

Predição de Corrosão Atmosférica em Materiais Metálicos Utilizando Aprendizado de Máquina

Vinicius Michelon Geremias

Programa de Pós-Graduação em Computação Aplicada
Universidade do Estado de Santa Catarina (UDESC)
Joinville, SC, Brasil
vmgeremias@hotmail.com

Thiago Brandenburg

Departamento de Ciência da Computação
Universidade do Estado de Santa Catarina (UDESC)
Joinville, SC, Brasil
thiagobranden@hotmail.com

Gustavo A. Fischer

ArcelorMittal
R&D Group
São Francisco do Sul, SC, Brasil
gustavo.fischer@amcontratos.com.br

Fabiano Miranda

ArcelorMittal
R&D Group
São Francisco do Sul, SC, Brasil
fmiranda.fabiano@gmail.com

José Francisco Silva Filho

ArcelorMittal
R&D Group
São Francisco do Sul, SC, Brasil
jose.francisco@arcelormittal.com.br

Rafael Stubs Parpinelli

Programa de Pós-Graduação em Computação Aplicada
Universidade do Estado de Santa Catarina (UDESC)
Joinville, SC, Brasil
rafael.parpinelli@udesc.br

Abstract—The study of the corrosion impact on metallic materials is of great importance to industries and the metallurgical market. Over the years, several formulas, known as Dose-Response Functions, have been developed with the aim of predicting corrosion based on environmental factors. Using data from the MICAT atmospheric corrosion study, this article proposes data-driven prediction models for four materials, named Carbon Steel, Aluminum, Copper, and Zinc. The models indicate that the Random Forest algorithm can be used to predict atmospheric corrosion of metallic materials with competitive results when compared to standard prediction functions.

Index Terms—Random Forest, Atmospheric Corrosion, MICAT

I. INTRODUÇÃO

A corrosão é a degradação de propriedades ou massa de um material ao longo do tempo devido a fatores ambientais [1]. Para os materiais metálicos, os processos de corrosão se dão por reações químicas e eletroquímicas que se passam na superfície do metal [2]. As principais reações químicas que acontecem por conta da corrosão de um metal podem ser explicadas pela reação de oxirredução, em que os elétrons liberados pela oxidação do metal são consumidos pela reação catódica. Uma condição necessária à existência deste processo é que o ambiente externo ao metal seja composto por um líquido condutor de íons (isto é, um eletrólito), seja aerado, e que esteja em contato direto com o metal [3].

Coloquialmente, o termo corrosão é associado ao termo ferrugem, que é um termo genérico para descrever os diferentes hidróxidos e óxidos, tais como $\text{Fe}(\text{OH})_2$, $\text{Fe}(\text{OH})_3$, $\text{FeO}(\text{OH})$, $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$, formados na corrosão do ferro. Apesar de, inicialmente a corrosão ter sido descrita somente para

materiais metálicos, hoje diversos materiais não metálicos são usados na engenharia, como concreto, polímeros, cerâmicas e semicondutores, portanto o termo pode ser estendido a outros materiais [4]. Neste artigo, o termo corrosão será usado para materiais metálicos.

A corrosão trás prejuízos financeiros para a economia global. De acordo com a *National Association of Corrosion Engineers (NACE)* a corrosão é responsável por perdas financeiras de 3.8% do Produto Interno Bruto americano [5]. O relatório da *World Custom Organization (WCO)* traz o custo das perdas por corrosão como 3-4% do PIB dos países industrializados, e que 30-50% desses custos poderiam ser evitados [6]. Essas perdas financeiras geradas pela corrosão podem ser classificadas como diretas e indiretas. As perdas diretas são causadas pelo custo de substituição de peças ou equipamentos que sofreram corrosão e os custos e manutenção de processos de proteção (proteção catódica, revestimentos metálicos e não metálicos, pinturas, etc). Os custos indiretos são causados por paralisações acidentais de processos produtivos industriais, perda de produtos por vazamentos em tubulações corroídas, perda de eficiência, contaminação de produtos e superdimensionamento em projetos, dentre outros [2].

Diante dos grandes prejuízos financeiros causados pela corrosão, justifica-se a necessidade de criar métodos de predição de corrosão de metais para o ambiente em que serão instalados. Para o desenvolvimento desses métodos, diversos projetos internacionais foram feitos com objetivo de construir bases de dados, tais como o *International Atmospheric Exposure Program (ISOCORRAG)* [7] e o Mapa da Ibero-América de

Corrosividade Atmosférica (MICAT) [8]. Atualmente, as normas internacionais ISO 9223 e a ISO 9224, contém fórmulas de predição de corrosão usando leis de potências e equações *dose-response* (DRF) desenvolvidas a partir de dados dessas bases [9] [10] [11].

Com o aperfeiçoamento da capacidade computacional ao longo das décadas, o uso algoritmos de aprendizado de máquina vem se demonstrando cada vez mais capacidade de resolver problemas com uso de bases de dados. Neste sentido, o objetivo deste trabalho é o desenvolvimento de modelos de aprendizado de máquina para a predição de corrosão atmosférica utilizando a base de dados do MICAT, a fim de comparar com as DRFs definidas pelas ISO e realizar uma análise comparativa. Desta forma, o problema abordado neste trabalho se caracteriza como sendo um problema de regressão univariada, onde a variável de interesse é a taxa de corrosão de diferentes materiais metálicos.

A estrutura deste trabalho é a seguinte: na seção II, será apresentado o embasamento teórico do artigo, abordando a corrosão atmosférica e os principais projetos relacionados a esse tema, incluindo a predição atmosférica tradicional e o uso de aprendizado de máquina, bem como trabalhos correlatos. A seção III descreve a metodologia utilizada e a estrutura da base de dados empregada nos modelos. Na quarta seção, são apresentados os protocolos de experimentos, juntamente com os resultados obtidos. Por fim, na seção V, são descritas as conclusões e os trabalhos futuros deste estudo.

II. REFERENCIAL TEÓRICO

A. Corrosão Atmosférica

Os ambientes corrosivos podem ser divididos em atmosférico, de águas naturais, de solo e de produtos químicos, e também em menor escala de alimentos, de substâncias fundidas, de solventes orgânicos, de madeira e de plásticos [2]. Dentre esses ambientes, vale destaque a corrosão atmosférica, a qual é responsável por cerca de metade do custo total das perdas por corrosão. Portanto, se fazem necessários meios de se prever o comportamento corrosivo de materiais nos diferentes ambientes atmosféricos.

O método mais utilizado para se entender o comportamento do processo de corrosão atmosférica são os ensaios de corrosão em campo, os quais consistem na exposição de espécimes (chapas retangulares do material metálico a ser testado, geralmente com dimensões de 150mmx100mm, com 1 à 3 mm de espessura [12]) em estações especificadas por normas, tais como a ASTM G50-20 [13] e a ISO 8565 [12]. Os ensaios seguem, com particularidades para cada norma, a seguinte ideia: espécimes são pesados e instalados em uma estação onde ficam expostos a ação do clima durante o período determinado (geralmente 1 ano). Após o término do período, as amostras são removidas e são submetidas a um processo de limpeza, onde a camada de óxidos formada é retirada utilizando metodologias especificadas por normas. Os espécimes então são pesados novamente, e cálculos de perda de espessura e de de massa são realizados. A Figura 1 mostra

uma estação utilizada em pesquisas realizadas pelo Instituto de Pesquisas Tecnológicas (IPT) [14].



Fig. 1. Estação de Corrosão usada na pesquisa do IPT.

Os ensaios são muito relevantes pois trazem a corrosividade do local de exposição de acordo com dados experimentais, entretanto para a maioria das aplicações tais ensaios são impraticáveis, seja pela capacidade técnica necessária para o manejo dos espécimes ou pelos longos períodos de exposição necessários para se obter resultados relevantes para aplicações comerciais. Portanto, surge a necessidade de prever a corrosão e/ou classificar a corrosividade do ambiente.

Nesse sentido, ensaios de campo desde o início do século XX foram realizados a fim de entender o comportamento dos metais na atmosfera. Percebeu-se que metais mais comumente utilizados (aço, zinco, cobre e alumínio) corroíam mais em ambientes poluídos. Descobriu-se que a combinação de alta umidade relativa (RH) com concentrações de SO₂, aumentava a corrosividade a partir de um valor crítico de RH [15]. Dentre todos os poluentes atmosféricos, cloretos provindos de aerossóis marinhos e dióxido de enxofre (SO₂) provenientes da queima de combustíveis fósseis são os mais agressivos agentes de corrosão da atmosfera [16].

Nas décadas 80 e 90 aconteceram grandes progressos na pesquisa de corrosão atmosférica, foram planejados e executados programas internacionais para a montagem de ensaios de campo e a coleta de dados climáticos, a fim de se criar bases de dados para definir relação entre as variáveis e a perda de espessura dos metais. Dentre os programas, vale destaque ao ISOCORRAG [7] e ao e ao projeto MICAT [8], os quais foram utilizados para a formação das normas ISO 9223, 9224, 9225 e 9226 [9] [10] [17] [18].

B. Projeto MICAT

O Projeto MICAT [19] [8], denominado Mapas de Ibero-América de Corrosividade Atmosférica, é um projeto do Programa Ibero-americano de Ciência e Tecnologia para o Desenvolvimento (CYTED). O MICAT foi realizado no período de 1989 à 1994, envolvendo 14 países (Argentina, Brasil, Colômbia, Costa Rica, Chile, Equador, México, Panamá, Peru, Portugal, Espanha, Uruguai e Venezuela) [20]. O projeto

tem como objetivo conseguir maior conhecimento acerca dos mecanismo de corrosão nas distintas atmosferas da Ibero-América e elaborar mapas de corrosividade atmosférica da Ibero-América. A Figura 2 mostra o mapa das estações do projeto.



Fig. 2. Mapa das Estações do Projeto MICAT [20]

Durante o programa foram coletados dados ambientais e de corrosão de 72 estações para 1, 2, 3 e 4 anos de exposição, com ocasionais estações expondo 5 anos. Os materiais expostos foram o Aço Carbono, Zinco, Cobre e Alumínio, para algumas estações foi também exposto o Aço Galvanizado, que constitui em um Aço Carbono protegido por uma camada externa de Zinco. Os dados ambientais coletados para cada estação foram a Temperatura Média Anual ($^{\circ}\text{C}$), Umidade Relativa Média (%), Tempo de Superfície Úmida (horas/ano em que Umidade está acima de %80 e temperatura acima de 0°C), Precipitação (mm/ano), Deposição de SO_2 e Cl^- ($\text{mg}/\text{m}^2\text{dia}$).

O projeto MICAT contribuiu profundamente para o estudo de corrosão atmosférica. Os dados do MICAT foram levados em consideração na construção da segunda versão da ISO 9223, tanto pela inserção de Temperatura e Umidade Relativa Média como parte da classificação das atmosferas [21], quanto a previsão de corrosão utilizando modelos desenvolvidos utilizando dados do MICAT [11].

C. Predição de Corrosão

A corrosão de metais e ligas não é constante ao longo do tempo e para a maioria dos metais ela diminui devido ao acúmulo de produtos de corrosão na superfície do metal exposto [10]. A predição de corrosão é feita utilizando uma combinação de duas normas: a ISO 9223 e a ISO 9224. A ISO 9224 utiliza a Equação 1 para descrever o comportamento logarítmico da corrosão a longo prazo.

$$D = r_{corr}t^b \quad (1)$$

Onde:

D = Ataque total ao material, em perda de massa ou espessura

t = Tempo de exposição em anos

r_{corr} = Taxa de corrosão para 12 meses (ISO 9223)

b = Constante metal-ambiente, geralmente menor que 1.

Como pode ser observado na Equação 1, é necessário o valores obtidos para 12 meses obtidos pela ISO 9223 para se realizar a predição a longo prazo. A ISO 9223 apresenta equações para a predição de taxa de corrosão para quatro materiais: aço carbono, zinco, alumínio e cobre [9]. As 4 equações fornecidas pela ISO foram desenvolvidas por Mikhailov e Tidblad [11] e são apresentadas *Dose Response Functions* (DRF). As Equações 2,3,4 e 5 representam as DRFs para Aço Carbono, Zinco, Cobre e Alumínio, respectivamente.

Aço Carbono

$$\begin{aligned} r_{corr} = & 1,77.P_d^{0,52}.exp(0,020.RH + f_{st}) + \\ & 0,102.S_d^{0,62}.exp(0,033.RH + 0,04T) \quad (2) \\ f_{st} = & 0,150.(T - 10) \text{ quando } T \leq 10^{\circ}\text{C}, \\ & \text{caso contrario } - 0,054.(T - 10) \end{aligned}$$

Zinco

$$\begin{aligned} r_{corr} = & 0,0129.P_d^{0,44}.exp(0,046.RH + f_{zn}) + \\ & 0,0017.S_d^{0,57}.exp(0,008.RH + 0,085T) \quad (3) \\ f_{zn} = & 0,038.(T - 10) \text{ quando } T \leq 10^{\circ}\text{C}, \\ & \text{caso contrario } - 0,071.(T - 10) \end{aligned}$$

Cobre

$$\begin{aligned} r_{corr} = & 0,005.P_d^{0,26}.exp(0,059.RH + f_{Cu}) + \\ & 0,010.S_d^{0,27}.exp(0,036.RH + 0,049.T) \quad (4) \\ f_{Cu} = & 0,126.(T - 10) \text{ quando } T \leq 10^{\circ}\text{C}, \\ & \text{caso contrario } - 0,080.(T - 10) \end{aligned}$$

Alumínio

$$r_{corr} = 0,0042.P_d^{0,73}.exp(0,025.RH + f_{Al}) + 0,0018.S_d^{0,60}.exp(0,020.RH + 0,094.T) \quad (5)$$

$$f_{Al} = 0,009.(T - 10) \text{ quando } T \leq 10^\circ C,$$

$$\text{caso contrario } - 0,043.(T - 10)$$

Onde:

r_{corr} = Corrosão do metal no primeiro ano ($\mu m/ano$)

T = Temperatura Média Anual (Celsius)

RH = Umidade Relativa Média Anual (%)

P_d = Deposição Anual Média de SO_2 ($mg/(m^2.d)$)

Cl = Deposição Anual Média de Cl ($mg/(m^2.d)$)

Diferentemente da abordagem definida pela ISO9223 com uso de equações é possível utilizar outras técnicas para a predição de corrosão, mais especificamente, a aplicação de algoritmos de Aprendizado de Máquina (*Machine Learning* - ML). ML é uma sub-área de Inteligência Artificial (IA) que desenvolve algoritmos que aprendem padrões em bases de dados de forma automatizada [22]. Essencialmente, estes algoritmos se propõem a resolver tarefas em dados de maneira supervisionada e não-supervisionada. No primeiro, os dados que serão utilizados para o aprendizado estão previamente catalogados e possuem uma ou mais variáveis alvo. Em contrapartida, no segundo, não existem rótulos nos dados, assim o modelo tentará encontrar padrões entre eles e, por exemplo, tentar agrupá-los de acordo com a similaridade e dissimilaridade dos dados [23].

Dentro do aprendizado supervisionado, há ainda uma subcategoria que abrange tarefas de regressão e de classificação. Problemas que envolvem a predição de um ou mais valores numéricos são determinados como regressão. Por outro lado, se o alvo do problema for uma ou mais classes, então tem-se uma tarefa de classificação [22].

Um exemplo de algoritmo de regressão é o *Random Forest* (RF). O RF é um algoritmo que combina várias árvores de regressão, treinadas em amostras aleatórias dos dados. Essa abordagem utiliza a técnica de *bagging*, em que cada árvore é construída com uma amostra de dados diferente das outras. Isso permite que cada árvore tome decisões independentes, explorando diferentes aspectos do espaço de busca. No final, os valores encontrados são agregados por meio de uma média para obter a predição final [24].

Essa abordagem tem a vantagem de evitar erros e predições incorretas que poderiam ocorrer se apenas uma árvore de regressão fosse utilizada. Ao combinar as previsões de várias árvores, o RF reduz a tendência de *overfitting* e melhora a estabilidade e a precisão geral do modelo de regressão.

Dessa forma, o RF se destaca como um método poderoso para a regressão, capaz de lidar com conjuntos de dados com-

plexos e variáveis de entrada correlacionadas. Sua capacidade de capturar relações não lineares e interações entre variáveis contribui para um desempenho mais robusto e confiável na tarefa de prever valores contínuos.

D. Trabalhos Correlatos

No trabalho de [25] é utilizado o algoritmo RF para a predição de um ano da corrosão atmosférica de aço carbono. O autor utilizou as bases de dados somente para o aço carbono dos projetos MICAT da publicação de 1995 [19], ISOCORRAG e ECE/UN para treinar o modelo e utilizou amostras russas para testar o modelo. Os dados do MICAT publicados em 1995 são uma média das 3 exposições de 1 ano do MICAT, enquanto na publicação de 1998 consta as 3 exposições separadamente, conforme a Figura 4. Também foi comparado com a DRF de aço carbono, onde o modelo atingiu melhores resultados na maiorias dos testes, demonstrando a eficácia do RF. A união das bases do MICAT, ISOCORRAG e ECE/UN apresentam problemas, pois o projeto ISOCORRAG não possui Umidade Relativa para todos os países participantes exceto a República Tcheca, e o ECE/UN não possui deposição de cloreto [16]. A integração da base de dados do MICAT e do ISOCORRAG depende da utilização do modelo de predição TOW-T-RH [26] e portanto tem um erro associado.

Utilizando uma base de dados chinesa, o trabalho apresentado em [27] desenvolveu modelos de RF e Máquina de Vetores de Suporte (SVM) para prever a corrosão atmosférica para aço carbono. Os seus resultados demonstraram que a RF quando os seus parâmetros estão ajustados, retornam melhores resultado que a SVM.

Em um estudo localizado no Japão, [28] apresenta uma análise de corrosão atmosférica do Aço Carbono, com aplicação de algoritmos de aprendizado de máquina. Diferentemente dos artigos citados, os modelos utilizaram variáveis de entrada adicionais, como a radiação solar e a velocidade do vento. Entre os resultados obtidos, os modelos baseados em árvores de decisão apresentaram o melhor desempenho. O artigo também teve como objetivo investigar as variáveis que exercem maior influência no processo corrosivo. Para essa análise, os autores utilizaram o algoritmo de RF, uma vez que ele é capaz de medir o peso de cada variável de forma precisa.

Dos trabalhos correlatos, a escolha do algoritmo RF se mostra interessante, tendo como base a disponibilidade limitada de dados para treinamento dos modelos. O algoritmo demonstra um desempenho adequado mesmo quando se dispõe de uma pequena quantidade de dados e, além disso, é capaz de lidar com relações não lineares entre as variáveis.

III. MODELO PROPOSTO E DESCRIÇÃO DA BASE DE DADOS

O modelo proposto é descrito no diagrama da Figura 3. Das etapas do projeto, primeiramente os dados do projeto MICAT foram coletados da referência bibliográfica [8]. Por ser um volume impresso e a base de dados não estar disponível digitalmente, a coleta de dados foi realizada manualmente e, posteriormente, revisada para garantir a precisão na coleta e

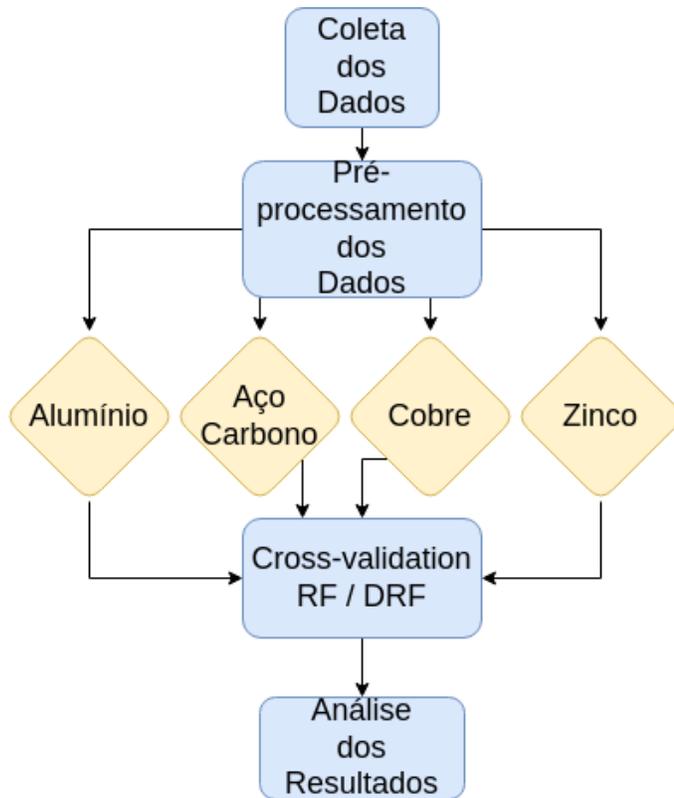


Fig. 3. Workflow do processo de aprendizado de máquina

evitar erros. A base de dados está disponível em download MICAT. Na próxima etapa do projeto, foi realizado pré-processamento dos dados, que envolveu a remoção de valores faltantes e a aplicação de transformações que serão detalhadas ainda nesta seção. A base de dados contém informações sobre os materiais: Aço Carbono, Alumínio, Cobre e Zinco. Portanto, os dados foram separados por material. Com a base consolidada, foi empregado o método de *cross-validation* para criar os modelos de RF e para obter os resultados das DRFs. Por fim, foi realizada uma análise comparativa entre os modelos RF e as DRFs.

Na base de dados do projeto MICAT encontram-se estudos que envolvem 14 diferentes países pertencentes a América Latina, com exceção de Portugal e Espanha. Para cada cidade que participou do projeto, 6 amostras foram coletadas, para quatro tipos de materiais: Aço carbono, Alumínio, Cobre e Zinco. Destas 6 coletas, foram utilizadas 3, visto que as demais coletas eram referentes a períodos de amostragem maiores que de 1 ano, como é mostrado na Figura 4.

Deste estudo, foram selecionadas amostras que possuíam temperatura, umidade relativa, níveis de cloreto no ambiente, níveis de sulfato no ambiente e valores de corrosão, contabilizando 587 amostras. Na Tabela I é possível observar a distribuição das variáveis. Todas estas variáveis foram incorporadas na base coletada e utilizadas para construção dos modelos de predição.

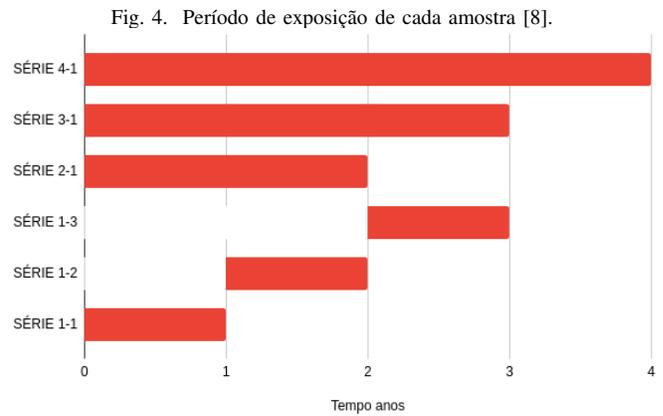


Fig. 4. Período de exposição de cada amostra [8].

TABLE I
SUMÁRIO DAS VARIÁVEIS DE AMBIENTE

Variável	Unidade	Média	Desvio Padrão
Temperatura (T)	°C	20.25	5.61
Umidade Relativa (RH)	%	74.44	10.42
Níveis de Cloreto	mg/m ² dia	32.21	66.30
Níveis de Sulfato	mg/m ² dia	14.87	16.52

Devido a dispersão dos valores de cloreto, sulfato e corrosão, transformações logarítmicas foram aplicadas a fim de reduzir estas distâncias. A fórmula adotada foi: $\ln(x + 1)$, onde x é o valor a ser transformado.

A Figura 5 apresenta o gráfico de dispersão dos dados de Aço Carbono para 12 meses envolvendo as variáveis: cloreto, sulfato e corrosão. No gráfico, o eixo x representa os valores de cloreto, o eixo y representa os valores de sulfato e a cor, que varia de roxo para amarelo, representa os valores de corrosão. Fica nítida a relação de aumento da corrosão quando existem altos valores de sulfato e cloreto. O mesmo comportamento acontece para os demais materiais, contudo fica mais evidente para o Aço Carbono por conta da alta variação dos valores de corrosão.

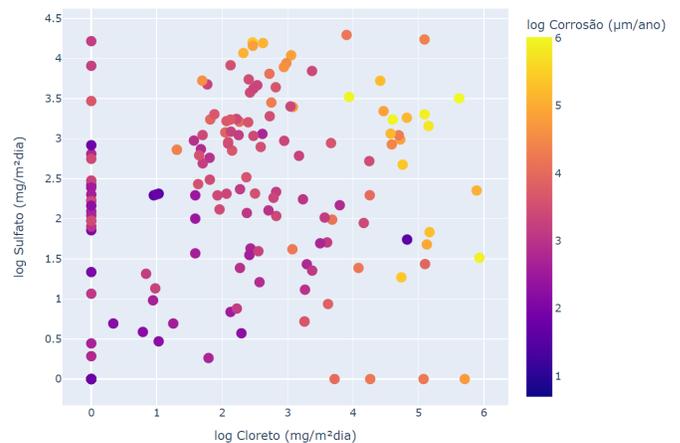


Fig. 5. Gráfico de Dispersão dos Dados para o Aço Carbono em 12 meses, em função de SO₂ e Cl-

IV. EXPERIMENTOS, RESULTADOS E ANÁLISES

A. Protocolo de Experimentos

Todos os experimentos foram conduzidos no Laboratório de Inteligência Computacional¹ utilizando ambiente de computação com sistema operacional Ubuntu 20.04, hardware com 16 GB RAM e processador AMD Ryzen 7 2700X. Os modelos foram desenvolvidos na linguagem Python 3.8.10 com o auxílio da biblioteca Scikit-learn [29].

A métrica avaliativa de erro médio absoluto (MAE) foi calculada. A Equação 6 traz o cálculo do MAE.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (6)$$

onde \hat{y} representa o valor predito, enquanto y representa o valor real. Essa métrica oferece uma visão geral do modelo, sem destacar erros significativos. O valor mínimo é 0, e quanto maior o valor, pior é o desempenho do modelo.

Foi aplicado *cross-validation* de 5 *folds* para criação e avaliação dos modelos de *Random Forest* e para avaliação das funções DRFs. Nesta abordagem, os dados são separados aleatoriamente em 5 partes, onde 4 partes são utilizadas para treinamento e a parte restante é utilizada para teste do modelo. Este processo é repetido 5 vezes modificando a partição de teste. Para cada *fold* de teste, são calculados os valores de MAE dos modelos de RF e das DRFs.

Para análise estatística dos resultados, o teste de DUNN foi aplicado. Neste trabalho, uma significância de 5% foi utilizada para aferir se os resultados obtidos pode ser considerados estatisticamente diferentes ou equivalentes entre os modelos de RF e as DRFs.

B. Resultados e Análises

Antes do treinamento dos modelos, foi realizada uma análise minuciosa do comportamento dos dados, focando na correlação entre as variáveis dependentes e independentes. Para avaliar essa correlação, foi utilizada a correlação de Pearson, uma medida estatística que indica o grau de associação linear entre duas variáveis. Os valores de correlação foram representados em matrizes para cada material, conforme ilustrado nas Figuras 6, 7, 8 e 9.

A escala dos valores de correlação de Pearson varia de -1 a 1. Quando o valor é 1, indica uma correlação positiva perfeita, o que significa que, à medida que uma variável aumenta, a outra também aumenta de forma proporcional. Por outro lado, quando o valor de correlação é -1, indica uma correlação negativa perfeita. Nesse caso, à medida que uma variável aumenta, a outra diminui de forma proporcional.

De acordo com as matrizes, as variáveis ambientais afetam os materiais de maneiras distintas no que se refere à corrosão. Não é viável atribuir um único valor a cada material e estabelecer uma correlação de corrosão geral. Um exemplo disso é a influência predominante do cloreto na corrosão do aço carbono, enquanto para outros metais outros fatores

também merecem destaque. As relações mencionadas foram estabelecidas exclusivamente para a base de dados MICAT.

Visto que os metais são distintos nos graus dos fatores corrosivos, foi desenvolvido um fluxo de *cross validation* para cada material. Na Tabela II é possível observar a média e o desvio padrão dos modelos para cada material.

TABLE II
MÉDIA DAS *folds* DO *cross validation* UTILIZANDO AS MÉTRICAS MAE E DUNN PARA OS MATERIAIS: AÇO CARBONO, ALUMÍNIO, COBRE E ZINCO, RESPECTIVAMENTE.

Aço Carbono	
Modelo	MAE
DRF	22.5192 ± 6.0659
RF	10.6191 ± 5.0689
DUNN	0.0282
Alumínio	
Modelo	MAE
DRF	0.3973 ± 0.0334
RF	0.0776 ± 0.0247
DUNN	0.0090
Cobre	
Modelo	MAE
DRF	0.6696 ± 0.085
RF	0.4195 ± 0.1067
DUNN	0.0162
Zinco	
Modelo	MAE
DRF	0.7822 ± 0.1538
RF	0.3883 ± 0.0782
DUNN	0.0090

Todos valores dos testes de DUNN da Tabela II apresentam valores menores que 5%, indicando que existem diferenças estatísticas entre os modelos RF e as DRFs. Para a métrica MAE, os modelos de RF mostraram melhores resultados quando comparado com as DRFs.

Os modelos de RF demonstram, de forma evidente, uma escala de erros consideravelmente menor em comparação com os DRFs. A média dos erros dos modelos de Aço Carbono e do Zinco são a metade quando comparados com as DRFs. Para o Alumínio o modelo se comporta ainda melhor, possuindo um quarto do erro quando comparado com a sua DRF.

Para uma melhor visualização, a Figura 10 apresenta em formato de diagrama de caixa os resultados dos *cross-validations*. Com esta visualização fica evidente que os resultados do modelo RF para o material alumínio se destacam significativamente em comparação com os da DRF.

¹LABICOM website: <https://labicom-udesc.github.io/>

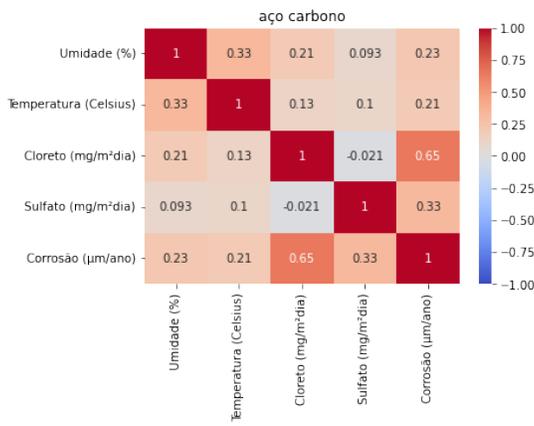


Fig. 6. Matriz de correlação de Pearson para Aço Carbono

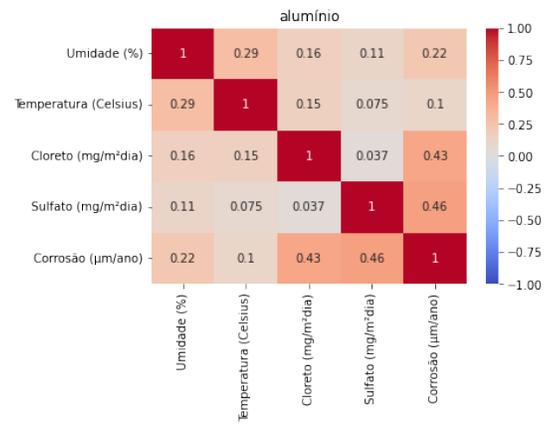


Fig. 7. Matriz de correlação de Pearson para Alumínio

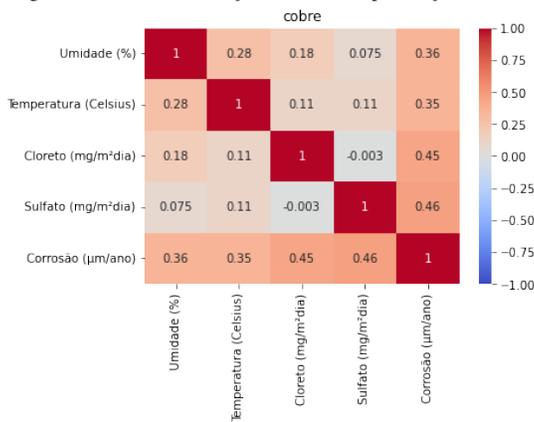


Fig. 8. Matriz de correlação de Pearson para Cobre

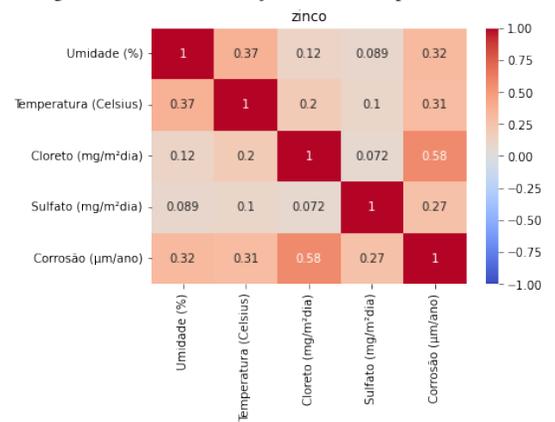


Fig. 9. Matriz de correlação de Pearson para Zinco

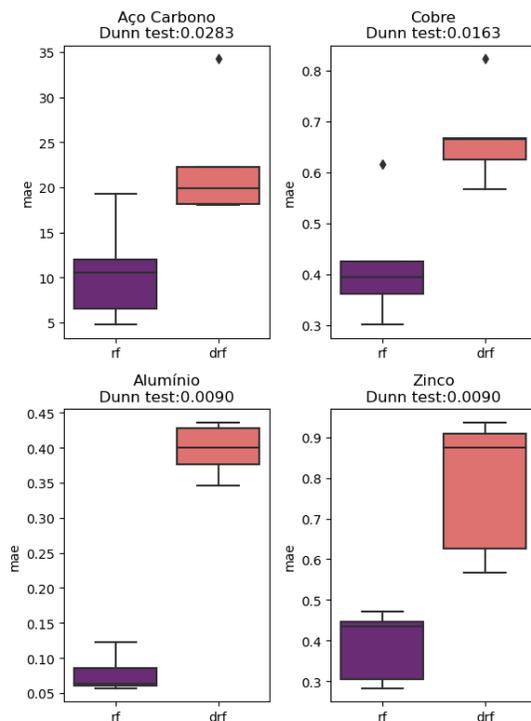


Fig. 10. Diagramas de caixa para os resultados obtidos nos materiais

V. CONCLUSÃO E TRABALHOS FUTUROS

A predição da corrosão atmosférica pode trazer benefícios significativos para as indústrias e empresas, permitindo a redução de prejuízos associados à criação e desenvolvimento de projetos metalúrgicos. Embora as DRFs empregadas pelas normas ISO forneçam uma previsão, elas não apresentam um nível de assertividade elevado. No entanto, este trabalho demonstrou que algoritmos de aprendizado de máquina, como o RF, são capazes de entregar resultados competitivos em comparação com as fórmulas DRFs.

Os modelos de RF utilizaram as variáveis de entrada descritas como as mais importantes pela ISO 9223. No entanto, este trabalho propõe explorar outras variáveis no futuro, como a pluviosidade anual, que podem ter um impacto significativo na predição da corrosão atmosférica.

Atualmente, apenas a base de dados do MICAT, que contém informações da América Latina, foi utilizada para desenvolver os modelos. No futuro, a inclusão de conjuntos de dados adicionais pode enriquecer o estudo, abrangendo diferentes regiões geográficas e permitindo uma análise mais abrangente dos fatores que influenciam a corrosão atmosférica. Isso con-

tribuirá para um modelo de previsão mais robusto e aplicável a diferentes condições atmosféricas.

Com um maior número de dados e o aumento dimensionalidade do problema com novas variáveis de entrada, outros algoritmos, como Redes Neurais Artificiais, também podem ser empregadas no desenvolvimento da abordagem proposta.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil), com apoio da agência FAPESC (Fundação de Amparo à Pesquisa e Inovação de Santa Catarina), com apoio da empresa ArcelorMittal S/A - Brasil e com apoio da CAPES - Brazil (PROAP/AUXPE - CONSOLIDA 3/4).

REFERENCES

- [1] P. A. S. PE, *Fundamentals of Metallic Corrosion: Atmospheric and Media Corrosion of Metals*. CRC Press, 2006.
- [2] V. Gentil, *Corrosão*. 6ª edição, 2011.
- [3] F. D. Pannoni, *Manual De Construção do Aço: Projeto e Durabilidade*. Instituto Aço Brasil, 2017.
- [4] A. A. M. T. Adikari, R. G. N. D. S. Munasinghe, and S. Jayatilake, "Prediction of atmospheric corrosion –a review," *Engineer: Journal of the Institution of Engineers, Sri Lanka*, vol. 47, no. 2, p. 75, Apr. 2014. [Online]. Available: <https://doi.org/10.4038/engineer.v47i2.6869>
- [5] G. H. Koch, M. P. Brongers, N. G. Thompson, Y. P. Virmani, J. H. Payer *et al.*, "Corrosion cost and preventive strategies in the united states," United States. Federal Highway Administration, Tech. Rep., 2002.
- [6] G. Schmitt, M. Schütze, G. F. Hays *et al.*, "Global needs for knowledge dissemination, research, and development in materials deterioration and corrosion control," *World Corrosion Organization*, vol. 38, p. 14, 2009.
- [7] D. Knotkova, K. Kreislova, and S. W. Dean, Eds., *ISOCORRAG International Atmospheric Exposure Program: Summary of Results*. ASTM International, May 2010. [Online]. Available: <https://doi.org/10.1520/ds71-eB>
- [8] M. Morcillo, E. Almeida, B. Rosales, J. Uruchurtu, and M. Marrocos, *Corrosion y Protección de Metales en las Atmósferas de Iberoamérica. Parte I — Mapas de Iberoamérica de Corrosividad Atmosférica (Proyecto MICAT, XVI/CYTED)*. CYTED, 1998.
- [9] ISO/9223, "Corrosion of metals and alloys - corrosivity of atmospheres -classification, determination and estimation," 2012.
- [10] ISO/9224, "Corrosion of metals and alloys — corrosivity of atmospheres — guiding values for the corrosivity categories," 2012.
- [11] A. A. Mikhailov, J. Tidblad, and V. Kucera, "The classification system of ISO 9223 standard and the dose-response functions assessing the corrosivity of outdoor atmospheres," *Protection of Metals*, vol. 40, no. 6, pp. 541–550, Nov. 2004. [Online]. Available: <https://doi.org/10.1023/b:prom.0000049517.14101.68>
- [12] I. 8565:2011, "Metals and alloys — atmospheric corrosion testing — general requirements," 2011.
- [13] ASTM/G50-20, "Standard practice for conducting atmospheric corrosion tests on metals," 2020.
- [14] N. L. de Almeida and Z. Panossian, *Corrosão Atmosférica 17 anos*. IPT, 1999.
- [15] C. Leygraf, I. O. Wallinder, J. Tidblad, and T. Graedel, *Atmospheric corrosion*. John Wiley & Sons, 2016.
- [16] B. Chico, I. Díaz, J. Simancas, and M. M. and, "Annual atmospheric corrosion of carbon steel worldwide. an integration of ISOCORRAG, ICP/UNECE and MICAT databases," *Materials*, vol. 10, no. 6, p. 601, May 2017. [Online]. Available: <https://doi.org/10.3390/ma10060601>
- [17] I. 9225:2012, "Corrosion of metals and alloys — corrosivity of atmospheres — measurement of environmental parameters affecting corrosivity of atmospheres," 2012.
- [18] I. 9226:2012, "Corrosion of metals and alloys — corrosivity of atmospheres — determination of corrosion rate of standard specimens for the evaluation of corrosivity," 2012.
- [19] M. Morcillo, "Atmospheric corrosion in ibero-america: The MICAT project," in *Atmospheric Corrosion*. ASTM International, 1995, pp. 257–257–19. [Online]. Available: <https://doi.org/10.1520/stp14924s>
- [20] M. Morcillo, B. Chico, D. de la Fuente, and J. Simancas, "Looking back on contributions in the field of atmospheric corrosion offered by the MICAT ibero-american testing network," *International Journal of Corrosion*, vol. 2012, pp. 1–24, 2012. [Online]. Available: <https://doi.org/10.1155/2012/824365>
- [21] M. Morcillo, E. Almeida, B. Chico, and D. de la Fuente, "Analysis of ISO standard 9223 (classification of corrosivity of atmospheres) in the light of information obtained in the ibero-american micat project," in *Outdoor Atmospheric Corrosion*. ASTM International, 2002, pp. 59–59–14. [Online]. Available: <https://doi.org/10.1520/stp10883s>
- [22] S. J. RUSSELL and P. NORVIG, *Artificial intelligence*, 3rd ed. New Jersey: Prentice Hall, 2010.
- [23] A. Geron, *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media, 2017.
- [24] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001. [Online]. Available: <https://doi.org/10.1023/A:1010933404324>
- [25] M. Gavryushina, A. Marshakov, and Y. Panchenko, "Application of the random forest algorithm to predict the corrosion losses of carbon steel over the first year of exposure in various regions of the world," *Corrosion Engineering, Science and Technology*, vol. 58, no. 3, pp. 205–213, 2023.
- [26] J. Tidblad, A. A. Mikhailov, and V. Kucera, "Model for the prediction of the time of wetness from average annual data on relative air humidity and air temperature," *Protection of Metals*, vol. 36, no. 6, pp. 533–540, 2000. [Online]. Available: <https://doi.org/10.1023/a:1026621009635>
- [27] Y. Zhi, Z. Jin, L. Lu, T. Yang, D. Zhou, Z. Pei, D. Wu, D. Fu, D. Zhang, and X. Li, "Improving atmospheric corrosion prediction through key environmental factor identification by random forest-based model," *Corrosion Science*, vol. 178, p. 109084, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0010938X20323659>
- [28] L. Yan, Y. Diao, Z. Lang, and K. Gao, "Corrosion rate prediction and influencing factors evaluation of low-alloy steels in marine atmosphere using machine learning approach," *Science and Technology of Advanced Materials*, vol. 21, no. 1, pp. 359–370, 2020, pMID: 32939161. [Online]. Available: <https://doi.org/10.1080/14686996.2020.1746196>
- [29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.