades linear e angular do navio para o rotulamento. Navios navegando apresentam velocidade linear, mas embarcações ancoradas e atracadas não tem o mesmo comportamento. Além disso, navios ancorados apresentam velocidade angular, pois rotacionam devido a ventos e correntes marítimas, enquanto navios atracados não apresentam ambas. Esse método se mostrou eficiente para paradas longas, mas para ancoragens curtas, onde o navio não rotacionou o suficiente, pode não ser confiável.

Na terceira abordagem foram apresentadas técnicas de *machine learning* para classificar os dados dinâmicos. Foram aplicados métodos de *clustering* não supervisionado (HDBSCAN) e algoritmos de classificação supervisionada (*CatBoost* e KNN). O melhor resultado foi obtido com a aplicação do *k-nearest neighbours* (KNN) utilizando um alto número de vizinhos para classificar os status.

Com os dados obtidos da classificação, foram extraídas as viagens dos navios baseada na identificação do navio e no tempo entre duas mensagens consecutivas de AIS. Para uma mesma embarcação, se o tempo for maior que 24 horas ou o intervalo entre duas mensagens for maior que 5 horas e o navio se moveu mais de 100 metros durante esse período, divide-se os dados em duas viagens diferentes.

A partir dos resultados, obteve-se as entradas e saídas do Porto de Piraeus, e foi pos-

Figura 2.5: Atrações por tipo de navio em Piraeus



Fonte: Martincic et al. (2020)

sível comparar a eficácia do método comparado com o boletim fornecido pela autoridade portuária. Foram calculados os números de atrações diárias, agregados por tipo de navio (cargueiros, tanques e navios de turismo) e o erro médio no agregado total foi de 4,46.

Na figura 2.5, é representado o número de atrações diárias segmentadas por tipo de navio (navios cargueiros, navios de passageiros e navios petroleiros), além de eventos de greve dos funcionários ("strikes") e de ventos fortes, para relacioná-los com os comportamentos observados. Observa-se que número de atrações para todos os tipos de navios apresentaram alta variância sem clara tendência clara durante o período observado. Além disso, através da inspeção visual dos dados, nota-se que os eventos de clima extremo e as greves impacto têm relação direta nas atrações, causando quedas bruscas nos números registrados. No caso dos navios de passageiros, também é possível visualizar uma diminuição das operações causada pela pandemia da COVID-19 entre Março e Maio de 2020, enquanto os navios cargueiros e tanques se mantiveram estáveis para o mesmo período.

2.4 Extração de trajetórias de embarcações

Liye, Qiang e Tien (2019) destacam o sistema de AIS como uma fonte de dados extremamente poderosa e confiável, que deve ser utilizada como ferramenta para extrair informações sobre o tráfego em áreas portuárias de forma exata, em comparação com ferramentas mais antigas, como observações visuais humanas, uso de radares e fotografias aéreas que consomem muito mais recursos humanos e financeiros.

Neste artigo, os autores propõe uma metodologia para analisar 182 milhões de registros coletados no porto de Cingapura. São apresentados dois principais módulos: um método de extração de trajetórias das embarcações a partir da qual realiza-se uma análise sobre

origem e destino (OD), e uma análise de estado tempo-espacial do tráfego, que consiste em um modelo para representar a variação do estado do tráfego de embarcações no porto.

Primeiramente é realizada a coleta e pré-processamento dos dados, com o objetivo de tornar os dados mais confiáveis, reduzindo o ruído, e preenchendo dados omissos com interpolação linear. É também aplicado um processamento utilizando *MapReduce* em um cluster *Hadoop* para lidar com o grande volume de dados.

Para a análise de demanda, a princípio apresenta-se um algoritmo para caracterizar a trajetória de uma embarcação de uma Zona de Tráfego Marítima (MTAZ) para outra. Uma trajetória (J_i) inclui dois tipos de atividades: as de viagem (T), que incluem a navegação, e as de recebimento (S) que incluem a carga e descarga de bens, suprimento de água e alimentos, etc. As definições de S e T são:

$$T_i = \{J_i(t_1^j, t_2^j) \mid j = 1, 2, \dots, n_T^i\} \quad (2.5)$$

$$S_i = \{J_i(t_1^j, t_2^j) \mid j = 1, 2, \dots, n_S^i\} \quad (2.6)$$

Onde n_T^i é o número de viagens realizadas pelo navio i , n_S^i é o número de recebimentos do mesmo navio, j é o index da viagem; $J_i(t_1^j, t_2^j)$ é a trajetória de número j do navio i realizada entre t_1 e t_2 . Assim, as atividades de uma embarcação podem ser definidas como:

$$A = \{A_i \mid A_i = T_i \cup S_i, i = 1, 2, \dots, n_a\} \quad (2.7)$$

Dado um navio i , para separar as trajetórias do conjunto J_i entre T_i e S_i , apresenta-se um algoritmo de extração de viagens. Caso o navio esteja em águas navegáveis, considera-se que a trajetória é uma atividade de navegação. Caso contrário, é analisada a condição 2.8, visto que uma embarcação também pode estar navegando em zonas de atracação com o objetivo de transitar para seu destino final em outra zona.

$$\text{mean_speed}(J(t_{\text{service_start}}, t_{\text{service_start}} + \tau_Z)) < \mu_\theta \quad (2.8)$$

Caso a velocidade média da trajetória definida entre o instante do início da trajetória e o instante do início da trajetória mais uma variável τ_Z for menor que μ_θ , considera-se que houve uma atividade de recebimento. O valor de μ_θ foi estabelecido em 1,5 nós e

representa o limite superior da velocidade média de navios que não estavam navegando (estabelecida em 1.5 nós). A variável τ_Z representa o limite inferior de tempo durante o qual os navios ficam em atividades de recebimento e foi estabelecida em 15 minutos.

As atividades T calculadas então são utilizadas no cálculo da Matriz OD, que utiliza as viagens T reconhecidas anteriormente e as distâncias entre todas as MTAZs para resultar em todas as rotas ligando pares de zonas de tráfego. O número de rotas entre as áreas de tráfego é o valor na matriz OD. O resultado deste método está representado na figura, onde os pares origem-destino entre 29 MTAZs foram traçados.

Figura 2.6: Rotas da Matriz OD para o Porto de Cingapura



Fonte: Liye, Qiang e Tien (2019)

Por meio desta análise foi possível identificar pares com maior demanda, como entre as MTAZs 17, 16, 12, 13, 0 e 9, pois são terminais ou locais de grandes ancoragens. Além disso, identificou-se que a demanda em 48,75%, quase metade das duplas, é zero, indicando que não houve viagens realizadas entre estes locais, e que a máxima demanda entre dois MTAZs é de 40. Durante o mês analisado, o comportamento da matriz OD se manteve estável ao longo dos dias.

Outra análise possibilitada pela identificação de viagens utilizando o método proposto é relacionada às rotas dos navios. Identificou-se que há uma diferença visível entre as rotas dos dois tipos de navios analisados. No caso dos navios de grande porte, foi possível visualizar rotas bem estabelecidas que variam pouco ao longo dos dias, onde uma grande proporção de navios de granel visitam as zonas de ancoragem 15, 17 e 28, enquanto navios de contêineres não visitam tais áreas. Já os navios rebocadores de pequeno porte

apresentam um comportamento muito mais complexo, navegando por todas as áreas e realizando diversas viagens.

Para a análise tempo-espacial, foi criado um modelo de autocorrelação onde o índice I de Moran global e o índice I de Moran local, são usados para descrever a velocidade dos navios e a densidade espacial do tráfego marítimo.

O índice global é obtido pela fórmula:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} s_{ij}}{\sigma^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \times (x_i - \bar{x})(x_j - \bar{x})}{\sigma^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \quad (2.9)$$

Onde x_i é o valor da observação i (que pode ser o número de navios ou a velocidade média dos navios em i) e \bar{x} é a média de todas as observações, $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ é a variância, s_{ij} é a diferença entre as observações i e j, w_{ij} é o peso espacial entre i e j (calculado como $\frac{1}{d_{ij}}$ onde d_{ij} é a diferença entre a célula i e a célula j) e n é o número total de observações,.

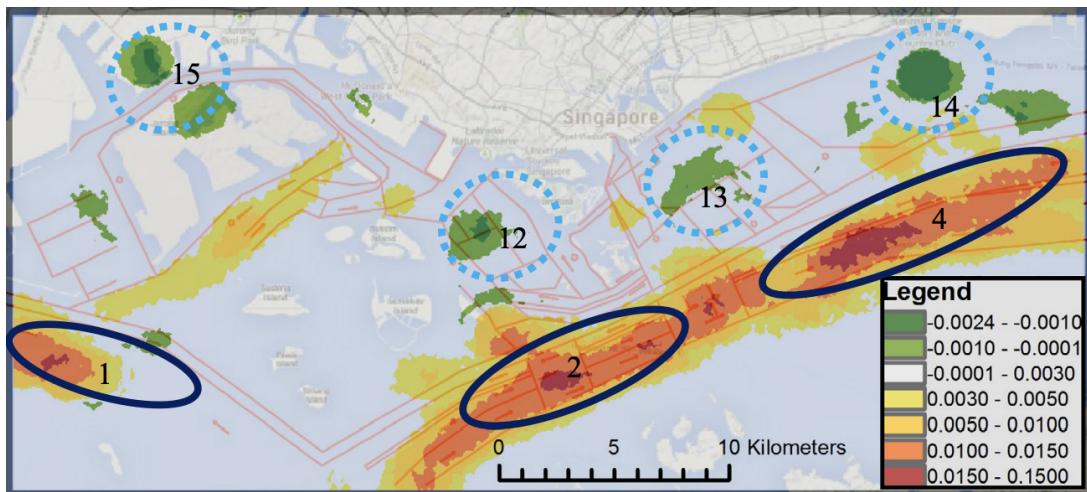
Neste caso, I é uma variável contínua no intervalo de [-1;1], onde valores próximos de 1 e -1 indicam correlações fortes e valores de I próximos de 0 indicam pouca relação entre os valores

O índice local de Moran da i-gésima observação I_i é definido como:

$$I_i = \frac{(n-1)(x_i - \bar{x}) \sum_{j=1}^n w_{ij} \cdot (x_j - \bar{x})}{\sum_{j=1, j \neq i}^n (x_j - \bar{x})^2} \quad (2.10)$$

Diferentemente do índice global, o índice local não precisa estar contido no intervalo [-1;1], mas valores próximos a zero indicam que x_i tem pouca relação com seus vizinhos.

Figura 2.7: Índices de Moran da velocidade das áreas do Porto de Cingapura



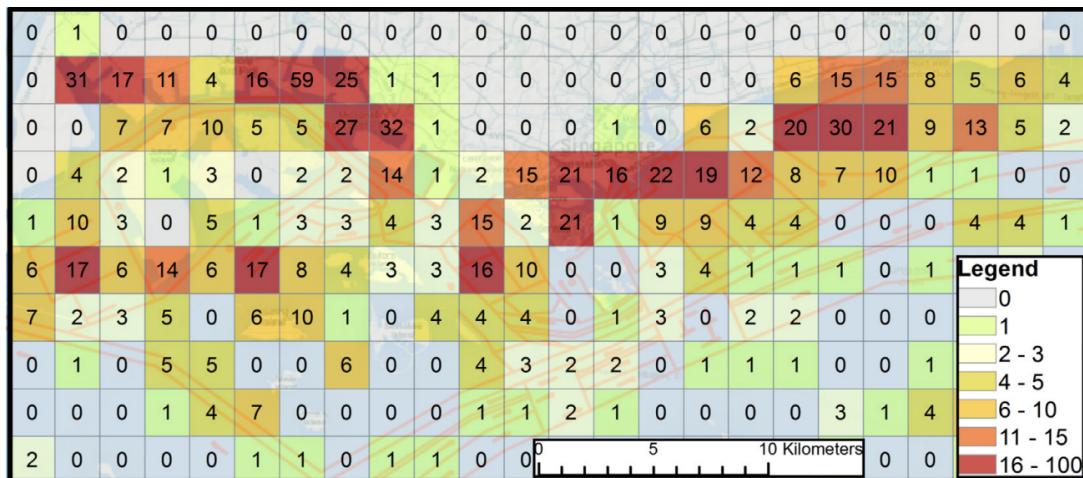
Fonte: Liye, Qiang e Tien (2019)

A área de estudo foi dividida em quadrados de 600 metros de lado, e a velocidade média e o número de navios foi calculada para cada quadrado. Identificou-se que os quadrados com altas velocidades médias se mantêm estáveis em alguns pontos de Porto ao longo do dia e entre dias diferentes.

Utilizando os índices I de Moran locais, foi possível identificar as áreas de fundeadouros e atracação pela alta variação de velocidade, resultando em um baixo índice I, e as áreas do Estreito de Cingapura pela alta e estável velocidade média, resultando em um alto índice I de Moran.

A análise tempo espacial da densidade do tráfego, utilizou uma divisão de quadrados de lado de 2000m, e indicou que a densidade de navios, bem como o índice de Moran local, são mais altos nos pontos de berços e ancoragem, e baixos no Estreito de Cingapura, conforme esperado. Estes padrões também se mantêm estáveis ao longo do dia e ao longo do mês.

Figura 2.8: Índices de Moran de densidade de navios das áreas do Porto de Cingapura



Fonte: Liye, Qiang e Tien (2019)

Já Zhao, Shi e Yang (2018), ao proporem um método de extração de trajetórias baseada em dados de AIS, destaca a importância de ter dados confiáveis e limpos. Para isso, faz um resumo dos principais erros que os dados de AIS apresentam, introduz um método de pré-processamento de dados que tratam destes erros, garantindo uma identificação satisfatória de trajetórias, e exemplifica o processo de preparação de dados em registros de trajetórias de embarcações nas ilhas *Zhoushan*.

Segundo o autor, alguns dos principais problemas que os dados brutos de AIS apresentam incluem:

- A integridade física dos dados, que inclui a confiabilidade da mensagem AIS e a completude do trajeto. A confiabilidade é medida pela coerência da mensagem recebida com os padrões estabelecidos. No caso do estudo de Zhao, foram identificadas mensagens com valores de latitude ou longitude fora dos limites estabelecidos entre 90 e 180 graus (erros nos dados dinâmicos) e registros com erros de tipo de navio e comprimento de navio (erros nos dados estáticos). A completude do trajeto é medida pelo número de registros na trajetória e pela integridade das informações estáticas da embarcação, como nome, tipo e comprimento.
- Integridade lógica espacial, definida pelo sentido lógico da relação tempo espacial dos dados dinâmicos. Há registros com mudança abrupta de posição, que não seria possível dada a diferença de tempo entre as mensagens e a velocidade alcançada por aquela embarcação. Além disso, por motivos diversos, incluindo erros humanos, há registros de um mesmo MMSI, que deveria funcionar como um registro único de

uma embarcação, sendo utilizado por mais de um navio.

- Precisão do horário do registro, pois o rótulo de data e tempo de um registro é realizado utilizando o momento registrado pela estação de recebimento (hora registrada) e o momento de geração da mensagem AIS (hora de geração). Como ambos os dados devem ser ou gerados, ou codificados pelas estações de recebimento, foram observadas inconsistências nas horas registradas de estações diferentes.

No caso dos erros relacionados à integridade física dos dados, sugere-se o descarte dos dados que não sejam compatíveis com os requisitos de projeto e em que as trajetórias tenham completude insuficiente. Embarcações com trajetórias com registros menores que um limite estabelecido (no caso do estudo, 100 registros) foram descartadas, e aquelas cujas informações estáticas relacionadas à identificação e caracterização do navio eram insuficientes ou incoerentes também foram eliminadas.

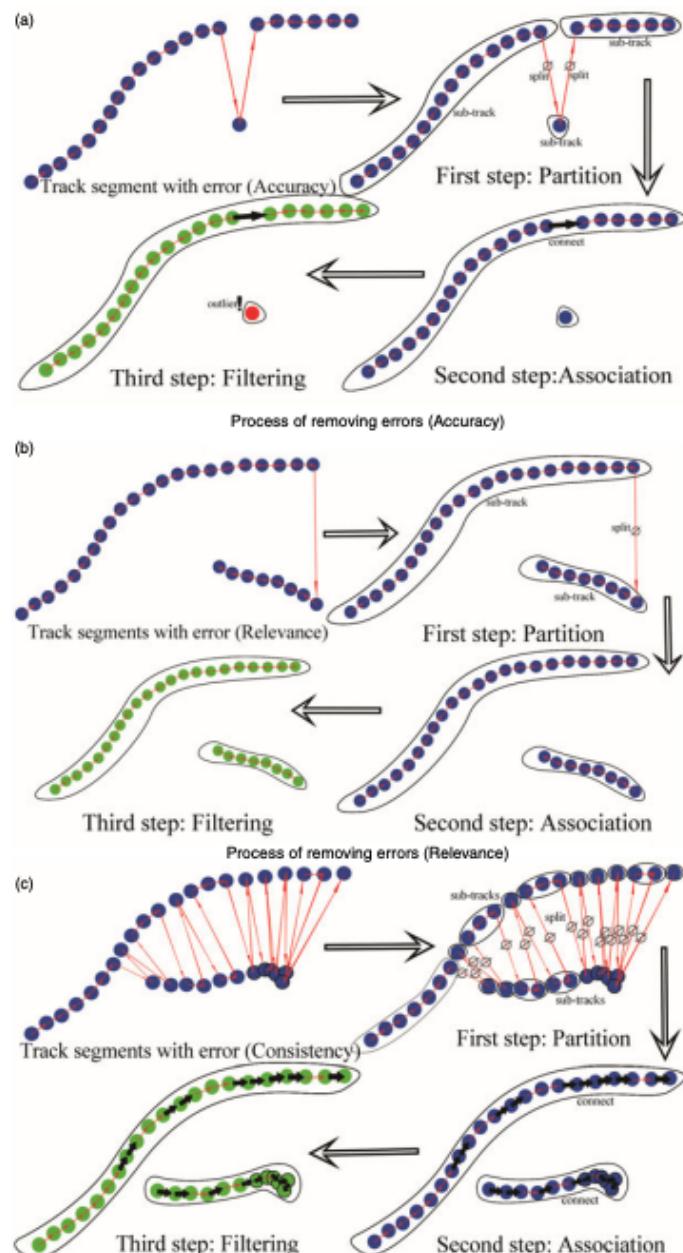
Para os erros relacionados à integridade lógica dos dados, é proposto um algoritmo que envolve três passos: Partição, Associação e Filtragem dos dados para criar trajetórias. Esses passos são aplicados primeiramente para os dados de tempo, a partir do qual são identificadas e eliminadas trajetórias *outliers* (trajetórias que não tem completude dado o limite mínimo de 100 registros por trajetória) e depois o mesmo processo é aplicado aos dados de posição para identificar sub trajetórias de uma mesma embarcação.

No passo da Partição, é calculada a diferença entre o tempo de dois registros. Se a diferença é menor que um limite estabelecido empiricamente em 10 minutos, os registros são considerados como parte de uma mesma sub trajetória. Por outro lado, se a diferença é maior, é criada uma nova sub trajetória (figura 2.9). O mesmo vale para o passo de Partição aplicado à posição: caso a diferença entre a velocidade registrada em um dado e a velocidade calculada para um mesmo registro seja maior que um limite experimental de 15 nós, cria-se uma nova trajetória.

É aplicado então o algoritmo de Associação. Para cada sub trajetória, calcula-se a diferença da mesma forma que no passo da Partição entre seu último ponto e o primeiro ponto de todas as demais sub trajetórias. Se a diferença entre os pontos é menor que o mesmo limite utilizado no passo anterior, as sub trajetórias são associadas.

Realiza-se então o processo de Filtragem, em que as sub trajetórias que não obedecem ao limite inferior de 100 registros por sub trajetória são marcados como *outliers* e são retirados da análise, como visto na figura 2.9

Figura 2.9: Passos para obtenção de trajetórias



(ZHAO; SHI; YANG, 2018)

2.5 Métricas de avaliação de modelos de classificação

Novakovic et al. (2017) define a tarefa de classificar como o processo de, por meio de semelhanças entre as características de objetos, categorizá-los entre classes pré determinadas e conhecidas. O autor reconhece a classificação como uma das principais tarefas de modelos, sejam amparados por inteligência artificial ou não, e discute algumas das principais métricas para mensurar a qualidade de modelos de classificação.

Uma primeira medida baseia-se na noção de erros, ou seja, considera-se que a qualidade de um modelo é tão maior quanto menos erros este comete ao tentar classificar um objeto novo. Esta métrica, conhecida como acurácia, é medida como a razão entre o número de exemplares corretamente classificados e o total de classificações realizadas.

$$\text{Acurácia} = \frac{\text{Classificações corretas}}{\text{Total de classificações}} \quad (2.11)$$

A acurácia é a principal medida utilizada para avaliar modelos de classificação devido à sua simplicidade de cálculo e entendimento e é especialmente interessante em casos de classificação não binária, em que os exemplares podem pertencer a somente uma de $n > 2$ classes. Nestes casos, os resultados podem ser representados em uma matriz quadrada de dimensão n , em que em um eixo temos a classificação conforme o modelo e no outro temos as classes reais às quais os objetos pertencem. A figura 2.11 representa esta estrutura para um modelo de classificação de emoções com $n = 5$ categorias.

Figura 2.10: Representação matricial dos resultados de classificação com 5 categorias

		Actual class				
		happy	sad	angry	gentle	frightened
Predicted class	happy	51	2	1	1	1
	sad	3	23	1	1	0
	angry	2	2	17	0	0
	gentle	0	1	2	9	1
	frightened	1	0	1	1	18

Fonte: Novakovic et al. (2017)

Seguindo esta estrutura, percebe-se que na diagonal principal estão os casos em que o modelo classificou corretamente os exemplares, enquanto todas as outras células representam os casos em que o modelo falhou. Porém não é possível explorar os tipos de erros nestes casos, uma vez que só há dois possíveis resultados: correto ou incorreto.

Nos casos em que a classificação pode ser representada de forma binária, ou seja,

as classificações possíveis podem ser "positivas" ou "negativas" possível calcular outras métricas de qualidade para o modelo de classificação, que consideram o tipo de erro cometido.

Utilizar outras métricas além da acurácia, como a precisão e o *recall*, que serão detalhados nesta seção, é especialmente importante quando a base de dados classificada é desbalanceada, com muito mais exemplares em uma categoria do que em outra. Nestes casos, é comum que a acurácia seja alta, levando a uma boa avaliação do modelo, porém isso não significa que o modelo tem uma boa performance em identificar os casos com poucas ocorrências, uma vez que estes têm pouco impacto no cálculo da acurácia. Nas aplicações em que identificar casos pouco frequentes é essencial, como no diagnóstico de doenças raras, utilizar somente a acurácia poderia levar à aprovação de modelos pouco eficientes para o fim mais importante.

Assim, é interessante transformar a classificação em um problema binário, de forma que os resultados só possam materializar quatro formas, listadas por Novakovic:

- **a (Verdadeiros negativos):** casos em que o modelo classificou corretamente um caso como sendo negativo
- **b (Falsos positivos):** casos em que o modelo classificou incorretamente um caso como sendo positivo
- **c (Falsos negativos):** casos em que o modelo classificou incorretamente um caso como sendo negativo
- **d (Verdadeiros positivos):** casos em que o modelo classificou corretamente um caso como sendo positivo

Figura 2.11: Representação matricial dos resultados de classificação binária

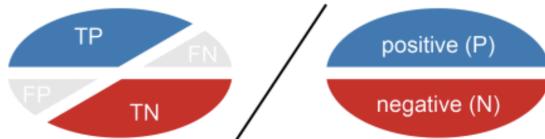
		Predicted class	
		Negatives	Positives
Actual class	Negatives	<i>a</i>	<i>b</i>
	Positives	<i>c</i>	<i>d</i>

Fonte: Vujovic (2021)

Segundo estas definições, é possível descrever a acurácia a partir da equação 2.12. Além disso, é possível calcular a taxa de cada tipo de resultado. Dentre as taxas, destacam-se duas a precisão e a acurácia:

$$Acurácia = \frac{a + d}{a + b + c + d} = \frac{\text{Verdadeiros positivos} + \text{Verdadeiros negativos}}{\text{Total de classificações}} \quad (2.12)$$

Figura 2.12: Representação gráfica da acurácia



Fonte: Vujovic (2021)

2.5.1 Precisão

É a razão entre os casos em que o modelo identificou corretamente um exemplar verdadeiro e todos os casos classificados como verdadeiro pelo modelo. É uma medida de confiabilidade, pois mede quanto confiável é um resultado positivo. No caso do exemplo de diagnóstico médico, a precisão responde à pergunta "Quantos dos diagnósticos positivos realmente eram de pacientes que tinham a doença?".

$$Precisão = \frac{d}{b + d} = \frac{\text{Verdadeiros positivos}}{\text{Falsos positivos} + \text{Verdadeiros positivos}} \quad (2.13)$$

Figura 2.13: Representação gráfica da precisão



Fonte: Vujovic (2021)

2.5.2 *Recall* ou Taxa de Verdadeiro positivos

É a razão entre os casos em que o modelo classificou corretamente um exemplar verdadeiro e todos os exemplares verdadeiros na base de dados. É uma medida de completude, pois mede quantos dos casos verdadeiros estão sendo identificados. No caso do exemplo de diagnóstico médico, o *recall* responde à pergunta "Quantos dos pacientes doentes foram corretamente diagnosticados como portadores da doença?".

$$Recall = \frac{d}{c + d} = \frac{\text{Verdadeiros positivos}}{\text{Falsos Negativos} + \text{Verdadeiros positivos}} \quad (2.14)$$

Figura 2.14: Representação gráfica do *recall*



Fonte: Vujovic (2021)

Assim, a precisão aumenta conforme os Falsos positivos diminuem, enquanto o *recall* melhora com a diminuição dos Falsos negativos. Ainda voltando ao exemplo médico, é possível compreender a importância de medir os tipos de erros, uma vez que um modelo com boa precisão mas baixo *recall* levaria muitos pacientes doentes a não serem diagnosticados como efermos, podendo ter consequências fatais.

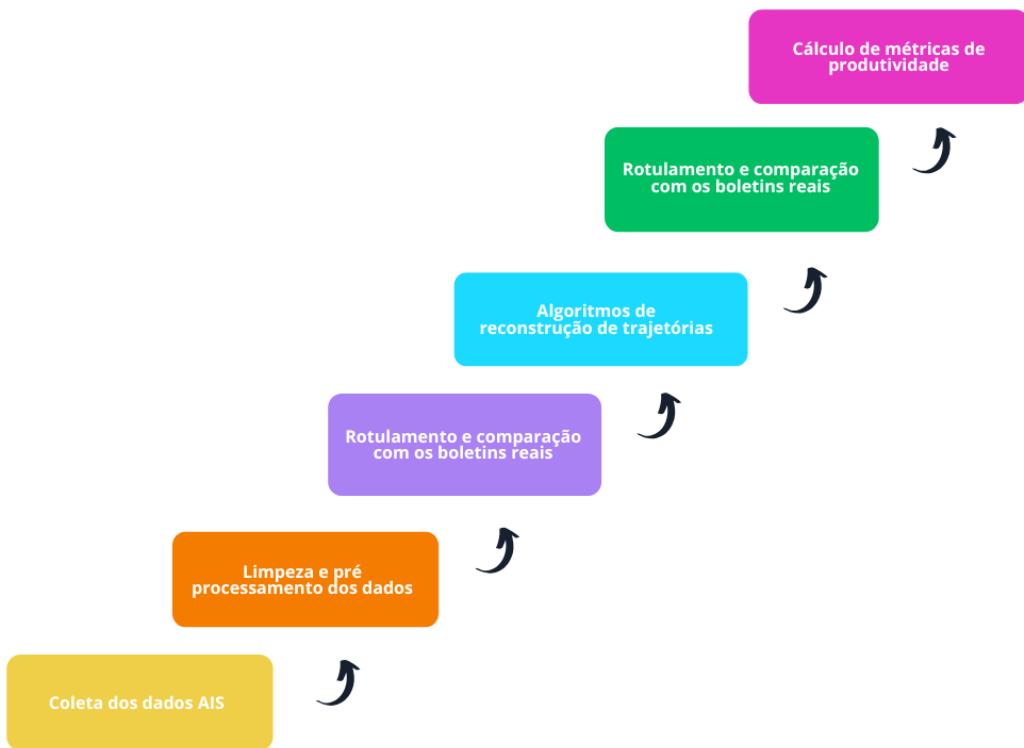
Dessa forma, ao escolher as métricas de qualidade e propor melhorias em modelos de classificação, é necessário considerar os custos, sejam financeiros, humanos, reputacionais, entre outros, que os erros causados irão gerar no contexto de aplicação.

3 METODOLOGIA

Para ilustrar a sequência lógica das operações aplicadas aos dados provenientes do sistema AIS (Sistema de Identificação Automática), foi elaborado o diagrama apresentado na figura 3.1. Neste diagrama, são descritas as etapas da metodologia proposta, a qual visa transformar os dados brutos coletados em uma estrutura organizada, confiável e de fácil utilização, permitindo que sejam empregados de maneira eficiente nas análises pertinentes aos objetivos do estudo.

O processo inicia-se com a coleta abrangente de dados AIS, que reúne informações sobre a posição, velocidade, curso, status de navegação e outras características das embarcações ao longo do tempo. Em seguida, realiza-se a etapa de pré-processamento e limpeza dos dados, que consiste na remoção de registros que apresentem inconsistências, erros ou ruídos. Esses registros, muitas vezes decorrentes de falhas no sinal ou de dados incompletos, devem ser filtrados para evitar que interfiram nos resultados das análises. Além disso, dados considerados irrelevantes para os objetivos do estudo são descartados, garantindo que o conjunto final de dados contenha apenas as informações interessantes.

Figura 3.1: Sequencia das principais operações de tratamentos dos dados



Fonte: Elaboração própria

Após o pré-processamento, é realizada uma primeira rotulação dos dados, onde as informações de posição dos navios são categorizadas em relação à sua localização no canal. A qualidade desses rótulos iniciais é então verificada por meio de uma comparação com registros oficiais fornecidos pela Autoridade Portuária.

Aplicam-se em seguida, três algoritmos de reconhecimento de trajetórias com o intuito de refinar o processo de rotulamento e melhorar a precisão das classificações sobre o comportamento das embarcações utilizando os dados dinâmicos coletados para identificar padrões de movimentação dos navios. Esses algoritmos auxiliam na reconstrução da linha do tempo das embarcações durante seu período de permanência no Porto de Santos. Ao identificar com maior precisão os momentos de atracação e movimentação das embarcações, esses algoritmos permitem uma segunda etapa de rotulação com qualidade superior, onde as atracções podem ser identificadas com um nível de acurácia mais elevado. Dessa forma, obtém-se um conjunto de dados estruturado e confiável para ser utilizado em análises que ofereçam *insights* detalhados sobre o tráfego portuário e as operações no porto.

Com os dados rotulados e organizados, procede-se finalmente ao cálculo das métricas de produtividade dos terminais. Essas métricas permitem uma avaliação do desempenho das operações no Porto de Santos, oferecendo uma visão sobre o fluxo de embarcações, o tempo médio de atracação e a utilização das instalações portuárias. A análise dessas métricas fornece um panorama do funcionamento atual do porto, identificando os pontos com boa eficiência e os possíveis gargalos nas operações.

Essas análises, além de retratarem a situação presente, servem como um ponto de partida para identificar onde há maior necessidade de intervenções que visem à melhoria da eficiência. Com base nos resultados das métricas, é possível destacar as áreas e processos que mais necessitam de ajustes ou otimizações, como a redução de tempos ociosos, a gestão de filas de espera e mudanças na alocação de recursos.

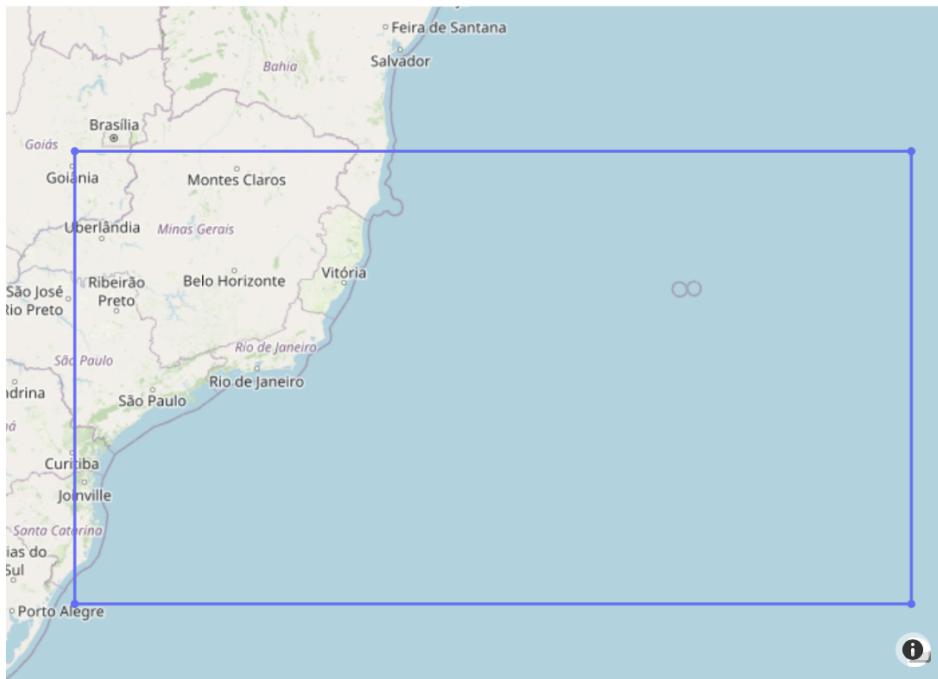
3.1 Coleta de dados

Atualmente há diversos serviços pagos e gratuitos que disponibilizam os dados no sistema de AIS, seja por meio de interfaces gráficas, por APIs, ou por exportação de dados.

Para o presente trabalho, foi utilizada a API da *AIS Stream*, um serviço que fornece dados de AIS via *WebSockets*, permitindo que os usuários utilizem linguagens programação como Python e Java para obter dados de navios ao redor do mundo. O serviço permite que os usuários definam quais tipos de dado querem receber (os tipos incluem dados estáticos do navio, dados de posição, dados de operações de busca e resgate, entre outros). Neste estudo foram recolhidos somente dados dos tipos *PositionReport* e *ShipStaticData*, que contém os dados de posição e dados estáticos do navio, respectivamente.

A *AIS Stream* também permite limitar as coordenadas geográficas dentro das quais deseja-se receber os dados, garantindo que somente a atividade dentro da área de interesse tenha seus dados recebidos, o que evita um volume grande de dados irrelevantes. No código implementado, foi definida inicialmente uma área mais ampla, cobrindo a região litorânea desde o sul de Santa Catarina, até o sul da Bahia. Dessa forma, os dados coletados podem ser utilizados como uma base histórica para outras análises e estudos no futuro com alvo em outras regiões além do Porto de Santos sem a necessidade de esperar meses de monitoramento da atividade.

Figura 3.2: Área total para a qual foi coletada dados AIS



Fonte: Elaboração própria

Por meio de uma implementação em Python, os dados foram coletados diariamente, sendo que as informações para cada dia foram salvas em arquivos csv separados. Juntando estes arquivos, no período de 14 de Maio de 2024 a 18 de Agosto de 2024 foram coletados 16.421.506 dados de posição por meio do sistema AIS utilizando o método apresentado.

Duas principais estruturas de dados foram criadas a partir das estruturas de *PositionReport* e *ShipStaticData* enviadas pela *AIS Stream*: a primeira são as tabelas que contém informações estáticas dos navios, como MMSID (aqui chamado de SHIP_ID), IMO, Nome, Dimensões, Destino, Tipo, etc. A segunda estrutura são as tabelas que contém as informações dinâmicas de navegação e identificadas pelo SHIP_ID.

Tabela 3.1: Campos da tabela de informações dinâmicas recebidas via AIS

Nome	Descrição	Tipo
SHIP_ID	MMSI* do navio	<i>int</i>
DATE	Data do registro de posição (yyyy-mm-dd)	<i>string</i>
TIME	Horário do registro de posição (HH:MM:SS)	<i>string</i>
SPEED	Velocidade do navio no registro (em nós)	<i>float</i>
LATITUDE	Latitude no horário do registro	<i>float</i>
LONGITUDE	Longitude no horário do registro	<i>float</i>

Fonte: Elaboração própria

Tabela 3.2: Campos da tabela de informações estáticas recebidas via AIS

Nome	Descrição	Tipo
SHIP_ID	MMSI do navio*	<i>int</i>
IMONUMBER	Número IMO do navio**	<i>string</i>
NAME	Nome do navio	<i>string</i>
TYPE	Código do tipo de navio	<i>int</i>
ETA	Horário estimado de chegada no porto de destino	<i>object</i>
MAXIMUM STATIC DRAUGHT	Profundidade máxima atingida pelo casco do navio abaixo da água	<i>float</i>
DESTINATION	Porto de destino	<i>string</i>

Fonte: Elaboração própria

* O MMSI (*Maritime Mobile Service Identity*)corresponde a um código de nove dígitos programado no AIS que deve ser único a cada embarcação (DGRM, 2014).

** O número IMO é o número do Registo Lloyd's, uma identificação única e obrigatória para embarcações com capacidade de mais de 100 toneladas ou que realizem viagens internacionais (Decreto-Lei n.º 295/2000, 2000)

Tais estruturas de dados foram as principais fundações para construção da base de dados que foi processada e analisada durante o presente estudo, sendo a compreensão de cada campo relevante para o resultado final.

3.2 Pré processamento de dados

Inicialmente as tabelas de dados dinâmicos e estáticos foram unidas em uma única entidade utilizando o SHIP_ID como chave primária, de forma que é possível visualizar os dados estáticos do navio para cada registro de posição, como nome, IMO e código de tipo de navio.

O resultado é então unido a uma terceira tabela, que contém as relações entre o código numérico de tipo de navio com o respectivo nome em extenso do tipo. Neste processo, a relação entre código e nome do tipo é "*many-to-one*", ou seja, diversos códigos podem ter o mesmo nome. As relações foram extraídas a partir do site da *MarineTraffic*, umas das principais plataformas de monitoramento de tráfego marítimo que fornece informações em tempo real sobre a localização e o movimento de embarcações, como forma de garantir que estão de acordo com os códigos dos dados coletados.

3.2.1 Limpeza dos dados

A limpeza de dados é um passo essencial para garantir a qualidade dos dados analisados, pois garantem que registros com erros ou dados que não são interessantes ao objetivo do estudo, sejam analisados. Assim, após a coleta dos dados, foram aplicadas uma série de transformações sobre os dados, tornando-os adequados para a análise.

3.2.1.1 Filtragem por tipo de Navio

Para o estudo em questão, somente navios cargueiros (*Cargo ships*), representados pelos códigos 70 a 79, foram analisados. Após este filtro restaram 6.725.124 registros na base.

3.2.1.2 Filtragem por área de interesse

Os dados coletados de posição abrangem uma região desde o sul da Bahia até o Paraná. Dessa forma, é necessário delimitar uma área mais específica de interesse, ou seja, aqueles com registros de longitude e latitude próximos do Porto de Santos. Para isso aplicou-se uma condição simples, que retira dados que estejam fora da área delimitada por um retângulo, cujos vértices estão disponíveis na tabela 3.3. A definição dos vértices buscou limitar o menor retângulo que contivesse toda a extensão do canal, bem como seus terminais e a área de fundeio, restando 734.477 registros de posição após esta filtragem.

Figura 3.3: Retângulo com a área filtrada para análise



Fonte: Elaboração própria

Tabela 3.3: Vértices da área filtrada para análise

Vértice da área analisada	Posição (Latitude, Longitude)
Vértice 1	-23.869000, -46.150561
Vértice 2	-24.277392, -46.150561
Vértice 3	-23.869000, -46.500000
Vértice 4	-24.277392, -46.500000

Fonte: Elaboração própria

3.2.1.3 Filtragem de Dados Inválidos

Foram retirados dados considerados equivocados, como de navios com IMO zero, visto que não é possível identificar estes navios.

Foram encontrados também registros com velocidade acima de 100 nós, alguns deles coletados poucos segundos depois de registros com velocidade zero. Muito provavelmente,

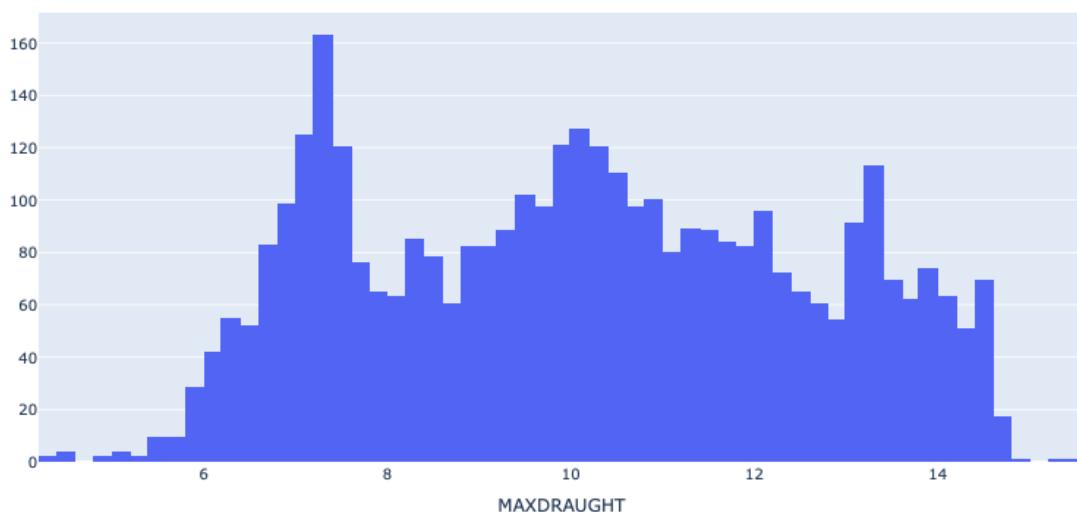
estes dados foram decorrentes de falhas no sistema de AIS e devem ser desconsiderados da análise. Após este filtro restaram 675.791 registros na base.

3.3 Análise exploratória inicial da base de navios

Após a aplicação dos filtros indicados, restando somente os dados válidos e interessantes para o presente estudo, foram identificados 898 navios distintos na base, que podem ter feito uma ou mais de uma viagem ao Porto de Santos durante o período em análise. Conforme o filtro de tipo de embarcação, 100% dos navios apresentam códigos de tipo entre 70 e 79, indicando que pertencem à categoria de navios cargueiros.

Quanto ao calado, medida que representa a profundidade com a qual o navio está submerso, calculou-se uma média de 10,1m, com um desvio padrão de 2,4m.

Figura 3.4: Distribuição do calado das embarcações



Fonte: Elaboração própria

Conforme observado no histograma da figura 3.4, a distribuição não é normal e há uma grande variação de valores mais baixos e altos. Isso ocorre pois um mesmo navio, ao adentrar no porto sem carga, está mais leve e apresenta um calado menor. Após atracar no terminal e ser carregado, o navio afunda mais, aumentando seu calado. Assim, é esperado que o calado seja distinto entre navios de acordo de portes diferentes, mas que também varie em um mesmo navio de acordo com o momento da viagem em que este se encontra pela carga transportada.

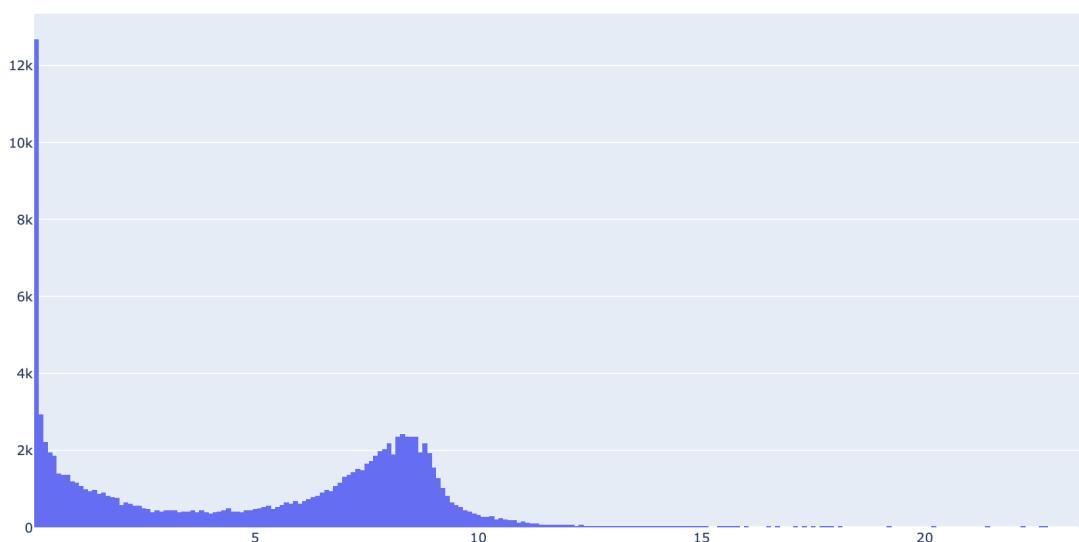
Analisando as velocidades coletadas, dos 675.791 dados de posição, 562.985 ou 83%

apresentaram velocidade nula, sendo os 17% restantes com velocidade maior que zero. Para os dados com velocidade não nula, a média foi de 5,03 nós, com desvio padrão de 3,55 nós.

Observando o histograma da distribuição das velocidades não nulas, percebe-se que há uma frequência muito alta de velocidades baixas pois os navios devem navegar de forma cautelosa enquanto realizam as manobras de atracação/desatração ou realizam pequenos ajustes enquanto atracados, sendo o tempo gasto de fato navegando no canal relativamente pequeno ao tempo total no Porto.

Além disso, há também uma concentração de valores próximos a 8,5 nós, representando a velocidade com a qual os navios efetivamente trafegam no canal para entrarem e saírem, sendo os valores intermediários menos frequentes pois são estados transitórios de aceleração e desaceleração.

Figura 3.5: Distribuição das velocidades não nulas dos dados em análise



Fonte: Elaboração própria

Embora não sejam frequentes, há ainda cerca de 10% dos registros com velocidade acima do limite estabelecido pela Autoridade Portuária de 9 nós para embarcações após o limiar da Fortaleza da Barra Grande (Marinha do Brasil, 2023), que podem ser explicados pelos navios que ainda estão na parte externa do canal, além de existirem *outliers*, valores altos superiores a 18,3 nós.

3.4 Algoritmo de Rotulamento com códigos

O rotulamento de dados é o processo de atribuir rótulos, categorias ou metadados a um conjunto de dados brutos. Esse processo visa organizar, estruturar e facilitar o entendimento e a manipulação dos dados, atribuindo-lhes significado ou classificações específicas. O rotulamento é uma técnica muito utilizada no campo de aprendizado de máquina supervisionado, uma vez que estes algoritmos necessitam de dados rotulados para "aprender" a categorizar novos dados, além de serem necessários para validar a qualidade dos resultados dos modelos. Entretanto, o rotulamento de dados pode ser utilizado em outros contextos além do *machine learning*, como no caso do presente trabalho, para o mapeamento de posições geográficas.

O método proposto para rotular os dados de posição coletados é baseado na localização e na velocidade das embarcações. Por meio dos rótulos é possível acompanhar a posição do navio, que é em sua forma bruta um dado contínuo de difícil compreensão sem o auxílio de uma representação cartográfica, de forma discreta, permitindo o acompanhamento da embarcação ao longo do seu percurso no canal e nos terminais com nomes que sejam familiares aos analistas.

3.4.1 Códigos de posição

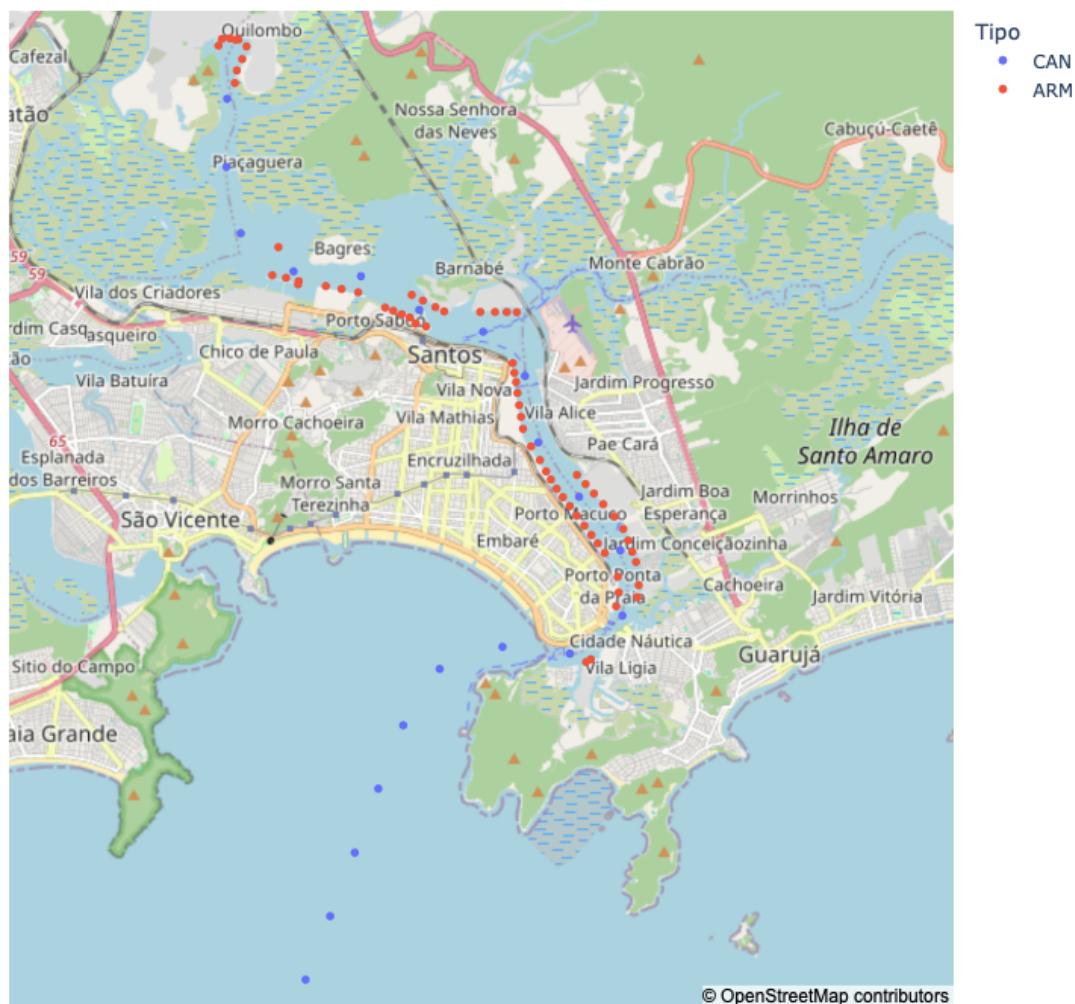
Para rotular as posições dos navios foram criados códigos, que são divididos em dois principais tipos de localizações no Porto de Santos:

- **Terminais do Porto:** com a posição sendo caracterizada pela Latitude e Longitude do centro do terminal. Foram levantados 65 códigos de terminal, pois no caso de terminais com mais de um berço foram gerados um rótulo para cada berço. Por exemplo, o terminal Tiplam tem 4 berços, identificados como Tiplam 1, Tiplam 2, e assim por diante.
- **Partições do Canal de navegação:** a partir de uma divisão do canal em trechos de um quilômetro e meio, foram gerados 20 rótulos de partição, cuja posição é representada pela Latitude e Longitude dos centros geográficos de cada partição.

Cada local neste conjunto de dados, seja um terminal ou uma partição do canal, recebe a denominação de "código", ou "rótulo". A Figura 3.6 ilustra as posições dos códigos ao longo do canal.

O objetivo das operações de rotulamento é corretamente identificar cada registro de posição de navios com um dos códigos propostos, para que seja possível distinguir em qual participação uma embarcação estava navegando ou em qual terminal estava operando em um determinado momento.

Figura 3.6: Geolocalização dos códigos utilizados para rotulação



Fonte: Elaboração própria

3.4.2 Cálculo das distâncias e atribuição do rótulo

Inicialmente foram calculadas as distâncias d entre as posições dos navios coletadas nos dados de AIS e cada um dos códigos gerados com auxílio da biblioteca NumPy para

calcular a Fórmula de Haversine (equação 3.1), equação aplicada em navegação para obter distâncias entre dois pontos de uma esfera a partir de suas latitudes e longitudes. O resultado da distância obtida é então multiplicada pelo raio estimado da Terra ($r = 6378,137\text{km}$ no equador) para se obter a distância em quilômetros.

$$d = 2r \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta\phi}{2} \right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2 \left(\frac{\Delta\lambda}{2} \right)} \right) \quad (3.1)$$

Onde:

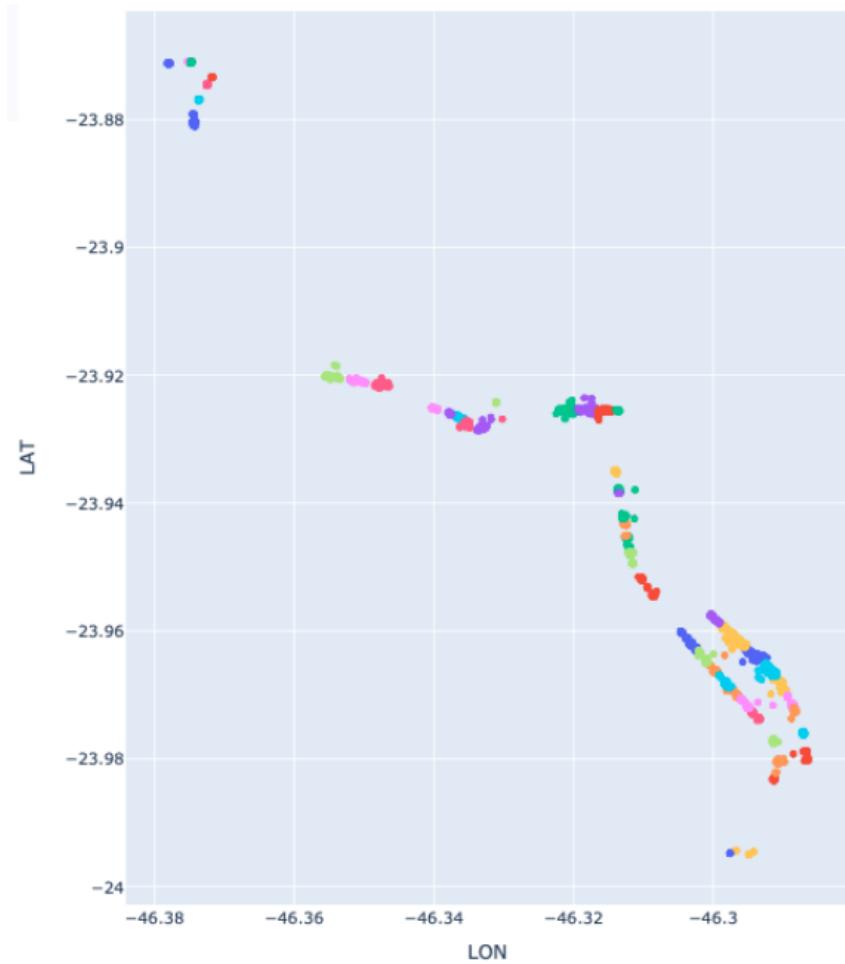
- d é a distância entre os dois pontos.
- r é o raio da esfera (por exemplo, o raio da Terra).
- ϕ_1 e ϕ_2 são as latitudes dos pontos 1 e 2 (em radianos).
- λ_1 e λ_2 são as longitudes dos pontos 1 e 2 (em radianos).
- $\Delta\phi = \phi_2 - \phi_1$ é a diferença de latitude.
- $\Delta\lambda = \lambda_2 - \lambda_1$ é a diferença de longitude.

É selecionado então o código de partição do canal com a menor distância em relação ao registro de posição do navio, e o código do terminal com a menor distância em relação à embarcação.

Caso o navio esteja parado (identificado com velocidade nula), considera-se que o navio estava atracado e o rótulo daquele dado será o código do terminal mais próximo. Caso o navio esteja se movimentando (velocidade maior que zero) sinalizando que a embarcação estava navegando, o dado será rotulado com o código de partição do canal mais próximo.

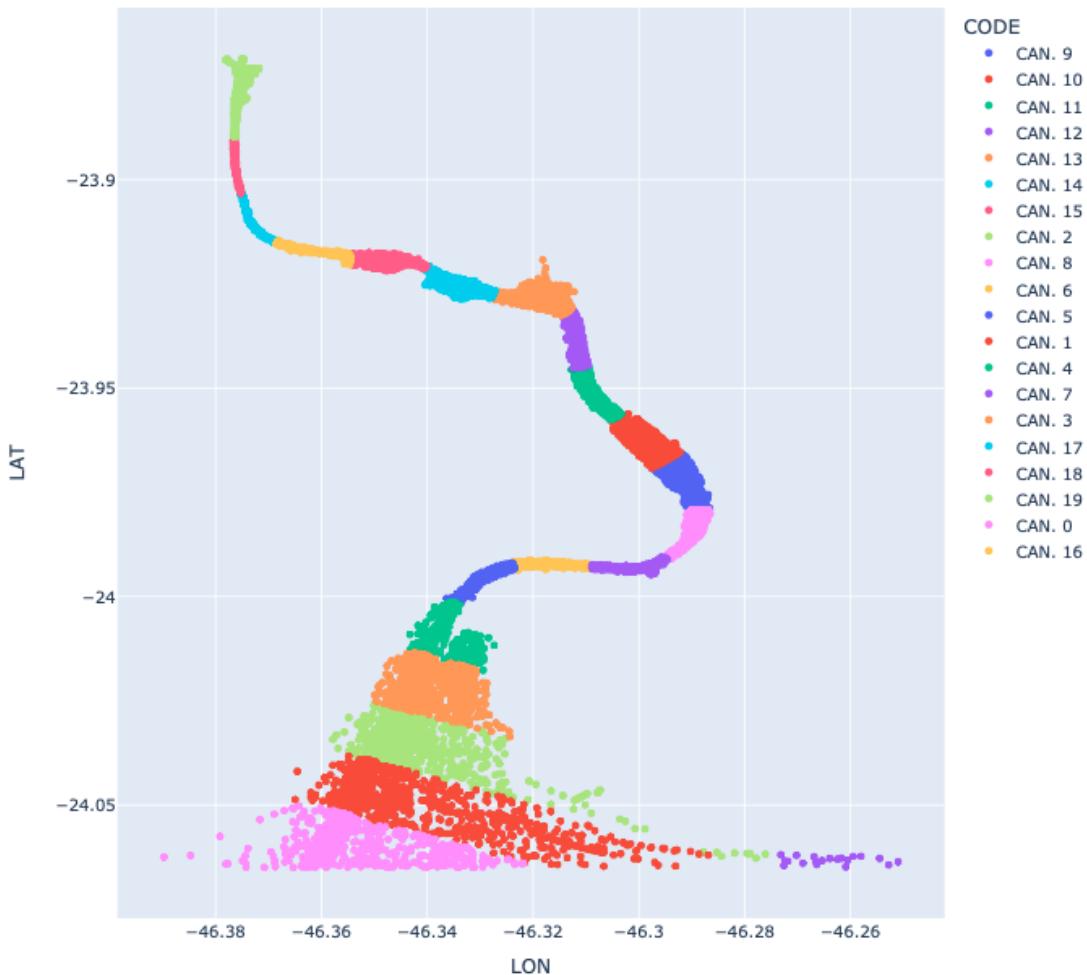
Os dados rotulados com códigos de terminal, e os dados rotulados com códigos de partição do canal para esta primeira aplicação estão disponíveis na figura 3.7 e 3.8, respectivamente

Figura 3.7: Rotulamento inicial dos dados de posição (terminais)



Fonte: Elaboração própria

Figura 3.8: Rotulamento inicial dos dados de posição (canal)



Fonte: Elaboração própria

3.4.3 Interpretações do Rotulamento inicial

Note que na figura 3.7, que representa os navios atracados separados por terminal de atracação, é visível que há registros de navios distantes dos demais pontos identificados com o mesmo rótulo, indicando que navios parados estavam invadindo a área de navegação do canal. Para tais acontecimentos são levantadas duas hipóteses:

1. **Há um erro na velocidade no registro**, e um navio que está em movimento foi erroneamente identificado como parado, fazendo com que o mesmo receba um código de terminal. Neste caso, o rótulo de terminal deve ser substituído por um de navegação;
2. **Há um erro na posição no registro**, e o ponto aparentemente no canal de

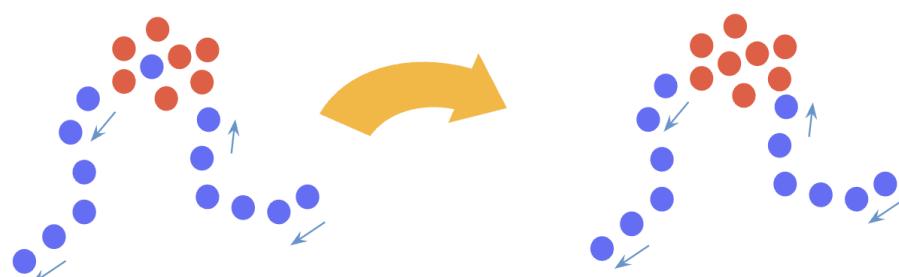
navegação na realidade estava próximo dos demais, dentro do terminal. Neste caso, apesar de a visualização parecer incorreta, o rótulo está correto, visto que o navio estava de fato ocupando o terminal naquele momento, e não é necessário fazer correções.

A primeira hipótese evidencia o principal problema do modelo de rotulação proposto: a decisão de atribuir um rótulo de terminal ou de canal a um dado de posição depende da velocidade instantânea daquele registro ser nula ou não. Isso gera erros pois a velocidade está sujeita a oscilações, tanto por falhas de envio quanto por fatores naturais, e como consequência os rótulos também oscilam entre terminal e canal mesmo que a embarcação não tenha mudado de posição ou realizado uma atracação e desatracação.

Para melhor esclarecer como a oscilação da velocidade afeta os rótulos, foram levantados os dois principais momentos em que ruídos da velocidade que causam erros na classificação: quando o navio já está atracado e quando o navio ainda está navegando, mas próximo de atracar.

- Caso 1:** O navio está atracado em um terminal, porém sua velocidade não é perfeitamente zero durante todo o período, pois correm pequenas oscilações devido a fatores climáticos, como ventanias e mudanças de maré, que fazem com que o navio apresente velocidade maior que zero nos registros. A variação também pode decorrer da necessidade de manobras no navio, ou ainda de erros na captação e transmissão de dados. Em todo caso, essas mudanças são insignificantes para o rotulamento, pois a embarcação estava de fato ocupando o terminal durante todo o período. Assim, é necessário que seja aplicado um tratamento no dado de navio movendo para que se torne um dado de navio parado, e consequentemente receba o rótulo correto de terminal.

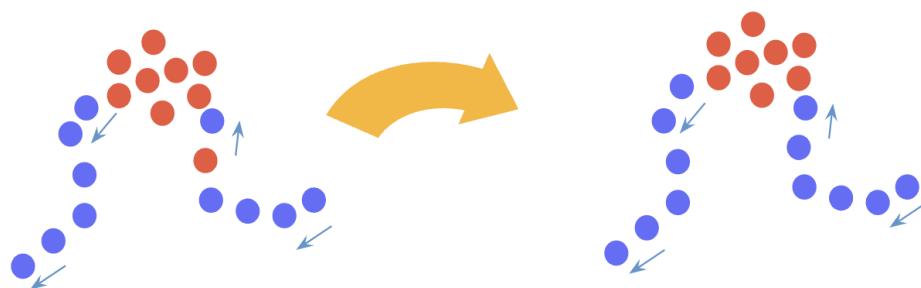
Figura 3.9: Caso 1: oscilações na velocidade enquanto o navio está atracado. Navio se movendo está representado em azul e parado em vermelho



Fonte: Elaboração própria

2. Caso 2: Durante o processo de atracação, o navio navegando diminui sua velocidade para manobrar e eventualmente parar em um terminal. Neste processo, ocorrem o envio de registros com velocidade zero seguidos de registros de movimento pois é natural que a velocidade varie durante a parada. Entretanto, não necessariamente toda oscilação para velocidade zero caracteriza a atracação em si, que só ocorre uma vez durante a viagem. Assim, é necessário corrigir os dados de navio parado para dados de navio movendo para garantir que o processo de atracação receba corretamente o rótulo de canal.

Figura 3.10: Caso 2: Oscilações na velocidade enquanto o navio está navegando. Navio se movendo está representado em azul e parado em vermelho



Fonte: Elaboração própria

Para evitar os dois problemas na classificação causados pelas oscilações nas velocidades que ocorrem nos momentos próximos e durante a atracação, é proposta uma série de operações sobre os dados de posição coletados com o objetivo de reconstruir as trajetórias de cada embarcação.

Neste contexto, uma trajetória é definida como um conjunto de registros contínuos no tempo nos quais o navio manteve-se com o mesmo estado (navio movendo ou navio atracado) durante uma visita ao Porto de Santos. Assim, por meio da reconstrução dessas entidades, é possível organizar o estado do navio de forma contínua e corretamente atribuir um rotulo de canal ou de movimentação.

3.5 Algoritmos de trajetórias

A partir da definição de trajetória apresentada anteriormente, considera-se que uma embarcação deve possuir, de forma geral, três trajetórias: uma primeira de movimentação (trajetória de chegada), quando o navio entra no canal e navega até o terminal, uma única trajetória de parada atracação, enquanto o navio está atracado no terminal, e uma última

de movimento, quando o mesmo se locomove para sair do canal (trajetória de saída).

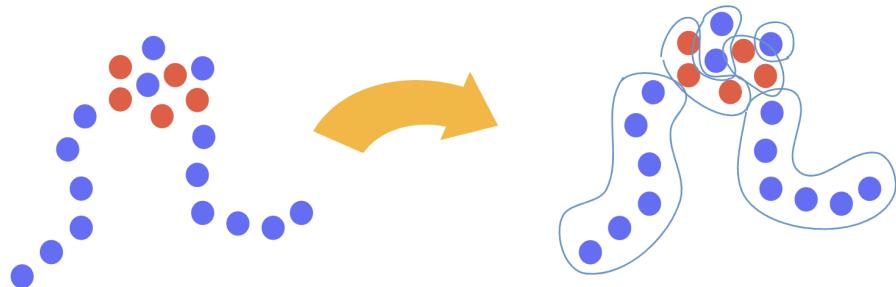
Esta aproximação de três trajetórias não é verdadeira em todos os casos, podendo variar para mais, caso um navio atraque em mais de um terminal antes de deixar o Porto, ou para menos, caso o sistema de AIS falhe em captar a navegação durante a entrada ou a saída da embarcação do canal. Apesar disso, entende-se que no caso geral um navio terá somente as três trajetórias apresentadas (trajetória de entrada, trajetória de parada e trajetória de saída). Dessa forma, números muito maiores que três trajetórias para um mesmo navio é um indicativo de erros na rotulação decorrentes dos casos 1 e 2 caracterizados na seção anterior.

Na presente seção são apresentados os três algoritmos utilizados para reconstruir as trajetórias dos navios cujos dados foram coletados. Por meio deles, espera-se corrigir as oscilações na velocidade e aplicar novamente a metodologia de rotulamento apresentada, obtendo um segundo rótulo melhorado. Além do uso das trajetórias para auxiliar no rotulamento, a reconstrução de trajetórias de navios a partir de dados de AIS é um tema de estudo com aplicações variadas como na previsão de percursos e prevenção de acidentes por meio de aprendizado de máquina.

O primeiro algoritmo aplicado é responsável por criar as trajetórias com base nos dados de posição, separando os registros de posição a partir de mudanças no estado do navio (de parado para movendo, ou vice e versa), e de intervalos longos de tempo entre dois registros recebidos para uma mesma embarcação. O Algoritmo 2 então utiliza os resultados do seu antecessor para ajustar as velocidades instantâneas a partir dos dados agregados das trajetórias. Com as velocidades ajustadas, são criadas novas trajetórias e o Algoritmo 3, por fim, agrupa as novas trajetórias decorrentes do Algoritmo 2 para criar as trajetórias finais que serão base para o novo rotulamento.

As figuras 3.11 e 3.12 retratam as transformações dos dados de acordo com a aplicação dos algoritmos. Nas figuras, os círculos azuis representam dados de movendo (velocidade maior que zero) e os círculos vermelhos representam dados de navio parado (velocidade igual a zero), e círculos unidos pelas indicações em azul pertencem à mesma trajetória.

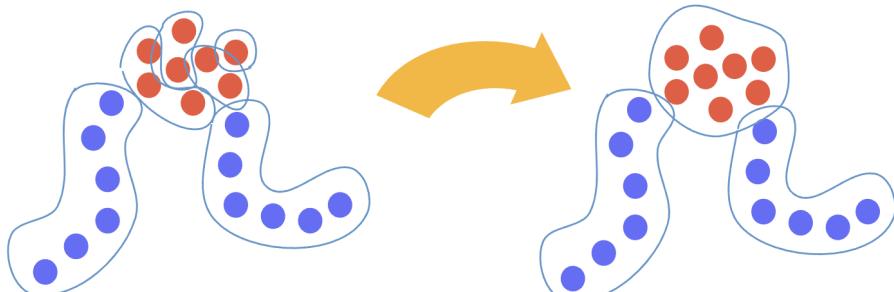
Figura 3.11: Representação da primeira parte da reconstrução de trajetórias



Fonte: Elaboração própria

Na figura 3.11 à esquerda, temos os dados antes das aplicações dos algoritmos, em que todos os círculos estão individualizados pois ainda não houve a reconstrução de trajetórias. Note que é possível observar o problema de oscilação mencionado anteriormente, pois há dados de navio movimentando misturados aos dados de navio parado durante a atracação no terminal. À direita temos o resultado da aplicação do Algoritmo 1, onde os dados ficam coloridos indicando a criação das trajetórias. Neste momento, os ruídos na velocidade resultam na geração de diversas quatro diferentes trajetórias no período em que o navio esteve parado no terminal com poucos dados em cada uma delas.

Figura 3.12: Representação da segunda e terceira partes da reconstrução de trajetórias



Fonte: Elaboração própria

A figura 3.12 à esquerda representa os dados após a aplicação do Algoritmo 2, que corrige as oscilações de velocidade durante a atracação no terminal, transformando os círculos azuis (dados de navio movendo) em vermelhos (dados de navio parado). Assim, não há mais a mistura de dados de movimentação e atracação, mas as trajetórias não são capazes de descrever este comportamento em somente 3 grandes momentos, pois a atracação ainda apresenta mais de uma trajetória.

O Algoritmo 3 resolve este problema, unindo as três trajetórias de atracação em uma

só. Assim, na direita da figura 3.12 está representado o estado esperado da reconstrução de trajetórias após a aplicação dos três algoritmos, em que as velocidades foram corrigidas e a viagem do navio foi reconstruída com três trajetórias somente.

A seguir, serão apresentados em detalhes os procedimentos dos três algoritmos utilizados, bem como as estruturas de dados utilizadas como entradas, e geradas como saídas das operações.

3.5.1 Principais estruturas de dados

As Tabelas de Trajetórias e de Dados de Posição são as principais estruturas de dados utilizadas para aplicações dos algoritmos neste estudo, servindo tanto como *input* como *output* para as operações realizadas. Ambas incluem informações sobre o estado das embarcações identificadas pelo número IMO, mas diferem na granularidade de registro, sendo portanto necessárias para descrever informações distintas.

A Tabela de Dados de Posição é resultante da união das tabelas recebidas pelo sistema AIS e com as características de navegação de um navio em um momento pontual, como velocidade instantânea, latitude e longitude. Dessa forma, essa tabela é útil quando deseja-se saber as propriedades de um navio em um instante específico.

Tabela 3.4: Estrutura das Tabelas de Trajetória

Nome	Descrição	Tipo
IMO	Número IMO do navio	int
NAME	Nome do navio	string
TINDEX	ID da trajetória	int
DATETIME	Horário do registro	datetime
SPEED	Velocidade do navio no registro (em nós)	float
LON	Longitude no horário do registro	float
LAT	Latitude no horário do registro	float

Fonte: Elaboração própria

Já a Tabela de Trajetórias é um agregado de dados de posição, tendo como granularidade (representada por uma linha da tabela) uma trajetória criada pelos algoritmos para uma embarcação. Assim, este conjunto de dados é útil para descrever as principais características de um conjunto de dados agrupados, utilizando métricas agregadas como velocidade média e deslocamento total no período.

Tabela 3.5: Estrutura das Tabelas de Trajetória

Nome	Descrição	Tipo
IMO	Número IMO do navio	int
NAME	Nome do navio	string
TINDEX	ID da trajetória	int
FIRST LAT	Latitude no ponto inicial da trajetória	float
LAST LAT	Latitude no ponto final da trajetória	float
FIRST LON	Longitude no ponto inicial da trajetória	float
LAST LON	Longitude no ponto final da trajetória	float
INIT DATETIME	Horário no início da trajetória	datetime
END DATETIME	Horário no final da trajetória	datetime
AVG SPEED	Velocidade média na trajetória (em nós)	float
ROW COUNT	Número de registros de dados para aquela trajetória	int
DISTANCE	Distância percorrida na trajetória	float

Fonte: Elaboração própria

O Algoritmo 1 utiliza uma tabela de dados de posição como *input*, pois busca comparar cada instante de tempo com o anterior para identificar mudanças no estado do navio. Os Algoritmos 2 e 3 utilizam tabelas de trajetórias como base, pois se fundamentam em informações agregadas, e não de um instante específico.

3.5.2 Algoritmo 1

No algoritmo 1 de criação de trajetórias, primeiramente os dados de posição são ordenados por navio e por horário, de forma que todos os registros para um navio aparecem sequencialmente antes do primeiro registro da próxima embarcação. Em seguida, é criada uma variável t_{index} , que será a identificação única de uma trajetória. Todos os registros de posição recebem umt_{index} para identificar a qual trajetória pertencem, e t_{index} tem seu valor incrementado, sinalizando a necessidade de uma nova trajetória, nas seguintes condições:

- **Caso 1:** o registro de um navio é diferente do anterior, pois dois navios diferentes não podem compartilhar uma mesma trajetória;
- **Caso 2:** há uma diferença de tempo superior a uma hora entre dois registros, indicando que houve um período no qual não foram coletados dados. Neste caso,

cria-se uma nova trajetória, pois neste intervalo entre os registros não é possível afirmar que o navio manteve o estado anterior. Este intervalo de tempo grande também pode ser decorrente de visitas diferentes realizadas pelo menos navio ao Porto de Santos;

- **Caso 3:** o navio mudou seu estado de movendo (velocidade maior que zero) para parado (velocidade igual a zero), ou vice e versa, indicando uma atracação ou uma desatracação;

Note que todos os casos comparam um registro de estado com o anterior, pois neste primeiro momento busca-se identificar características que justifiquem uma nova trajetória. Os algoritmos seguintes têm como objetivo refinar essa busca, tornando as trajetórias medidas mais fiéis aos verdadeiros movimentos das embarcações e reduzindo o erro causado pelas oscilações de velocidade.

Como resultado destas operações são geradas duas principais tabelas: uma tabela de dados de posição com rótulo de trajetória t_{index} e uma tabela de trajetórias utilizando t_{index} .

Algorithm 1 Iteração para rotulamento das posições com índice de trajetória

```

1:  $t_{\text{index}} \leftarrow 1$                                  $\triangleright$  Inicializa o índice de trajetória
2:  $n \leftarrow$  tamanho total dos dados de posição
3: for  $i \leftarrow 1$  to  $n$  do                       $\triangleright$  Itera sobre todas as linhas de dados
4:   if  $IMO(i) \neq IMO(i - 1)$  then
5:      $t_{\text{index}} \leftarrow L_{\text{index}} + 1$            $\triangleright$  Mudança de navio, nova trajetória
6:   end if
7:   if  $T(i) - T(i - 1) > 1$  hora then
8:      $t_{\text{index}} \leftarrow t_{\text{index}} + 1$        $\triangleright$  Diferença de tempo maior que 1 hora, nova trajetória
9:   end if
10:  if  $v(i) = 0$  and  $v(i - 1) > 0$  then
11:     $t_{\text{index}} \leftarrow t_{\text{index}} + 1$            $\triangleright$  Mudança de estado, nova trajetória
12:  end if
13:   $t_{\text{index}}(i) \leftarrow t_{\text{index}}$              $\triangleright$  Rotula a linha atual com a trajetória
14: end for

```

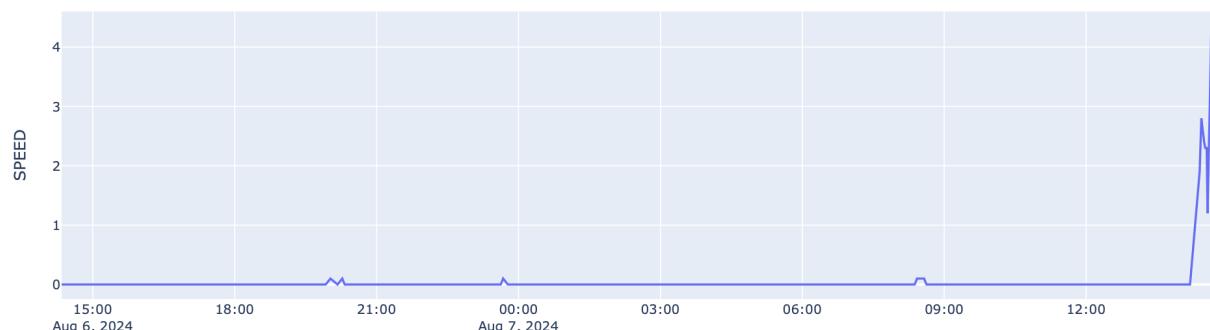
É importante destacar, entretanto, que esta primeira iteração do Algoritmo 1 por si só não é muito útil para descrever os verdadeiros roteiros das embarcações, pois gera um número muito alto de trajetórias com poucos registros de posição cada uma, enquanto o esperado seriam somente três trajetórias descrevendo um movimento longo e contínuo. Esse desalinhamento ocorre pois, da mesma forma que o rotulamento, as trajetórias dependem da velocidade, sendo que ruídos neste aspecto causam erros.

Assim, faz-se necessário a aplicação de mais algoritmos para tornar as trajetórias geradas nos objetos desejados para a análise. O Algoritmo 2, responsável pelo ajuste de velocidades, suaviza as variações da velocidade instantânea. Em seguida, o Algoritmo 3 agrupa os dados processados pelo Algoritmo 2, unindo os trechos contínuos das trajetórias.

3.5.3 Algoritmo 2

Na figura 3.13, temos o perfil da velocidade do navio STAR CRIMSON entre as 15 horas do dia 06 de agosto e 14 horas do dia 07 do mesmo mês. Percebe-se que durante o período em que a embarcação estava majoritariamente parada, indicando uma atracação, houveram pequenas oscilações de velocidade entre 0 e 0.1, antes do navio efetivamente desatracar, às 14 horas, quando ganha uma velocidade maior.

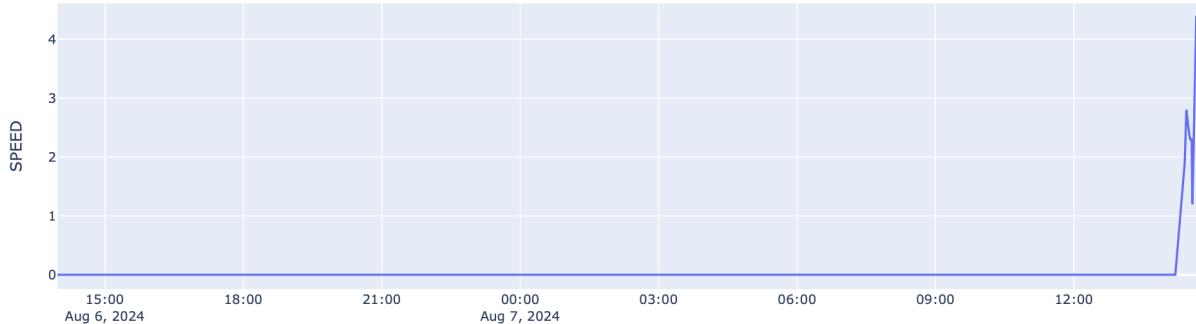
Figura 3.13: Perfil da velocidade original do STAR CRIMSON



Fonte: Elaboração própria

O Algoritmo 2 é responsável por corrigir estas oscilações de velocidade que ocorrem principalmente nos momentos em que o navio está atracado, próximo de atracar ou desatracar. A figura 3.14 apresenta o perfil atualizado da velocidade do STAR CRIMSON após a aplicação do Algoritmo 2, onde é perceptível que o ruído na velocidade durante a atracação foi eliminado.

Figura 3.14: Perfil da velocidade ajustada do STAR CRIMSON



Fonte: Elaboração própria

O algoritmo se inicia ordenando as trajetórias por navio e por data de forma crescente, de forma que todas as trajetórias de um mesmo navio apareçam sequencialmente antes das trajetórias da próxima embarcação. Em seguida, cada trajetória gerada no Algoritmo 1 é analisada e tem sua velocidade média ajustada para zero caso atenda às seguintes condições:

- **Condição 1:** o navio não tenha tido um deslocamento considerável durante a trajetória (deslocamento menor que 8m);
- **Condição 2:** velocidade média da trajetória tenha sido menor que 2 nós;

Nos demais casos, em que houve uma distância percorrida considerável ou em que a velocidade média da trajetória seja maior que dois nós, não é realizado ajuste na velocidade.

Algorithm 2 Iteração para correção das velocidades instantâneas

```

1:  $s \leftarrow$  tamanho total dos dados de trajetória
2: for  $i \leftarrow 1$  to  $s$  do ▷ Itera sobre todas as linhas de dados
3:    $D \leftarrow d(P_F(s), P_I(s_i))$ 
4:   if  $D < 0.008km$  then
5:     if  $\bar{v}(s) > 2$  nós then
6:        $\bar{v}'(s) = 0$  ▷ Ajusta a velocidade média da trajetória para zero
7:     end if
8:   else
9:      $\bar{v}'(s) = \bar{v}(s)$ 
10:  end if
11: end for

```

Onde:

- $P_F(s_i)$ é a última posição da trajetória s_i
- $P_I(s_i)$ é a primeira posição da trajetória s_i
- $d(a, b)$ é a distância entre os pontos a e b , calculada utilizando a fórmula de Haversine
- $\bar{v}(s)$ é a velocidade média da trajetória s_i
- $\bar{v}'(s)$ é a média média corrigida da trajetória s_i

Os parâmetros velocidade média e deslocamento durante a trajetória foram escolhidos para identificar trajetórias inválidas pois são os que melhor descrevem os eventos de oscilação de velocidade: considerou-se que caso o navio tenha se deslocado pouco, e com uma velocidade baixa, essa trajetória caracteriza uma oscilação, e não uma movimentação válida.

Para calcular o deslocamento percorrido durante a trajetória foi aplicada novamente a equação baseada na função de Haversine, para calcular a distância entre o primeiro e o último registro daquela trajetória e o valor limite de 8m foi estabelecido a partir do percentil 75% do deslocamento de navios com velocidade nula.

Após a aplicação do Algoritmo 2, temos uma tabela de trajetórias com as velocidades médias das trajetórias ajustadas. Juntamos então esta tabela de trajetórias com a tabela de posição utilizada inicialmente, de forma que os registros cujas trajetórias tiveram sua velocidade média ajustada tenham suas velocidades instantâneas também ajustadas para zero e os demais registros mantenham suas velocidades instantâneas originais. Desta forma obtemos, uma tabela de dados de posição com velocidade instantânea ajustada.

A esta tabela de dados de posição ajustada, é aplicado então novamente o Algoritmo 1 para criar uma tabela de trajetórias com identificações atualizadas (t_{index2}) a partir das novas velocidades.

Na figura 3.15 estão representadas as trajetórias do navio STAR CRIMSON criadas pelo algoritmo 1. Nota-se que a velocidade estava oscilando entre 0.1 nós e 0 nós gerando novas trajetórias (t_{index}), embora na realidade a embarcação estava atracada durante todo o período.

Figura 3.15: Trajetórias geradas após a aplicação do algoritmo 1 para o STAR CRIMSON

	NAME	IMO	TINDEX	INIT	END	Avg_Speed
0	STAR CRIMSON	9758870.0	14994	2024-08-05 00:08:42	2024-08-05 01:12:50	8.384848
1	STAR CRIMSON	9758870.0	14995	2024-08-05 19:49:31	2024-08-05 20:37:31	0.000000
2	STAR CRIMSON	9758870.0	14996	2024-08-06 05:34:29	2024-08-06 09:25:29	0.000000
3	STAR CRIMSON	9758870.0	14997	2024-08-06 17:55:29	2024-08-06 19:55:29	0.000000
4	STAR CRIMSON	9758870.0	14998	2024-08-06 20:01:29	2024-08-06 20:01:29	0.100000
5	STAR CRIMSON	9758870.0	14999	2024-08-06 20:10:29	2024-08-06 20:10:29	0.000000
6	STAR CRIMSON	9758870.0	15000	2024-08-06 20:16:29	2024-08-06 20:16:29	0.100000
7	STAR CRIMSON	9758870.0	15001	2024-08-06 20:19:31	2024-08-06 23:37:29	0.000000
8	STAR CRIMSON	9758870.0	15002	2024-08-06 23:40:29	2024-08-06 23:40:29	0.100000
9	STAR CRIMSON	9758870.0	15003	2024-08-06 23:46:29	2024-08-07 08:22:29	0.000000
10	STAR CRIMSON	9758870.0	15004	2024-08-07 08:25:31	2024-08-07 08:34:29	0.100000
11	STAR CRIMSON	9758870.0	15005	2024-08-07 08:37:30	2024-08-07 14:11:46	0.000000

Fonte: Elaboração própria

Com a aplicação do algoritmo 2 e o recalcular das trajetórias, foi possível diminuir o número de trajetórias de 12 para somente 3 (t_{index2}), a partir da eliminação dos ruídos de velocidade nos dados de posição do STAR CRIMSON.

Figura 3.16: Trajetórias geradas após a aplicação do algoritmo 2 para o STAR CRIMSON

	NAME	IMO	TINDEX2	INIT	END	Avg_Speed
0	STAR CRIMSON	9758870.0	4705	2024-08-05 00:08:42	2024-08-05 01:12:50	8.384848
1	STAR CRIMSON	9758870.0	4706	2024-08-05 19:49:31	2024-08-05 20:37:31	0.000000
2	STAR CRIMSON	9758870.0	4707	2024-08-06 05:34:29	2024-08-06 09:25:29	0.000000
3	STAR CRIMSON	9758870.0	4708	2024-08-06 17:55:29	2024-08-07 14:11:46	0.000000
4	STAR CRIMSON	9758870.0	4709	2024-08-07 14:23:59	2024-08-07 16:35:10	7.501724

Fonte: Elaboração própria

Note que apesar das operações do Algoritmo 2 já diminuírem o número de trajetórias, há casos em que seria possível unir trajetórias adjacentes, pois duas trajetórias seguidas e identificadas com indexes distintos podem estar descrevendo uma mesma atração. O Algoritmo 3 tem como objetivo realizar a união destas trajetórias em uma única trajetória que descreve todo o tempo em que o navio esteve atracado em um mesmo terminal.

3.5.4 Algoritmo 3

O Algoritmo 3 busca agregar trajetórias consecutivas em que o navio estava parado, representando em uma única trajetória, o tempo de atração em um terminal. A Figura

3.17 mostra uma tabela de trajetórias, produto do Algoritmo 2, e exemplifica o caso em que o Algoritmo 3 é necessário.

Para o navio SOL DO BRASIL, nas linhas 30 a 41 temos 11 trajetórias de navio atracado (velocidade zero) que não apresentaram uma grande distância percorrida, sinalizando que o navio esteve atracado sem se deslocar significativamente durante este tempo. Desta forma, é interessante unir as 11 trajetórias em uma só que represente todo o tempo em que o navio esteve atracado.

Figura 3.17: Trajetórias do navio SOL DO BRASIL antes da aplicação do Algoritmo 3

	NAME	IMO	TINDEX2	INIT	END	Avg_Speed	ROWS	DIST
29	SOL DO BRASIL	9018658	30	2024-06-05 09:03:23	2024-06-05 09:58:22	3.746153846153846	39	2.6831935222222962
30	SOL DO BRASIL	9018658	31	2024-06-05 10:01:32	2024-06-05 17:11:50	0.0	41	2.22222224264543e-07
31	SOL DO BRASIL	9018658	32	2024-06-05 18:29:51	2024-06-06 13:20:49	0.0	91	2.2222222252119334e-08
32	SOL DO BRASIL	9018658	33	2024-06-06 14:26:49	2024-06-06 15:50:48	0.0	5	5.55555551187455e-08
33	SOL DO BRASIL	9018658	34	2024-06-06 17:29:49	2024-06-06 17:29:49	0.0	1	0.0
34	SOL DO BRASIL	9018658	35	2024-06-06 19:20:48	2024-06-07 00:26:48	0.0	8	1.3611111116988596e-07
35	SOL DO BRASIL	9018658	36	2024-06-07 01:32:47	2024-06-07 07:41:48	0.0	38	2.5000000005212544e-07
36	SOL DO BRASIL	9018658	37	2024-06-07 08:59:48	2024-06-07 13:50:48	0.0	29	3.61111113008797e-08
37	SOL DO BRASIL	9018658	38	2024-06-07 16:20:48	2024-06-07 23:08:46	0.0	29	3.4722222220974825e-07
38	SOL DO BRASIL	9018658	39	2024-06-08 01:02:47	2024-06-08 01:08:46	0.0	2	5.555555539345076e-09
39	SOL DO BRASIL	9018658	40	2024-06-08 02:59:46	2024-06-08 04:35:46	0.0	7	2.77777777962203e-08
40	SOL DO BRASIL	9018658	41	2024-06-08 05:53:46	2024-06-08 06:35:45	0.0	4	2.7777777744094893e-08
41	SOL DO BRASIL	9018658	42	2024-06-08 07:47:46	2024-06-08 07:50:46	0.0	2	1.111111149744425e-08
42	SOL DO BRASIL	9018658	43	2024-06-08 08:04:22	2024-06-08 08:15:53	6.9666666666666667	9	0.5831012500000827
43	SOL DO BRASIL	9018658	44	2024-06-08 09:37:58	2024-06-08 10:18:40	14.658333333333333	12	25.664285680553817

Fonte: Elaboração própria

O Algoritmo 3 é muito similar ao Algoritmo 1, pois analisa os registros sequencialmente para criar trajetórias a partir de critérios determinados. Entretanto, análogo ao Algoritmo 2 e destoando do Algoritmo 1, utilizamos a caracterização agregada das trajetórias como critério para ajustar e criar novas trajetórias, ao invés dos dados instantâneos de posição e velocidade. Assim, o Algoritmo 3 utiliza a tabela de trajetórias gerada no final do Algoritmo 2 para calcularas trajetórias finais.

O procedimento consiste em atribuir às trajetórias uma nova identificação t_{index3} a partir de novos critérios para criar uma trajetória:

- **Caso 1:** o registro de um navio é diferente do anterior, pois dois navios diferentes não podem compartilhar uma trajetória;
- **Caso 2:** há uma mudança na velocidade média em relação ao registro anterior. Fazemos isto pois assim registros com velocidade não nula devem pertencer a trajetórias diferentes, mas dois registros seguidos em que a velocidade é zero podem pertencer à mesma trajetória;

- **Caso 3:** a trajetória tem velocidade média zero e a distância entre registros seguidos é maior que 0.04km, pois nestes casos houve um deslocamento significativo mesmo que a velocidade tenha se mantido nula. Neste caso é necessário criar uma nova trajetória pois houve uma quebra relevante no posicionamento;

Note que desta vez rotulamos as trajetórias já criadas com t_{index3} , e não os registros de posição individualmente. Dessa forma, ao final da aplicação do Algoritmo 3, temos uma tabela de trajetória com duas identificações: as trajetórias originais t_{index2} , provenientes do algoritmo 2, e as trajetórias ajustadas (t_{index3}).

As trajetórias originais são descartadas e as trajetórias ajustadas são utilizadas para realizar agregações e obter uma última tabela de trajetórias.

Algorithm 3 Iteração para juntar trajetórias de atracação

```

1:  $t_{\text{index}} \leftarrow 1$                                      ▷ Inicializa o índice de trajetória
2:  $s \leftarrow$  tamanho total dos dados de posição
3: for  $i \leftarrow 1$  to  $s$  do                                ▷ Itera sobre todas as linhas de dados
4:   if  $IMO(s_i) \neq IMO(s_{i-1})$  then
5:      $t_{\text{index}} \leftarrow t_{\text{index}} + 1$           ▷ Mudança de navio, nova trajetória
6:   end if
7:   if  $\bar{v}(s) \neq \bar{v}(s_{i-1})$  then
8:      $t_{\text{index}} \leftarrow t_{\text{index}} + 1$           ▷ Mudança de tipo de trajetória, nova trajetória
9:   end if
10:   $D \leftarrow d(P_F(s), P_F(s_{i-1}))$ 
11:  if  $\bar{v}(s) = 0$  and  $D > 0.04\text{km}$  then
12:     $t_{\text{index}} \leftarrow t_{\text{index}} + 1$       ▷ Distância percorrida em trajetória de atracação, nova
        trajetória
13:  end if
14:   $t_{\text{index}}(s) \leftarrow t_{\text{index}}$            ▷ Rotula a linha atual com a trajetória
15: end for

```

Na Figura 3.18 temos o produto do Algoritmo 3 para o navio SOL DO BRASIL, em que as 11 trajetórias identificadas anteriormente foram agregadas em somente uma, que resume o tempo em que o navio esteve parado em uma mesma posição. Note que, embora o número do ID das trajetórias tenham mudado, as características das trajetórias de navegação não foram alteradas.

Figura 3.18: Trajetórias do navio SOL DO BRASIL após da aplicação do Algoritmo 3

	NAME	IMO	TINDEX3	INIT	END	AVG_SPEED	ROWS	DIST
20	SOL DO BRASIL	9018658	21	2024-06-05 09:03:23	2024-06-05 09:58:22	3.746153846153846	39	2.683193522222962
21	SOL DO BRASIL	9018658	22	2024-06-05 10:01:32	2024-06-08 07:50:46	0.0	257	8.02777782242436e-07
22	SOL DO BRASIL	9018658	23	2024-06-08 08:04:22	2024-06-08 08:15:53	6.966666666666667	9	0.5831012500000827
23	SOL DO BRASIL	9018658	24	2024-06-08 09:37:58	2024-06-08 10:18:40	14.658333333333331	12	25.664285680553817

Fonte: Elaboração própria

Dessa forma, o modelo de reconhecimento de trajetórias é capaz trajetórias de atração que foram produzidas no passo anterior, unindo todo o período em que o navio esteve atracado em uma única trajetória que poderá ser utilizado para identificar os eventos de atração e desatração de uma embarcação, sem interferir nas trajetórias de movimentação que estavam corretamente rotuladas.

4 RESULTADOS

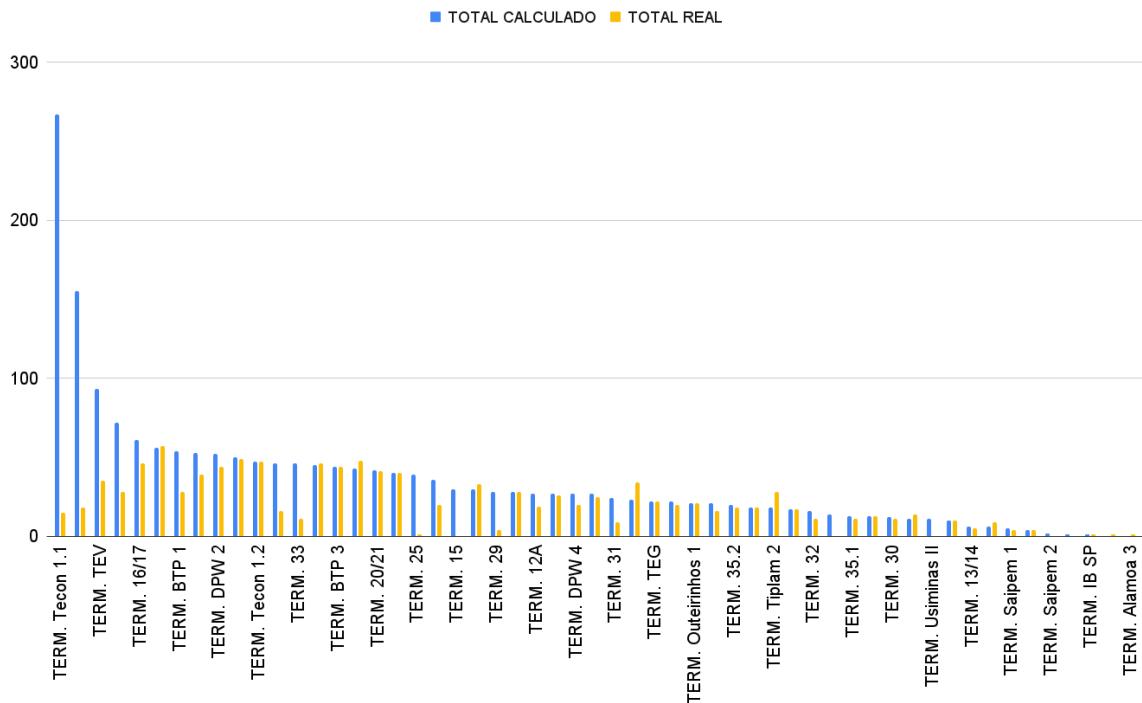
4.1 Qualidade dos rótulos preliminares

A partir dos dados de posição com os rótulos iniciais, antes da aplicação dos algoritmos de reconstrução de trajetória, foi criada uma tabela com os registros de horários de entrada e saída de cada navio dos terminais e os resultados foram comparados com boletins de atracações fornecidos pela autoridade portuária. Os boletins, apesar de gerados manualmente, são registros oficiais, e foram utilizados como "fonte da verdade" para estimar a qualidade dos cálculos de atracação.

A partir desta primeira rotulação, foram calculadas 1.896 operações de atracação no período em análise (entre os dias 14 de maio de 2024 e 19 de agosto de 2024), realizadas por 864 navios diferentes em 53 dos terminais do Porto de Santos, sendo o terminal com mais atracações o Terminal Tecon1.1, com 270 atracações.

Estes resultados destoam fortemente dos dados dos boletins, que registram 1.126 operações de 869 navios em 50 terminais, sendo o Terminal Tecon 2 o com mais atracações, com 57 operações. A figura 4.1 evidencia que a diferença entre o número de atracações calculadas e as reais é alto em quase todos os terminais. Além disso, percebe-se que as distribuições também são distintas, sendo que uma grande parte dos terminais com muitas atracações identificadas pelo rotulamento na realidade tiveram poucas operações registradas no período pelos boletins.

Figura 4.1: Comparação entre o número de atracações calculadas a partir do rotulamento e as reais, por terminal



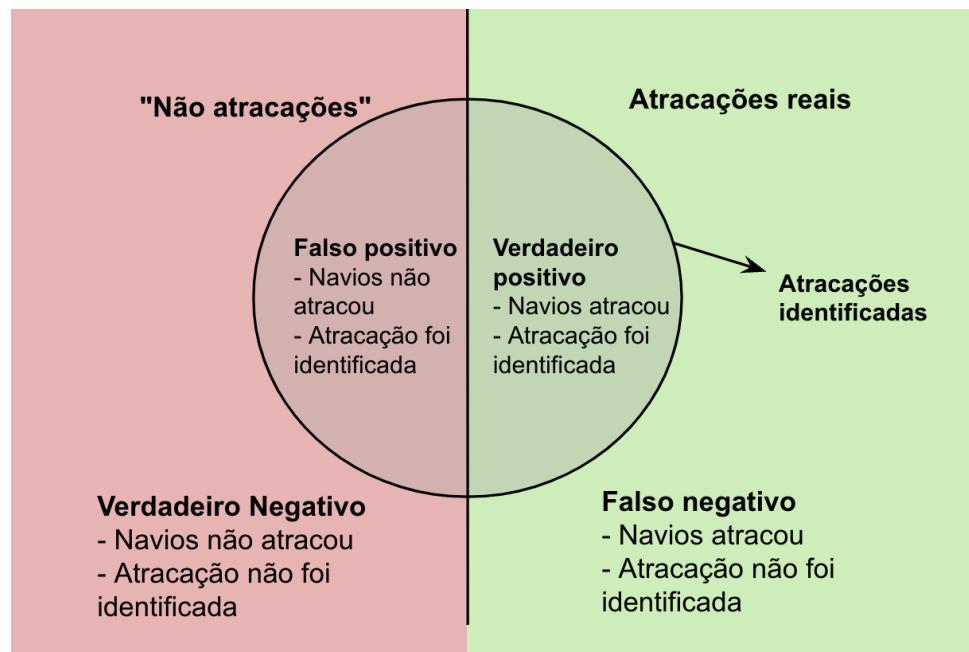
Fonte: Elaboração própria

Para melhor quantificar a qualidade das atracações estimadas, estas foram divididas segundo as seguintes definições:

- **Falsos Negativos:** casos em que a atracação estava presente nos boletins (atração real), mas não na tabela calculada de atracação (atração não identificada);
- **Falsos Positivos:** casos em que a atracação estava presente na tabela calculada de atracação, mas não nos boletins;
- **Verdadeiros Positivos:** casos em que a atracação estava presente tanto na tabela calculada de atracação quanto nos boletins;

A figura 4.2 representa a relação lógica entre as definições de resultados. Note que, neste caso, não há sentido em calcular os Verdadeiros Negativos, uma vez que não há como quantificar as "não atracações" tanto dos eventos reais quanto dos calculados.

Figura 4.2: Representação gráfica da relação dos resultados



Fonte: Elaboração própria

A partir destas definições, foi possível chegar à contagem da tabela 4.1.

Tabela 4.1: Resultados das atrações por tipo a partir do rotulamento inicial

Atrações (Rotulamento inicial)	%
Falsos Negativos	3.16%
Falsos Positivos	43.48%
Verdadeiros Positivos	53.36%
Total	1992
	100%

Fonte: Elaboração própria

Com estes números foram calculadas métricas para medir a qualidade de modelos de classificação. A partir delas é possível entender a performance do método de classificação e quais os principais tipos de erros gerados para então adequar o procedimento.

Tabela 4.2: Métricas de qualidade para o rotulamento inicial

Atracções (Rotulamento inicial)	
Acurácia	54,26%
Precisão	56,06%
Recall	94,40%

Fonte: Elaboração própria

Percebe-se que o *recall* inicial está alto, indicando que o método está eficientemente identificando as atracções que ocorrem nos terminais, visível pelo baixo número de Falsos Negativos em relação aos Verdadeiros Positivos. Por outro lado, a precisão é de apenas 56%, indicando um alto número de Falsos Positivos em relação aos casos de Verdadeiros Positivos pois o método estava inicialmente identificando mais atracções do que as que ocorrem na realidade.

Essas "falsas atracções" se traduziam no alto número de trajetórias geradas no primeiro momento, e por este motivo, durante a tratativa por meio dos algoritmos de trajetória, o foco foi em diminuir os casos de Falsos Positivos para melhorar a precisão, e por consequência a acurácia total do modelo.

Calcular as métricas de qualidade para a primeira tentativa de rotulação foi importante para identificar que os rótulos não estavam confiáveis, como se evidencia na acurácia, sendo necessários mais tratamentos que melhorassem o processo antes de utilizar os rótulos como dados para analisar a eficiência do Porto de Santos. Utilizar os rótulos com a baixa acurácia poderia levar a conclusões imprecisas, gerando possíveis retrabalhos e má alocação de recursos.

4.2 Resultados dos algoritmos de trajetória

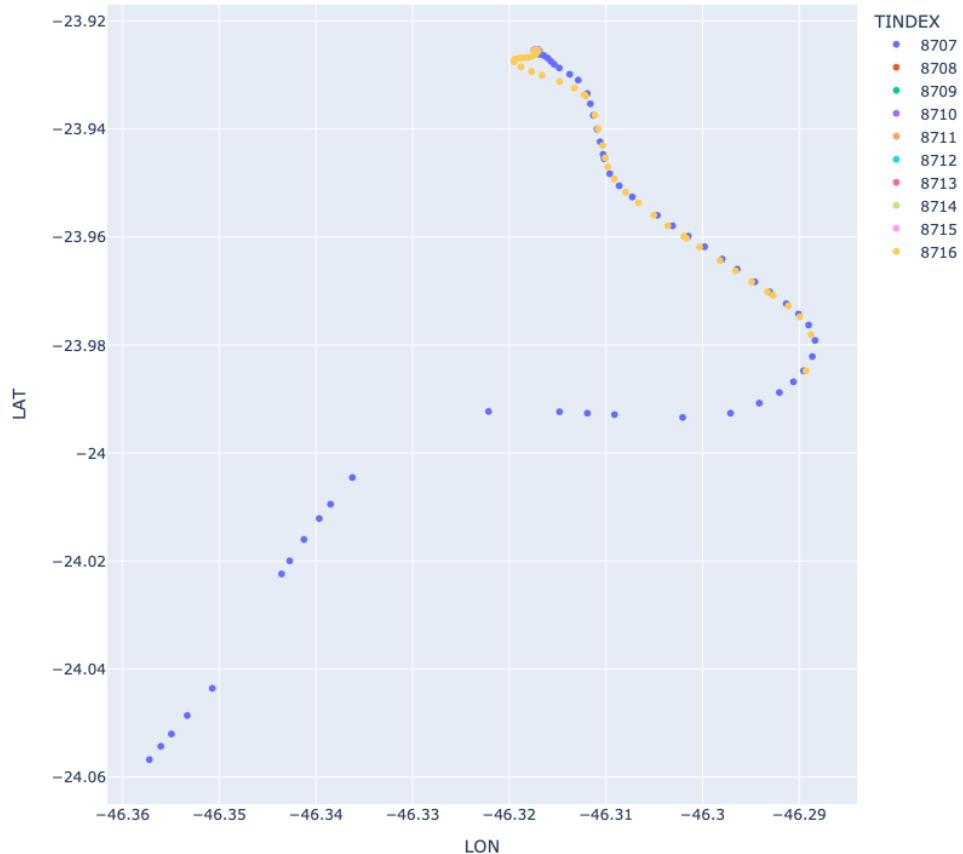
Conforme mencionado anteriormente, o objetivo dos algoritmos de trajetória foi de auxiliar o processo de rotulação por meio da separação dos dados de posição do navio em três grandes períodos: navegação para entrada, atracção (ou operação) e navegação para saída. Os dados foram inicialmente quebrados em inúmeras trajetórias curtas e incompletas, porém os algoritmos seguintes foram capazes de corrigir a identificação das trajetórias, o que se traduz na diminuição do número de trajetórias e no aumento do

número de dados por trajetória.

O Algoritmo 1 cria as trajetórias a partir da base de dados com as posições e velocidades, aos quais não foi aplicado nenhum tratamento ainda. Assim, neste primeiro momento foi observado que um mesmo navio com somente uma parada apresentava diversas trajetórias com poucos dados cada uma, devido às oscilações de velocidade que causam quebras de trajetória conforme os critérios do Algoritmo 1.

Na figura 4.3 é exemplificado este fenômeno pelas trajetórias criadas pelo Algoritmo 1 para o navio APL TURKEY. O algoritmo dividiu os 430 registros de posição, enviados para o sistema de AIS entre os dias 08 e 09 de Junho de 2024 em 10 diferentes trajetórias. As trajetórias 8707 e 8716 são de movimentação, representando a entrada e a saída do canal, respectivamente. As trajetórias 8708 a 87015 são 8 trajetórias de atracação, representando o tempo que o navio esteve parado em um único ponto.

Figura 4.3: Trajetórias do navio APL TURKEY criadas pelo Algoritmos 1



Fonte: Elaboração própria

A figura 4.4 é uma ampliação somente com as trajetórias de parada, onde é possível melhor visualizar que os dados de uma mesma atracação estão divididos em diversas trajetórias.

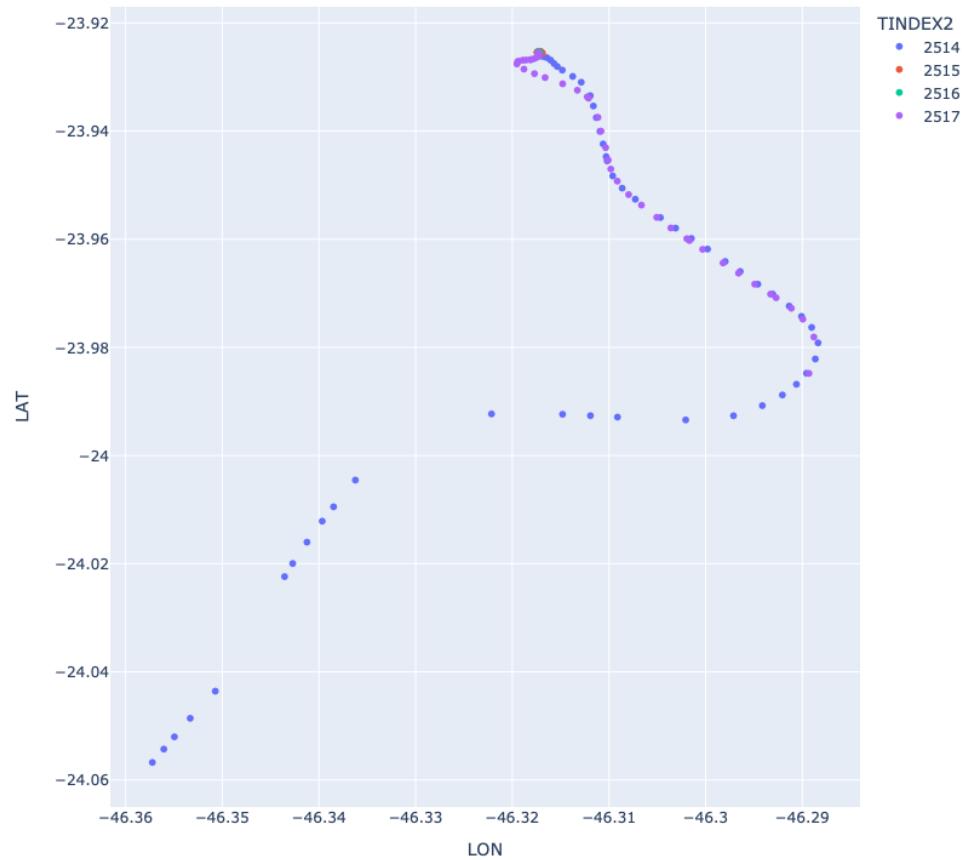
Figura 4.4: Ampliação das trajetórias somente de parada do navio APL TURKEY criadas pelo Algoritmos 1



Fonte: Elaboração própria

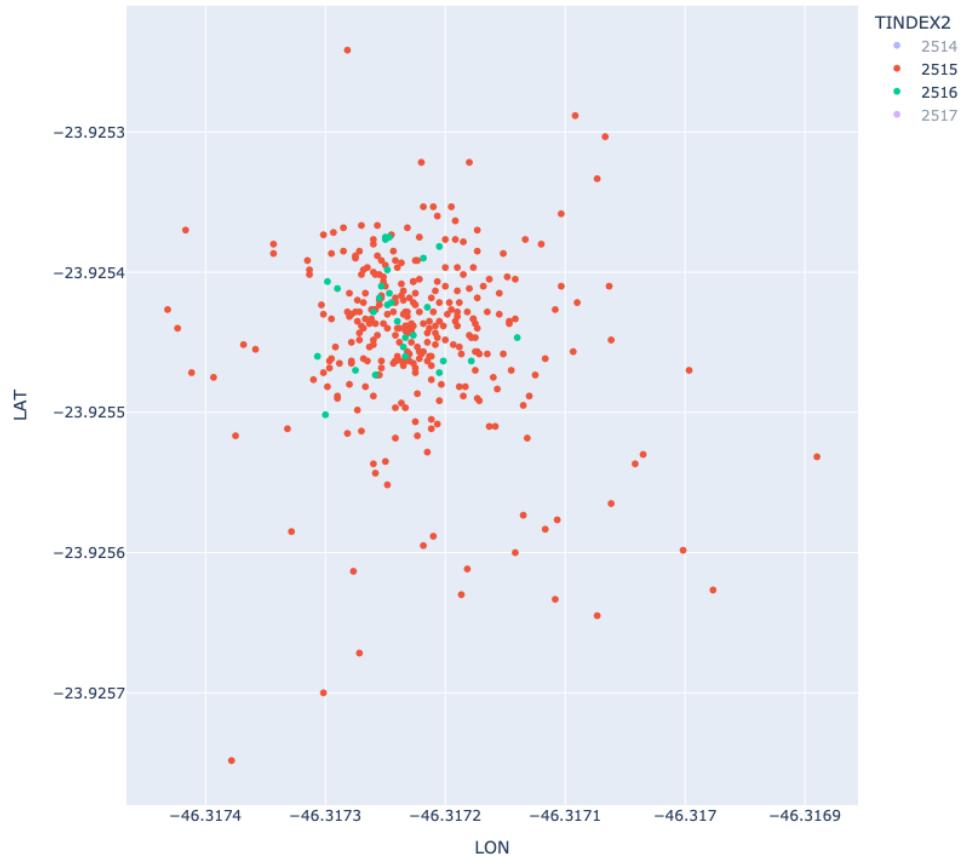
Com a aplicação do Algoritmo 2, as velocidades são ajustadas para eliminar as oscilações dos registros de velocidade enquanto o navio está atracado, eliminando trajetórias de movimentação inválidas. A figura 4.5 representa esta mudança: o número de trajetórias de parada diminuiu de 10 para somente 4. As trajetórias 2514 e 2517 agora representam as trajetórias de ida e volta no canal, cujos dados permanecem intactos em relação ao passo anterior, mas, na figura 4.6, focando novamente nas trajetórias de atracação, observamos que agora há somente duas trajetórias.

Figura 4.5: Trajetórias do navio APL TURKEY criadas pelo Algoritmos 1



Fonte: Elaboração própria

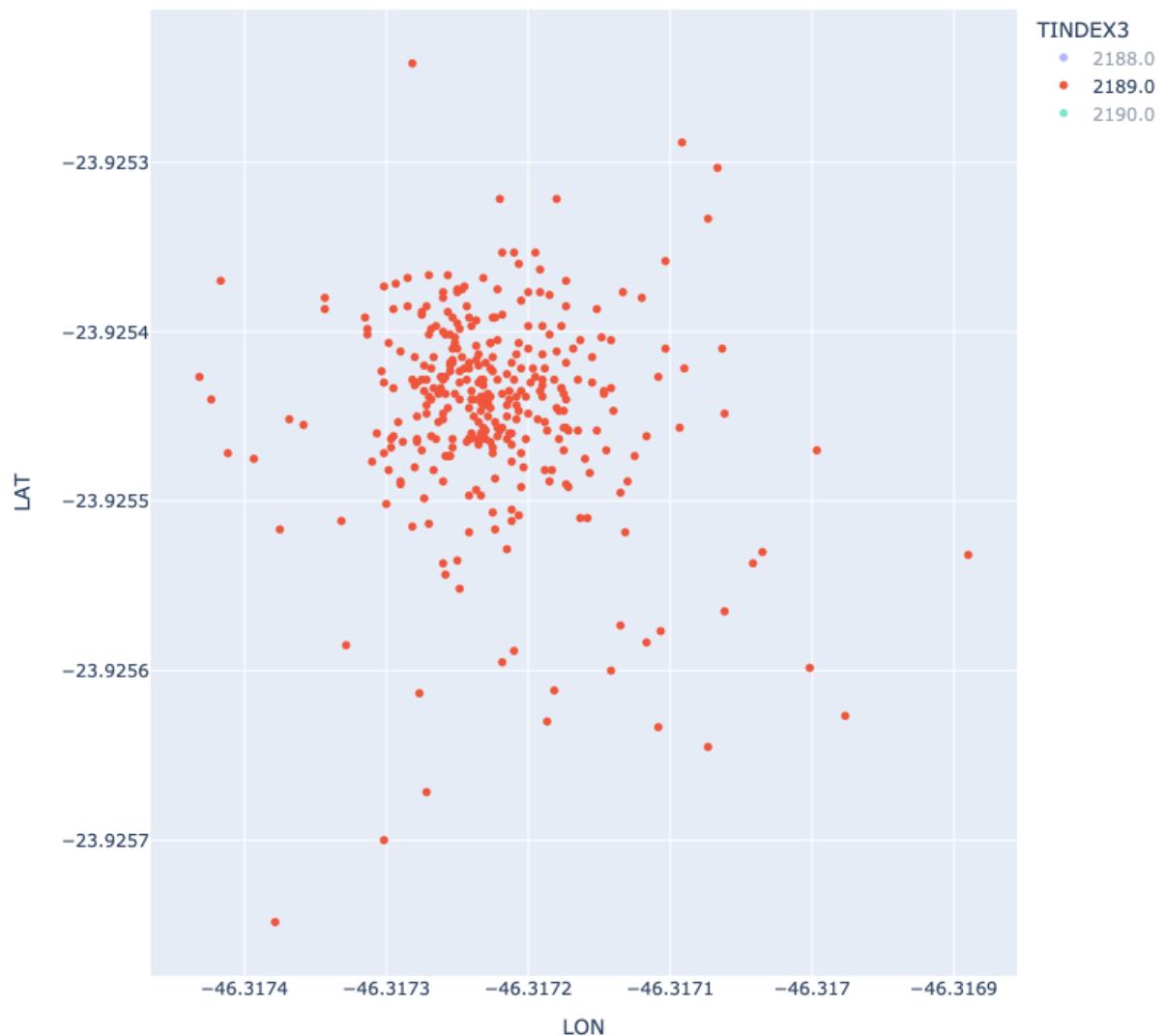
Figura 4.6: Ampliação das trajetórias somente de parada do APL TURKEY criadas após os ajustes do Algoritmos 2



Fonte: Elaboração própria

Por fim, com o Algoritmo 3, é possível unir as duas trajetórias consecutivas nas quais o navio esteve parado no mesmo terminal em uma única trajetória que contempla todo o período de operação. Na figura 4.7, observa-se que as duas trajetórias de atracação, que não haviam sido unidas anteriormente devido ao delta de tempo, passam a ser uma única trajetória (2189), pois o Algoritmo 3 identificou que não houve uma mudança de posição significativa entre elas.

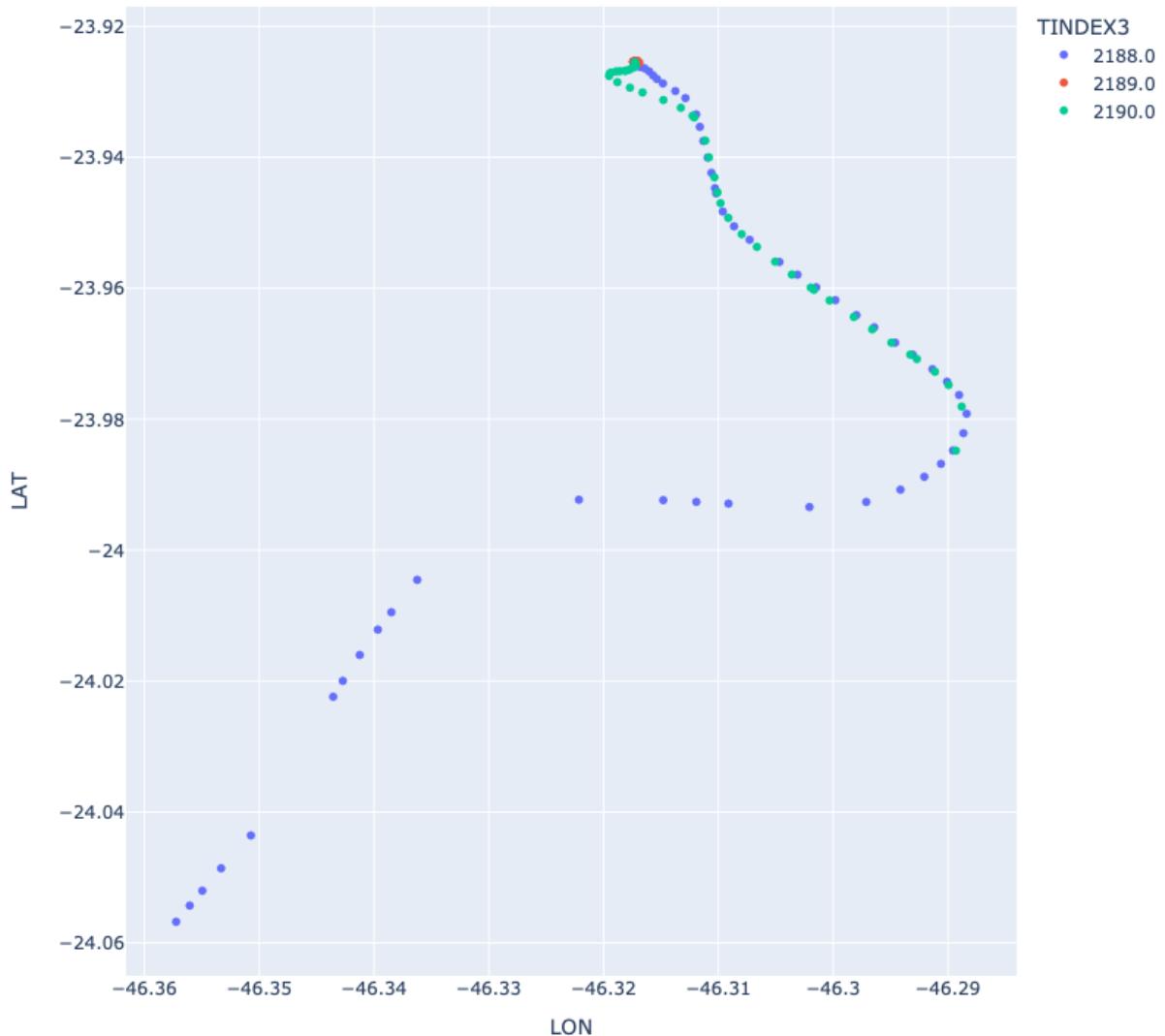
Figura 4.7: Trajetória única de parada do APL TURKEY após aplicação do Algoritmo 3



Fonte: Elaboração própria

Assim, por meio da aplicação das operações descritas, foi possível organizar os dados de posição do navio APL TURKEY em sua viagem ao longo do canal, dividindo-os somente entre as três trajetórias propostas: entrada, atracação e saída, conforme visto na figura 4.8.

Figura 4.8: Trajetórias do navio APL TURKEY criadas pelo Algoritmo 3



forma mais contínua por meio da eliminação de ruídos na velocidade instantânea.

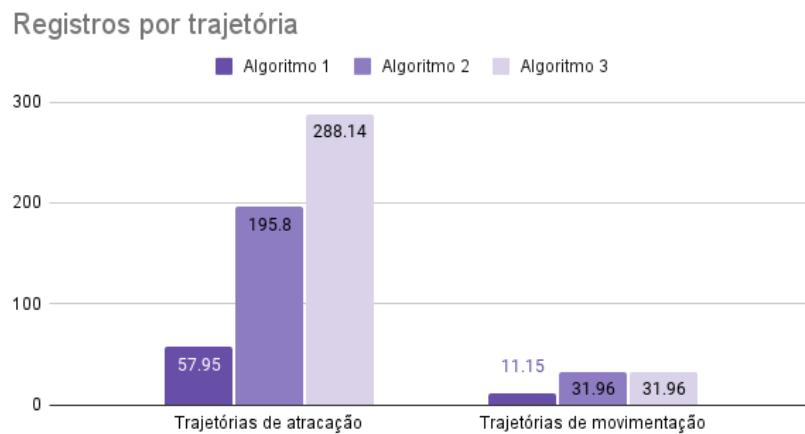
Já comparando as trajetórias do Algoritmo 3 com as do Algoritmo 2, há uma divergência de resultados quando separamos as trajetórias de atracação das de navegação. As de atracação apresentaram resultado similar ao anterior, com diminuição do número de trajetórias e da média de trajetórias por navio e aumento do número de registros por trajetória. Este resultado é esperado, pois o Algoritmo 3 agrupa trajetórias de atracação subjacentes. Já as trajetórias de navegação não apresentaram diferença em relação às do Algoritmo 2, o que também é lógico, visto que o Algoritmo 3 associa apenas trajetórias de atracação.

Figura 4.9: Número total de trajetórias geradas após a aplicação de cada algoritmo



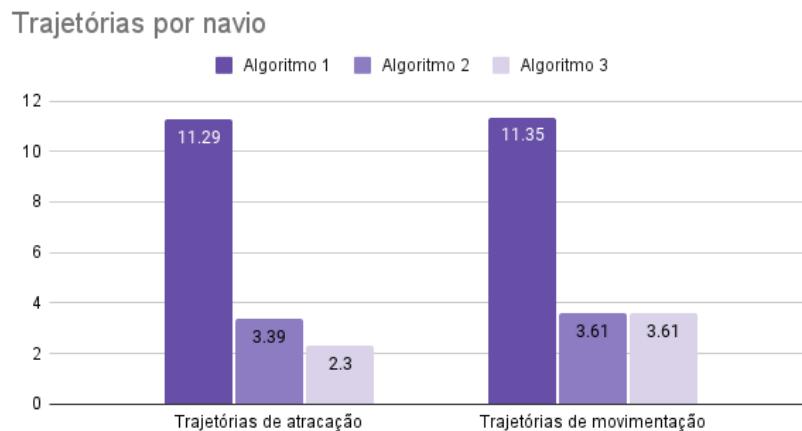
Fonte: Elaboração própria

Figura 4.10: Número médio de dados registros por trajetória após a aplicação de cada algoritmo



Fonte: Elaboração própria

Figura 4.11: Número médio de trajetória por navio na base após a aplicação de cada algoritmo



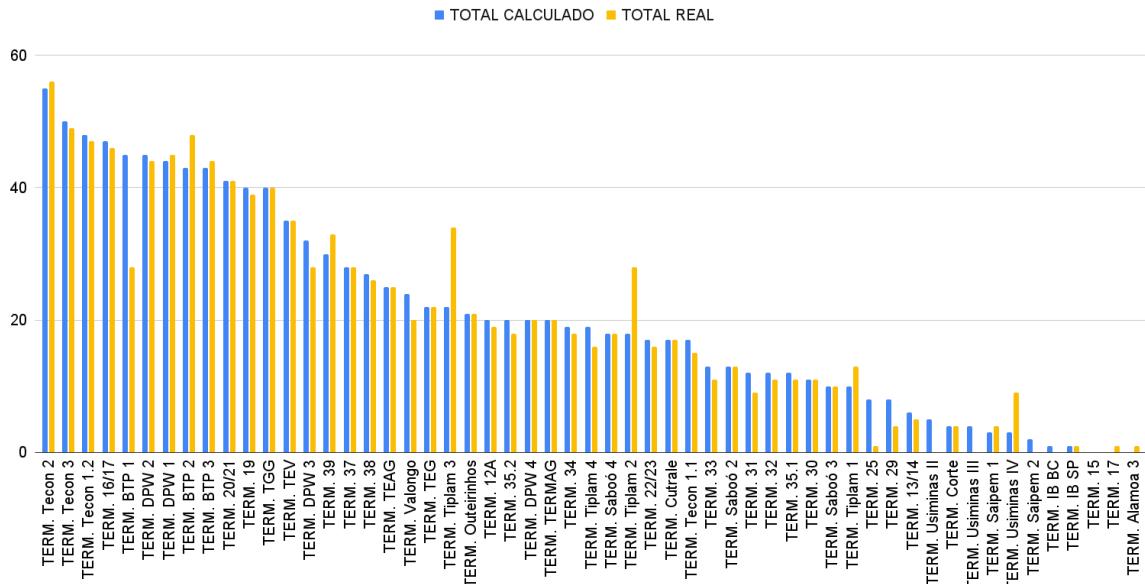
Fonte: Elaboração própria

4.3 Resultados da rotulação com os dados tratados

Uma vez desenvolvido e validado o processo de identificação de trajetórias, foi aplicado mais uma vez o algoritmo de rotulamento de códigos, desta vez aos dados com as trajetórias resultantes do Algoritmo 3, e uma segunda tabela com os registros de horários de entrada e saída de cada navio dos terminais foi gerada. As operações foram comparadas com os boletins oficiais da mesma forma que anteriormente e calculou-se as métricas de qualidade dos novos rótulos.

Foram calculadas 1.150 operações no período analisado, realizadas por 861 navios diferentes em 51 dos 65 berços do Porto de Santos em análise. A relação de atracções por terminal está na figura 4.12, sendo o terminal com mais atracções o Terminal Tecon 2, com 55 atracções. É possível observar que, comparado com a primeira tentativa de identificar atracções, a distribuição entre os terminais e o número de atracções é muito mais similar ao números de referência dos boletins.

Figura 4.12: Comparação entre o número de atracações calculadas a partir do rotulamento e as reais, por terminal



Fonte: Elaboração própria

Em relação à classificação das atracações calculadas quanto à sua relação com as atracações reais, foi obtida a tabela 4.3

Tabela 4.3: Resultados das atracações por tipo a partir do rotulamento após os tratamentos

Atracações (Rotulamento após tratamentos)		%
Falsos Negativos	64	5.27%
Falsos Positivos	91	7.49%
Verdadeiros Positivos	1059	87.24%
Total	1214	100%

Fonte: Elaboração própria

As métricas de qualidade resultantes destes valores estão disponíveis na tabela 4.4:

Tabela 4.4: Métricas de qualidade para o rotulamento após os tratamentos

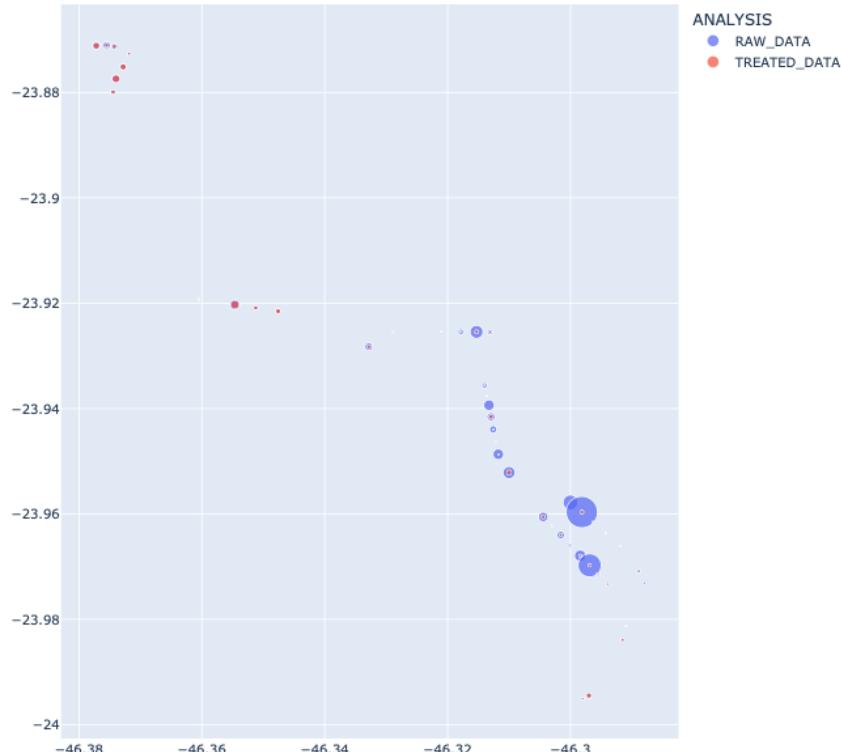
Atrações (Rotulamento inicial)	
Acurácia	87,23%
Precisão	92,09%
Recall	94,30%

Fonte: Elaboração própria

Embora a melhora nos três indicadores seja expressiva, destaca-se a melhora na precisão, demonstrando que o método foi eficiente em diminuir os casos de falso positivo causados pelas oscilações de velocidade.

A figura 4.13 é uma representação geográfica dos erros de falso positivo para os terminais do Porto de Santos, podendo-se observar uma diminuição expressiva de erros em praticamente todos os locais.

Figura 4.13: Comparaçāo entre os erros de falso positivo antes e depois do tratamento

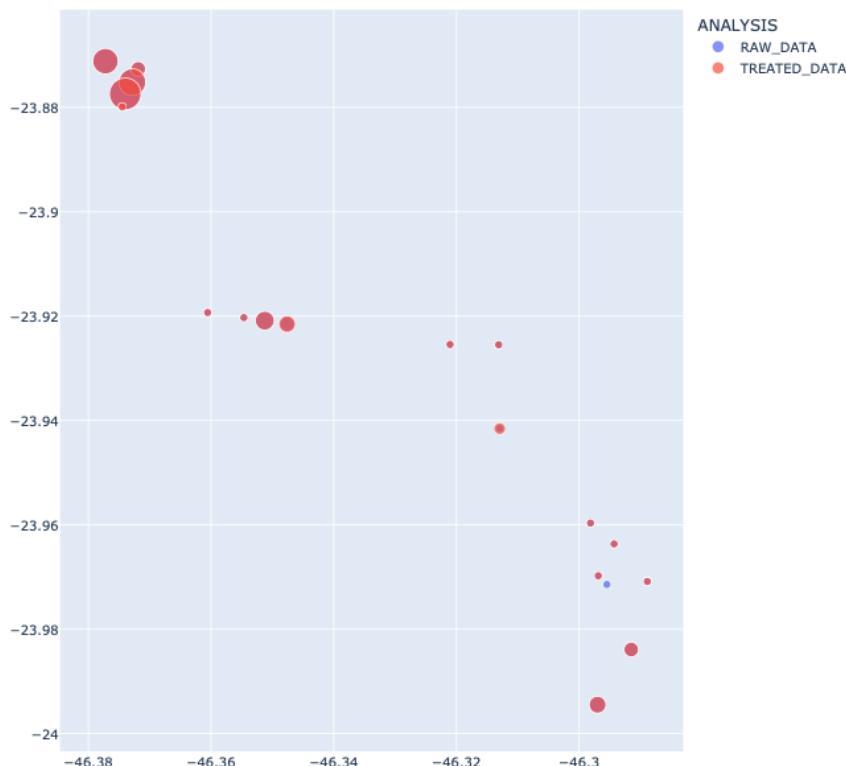


Fonte: Elaboração própria

A figura 4.14 é a mesma representação, porém comparando os casos de falso negativo

de cada terminal antes e depois da aplicação dos algoritmos de reconhecimento de trajetória, onde é possível notar que há uma sobreposição entre os círculos cujo tamanho é proporcional ao número de erros para os dados tratados e os dados do rotulamento inicial, significando que houve pouca melhorar nestes casos.

Figura 4.14: Comparação entre os erros de falso negativo antes e depois do tratamento



Fonte: Elaboração própria

Assim, é possível concluir a aplicação dos algoritmos de trajetória e a nova rotulação foram efetivos em melhorar a qualidade dos rótulos de posição, tornando os dados gerados mais confiáveis para serem utilizados em análises da eficiência portuária.

Devido ao método construído, que foca em corrigir as velocidades não nulas durante a atracação, a diminuição dos erros está ligada ao tipo de erro predominante em um terminal (maior diminuição nos terminais em que haviam mais falsas atracações), e aparenta ter pouca relação com a localização geográfica dos terminais.

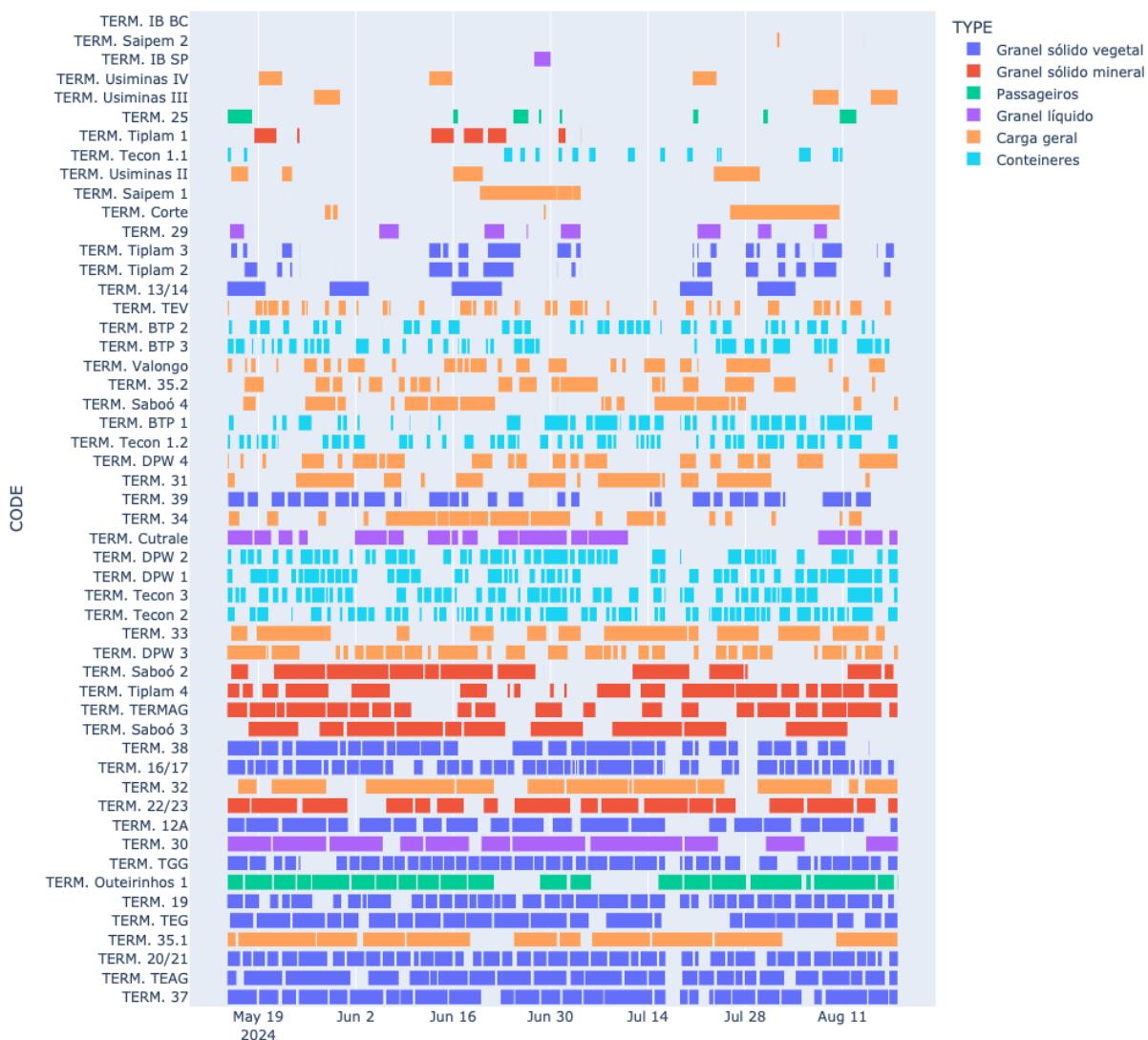
4.4 Análise das atracações e eficiência dos terminais

A partir do método apresentado e validado anteriormente foi possível visualizar as atracações que ocorreram no período em análise e calcular algumas métricas de uso dos

terminais que podem auxiliar na melhor compreensão do funcionamento e da eficiência atual do Porto de Santos, e no estudos de melhorias. Foram calculados o tempo de permanência média, a taxa de ocupação e o tempo médio entre operações para os 51 terminais nos quais foram identificadas atracções, além das métricas gerais para o Porto.

Os resultados estão disponíveis na íntegra no Anexo B, porém diversas conclusões já podem ser extraídas da figura 4.15, onde é apresentado o gráfico de ocupação dos terminais do Porto de Santos. Neste tipo de gráfico, as barras representam os momentos em que um navio esteve atracado no terminal, indicando que o mesmo estava operando. Barras distintas indicam atracções diferentes e as cores indicam o tipo de carga movimentada em um terminal. Dessa forma, a partir da análise do tamanho e da quantidade de espaço ocupado pelas barras, é possível extrair informações relevantes a respeito da taxa de ocupação, do tempo de permanência e do tempo entre operações de cada localização.

Figura 4.15: Gráfico de ocupação dos terminais no Porto de Santos



Fonte: Elaboração própria

Percebe-se que há uma alta variação na ocupação do espaço pelas barras, sendo que em alguns locais quase não foram identificadas operações, enquanto em outras o terminal esteve ocupado em quase todo o tempo, sinalizando que há um desbalanço na eficiência e/ou demanda entre os terminais. Outro fenômeno interessante observado no gráfico é a diferença no tamanho das barras pois há terminais com atracções mais curtas (barras menores), e outros barras mais compridas, representando operações mais longas. Em um mesmo terminal, a variação no tamanho das barras não é tão grande, indicando que a duração da atracção em um mesmo terminal é estável de forma geral.

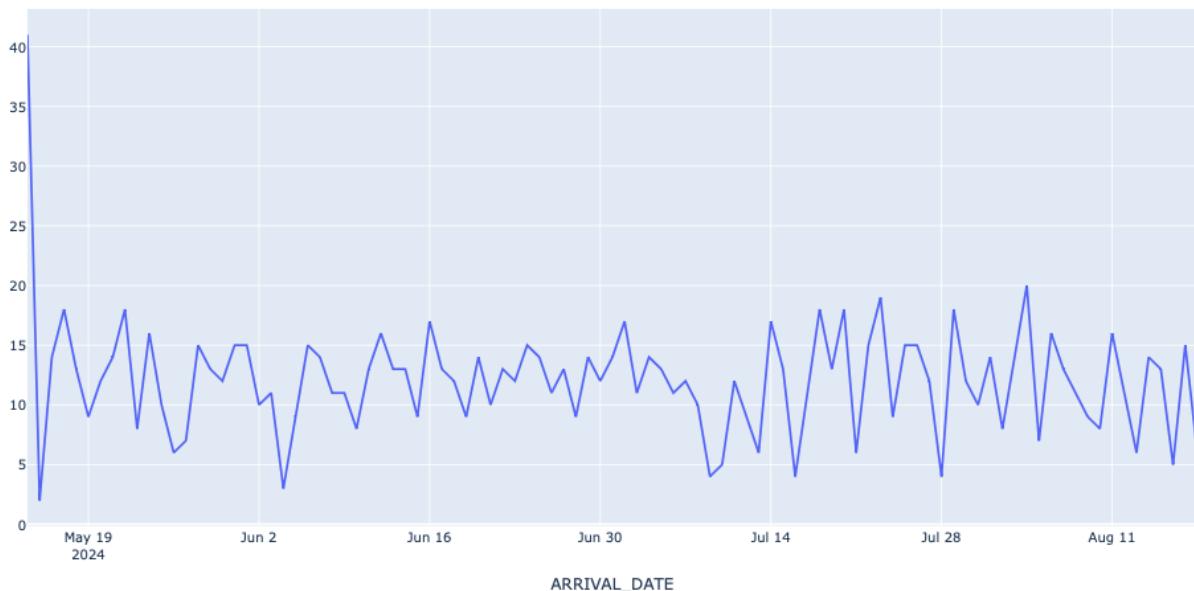
Além disso, a partir das cores é possível associar os terminais de determinados tipos de carga a maiores ocupações: os terminais de contêineres e granel sólido vegetal apresentam menores ocupações que os de granel sólido vegetal, por exemplo.

A seguir é apresentada uma análise mais detalhada da eficiência geral para o agregado de todas as operações identificadas no Porto de Santos, da eficiência por terminal, e da eficiência por tipo de carga.

4.4.1 Análise geral das atracações no Porto de Santos

Observando primeiramente o número de atracações por dia, não é possível extrair uma sazonalidade ou tendência, indicando que as entradas se mantiveram estáveis ao longo do tempo. No primeiro dia em análise (14 de maio de 2024), é notável um registro maior de atracações, pois os navios que já se encontravam atracados e operando em terminais nesta data foram contabilizados como atracações no dia 14, mesmo que não tenham efetivamente realizado a operação de atracar neste dia.

Figura 4.16: Atracações por dia no Porto de Santos

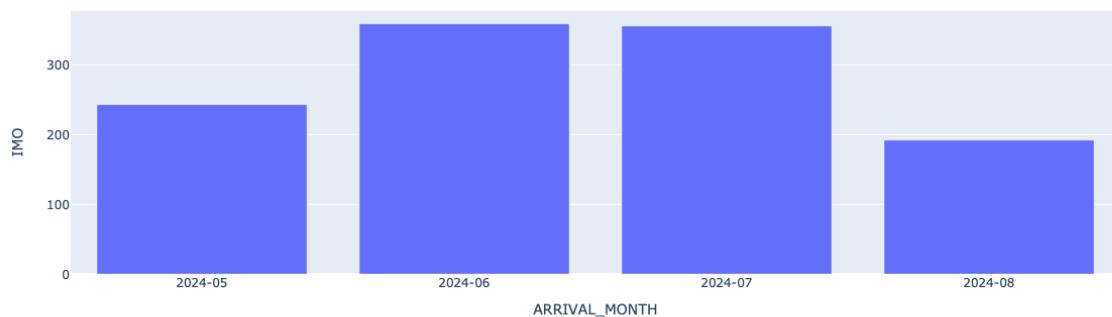


Fonte: Elaboração própria

Observando o número de atracações por mês, a quantidade de atracações também se manteve estável nos dois meses completos nos quais foram coletados dados (Junho e Julho de 2024). Os meses de Maio e Agosto apresentam menos atracações, porém os dados não estão completos para estes meses (em maio foram coletados dados a partir do dia 14 e em

agosto até o dia 19), sendo portanto a diminuição de operações identificadas um fenômeno esperado.

Figura 4.17: Atracções por mês no Porto de Santos

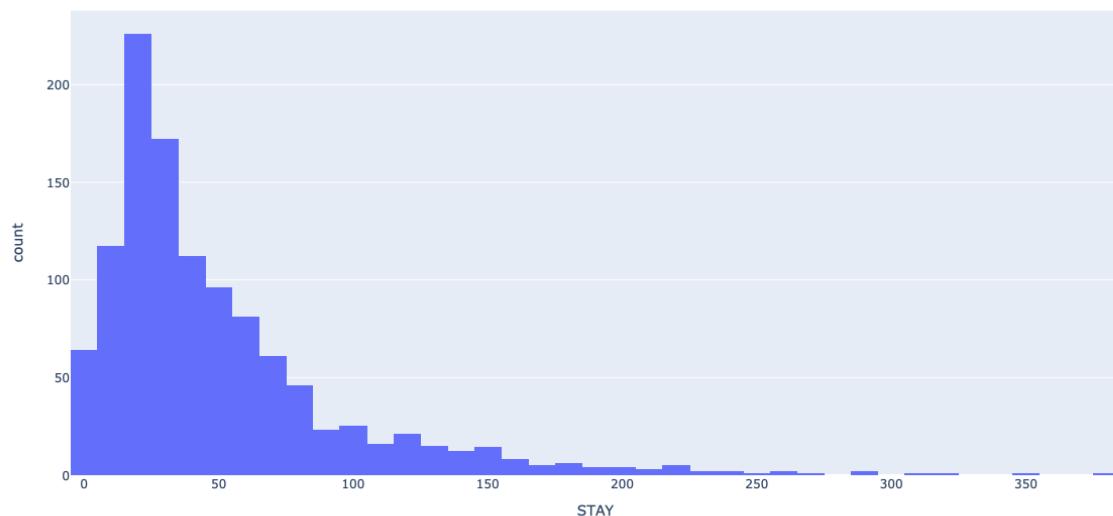


Fonte: Elaboração própria

Passando para a análise das métricas de uso para o agregado dos terminais nos quais foram identificadas atracções no período analisado, foram acumuladas 57.811 horas de operação, resultando em uma taxa de ocupação geral TO=48,0%. Embora taxas de ocupação muito altas sejam indicativos de congestionamentos, a taxa observada para o Porto de Santos está aquém do esperado (<50%) (MWASENGA, 2012), indicando uma subutilização dos recursos e evidenciando que há espaço para melhorias nas operações.

Quanto ao tempo de permanência médio das embarcações nos terminais, obteve-se uma média de 50 horas de operação, a partir do momento em que o navio atraca (velocidade diminui para zero), até a desatracção (ganha velocidade maior que zero), com um desvio padrão de 48 horas.

Figura 4.18: Distribuição do tempo de permanência das embarcações no Porto de Santos



Fonte: Elaboração própria

Observando o histograma do tempo de permanência, há uma assimetria à direita, sendo que 75% dos valores é menor que 65 horas, indicando que atracações muito longas não são tão frequentes, embora existam operações *outliers* de até 379 horas.

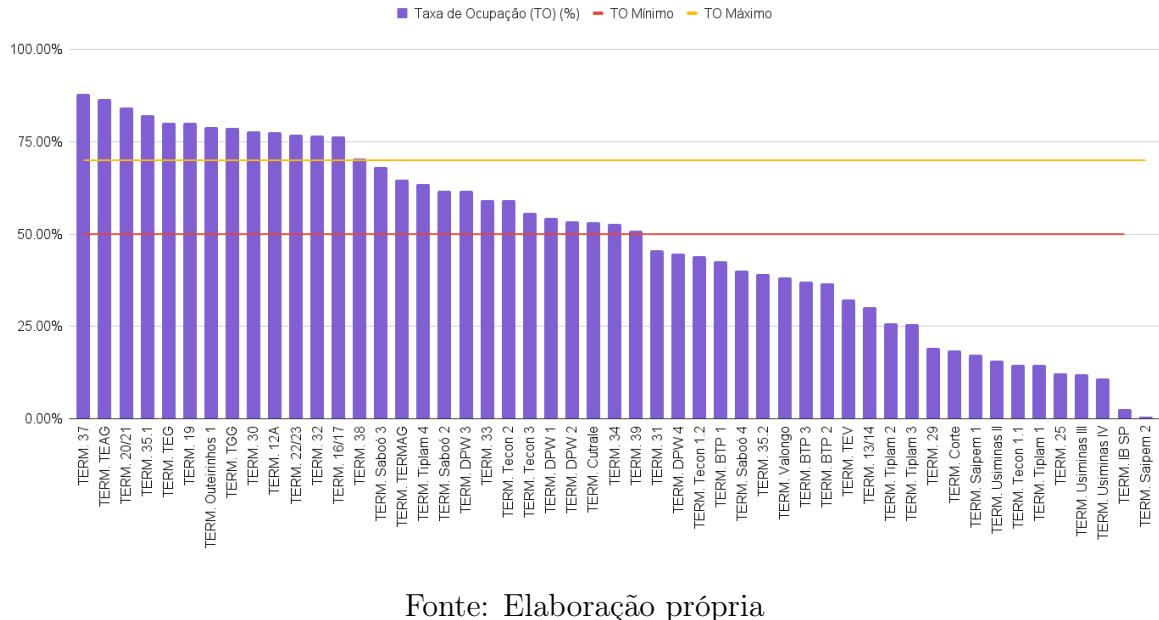
O tempo médio entre operações para o agregado dos terminais foi de 54,5 horas, ou seja, houve um intervalo médio de mais de dois dias sem operações entre uma atracação e a próxima nos terminais, indicando uma subutilização dos recursos.

4.4.2 Análise da eficiência por terminal

4.4.2.1 Taxa de Ocupação (TO)

Na figura 4.19, percebe-se que a TO variou bastante entre os terminais, sendo o Terminal 37 o com a maior taxa de ocupação (88%), e o terminal Saimpem 2 o com menor TO (0,6%, decorrente de somente duas operações), indicando que há terminais com maior eficiência operacional, e/ou maior demanda.

Figura 4.19: Taxa de Ocupação dos terminais do Porto de Santos



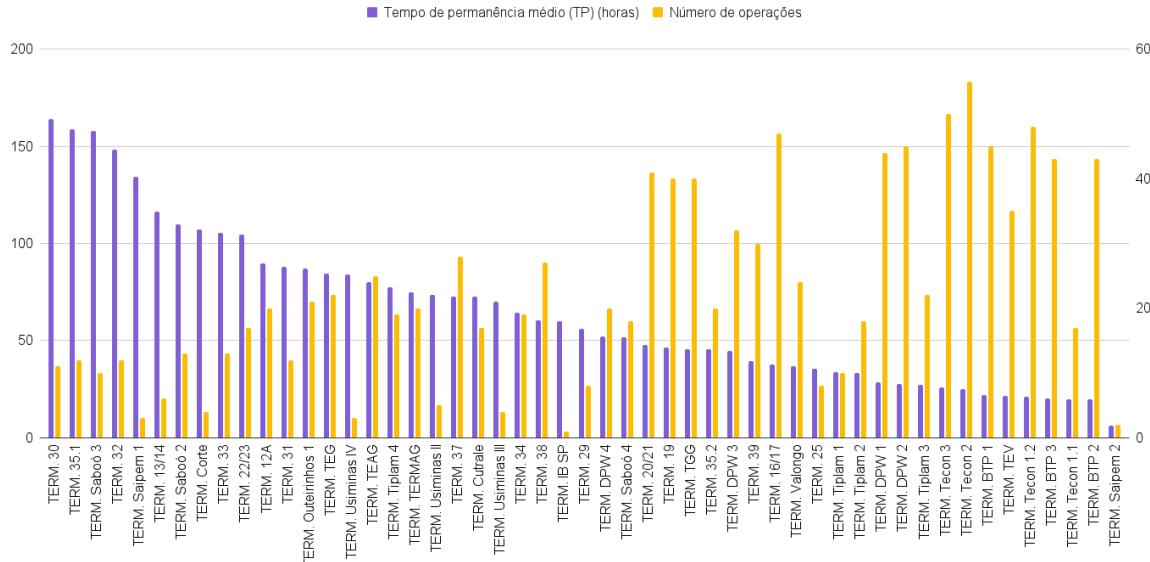
Fonte: Elaboração própria

Somente 14 dos 51 terminais tiveram uma taxa de ocupação na faixa recomendada em Mwasenga (2012), sendo que 13 tiveram taxas superiores, indicando possíveis congestionamentos e problemas no nível de serviço, e os 24 restantes apresentaram taxas inferiores, indicando uma possível demanda baixa ou operações pouco eficientes.

4.4.2.2 Tempo de Permanência Médio (TP)

Além disso, o tempo de permanência médio também apresentou variação alta entre o terminal com maior média (Terminal 30, com 164 horas de operação em média) e o terminal com menor média (Terminal Saipem 2, com 6,4 horas). Na figura 4.20, observa-se uma relação inversa entre o TP e o número de atracações, ou seja, terminais com menos atracações apresentam tempos de permanência médio maiores pois as operações realizadas eram mais complexas e demoradas, enquanto terminais com alto número de atracações têm TP baixo, pois as operações eram mais simples e rápidas. Embora haja uma tendência a este comportamento, uma exceção é o Terminal Saipem 2, onde foram identificadas somente duas atracação que totalizaram 6,4 horas, resultando em baixo número de operações e baixo TP.

Figura 4.20: Relação entre TP e número de atracções dos terminais do Porto de Santos



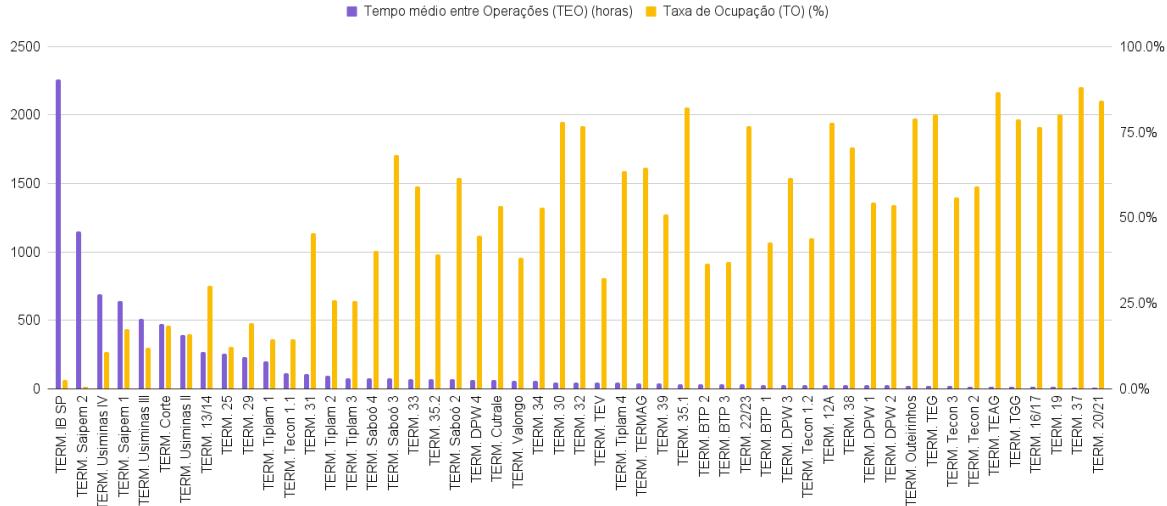
Fonte: Elaboração própria

4.4.2.3 Tempo médio Entre Operações (TEO)

Quanto ao tempo médio entre operações, o menor valor foi do Terminal 20/21 com 8,92 horas, e o maior foi novamente do Terminal IB SP, onde somente uma operação foi registrada, resultando em um TEO igual a todo o tempo sem operações, de 2.256 horas. Estes terminais com poucas e curtas operações causaram que a distribuição do TEO fosse muito desbalanceada, com alguns casos de TEO muito altos devido à baixa utilização destes terminais.

Isso se observa na figura 4.21, onde o TEO se relaciona inversamente com a TO de forma que os terminais com maiores intervalos entre operações de forma geral tem ocupações menores. Ambas as métricas são indicativas que estes locais estão com baixa demanda ou graves problemas de eficiência que causaram que as operações fossem muito escassas.

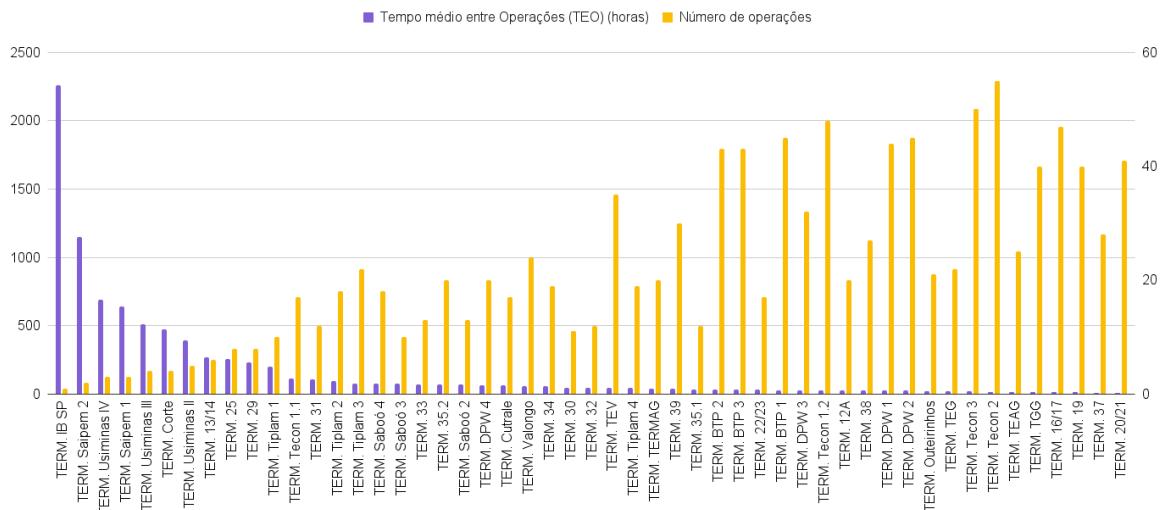
Figura 4.21: Relação entre TEO e TO dos terminais do Porto de Santos



Fonte: Elaboração própria

O mesmo se evidencia na relação entre o TEO e o número de atracções presente na figura 4.22, onde nota-se que os maiores TEOs estão nos terminais com menos operações, indicando a baixa utilização de alguns recursos.

Figura 4.22: Relação entre TEO e número de operações dos terminais do Porto de Santos



Fonte: Elaboração própria

Neste caso é interessante realizar uma análise mais detalhada destas localizações mais problemáticas para entender as causas dos resultados. Algumas hipóteses incluem baixa

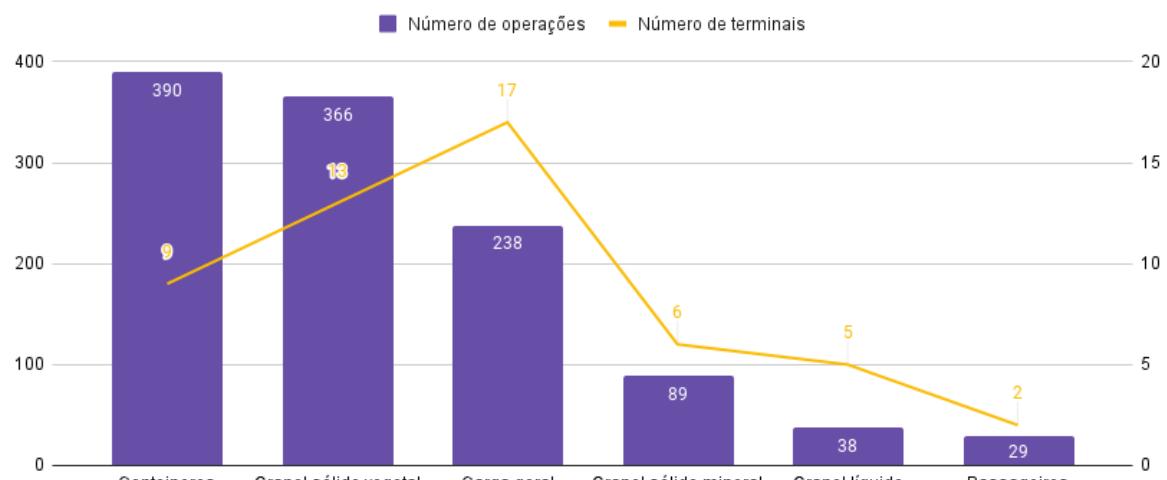
demandas por estes terminais e baixa disponibilidade devido à interdições para manutenção e reparos.

4.4.3 Análise da eficiência por tipo de carga

Além de analisar o Porto de forma geral e os terminais individualmente, outro aspecto interessante é observar como a eficiência das operações varia entre os diferentes tipos de carga movimentados pelos terminais. Para este estudo, os terminais foram classificados em seis categorias de acordo com o tipo de carga que operam: granel sólido vegetal, granel sólido mineral, granel líquido, passageiros, carga geral e contêineres.

As relações detalhadas entre cada terminal e seu respectivo tipo de carga estão disponíveis no Anexo A, e é importante destacar que o número de terminais por tipo de carga não é uniforme: há 17 terminais que movimentam carga geral e 13 que movimentam granel sólido vegetal, enquanto somente 2 terminais de passageiros e 5 de granel sólido foram identificados.

Figura 4.23: Número de operações registradas por tipo de carga



Fonte: Elaboração própria

Analizando o número de operações por tipo de carga, os tipos com menos terminais apresentaram menor número de operações (passageiros, granel sólido mineral e granel líquido). Já contêineres e granel sólido vegetal apresentaram o maior número de operações no período, embora carga geral seja o tipo com maior número de terminais associados.

As métricas de eficiência para cada tipo de carga foram calculadas de duas maneiras distintas: primeiro, pela média das métricas já calculadas para os terminais que operam

aquele tipo de carga (média dos terminais); e segundo, pela média de todas as operações realizadas para o tipo de carga sem distinção do terminal (média das operações).

Enquanto a média dos terminais reflete o desempenho geral dos terminais envolvidos em um determinado tipo de carga, diminuindo o peso do número de operações realizadas por aquele terminal, a média das operações oferece uma visão global de todas as operações do tipo de carga, sem considerar em qual terminal a operação aconteceu. Comparando as duas métricas, é possível identificar caso haja um desbalanço muito grande entre as operações dos terminais de um mesmo tipo de carga.

4.4.3.1 Taxa de Ocupação (TO)

Analizando a definição da taxa de ocupação (tempo total de operação sobre tempo total disponível), percebe-se que o resultado será igual nos dois métodos de cálculo (média dos terminais e médias das operações).

$$TO_T = \frac{\sum_{\theta=1}^{n_T} O_\theta^T}{t_T} \quad (4.1)$$

$$TO_T = \sum_{i=1}^{m_T} \left(\frac{\sum_{\theta=1}^{n_i} O_\theta^i}{t_{\text{análise}}} \right) \quad (4.2)$$

Onde:

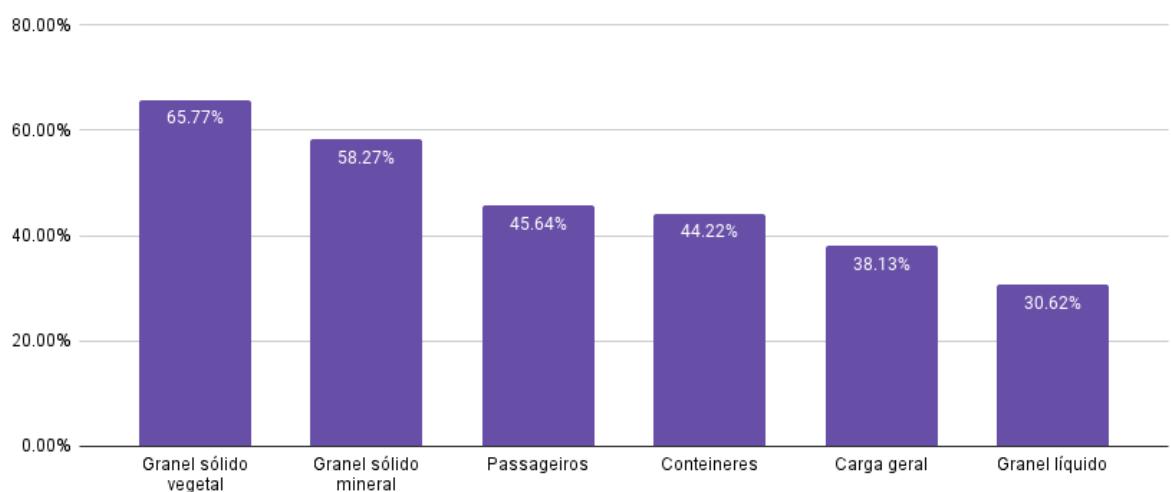
- O_θ^T é o tempo de operação gasto na θ -ésima operação do tipo de carga T
- O_θ^i é o tempo de operação gasto na θ -ésima operação do terminal i
- $t_{\text{análise}}$ é o tempo em análise do estudo (tempo disponível para um terminal)
- t_T é o tempo total disponível para as cargas do tipo T
- n_T é o número de operações identificadas para cargas do tipo T
- n_i é o número de operações identificadas no terminal i
- m_T é o número de terminais que operam cargas do tipo T

Note que $t_T = t_{\text{análise}} \times m_T$, ou seja, o tempo total disponível para o tipo de carga é diretamente proporcional ao número de terminais que operam o tipo de carga. Além disso, claramente o tempo total de operação para todas as operações de um tipo de carga

é igual à soma do tempo de operação de todos os terminais que operam este tipo de carga ($\sum_{\theta=1}^{n_T} O_{\theta}^T = \sum_{i=1}^{m_T} (\sum_{\theta=i}^{n_i} O_{\theta}^i)$). Assim, a igualdade das equações é válida.

O tipo de carga cujos terminais apresentaram maior ocupação foi o de granel sólido vegetal com 65,77%, indicando que os terminais deste tipo de carga são em média mais ocupados, enquanto a menor ocupação foi a de granel líquido, com 30,62%, apontando uma possível menor demanda por este tipo de produto.

Figura 4.24: Taxa de Ocupação (TO) por tipo de carga no Porto de Santos



Fonte: Elaboração própria

Este resultado está de acordo com o gráfico de ocupação da figura 4.15, onde os terminais com maiores taxas de ocupação tem mais espaço preenchido pelas barras e estão mais abaixo no gráfico, enquanto os tipos com menores TOs estão mais acima do gráfico com menos espaço preenchido.

4.4.3.2 Tempo de Permanência Médio (TP)

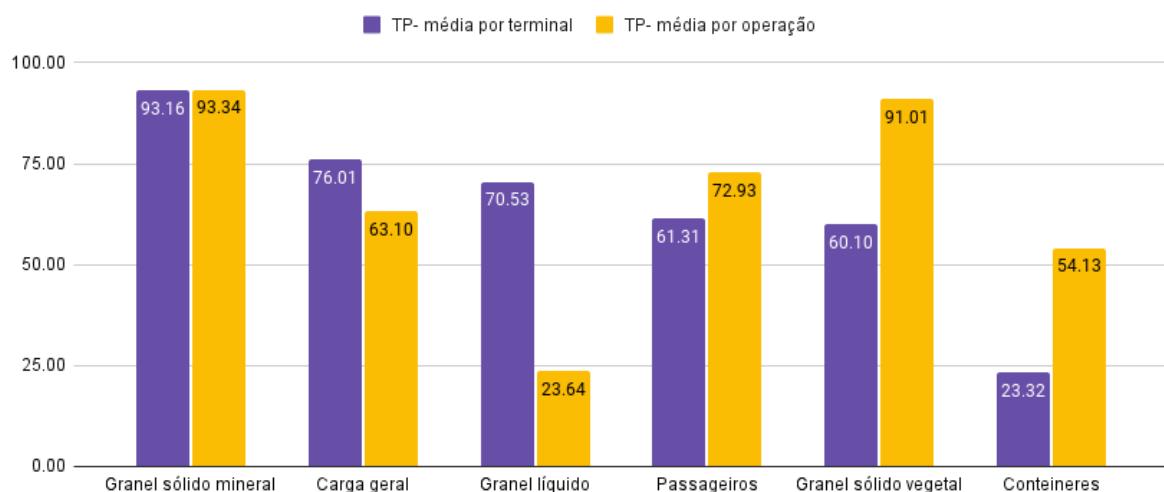
Diferentemente da Taxa de Ocupação, o tempo de permanência médio terá resultados diferentes utilizando a média por terminal (equação 4.3) e por operação (equação 4.4).

$$TP_T = \frac{\sum_{\theta=1}^{n_T} O_{\theta}^T}{n_T} \quad (4.3)$$

$$TP_{(terminais)} = \sum_{i=1}^{m_T} \left(\frac{\sum_{\theta=i}^{n_i} O_{\theta}^i}{n_i} \right) \quad (4.4)$$

Analizando primeiramente a média por terminal, o maior tempo de permanência foi identificado para os terminais que movimentam granel sólido mineral, e o menor foi para os terminais de contêineres, indicando que os terminais de alguns tipos de carga são mais eficientes do que de outros. Já analisando a média por operação, o granel sólido mineral mantém o maior TP, possivelmente pela maior complexidade em lidar com este tipo de produto, porém o menor TP passa a ser do granel líquido, indicando que as operações individuais deste tipo de carga são mais velozes que para os demais.

Figura 4.25: Relação entre TP e o tipo de carga movimentado para os dois métodos de cálculo



Fonte: Elaboração própria

Nota-se que para a maior parte dos tipos, houve uma discrepância entre os resultados calculados pelos dois métodos. No caso do granel líquido, a média por terminal foi maior que a por operação, indicando que os TPs dos terminais de um mesmo tipo de carga não são uniformes entre si, possivelmente com uma parcela dos terminais com TPs altos inflando a média dos terminais.

Os tipos granel sólido vegetal e contêineres apresentaram comportamento contrário, a média por operação é maior que a média por terminal, indicando que uma parte pequena dos terminais teve TP muito baixo, diminuindo a média dos terminais. No caso do granel sólido vegetal, carga geral e passageiros, as duas médias foram similares, indicando uma distribuição mais uniforme entre os terminais que operam cargas destes tipos.

4.4.3.3 Tempo médio Entre Operações (TEO)

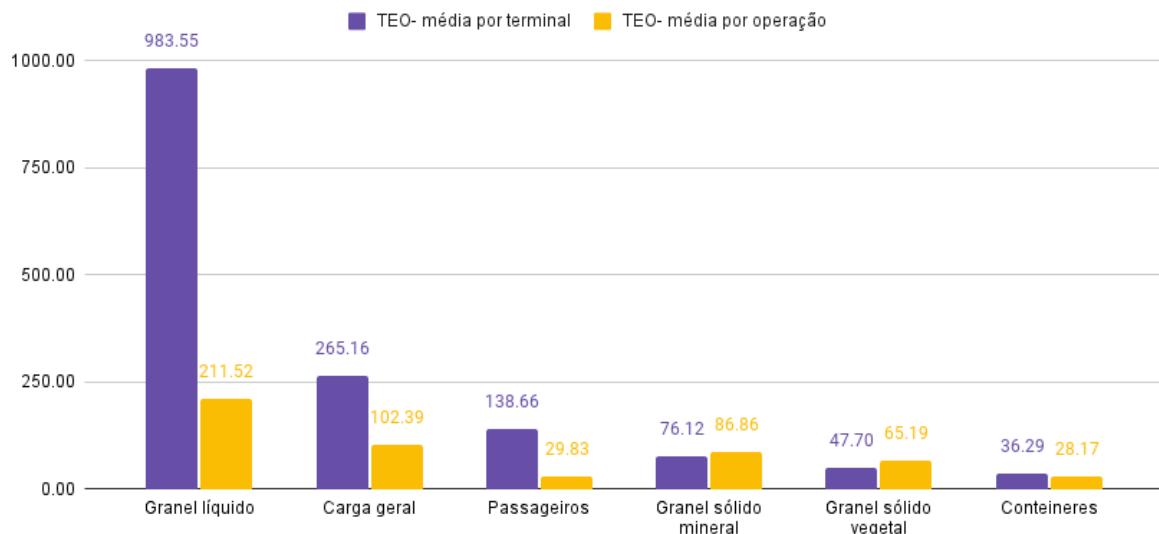
O resultado do tempo entre operações também difere utilizando as métricas dos terminais como base da média (equação 4.6), ou utilizando as operações individualmente (equação 4.5)para calcular a métrica do tipo de carga.

$$TEO_T = \frac{t_T - \sum_{\theta=1}^{n_T} O_\theta^T}{n_T} \quad (4.5)$$

$$TEO_T = \sum_{i=1}^{m_T} \frac{\left(\frac{t_{\text{analise}} - \sum_{\theta=i}^{n_i} O_\theta^i}{n_i} \right)}{n_T} \quad (4.6)$$

Nos dois métodos de cálculo, o tipo de carga com maior tempo entre operações foi o granel líquido, indicando que possivelmente há menor demanda para este tipo de carga, ou que as operações estão com mais gargalos na preparação do terminal para receber um novo navio (o tempo de permanência para granel líquido foi baixo, indicando que a operação uma vez que o navio já atracou, é ágil). Ainda pelos dois métodos, terminais de contêineres apresentaram o menor TEO, indicando que tanto analisando por operação, quanto por terminal, este tipo de carga tem maior agilidade entre uma operação e outra.

Figura 4.26: Relação entre TEO e número de operações dos terminais do Porto de Santos



Fonte: Elaboração própria

Na figura 4.26, fica clara a diferença entre as médias para dois métodos de calculo no caso dos tipo Granel líquido, carga geral e passageiros, indicando que a TEO teve uma

grande variação entre os terminais do mesmo tipo (conforme é possível observar no gráfico de ocupação da figura 4.15). Já os terminais que operam granel sólido mineral e vegetal e de containeres apresentaram valores de TEO mais similares nos dois métodos, pois os terminais apresentaram comportamentos mais uniformes entre si.

5 CONCLUSÃO

Com base no estudo realizado, destaca-se o papel estratégico que o Porto de Santos desempenha na economia brasileira, atuando como um dos principais centros de movimentação de cargas e atuando como ponto de ligação para rotas comerciais nacionais e internacionais.

Neste contexto, a partir da revisão bibliográfica, evidencia-se a relevância de implementar tecnologias de coleta, análise e integração de dados para aprimorar a gestão das operações portuárias, especialmente no que tange à identificação de atracções e ao monitoramento da ocupação dos terminais. Por meio da coleta e processamento de dados públicos do sistema AIS, o presente trabalho criou um sistema de rotulamento dos dados de posição baseado na reconstrução das trajetórias das embarcações, permitindo uma visualização dos padrões de atracções e dos tempos de permanência.

Foi identificado que, para o período em análise, o número de atracções diárias e mensal se mantiveram estáveis, embora a taxa de ocupação dos terminais do Porto de Santos esteja aquém do esperado (MWASENGA, 2012), indicando a possível necessidade de revisão de seus sistemas de *scheduling*, uma baixa velocidade nas operações de carga e descarga, ou ainda um alto número de incidentes e imprevistos.

Analizando a ocupação dos terminais, percebe-se que há localizações com operações mais críticas que outras, evidenciada pela variação na taxa de ocupação entre 88% e 0,6%. Essa variabilidade conversa com variação no número de operações por terminal e do tempo entre operações.

Analizando não os terminais individualmente, mas agregados pelo tipo de carga com o qual operam, também é observado a presença de comportamentos diferentes. Concluiu-se que os terminais de granel sólido vegetal apresentam em média maiores ocupações (65,8%), e os de granel líquido apresentam as menores ocupações (30,6%), com os terminais de granel sólido mineral, passageiros, contêineres e carga geral apresentando taxas intermediárias.

A utilização de dados de AIS, como explorado no trabalho, oferece uma abordagem prática e acessível para alcançar maior eficiência, identificando as áreas onde melhorias operacionais são mais urgentes, o que contribuirá para manter o porto competitivo frente a outros grandes portos mundiais que já implementam tecnologias integradas de análise de dados. Esse avanço tecnológico representa uma contribuição significativa para o desenvolvimento do Porto de Santos, ao mesmo tempo que o posiciona como um modelo de inovação e eficiência operacional dentro da indústria portuária nacional.

5.1 Próximos Passos

Quanto aos próximos passos do projeto, destaca-se a possibilidade de melhorias no método de reconstrução de trajetórias para que seja mais eficiente em diminuir os erros de "falsos negativos" gerados. Mesmo que este tipo de erro de classificação seja menos frequente que os de "falso positivo", que foram o foco do tratamento deste estudo, é importante entender a causa destes problemas, sejam elas relacionadas à natureza do sistema AIS em captar e enviar informações ou pelo próprio método proposto de tratamento, como forma de garantir um melhor domínio dos dados coletados e maior confiabilidade das análises subsequentes.

Também é encorajado que outros métodos de correção da velocidade sejam explorados. No presente estudo foi utilizada uma heurística simples baseada nos dados agregados de posição do navio para estabelecer se a velocidade do navio deveria ser corrigida, porém utilizar técnicas já conhecidas no contexto de rotulação de dados seria interessante, pois possibilitaria a comparação de diferentes métodos e identificação do mais eficaz. Poderiam ser utilizadas desde algoritmos mais simples, como KNN (*k-nearest-neighbours*), até técnicas mais robustas como redes neurais (CHEN et al., 2020).

Outro campo a ser explorado é o uso das informações geradas para basear estudos detalhados com propostas de melhorias nas operações do Porto de Santos. Embora já tenha sido realizada uma análise inicial da eficiência do Porto de Santos, por meio dos dados de posição do sistema AIS eficientemente rotulados gerados neste estudo, é possível efetivamente gerar modelos e políticas que melhorem o funcionamento da instituição. Uma sugestão é levantar modelos de sequenciamento de navios para a entrada no canal e atracação nos terminais, diminuindo o tempo de espera de embarcações e aprimorando a eficiência com a qual o porto atualmente funciona.

Por fim, enfatiza-se a possibilidade e necessidade de estudar outros eventos por meio

do tratamento dos dados de AIS, que incluem a análise do dos navios nas áreas de fundeio, a identificação de padrões nas filas de entrada e saída, e predição de atrasos e acidentes.

REFERÊNCIAS

- BICHOU, K.; GRAY, R. A logistics and supply chain management approach to port performance measurement. *Maritime Policy & Management*, v. 31, n. 1, p. 47–67, 2004. Disponível em: <<https://doi.org/10.1080/0308883032000174454>>.
- UNCTAD. *Review of Maritime Transport*. 2024. Accessed: 2024-11-04. Disponível em: <<https://unctad.org/topic/transport-and-trade-logistics/review-of-maritime-transport>>.
- Autoridade Portuária de Santos. *Fatos e Dados 2023*. [S.l.], 2023. Acesso em: 4 nov. 2024. Disponível em: <<https://www.portodesantos.com.br/wp-content/uploads/Fatos-e-Dados-2023-FF.pdf>>.
- Autoridade Portuária de Santos. *Conheça o Porto: História*. 2024. Acesso em: 4 nov. 2024. Disponível em: <<https://www.portodesantos.com.br/>>.
- DALANDI, C. et al. Efficiency analysis and benchmarking of container ports operating in lower-middle-income countries: a dea approach. *Journal of Shipping and Trade*, 2024.
- CHUNG, K. C. *Port Performance Indicators*. [S.l.], 2013. Accessed: 2024-11-04. Disponível em: <<https://documents1.worldbank.org/curated/en/303501468337289364/pdf/816090BRI0Infr00Box379840B00PUBLIC0.pdf>>.
- NOTTEBOOM, T.; PALLIS, A.; RODRIGUE, J.-P. *Port Economics, Management and Policy*. Routledge, 2022. Accessed: 2024-11-04. Disponível em: <<https://www.taylorfrancis.com/books/mono/10.4324/9780429318184/port-economics-management-policy-theo-notteboom-athanasios-pallis-jean-paul-rodrigue>>.
- MWASENGA, A. L. *Port and Shipping Services in Tanzania*. [S.l.], 2012. Accessed: 2024-11-04. Disponível em: <https://unctad.org/system/files/non-official-document/dtl__ttl__2012d11__Mwasenga.pdf>.
- Comptroller and Auditor General of India. *Union Performance Civil: Functioning of Major Port Trusts*. [S.l.], 2009. Acesso em: 4 nov. 2024. Disponível em: <https://www.cag.gov.in/uploads/download__audit__report/2009/Union__Performance__Civil__Functioning__Major__Port__Trust__3__2009__chapter__7.pdf>.
- YANG, D. et al. How big data enriches maritime research – a critical review of automatic identification system (ais) data applications. *Transport Reviews*, v. 39, n. 6, p. 755–773, 2019. Disponível em: <<https://doi.org/10.1080/01441647.2019.1649315>>.
- JOHNSON, G. et al. Improving meteorological models using ships' weather data communicated via ais. In: . [S.l.: s.n.], 2021. p. 485–501.
- RAJABI, A. et al. Towards smart port: An application of ais data. In: *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE*

- 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*. [S.l.: s.n.], 2018. p. 1414–1421.
- MARTINCIC, T. et al. Vessel and port efficiency metrics through validated ais data. In: *Global Oceans 2020: Singapore – U.S. Gulf Coast*. [S.l.: s.n.], 2020. p. 1–6.
- LIYE, Z.; QIANG, M.; TIEN, F. Big ais data based spatial-temporal analyses of ship traffic in singapore port waters. *Transportation Research Part E: Logistics and Transportation Review*, v. 129, p. 287–304, 2019. ISSN 1366-5545. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1366554516309516>>.
- ZHAO, L.; SHI, G.; YANG, J. Ship trajectories pre-processing based on ais data. *Journal of Navigation*, v. 71, n. 5, p. 1210–1230, 2018.
- NOVAKOVIC, J. D. et al. Evaluation of classification models in machine learning. In: *Theory and Applications of Mathematics Computer Science* 7. [s.n.], 2017. p. 29–46. Accessed: 2024-11-04. Disponível em: <<https://typeset.io/pdf/evaluation-of-classification-models-in-machine-learning-1u2pog86m5.pdf>>.
- VUJOVIC, Z. D. Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, The Science and Information Organization, v. 12, n. 6, 2021. Disponível em: <<http://dx.doi.org/10.14569/IJACSA.2021.0120670>>.
- DGRM. *Circular N.º 14: Identidades do Serviço Móvel Marítimo - MMSI*. 2014. Acessado em: 16 de novembro de 2024. Disponível em: <<https://www.dgrm.pt/documents/20143/86274/C\14\PT.pdf/f22f44df-12fb-954e-14a0-0ae334442306>>.
- Decreto-Lei n.º 295/2000. *Decreto-Lei n.º 295/2000, de 17 de novembro*. 2000. Acessado em: 16 de novembro de 2024. Disponível em: <<https://diariodarepublica.pt/dr/detalhe/decreto-lei/295-534474>>.
- Marinha do Brasil. *Portaria 100/2023*. [S.l.], 2023. Acesso em: 4 nov. 2024. Disponível em: <https://www.marinha.mil.br/cpsc/sites/www.marinha.mil.br.cpsc/files/Port\%20100_2023.pdf>.
- CHEN, X. et al. A ship movement classification based on automatic identification system (ais) data using convolutional neural network. *Ocean Engineering*, v. 218, p. 108182, 2020. ISSN 0029-8018. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0029801820311124>>.

ANEXO A – RELAÇÃO DE TERMINAIS E TIPOS DE CARGA

Tipo de carga	Terminais
Carga geral	Usiminas V, Usiminas IV, Usiminas III, Usiminas II, Usiminas I, Saboó 4, Corte, Valongo, DPW 3, DPW 4, Terminal 31, Terminal 32, Terminal 33, Terminal 34, Terminal 35.1, Terminal 35.2, TEV, Saipem 1, Saipem 2
Granel sólido mineral	Tiplam 1, Tiplam 4, Saboó 1, Saboó 2, Saboó 3, Terminal 22/23, TERMAG
Granel sólido vegetal	Tiplam 2, Tiplam 3, Terminal 12A, Terminal 13/14, Terminal 15, Terminal 16/17, Terminal 19, Terminal 20/21, Outeirinhos 3, TGG, TEAG, TEG, Terminal 37, Terminal 38, Terminal 39
Granel líquido	Alamoá 1, Alamoá 2, Alamoá 3, Alamoá 4, AGEÓ SP, IB SP, IB BC, Terminal 29, Terminal 30, DOW, Cutrale, AGEÓ SP, TRSP
Conteineres	BTP 1, BTP 2, BTP 3, DPW 1, DPW 2, Tecon 1, Tecon 1, Tecon 2, Tecon 3
Passageiros	Terminal 25, Outeirinhos 1, Outeirinhos 2

ANEXO B – MÉTRICAS DE USO DOS TERMINAIS

Terminal	Tempo de permanência médio (TP)	Número de operações	Taxa de Ocupação (TO)	Tempo médio entre Operações (TEO)
TERM. 12A	89.93	20	77.6%	25.91
TERM. 13/14	116.49	6	30.2%	269.67
TERM. 16/17	37.68	47	76.4%	11.62
TERM. 19	46.40	40	80.1%	11.52
TERM. 20/21	47.60	41	84.2%	8.92
TERM. 22/23	104.71	17	76.8%	31.58
TERM. 25	35.39	8	12.2%	254.23
TERM. 29	55.86	8	19.3%	233.76
TERM. 30	164.07	11	77.9%	46.56
TERM. 31	88.00	12	45.6%	105.07
TERM. 32	148.15	12	76.7%	44.92
TERM. 33	105.60	13	59.3%	72.62
TERM. 34	64.46	19	52.9%	57.48
TERM. 35.1	158.70	12	82.2%	34.38
TERM. 35.2	45.39	20	39.2%	70.45
TERM. 37	72.85	28	88.0%	9.89
TERM. 38	60.45	27	70.4%	25.37
TERM. 39	39.35	30	50.9%	37.88
TERM. BTP 1	21.95	45	42.6%	29.53
TERM. BTP 2	19.72	43	36.6%	34.16
TERM. BTP 3	20.00	43	37.1%	33.89
TERM. Corte	107.30	4	18.5%	471.93

TERM. Cutrale	72.67	17	53.3%	63.62
TERM. DPW 1	28.69	44	54.5%	23.96
TERM. DPW 2	27.58	45	53.6%	23.91
TERM. DPW 3	44.63	32	61.6%	27.77
TERM. DPW 4	51.94	20	44.8%	63.91
TERM. IB SP	60.06	1	2.6%	2256.87
TERM. Outeiri-nhos 1	87.23	21	79.1%	23.10
TERM. Saboó 2	109.91	13	61.7%	68.31
TERM. Saboó 3	158.11	10	68.2%	73.58
TERM. Saboó 4	51.71	18	40.2%	77.01
TERM. Saipem 1	134.17	3	17.4%	638.14
TERM. Saipem 2	6.40	2	0.6%	1152.06
TERM. TEAG	80.32	25	86.7%	12.36
TERM. Tecon 1.1	19.88	17	14.6%	116.41
TERM. Tecon 1.2	21.22	48	44.0%	27.05
TERM. Tecon 2	24.94	55	59.2%	17.19
TERM. Tecon 3	25.86	50	55.8%	20.48
TERM. TEG	84.37	22	80.1%	20.94
TERM. TER-MAG	74.90	20	64.7%	40.94
TERM. TEV	21.35	35	32.3%	44.84
TERM. TGG	45.61	40	78.7%	12.32
TERM. Tiplam 1	33.75	10	14.6%	197.95
TERM. Tiplam 2	33.21	18	25.8%	95.51
TERM. Tiplam 3	27.08	22	25.7%	78.23
TERM. Tiplam 4	77.60	19	63.6%	44.34
TERM. Usiminas II	73.50	5	15.9%	389.89
TERM. Usiminas III	70.04	4	12.1%	509.19
TERM. Usiminas IV	83.84	3	10.9%	688.47
TERM. Valongo	36.91	24	38.2%	59.62