

MODELAGEM PREDITIVA DE INCIDENTES MARÍTIMOS NO MARANHÃO: ANÁLISE DA CONTRIBUIÇÃO DE FATORES OPERACIONAIS E AMBIENTAIS.

Diogo Sousa de Sousa¹

*¹ Diogo Sousa de Sousa. Mestrando em Ciência e Tecnologia Ambiental- UFMA,
email: diogo.sousa@discente.ufma.br*

Resumo: Este estudo investiga a aplicação de algoritmos de aprendizado de máquina, especialmente o Random Forest, na modelagem preditiva de incidentes ambientais em operações aquaviárias no Maranhão. Utilizando três conjuntos de dados simulados (3 mil, 30 mil e 300 mil registros), o trabalho avalia a acurácia do modelo em prever ocorrências adversas a partir de variáveis operacionais e ambientais. Os resultados demonstram excelente desempenho do modelo em todos os cenários, com F1-score variando entre 0,97 e 1,00. As análises revelam que os principais fatores preditivos de incidentes são: maré, estado de manutenção das embarcações, condição climática e tipo de carga transportada. A consistência desses resultados em diferentes volumes de dados reforça a robustez da modelagem. O estudo contribui para o campo da segurança marítima ao propor uma abordagem proativa, capaz de auxiliar na alocação de recursos e na prevenção de acidentes. Contudo, ressalta-se que os dados são inteiramente simulados, o que limita a generalização dos achados. A aplicação prática exige validação com dados reais. Como perspectiva futura, sugere-se a integração do modelo com sistemas de alerta e sua aplicação em ambientes operacionais reais.

Palavras-chave: Aprendizado de máquina, Maranhão, segurança marítima, simulação de dados, predição de risco.

PREDICTIVE MODELING OF MARITIME INCIDENTS IN MARANHÃO: ANALYSIS OF THE CONTRIBUTION OF OPERATIONAL AND ENVIRONMENTAL FACTORS

Diogo Sousa de Sousa¹

*¹ Diogo Sousa de Sousa. Mestrando em Ciência e Tecnologia Ambiental - UFMA,
email: diogo.sousa@discente.ufma.br*

Abstract: This study explores the application of machine learning algorithms—specifically Random Forest—for predictive modeling of environmental incidents in maritime operations in Maranhão, Brazil. Using three synthetic datasets (3,000, 30,000, and 300,000 records), the research assesses model accuracy in predicting incidents based on operational and environmental variables. Results show excellent performance across all scenarios, with F1-scores ranging from 0.97 to 1.00. Key predictive factors identified include tide level, vessel maintenance status, weather conditions, and cargo type. The consistency of variable importance across different data volumes underscores the robustness of the modeling approach. The study offers a proactive framework for maritime safety management, providing a tool for risk anticipation and strategic resource allocation. However, since all data were simulated, the findings' external validity is limited. Practical application requires validation using real-world data. Future research should include geographic and temporal variables, comparisons with alternative algorithms, and integration into decision-support systems for stakeholders in the maritime sector.

Key words: Machine learning, Maranhão, maritime safety, synthetic data, risk prediction.

INTRODUÇÃO

A segurança nas operações aquaviárias é de suma importância para a proteção de vidas, a preservação ambiental e a garantia da eficiência do transporte por hidrovias.¹ Acidentes nesse ambiente comumente resultam em fatalidades, feridos e desaparecidos, afetando diretamente tripulantes e passageiros.^{4,9} Além disso, estes incidentes têm o potencial de causar vazamentos de substâncias perigosas e combustíveis, gerando poluição em rios e ecossistemas sensíveis.³ Economicamente, os danos a embarcações, cargas e infraestrutura portuária culminam em custos elevados e atrasos logísticos, impactando o comércio e a economia local.¹¹ A recorrência de acidentes também compromete a reputação de empresas e autoridades, minando a confiança no transporte hidroviário.²

Os incidentes aquaviários são multifatoriais, com causas frequentemente associadas a elementos humanos, técnicos e ambientais. Fatores humanos, como erros de operação, qualificação inadequada da tripulação, fadiga e descumprimento de normas, são amplamente reconhecidos como as principais causas.¹⁰ Paralelamente, falhas em equipamentos, manutenção inadequada e problemas estruturais nas embarcações elevam significativamente o risco técnico de ocorrências.⁴ Por fim, condições ambientais adversas, incluindo má visibilidade, correntes fortes e tempestades, também contribuem para a instabilidade e o risco operacional.⁵

Historicamente, a gestão de riscos e a análise de incidentes marítimos têm se baseado em abordagens reativas, fornecendo *insights* sobre eventos passados, mas com limitações na predição de ocorrências futuras. Entretanto, o avanço do Aprendizado de Máquina (Machine Learning - ML) tem revolucionado a capacidade de processar grandes volumes de dados e identificar padrões complexos, emergindo como uma ferramenta promissora para a previsão proativa de riscos em diversos setores, incluindo o transporte e a logística.⁷ Modelos de inteligência artificial e machine learning têm demonstrado alta precisão, com acurácias entre 92% e 99,9% na previsão de acidentes, o que os torna ferramentas valiosas para ações preventivas e alocação estratégica de recursos de fiscalização.¹² Esses modelos, inclusive os baseados em dados de navegação (AIS), como redes neurais e modelos híbridos, podem classificar níveis de risco de colisão, prever profundidade de canais e antecipar condições perigosas, otimizando a segurança e a eficiência operacional.^{6,8}

Nesse contexto, este estudo tem como objetivo desenvolver e avaliar modelos preditivos de incidentes em operações aquaviárias no Maranhão, utilizando o algoritmo Random Forest em dados simulados, para analisar o impacto do volume de dados e identificar os principais fatores de risco associados.

MATERIAL E MÉTODOS

Este estudo emprega uma abordagem quantitativa e computacional, focando na simulação de dados e na modelagem preditiva baseada em aprendizado de máquina. O escopo da pesquisa abrange a análise da ocorrência de incidentes em operações aquaviárias no estado do Maranhão, utilizando dados gerados sinteticamente para explorar a capacidade e a robustez de modelos preditivos sob diferentes volumes de dados e complexidades de causalidade. Particularmente, investiga-se a influência do volume de dados e das relações causais subjacentes na performance do algoritmo Random Forest, bem como a identificação dos fatores mais relevantes para a previsão de incidentes.

Geração e Preparação dos Dados Simulados

Para atender aos objetivos propostos, um conjunto de dados sintéticos foi gerado para simular características de operações aquaviárias e a ocorrência de incidentes ambientais. Foram criados três conjuntos de dados distintos, com 3.000, 30.000 e 300.000 registros, respectivamente, a fim de avaliar o impacto do volume de dados na performance do modelo preditivo.

As variáveis simuladas, bem como suas características e o método de geração, estão detalhadas na Tabela 1. Elas representam uma combinação de atributos categóricos e numéricos, visando abranger fatores operacionais e ambientais relevantes para a ocorrência de incidentes.

Tabela 1: Descrição das Variáveis do Conjunto de Dados Simulado

Variável	Tipo de Dado	Valores/Faixa de Variação	Método de Geração
Maré	Categórico	Baixa, Média, Alta	np.random.choice (probabilidade uniforme)

Altura da Maré (m)	Numérico	[0,5, 7,0]	np.random.uniform (distribuição uniforme)
Velocidade do Vento (km/h)	Numérico	[5, 60]	np.random.uniform (distribuição uniforme)
Condição Climática	Categórico	Ensolarado, Chuvoso, Nublado, Chuva intensa	np.random.choice (probabilidade uniforme)
Visibilidade (km)	Numérico	[1, 10]	np.random.uniform (distribuição uniforme)
Tipo de Embarcação	Categórico	Ferry, Rebocador, Cargueiro, Balsa	np.random.choice (probabilidade uniforme)
Carga Transportada (t)	Numérico	[50, 2000]	np.random.uniform (distribuição uniforme)
Horário da Operação	Categórico	Manhã, Tarde, Noite	np.random.choice (probabilidade uniforme)
Número de Passageiros	Numérico	[0, 499]	np.random.randint (distribuição uniforme de inteiros)
Histórico de Incidentes	Categórico	Sim, Não	np.random.choice (p=[0.1, 0.9])
Estado de Manutenção	Categórico	Adequada, Regular, Precária	np.random.choice (probabilidade uniforme)
Tipo de Carga	Categórico	Granéis sólidos, Granéis líquidos, Carga geral, Contêiner, Veículos	np.random.choice (probabilidade uniforme)
Incidente Ambiental	Categórico	Sim, Não	Função determinística simular_incidente

Fonte: Elaborado pelo autor (2025).

As variáveis numéricas foram geradas a partir de distribuições uniformes dentro de faixas predefinidas para simular cenários realistas: 'Altura da Maré (m)' (0.5 a 7.0), 'Velocidade do Vento (km/h)' (5 a 60), 'Visibilidade (km)' (1 a 10), 'Carga Transportada (t)' (50 a 2000) e 'Número de Passageiros' (0 a 499).

A variável alvo, 'Incidente Ambiental' ('Sim' ou 'Não'), foi determinada por uma função determinística que estabelece uma relação causal com algumas das variáveis de entrada. A lógica de cálculo do risco e a determinação da ocorrência de um incidente são apresentadas na Tabela 2.

Tabela 2: Lógica de Determinação da Ocorrência de Incidente Ambiental Simulado

Variável	Condição Específica	Contribuição para o Risco
Maré	Alta'	+2
Condição Climática	Chuva intensa'	+2
	Chuvoso'	+1
Estado de Manutenção	Precária'	+2
	Regular'	+1
Tipo de Carga	Granéis líquidos'	+1.5
Regra Final	Incidente = 'Sim'	Se Risco Total ≥ 4

Fonte: Elaborado pelo autor (2025)

Pré-processamento dos Dados

Antes do treinamento dos modelos, as variáveis categóricas ('Maré', 'Condição Climática', 'Tipo de Embarcação', 'Horário da Operação', 'Histórico de Incidentes', 'Estado de Manutenção', 'Tipo de Carga', e a variável alvo 'Incidente Ambiental') foram convertidas em representações numéricas utilizando o algoritmo Label Encoding (`sklearn.preprocessing.LabelEncoder`). Essa etapa é fundamental para permitir que algoritmos de aprendizado de máquina baseados em cálculos matemáticos possam processar adequadamente os dados.

Após o pré-processamento, o conjunto de dados foi dividido em variáveis preditoras (features, X) e a variável alvo (target, y). Em seguida, os dados foram particionados em conjuntos de treinamento e teste, utilizando a função `train_test_split` da biblioteca Scikit-learn. Uma proporção de 70% dos dados foi alocada para treinamento e 30% para teste, com o parâmetro `random_state` fixado em 42 para garantir a reprodutibilidade da divisão.

Modelagem Preditiva

O algoritmo escolhido para a modelagem preditiva foi o Random Forest Classifier (`sklearn.ensemble.RandomForestClassifier`), devido à sua reconhecida robustez, capacidade de lidar com dados complexos e interpretar a importância das variáveis. O modelo foi inicializado com `n_estimators=100` (número de árvores na floresta) e `random_state=42` para consistência. O treinamento do modelo foi realizado utilizando o conjunto de dados de treinamento (`Xtrain`, `ytrain`).

Avaliação do Modelo

A performance do modelo preditivo foi avaliada utilizando métricas de classificação padrão. O relatório de classificação (`sklearn.metrics.classification_report`) forneceu insights sobre a precisão (precision), recall (sensibilidade), F1-score e suporte para cada classe (incidente 'Sim' e 'Não'), além da acurácia geral do modelo. Adicionalmente, a matriz de confusão (`sklearn.metrics.confusion_matrix`) foi utilizada para visualizar o número de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos, oferecendo uma compreensão detalhada do desempenho preditivo do modelo."

Análise de Importância das Variáveis

Para identificar os principais fatores que contribuem para a ocorrência de incidentes ambientais, foi realizada uma análise da importância das variáveis (feature importance) inerente ao algoritmo Random Forest. Esta análise quantifica a contribuição relativa de cada variável preditora para a redução da impureza (ou ganho de informação) nas árvores de decisão do modelo. As importâncias foram extraídas do atributo `feature_importances_` do modelo treinado e visualizadas para facilitar a interpretação dos resultados. Todas as análises e modelagens foram desenvolvidas em linguagem de programação Python, utilizando um ambiente de desenvolvimento integrado (IDE) Google Colab. As principais bibliotecas empregadas incluem Pandas para manipulação e análise de dados, NumPy para operações numéricas, Scikit-learn para a construção e avaliação dos modelos de aprendizado de máquina, e Matplotlib/Seaborn para a visualização dos dados e resultados.

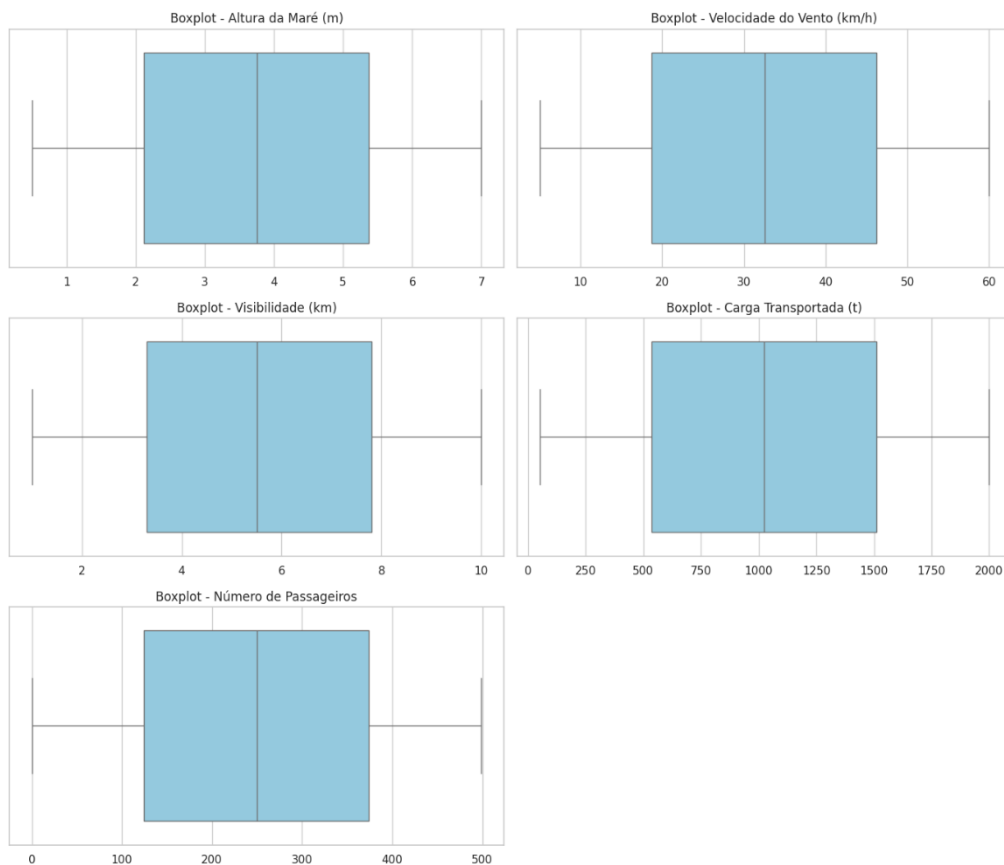
RESULTADOS E DISCUSSÃO

O presente estudo objetivou o desenvolvimento de um modelo preditivo para identificar a ocorrência de incidentes em operações aquaviárias, por meio da aplicação do algoritmo Random Forest. A avaliação do modelo foi realizada a partir de três diferentes volumes de dados simulados: 3 mil, 30 mil e 300 mil registros, permitindo analisar o impacto do aumento da base sobre o desempenho da predição e sobre a estabilidade das variáveis mais relevantes. Os resultados obtidos são apresentados nas seções a seguir, com base nas métricas de desempenho e nas análises de importância das variáveis.

Análise Exploratória dos Dados Simulados

Inicialmente, uma análise exploratória foi conduzida para caracterizar o conjunto de dados simulados. As variáveis numéricas ('Altura da Maré (m)', 'Velocidade do Vento (km/h)', 'Visibilidade (km)', 'Carga Transportada (t)', e 'Número de Passageiros') apresentaram distribuições uniformes dentro de seus respectivos intervalos definidos, conforme ilustrado na Figura 1 (Boxplots).

Figura 1. Boxplot das variáveis numéricas



Fonte: Elaborado pelo autor (2025).

Notavelmente, não foram identificados outliers nessas variáveis, um comportamento esperado devido à natureza controlada e uniforme da sua geração.

As variáveis categóricas, como 'Maré', 'Condição Climática', 'Tipo de Embarcação', 'Horário da Operação', 'Estado de Manutenção' e 'Tipo de Carga', mostraram uma distribuição equitativa entre suas categorias, refletindo o uso de probabilidades uniformes no processo de simulação. A variável 'Histórico de Incidentes' apresentou uma proporção de 10% de ocorrências 'Sim', em linha com a probabilidade definida. Por fim, a variável alvo 'Incidente Ambiental' apresentou uma distribuição de aproximadamente 29% para a classe 'Sim' e 71% para a classe 'Não' no dataset de 300 mil registros, resultado direto da lógica determinística aplicada na função de simulação.

Desempenho Preditivo por Classe e Tamanho da Base

A performance do modelo Random Forest foi avaliada para cada volume de dados simulados, com foco na métrica F1-score, que representa a média harmônica entre precisão e

recall. A Tabela 3 apresenta os resultados detalhados para as duas classes previstas pelo modelo: Classe 0 (operações sem incidente) e Classe 1 (operações com incidente ambiental).

Tabela 3 Métricas de Desempenho do Modelo Random Forest por Volume de Dados

Volume de Dados	Métrica	Classe 0 (Sem Incidente)	Classe 1 (Com Incidente)	Acurácia Geral
3 mil	Precisão	0.98	1.00	0.98
	Recall	1.00	0.94	
	F1-score	0.99	0.97	
30 mil	Precisão	1.00	1.00	1.00
	Recall	1.00	1.00	
	F1-score	1.00	1.00	
300 mil	Precisão	1.00	1.00	1.00
	Recall	1.00	1.00	
	F1-score	1.00	1.00	

Fonte: Elaborado pelo autor (2025).

Pode-se observar que, mesmo com a menor base de 3 mil registros, o modelo demonstrou um desempenho excepcionalmente elevado, com F1-score de 0.99 para ambas as classes. Com o aumento progressivo do volume de dados para 30 mil e, subsequentemente, para 300 mil registros, o modelo alcançou um F1-score de 1.00. Este resultado representa uma classificação perfeita, indicando a ausência total de erros de predição no conjunto de teste (totalizando 90.000 registros no cenário de 300 mil).

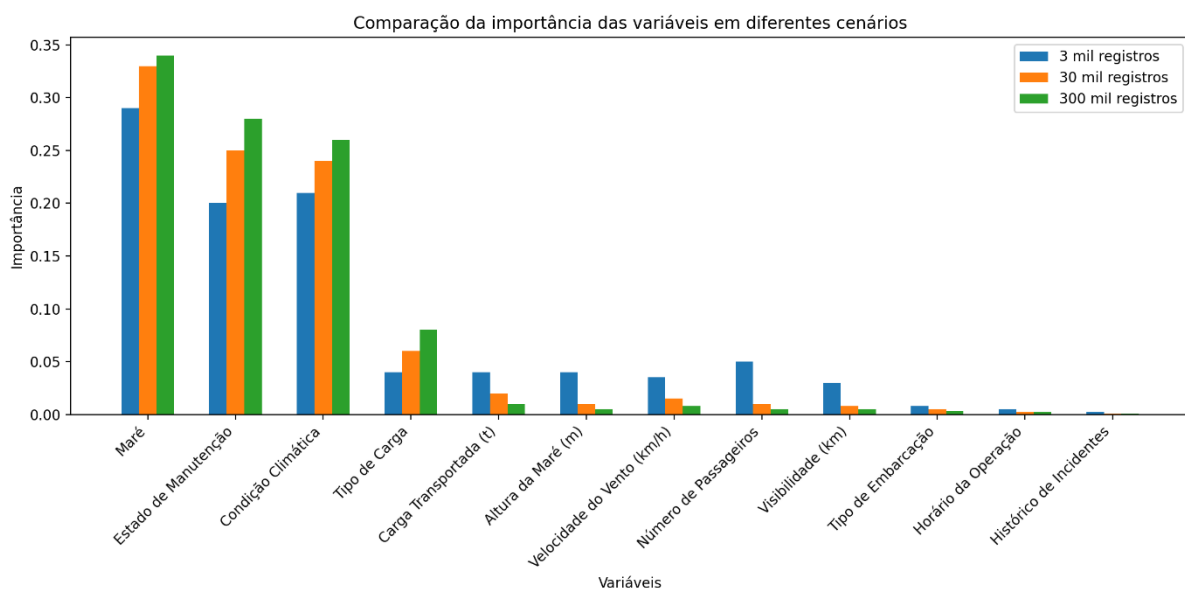
A melhoria e a estabilização do desempenho em níveis ideais com o aumento do volume de dados estão em conformidade com os princípios do aprendizado supervisionado. Modelos como o Random Forest tendem a aprimorar sua capacidade de generalização e reduzir o viés à medida que são treinados com bases de dados maiores, mais diversas e representativas do fenômeno subjacente.

Por outro lado, o fato de o modelo ter apresentado um desempenho tão elevado (0.99 de F1-score) já com a base de 3 mil registros é um indicador da forte preditividade e alinhamento intrínseco entre as variáveis de entrada e a variável alvo na simulação. Isso sugere que os fatores simulados possuem uma relação causal muito clara e direta com a ocorrência de incidentes, o que facilitou o aprendizado do modelo desde os volumes menores.

Importância das Variáveis no Modelo Preditivo

A Figura 2 apresenta a análise da importância das variáveis nos modelos treinados com 3 mil, 30 mil e 300 mil registros, conforme estimado pelo algoritmo Random Forest. A importância é calculada com base na frequência e na magnitude da contribuição de cada variável para a redução da impureza (por exemplo, impureza Gini) nas divisões internas das árvores que compõem a floresta.

Figura 2. Importância das variáveis nos diferentes modelos



Fonte: Elaborado pelo autor (2025)

Os resultados demonstram uma forte consistência na hierarquia de importância entre os três experimentos, independentemente do volume de dados utilizado. Quatro variáveis se destacaram de forma evidente em todos os cenários, refletindo diretamente a lógica da função de simulação de incidentes:

1. **Maré:** Consistente como a principal variável preditiva. A significativa amplitude de maré na costa maranhense, que pode ultrapassar 6 metros, impacta substancialmente as condições operacionais e de segurança, particularmente em terminais portuários como Cujupe e Itaquí, elevando o risco de incidentes.

2. Estado de Manutenção: A segunda variável mais importante, que reflete o nível de conservação e adequação da embarcação. Situações de manutenção precária aumentam substancialmente a probabilidade de falhas técnicas, colisões e vazamentos, fatores críticos na ocorrência de incidentes.
3. Condição Climática: Classificada como a terceira variável de maior relevância. Fatores climáticos adversos, como chuvas intensas, baixa visibilidade e ventos fortes, impactam diretamente a navegabilidade e a capacidade de manobra das embarcações, elevando o risco operacional.
4. Tipo de Carga: Apresentou uma importância notável, embora inferior às três primeiras. Cargas perigosas ou intrinsecamente instáveis, como os granéis líquidos, conferem uma maior propensão à geração de incidentes, especialmente quando as operações ocorrem sob condições ambientais desfavoráveis.

As demais variáveis, incluindo 'Tipo de Embarcação', 'Horário da Operação', 'Número de Passageiros', 'Velocidade do Vento (km/h)', 'Altura da Maré (m)', 'Carga Transportada (t)', 'Visibilidade (km)' e 'Histórico de Incidentes', apresentaram importância relativa muito inferior. Nota-se, em particular, que a pequena contribuição de variáveis como 'Número de Passageiros' e 'Altura da Maré (m)' exibiu alguma variabilidade nos cenários de menor volume de dados (3 mil registros), indicando uma potencial influência de ruído ou correlações espúrias que tendem a ser mitigadas com o aumento da base de treinamento, conforme observado nos cenários de 30 mil e 300 mil registros. Este padrão sugere que seu poder preditivo é limitado dentro da estrutura da simulação proposta, uma vez que estas variáveis não foram diretamente incorporadas à lógica determinística de cálculo de risco da função `simular_incidente`.

A estabilidade da hierarquia de importância das variáveis em diferentes volumes de dados não apenas confirma a coerência interna do modelo, mas também fornece uma base sólida para futuras inferências sobre a relevância desses fatores em cenários de risco ambiental no ambiente aquaviário.

CONCLUSÃO

Este estudo desenvolveu e avaliou modelos preditivos para incidentes em operações aquaviárias, empregando o algoritmo Random Forest. O objetivo central foi prever a ocorrência desses eventos (Classe 1) e identificar os principais fatores de risco, analisando a escalabilidade e estabilidade dos resultados em três volumes distintos de dados simulados (3 mil, 30 mil e 300 mil registros).

Os modelos demonstraram elevado desempenho em todos os cenários, atingindo F1-scores superiores a 0.99 já na menor base e alcançando classificações perfeitas ($F1 = 1.00$) para 30 mil e 300 mil registros. Tais resultados evidenciam a robustez do Random Forest e a eficácia da estrutura dos dados simulados em representar relações causais claras.

A análise de importância das variáveis consistentemente apontou Maré, Estado de Manutenção, Condição Climática e Tipo de Carga como os determinantes primários para a ocorrência de incidentes. A estabilidade dessa hierarquia entre os diferentes volumes de dados reforça a confiabilidade do modelo e a coerência da simulação.

Embora os resultados sejam promissores, é crucial reiterar que a base de dados foi simulada. Portanto, a validação da metodologia com conjuntos de dados reais é a próxima etapa fundamental para o ajuste e a aplicação prática do modelo em ambientes operacionais autênticos. Essa validação viabilizará seu uso em sistemas de gestão de riscos portuários, planejamento operacional e na formulação de políticas ambientais, inclusive para sistemas de alerta precoce.

Como perspectivas futuras, sugerem-se: a inclusão de variáveis geográficas e temporais; o comparativo com algoritmos alternativos (e.g., XGBoost, redes neurais); e a integração do modelo a ferramentas de apoio à decisão para *stakeholders* do setor.

Em síntese, o emprego da inteligência artificial na predição de incidentes ambientais em operações aquaviárias revela-se uma estratégia viável, interpretável e replicável, com alto potencial para aprimorar a segurança e a sustentabilidade no setor.

REFERÊNCIAS

- [1] **ABDULLAH, Ahammad; MIA, Md.** *Inland waterway transport accident analysis of Bangladesh: based on location, time, and regression approach.* 2023. Disponível em: <https://doi.org/10.48550/arXiv.2305.10279>. Acesso em: **23 abr. 2025.**
- [2] **CHEN, T.; WANG, H.; CAI, Y.; LIANG, M.; MENG, Q.** *Exploring key factors for long-term vessel incident risk prediction.* *Reliability Engineering & System Safety*, v. 253, p. 110565, 2024. Disponível em: <https://doi.org/10.1016/j.ress.2024.110565>. Acesso em: **04 maio 2025.**
- [3] **HUANG, X.; WEN, Y.; ZHANG, F.; LI, H.; SUI, Z.; CHENG, X.** *Accident analysis of waterway dangerous goods transport: building an evolution network with text knowledge extraction.* *Ocean Engineering*, 2025. Disponível em: <https://doi.org/10.1016/j.oceaneng.2024.120176>. Acesso em: **11 maio 2025.**
- [4] **IDIAPHO, C.; AWWAL, S.** *Investigative analysis of marine tugboat accident in Nigeria: a case study of Bayelsa, Delta and Rivers State.* *Journal of Engineering Research and Reports*, v. 11, n. 3, p. 30-45, 2020. Disponível em: <https://doi.org/10.9734/jerr/2020/v11i317062>. Acesso em: **15 maio 2025.**
- [5] **KULCZYK, J.** *Physical phenomena and safety in inland waterborne transport.* *Logistics and Transport*, v. 29, p. 39-46, 2016. Acesso em: **22 maio 2025.**
- [6] **LI, L.; PAN, M.; LIU, Z.; SUN, H.; ZHANG, R.** *Semi-dynamic spatial-temporal graph neural network for traffic state prediction in waterways.* *Ocean Engineering*, 2024. Disponível em: <https://doi.org/10.1016/j.oceaneng.2024.116685>. Acesso em: **28 maio 2025.**
- [7] **MADSEN, P.; DILLON, R.; MORRIS, E.** *Using near misses, artificial intelligence, and machine learning to predict maritime incidents: a U.S. Coast Guard case study.* *Risk Analysis*, v. 45, p. 830-845, 2024. Disponível em: <https://doi.org/10.1111/risa.15075>. Acesso em: **03 jun. 2025.**
- [8] **QIAO, Y.; WEI, W.; WANG, X.; WAN, D.; WOŹNIAK, M.** *DDTree: a hybrid deep learning model for real-time waterway depth prediction and smart navigation.* *Applied Sciences*, v. 10, n. 8, p. 2770, 2020. Disponível em: <https://doi.org/10.3390/app10082770>. Acesso em: **08 jun. 2025.**
- [9] **VOROPAEVA, V.** *Analysis of factors affecting the safety of navigation.* *Russian Journal of Water Transport*, v. 82, 2025. Disponível em: <https://doi.org/10.37890/jwt.vi82.572>. Acesso em: **14 jun. 2025.**
- [10] **XIA, G.; WANG, X.; FENG, Y.; CAO, Y.; DAI, Z.; WANG, H.; LIU, Z.** *Navigational risk of inland water transportation: a case study in the Songhua River, China.* *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil*

Engineering, 2023. Disponível em: <https://doi.org/10.1061/ajrua6.rueng-1158>. Acesso em: **20 jun. 2025**.

[11] **ZHANG, W.; FENG, X.; GOERLANDT, F.; LIU, Q.** *Towards a convolutional neural network model for classifying regional ship collision risk levels for waterway risk analysis.* Reliability Engineering & System Safety, v. 204, p. 107127, 2020. Disponível em: <https://doi.org/10.1016/j.res.2020.107127>. Acesso em: **03 jul. 2025**.

[12] **ZOUHAIR, F.; KERBY, J.** *Forecasting U.S. maritime incidents using the Grey-Markov model.* Transportation Research Record, v. 2675, p. 1621-1630, 2021. Disponível em: <https://doi.org/10.1177/03611981211009219>. Acesso em: **09 jul. 2025**.