



INSTITUTO FEDERAL DE CIÊNCIA E TECNOLOGIA DE
PERNAMBUCO

Campus Recife

CACSEM

Curso de Bacharelado em Engenharia Mecânica

FÁBIO HENRIQUE PALMEIRA SILVA

**DESENVOLVIMENTO DE MODELO COMPUTACIONAL BASEADO EM
MACHINE LEARNING PARA PREDIÇÃO DE PROPRIEDADES
ANTICORROSIVAS DE COMPÓSITOS POLIMÉRICOS ADITIVADOS COM
GRAFENO**

Recife

2023

FÁBIO HENRIQUE PALMEIRA SILVA

**DESENVOLVIMENTO DE MODELO COMPUTACIONAL BASEADO EM
MACHINE LEARNING PARA PREDIÇÃO DE PROPRIEDADES
ANTICORROSIVAS DE COMPÓSITOS POLIMÉRICOS ADITIVADOS COM
GRAFENO**

Trabalho de conclusão de curso apresentado ao Departamento de Engenharia Mecânica do Instituto Federal de Ciência e Tecnologia de Pernambuco, como requisito para a obtenção do título de Bacharel em Engenharia Mecânica.

Orientador: Prof. Dr. Frederico Duarte de Menezes

Recife
2023

S586d

2023

Silva, Fábio Henrique Palmeira.

Desenvolvimento de modelo computacional baseado em Machine Learning para predição de propriedades anticorrosivas de compósitos poliméricos aditivados com grafeno / Fábio Henrique Palmeira Silva. --- Recife: O autor, 2023.
52f. il. Color.

TCC (Curso de Engenharia Mecânica) – Instituto Federal de Pernambuco, Recife, 2023.

Inclui Referências.

Orientador: Professor Dr. Frederico Duarte de Menezes.

1. Corrosão. 2. Nanocompósitos poliméricos. 3. Machine Learning. I. Título. II. Menezes, Frederico Duarte de (orientador). III. Instituto Federal de Pernambuco.

CDD 620.1 (22ed.)

**DESENVOLVIMENTO DE MODELO COMPUTACIONAL BASEADO EM
MACHINE LEARNING PARA PREDIÇÃO DE PROPRIEDADES
ANTICORROSIVAS DE COMPÓSITOS POLIMÉRICOS ADITIVADOS COM
GRAFENO**

_____. Recife, _____.

Professor Orientador

Convidado 1

Convidado 2

Recife

2023

RESUMO

Os males diretos e indiretos causados pelo fenômeno corrosivo na deterioração de metais representam gastos e consequências muitas vezes incomensuráveis, afetando não só os setores públicos, comerciais e industriais, mas também representando problemas de cunho ambiental e saúde pública. Uma abordagem que tem cada vez mais despertado o interesse do meio científico e industrial é a utilização de nanocompósitos poliméricos com propriedades anticorrosivas, atuando como barreira entre o substrato metálico e o meio corrosivo. Estes nanocompósitos são desenvolvidos a partir da combinação de uma matriz polimérica e nanopartículas (fase dispersa) de propriedades desejáveis, que resultarão num material com propriedades otimizadas. Contudo, devido à amplitude de combinações possíveis entre compostos nas mais diversas concentrações, torna-se inviável analisar, por meio de métodos convencionais, quais combinações são ideias para uma aplicação específica. Para contornar este problema pode-se recorrer ao uso de modelos preditivos baseados em aprendizado de máquina (*Machine Learning*). Em vista disso, neste trabalho, foram estudados modelos computacionais baseados em *Machine Learning*, para a predição de características anticorrosivas de nanocompósitos poliméricos. Foram catalogados 100 artigos da literatura, resultando em um conjunto de dados com 40 instâncias. Os modelos de regressão para prever o potencial de corrosão do aço carbono recoberto com os nanocompósitos selecionados foram treinados através da abordagem de aprendizado de máquina automatizado (*AutoML*), sendo utilizadas ferramentas de quimioinformática para adequar a representação dos polímeros constituintes de cada compósito. Quando a variável resposta para o teste do modelo foi representada pelo potencial de corrosão obteve-se um valor da métrica de avaliação relativamente baixo para o erro quadrático médio ($R^2 = 0,143$). No entanto, pôde-se observar um melhor resultado para o erro quadrático médio ($R^2 = 0,429$) no teste do modelo quando a variável resposta é representada pela diferença entre potenciais de corrosão, com e sem a presença do óxido de grafeno nos nanocompósitos.

Palavras-chave: Corrosão; Nanocompósitos; *Machine Learning*; Quimioinformática. AutoML.

ABSTRACT

The direct and indirect damages caused by the corrosive phenomenon in the deterioration of metals represent costs and consequences that are often incalculable, affecting not only the public, commercial and industrial sectors, but also posing environmental and public health problems. An approach that has increasingly attracted the interest of the scientific and industrial community is the use of polymeric nanocomposites with anticorrosive properties that act as a barrier between the metallic substrate and the corrosive environment. These nanocomposites are developed from the combination of a polymer matrix and nanoparticles (dispersed phase) with desirable properties, resulting in a material with optimized properties. However, due to the range of possible combinations between compounds in different concentrations, it becomes impractical to analyze by conventional methods which combinations are ideal for a specific application. To circumvent this problem, one can resort to the use of predictive models based on machine learning. With this in mind, this work investigated computational models based on machine learning for predicting the anticorrosive properties of polymeric nanocomposites. 100 literature articles were cataloged, resulting in a data set of 40 instances. Regression models to predict the corrosion potential of covered carbon steel with selected nanocomposites were trained through the Automated Machine Learning (*AutoML*) approach, using chemoinformatics tools to adjust the representation of the constituent polymers of each composite. When the response variable for the test model was represented by the corrosion potential, a relatively low evaluation metric value was obtained for the mean quadratic error ($R^2 = 0,143$). However, a better result could be observed for the mean quadratic error ($R^2 = 0,429$) in the test model when the response variable is represented by the difference between corrosion potentials, with and without the presence of graphene oxide in the nanocomposites.

Keywords: Corrosion; Nanocomposites; Machine Learning; Chemoinformatics; AutoML.

LISTA DE FIGURAS

Figura 1 - Classificação dos tipos de inibidores de corrosão quanto a sua natureza química e comportamento nas reações	16
Figura 2 - Incidência do fenômeno corrosivo numa instalação industrial	17
Figura 3 - Mecanismo da corrosão eletroquímica num eletrodo aço carbono em meio aquoso	19
Figura 4 - Curvas de polarização representando de forma esquemática a corrosão de um metal em solução aquosa desaerada.....	21
Figura 5 - Curvas de polarização anódica e catódica (a) Curva de polarização resultante (b)	24
Figura 6 - Esquema da corrosão de um metal cujas reações catódica e anódica são controladas pela polarização de ativação na região de <i>Tafel</i>	25
Figura 7 - Representação da equação de <i>Wagner-Traud</i> em um gráfico E vs. $\log \Delta i $	27
Figura 8 - Classificações possíveis para os compósitos baseando-se na estrutura da fase dispersa.....	29
Figura 9 – Representação geométrica das formas alotrópica do carbono	31
Figura 10 - Representações para a estrutura molecular da Aspirina	39
Figura 11 - Etapas realizadas durante o desenvolvimento do trabalho	41

LISTA DE TABELAS

Tabela 1 - Modelo de Catalogação dos dados referentes as propriedades anticorrosivas dos nanocompósitos	42
Tabela 2 - Valores das métricas de avaliação dos modelos de regressão obtidos, tendo como variável de resposta o valor de E_{corr}	46
Tabela 3 - Valores das métricas de avaliação dos modelos de regressão obtidos, tendo como variável de resposta o ΔE_{corr}	46

LISTA DE ABREVIATURAS

GO	Óxido de grafeno
GF	Grafeno
<i>ML</i>	<i>Machine Learning</i>
NC	Nanocompósito

LISTA DE SÍMBOLOS

b_a	Coeficiente anódico do declive de <i>Tafel</i>
b_c	Coeficiente catódico de declive de <i>Tafel</i>
E	Potencial
E_{corr}	Potencial de Corrosão
ΔE_{corr}	Diferença entre potenciais de corrosão
i	Densidade de corrente
i_a	Densidade de corrente anódica
i_c	Densidade de corrente catódica
i_{corr}	Densidade de corrente de corrosão

SUMÁRIO

1	INTRODUÇÃO	12
1.1	CONTEXTUALIZAÇÃO	12
1.2	OBJETIVOS	14
2	REVISÃO BIBLIOGRÁFICA	15
2.1	AÇO CARBONO	15
2.2	INIBIDORES DE CORROSÃO	15
2.3	O MECANISMO DA CORROSÃO	17
2.4	POTENCIAL DE CORROSÃO	20
2.5	VELOCIDADE DE CORROSÃO	22
2.6	CURVAS DE POLARIZAÇÃO	23
2.6.1	Equação de <i>Wagner-Traud</i> e Extrapolação de <i>Tafel</i>	24
2.8	MATERIAIS COMPÓSITOS E NANOMATERIAIS	27
2.8.1	Nanocompósitos e suas particularidades	28
2.8.2	Grafeno	30
2.8.3	Óxido de grafeno	32
2.9	MACHINE LEARNING: FUNDAMENTOS	34
2.9.1	Aprendizado supervisionado	35
2.9.2	Aprendizado não supervisionado	36
2.9.3	Abordagem do <i>Auto Machine Learning</i>	36
2.9.4	Ferramentas de quimioinformática	38
2.9.4.1	Representação gráfica: <i>SMILES</i>	38
2.9.4.2	Criação de descritores: <i>mol2vec</i>	39
3	METODOLOGIA	41
4	ANÁLISE DOS RESULTADOS	44
4.1	LEVANTAMENTO BIBLIOGRÁFICO	44
4.2	CONSTRUÇÃO DE CONJUNTO DE TREINAMENTO PARA <i>ML</i>	44

4.3	CRIAÇÃO DE DESCRITORES	45
4.4	TREINAMENTO DE MODELOS DE REGRESSÃO.....	46
5	CONSIDERAÇÕES FINAIS	48
	REFERÊNCIAS.....	49

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

A utilização de técnicas metalúrgicas para a extração e confecção de metais remontam a antiguidade, o domínio de tecnologias metalúrgicas fora tão importante para o progresso da humanidade que se tornaram marcos históricos (MOURÃO et al., 2007).

A necessidade de metais com melhores propriedades mecânicas culminou em avanços tecnológicos expressivos no campo da metalurgia e siderurgia durante a Primeira e Segunda Revolução Industrial, nos séculos XVIII e XIX, respectivamente. O final do século XIX foi marcado pelo desenvolvimento de processos inovadores na conversão de ferro gusa, um dos fatores que proporcionou um aumento considerável na produção do aço. Entre os anos de 1870 e 1900 o aumento da produção do aço foi de cerca de 56 vezes em comparação as décadas anteriores, atingindo uma produção mundial de 28 milhões de toneladas na virada do século. Este crescimento na produção promoveu uma expressiva expansão na indústria da metalmecânica, que se tornou apta a prover metal de qualidade para suprir os setores industriais, da construção civil e ferroviário (MOURÃO, et al., 2007).

Atualmente, países fortemente industrializados têm uma elevadíssima demanda por aço e suas ligas, seja para satisfazer a progressiva mecanização industrial como para o desenvolvimento de componentes que exigem materiais que representam o estado da arte metalúrgica. Pode-se assumir então que, quanto mais desenvolvido tecnologicamente for uma nação, maior será sua necessidade por metais de alta qualidade, e conseqüentemente, maior serão os gastos causados por problemas aos quais as ligas metálicas estão sujeitas, sendo o principal deles, a corrosão (GENTIL, 2011).

Os males diretos e indiretos causados pela corrosão afetam diversos setores de um país, provocando desde perdas econômicas massivas a danos à saúde e ao meio ambiente (GENTIL, 2011).

Para a indústria, custos relacionados às conseqüências indiretas da corrosão são muitas vezes difíceis de se avaliar. Estruturas metálicas, máquinas e tubulações são continuamente acometidos pela corrosão, de forma que eventualmente sua

funcionalidade pode acabar sendo comprometida. Por conseguinte, torna-se indispensável a adoção, ainda que em fases iniciais de projetos de máquina ou instalações industriais, a adoção de medidas que evitem ou retardem ao máximo os danos causados pelo processo corrosivo (GENTIL, 2011).

Uma das formas de evitar a deterioração de superfícies metálicas é através do revestimento da superfície metálica por inibidores corrosivos. A utilização de nanocompósitos (NCs) aditivados com grafeno (GF) ou óxido de grafeno (GO) para atuarem como inibidores corrosivos desempenham um importante papel no combate à corrosão. A incorporação do GF ou GO pode promover um significativo aumento das propriedades anticorrosivas de uma matriz polimérica, produto de suas excelentes propriedades eletrônicas, mecânicas e microestruturais (AMUDHA A et al., 2019).

O desenvolvimento otimizado de novos materiais é um dos grandes desafios da atualidade. Dispor de processos automatizados que permitam a tomada de decisões rápidas, sobre que material deve ser desenvolvido para uma dada aplicação, são cruciais para a economia de tempo e recursos financeiros no desenvolvimento destes materiais. Neste contexto, ferramentas computacionais baseados em Aprendizado de Máquina ou *Machine Learning (ML)* (GU et al., 2018; HAMIDI et al., 2020), vem sendo cada vez mais utilizadas como primeira etapa no desenvolvimento de novos materiais, pois permitem a obtenção de resultados rápidos, com baixo custo e que promovem suporte no direcionamento das etapas experimentais de produção do material (GU et al., 2018; HAMIDI et al., 2020)

Contudo, diferentemente problemas clássicos resolvidos através do uso de técnicas de *ML*, o estudo de predição ou seleção de materiais torna-se uma tarefa complexa, pois diferentes moléculas possuem diferentes quantidades de átomos constituintes, assim como número e tipos de ligações entre estes átomos, tornando a tarefa de representar estas espécies químicas de forma tabular, isto é, com o mesmo número de descritores representando cada espécie química, um tanto difícil (LOUIS et al., 2020).

Para contornar este obstáculo, redes neurais baseadas em grafos, ou *graph neural networks (GNN)*, vem sendo estudadas e aplicadas com êxito para problemas na área de ciência dos materiais. Desta forma, pode-se construir novos conjunto de dados subdividindo-os em conjuntos de treinamento e teste, para a construção de modelos de classificação ou regressão para a resolução de problemas na área de materiais (LOUIS et al., 2020).

No contexto dos materiais compósitos, a substituição da fase dispersa em escala macro ou microscópica para materiais em escala nanométrica resulta em uma nova classe de materiais denominados nanocompósitos, com significantes melhorias nas mais diversas propriedades do material obtido (KUMAR; KRISHNAMOORTI, 2010). Ciente da importância do desenvolvimento de tecnologias para agilizar o processo inicial do desenvolvimento de materiais, manifestou-se o interesse no tema do presente trabalho, que teve como foco o estudo e desenvolvimento de um modelo computacional baseado em *Machine Learning* para atuar como ferramenta de suporte a decisão na seleção de nano-compósitos poliméricos aditivados com grafeno ou óxido de grafeno, sendo esta seleção baseada na predição das propriedades anticorrosivas destes nanomateriais.

1.2 OBJETIVOS

O principal objetivo deste trabalho foi desenvolver um modelo computacional, baseado em *Machine Learning*, para atuar na predição das propriedades anticorrosivas de polímeros aditivados com grafeno ou óxido de grafeno.

Para proporcionar um melhor aproveitamento ao longo da execução do trabalho foi estabelecido os seguintes objetivos específicos:

- Criar uma base de dados no tocante a aplicação de nanocompósitos poliméricos aditivados com grafeno, óxido de grafeno ou derivados, aplicados em revestimentos anticorrosivos;
- Realizar estudos de modelagem computacional, através de *Machine Learning*, para criação do modelo de predição de propriedades anticorrosivas dos nanocompósitos;
- Analisar os resultados obtidos para identificar quais parâmetros para o treinamento dos modelos resulta nas melhores taxas de predição;
- Discutir, baseando-se nos resultados obtidos, a eficiência do modelo computacional adotado, avaliando do ponto de vista quantitativo sua aplicabilidade e possíveis melhorias em sua construção.

2 REVISÃO BIBLIOGRÁFICA

Para garantir a inteligibilidade relativa a problemática deste trabalho, foi realizado um levantamento bibliográfico a respeito dos principais temas que o abrange, sendo eles o fenômeno da corrosão e seus efeitos em materiais metálicos; a natureza dos inibidores, em especial, nanocompósitos poliméricos aditivados com grafeno e óxido de grafeno; modelos computacionais e aplicações de *Machine Learning* voltadas para a predição de propriedades de materiais.

2.1 AÇO CARBONO

O aço carbono, uma liga metálica que contém entre 0,008% e 2,11% de concentração de carbono, é um material obtido através de processos siderúrgicos e de vasta aplicação. Suas propriedades são fortemente afetadas pelo teor de carbono presente em sua constituição e podem ser alteradas através da aplicação de elementos liga (ZOLIN, 2011).

As excelentes propriedades do aço fazem com que seu uso seja direcionado para a confecção dos mais diversos produtos, abrangendo desde setores do segmento automotivo, agrícola e da construção civil e até na produção de simples acessórios domésticos.

Devido a esta extensa gama de aplicações manifesta-se a seguinte problemática: como prover ao metal uma proteção adequada ao ambiente o qual ele será destinado, buscando sempre alternativas que proporcionem alta eficiência com um bom custo benefício, além de serem integradas por tecnologias verdes, ou seja, apresentando biocompatibilidade (GROSSER, 2015).

2.2 INIBIDORES DE CORROSÃO

Segundo Gentil (2011), em determinados casos, o fenômeno da corrosão pode ser considerado como o inverso do processo siderúrgico, do qual obtêm-se o metal nas propriedades desejadas a partir do material extraído da natureza; enquanto o processo corrosivo age oxidando a superfície deste metal, gerando, muitas vezes, um

substrato metálico de natureza bastante semelhante ao material original encontrado na natureza.

Os inibidores de corrosão compõem uma classe de substâncias, muitas vezes múltiplas substâncias incorporadas numa matriz polimérica, que quando desenvolvidos nas concentrações ideais para dada aplicação, promovem a desaceleração ou até impedem a ação de um agente corrosivo em dado material. (GROSSER, 2015).

Na maior parte dos casos estes inibidores são de adsorção, sua presença na superfície desejada e cria uma fina e persistente película que diminui a taxa de corrosão, posto que enfraquece a ação corrosiva das reações anódica, catódica ou ambas simultaneamente (FRAUCHES-SANTOS, 2014).

Estes inibidores podem ser classificados quanto à sua composição (orgânicos ou inorgânicos) ou pelo seu comportamento na reação eletroquímica (oxidantes, não-oxidantes, anódicos, catódicos ou mistos), como pode ser observado na figura abaixo: (FRAUCHES-SANTOS, 2014).

Figura 1 - Classificação dos tipos de inibidores de corrosão quanto a sua natureza química e comportamento nas reações



Fonte: FERNANDES (2018)

Atualmente, a constante busca por materiais biocompatíveis, e com excelentes propriedades anticorrosivas, é de profundo interesse para a indústria, proporcionando

aos pesquisadores deste campo de estudo, excelentes oportunidades para atuar conjuntamente às empresas do setor industrial, no desenvolvimento destes materiais. Almeja-se cada vez mais a utilização de inibidores não tóxicos, evitando o uso de compostos sintéticos como os desenvolvidos a base de nitratos, cromatos, derivados de imidazol e tioureia (GROSSER, 2015).

Tubulações, estruturas metálicas e maquinários industriais são exemplos de casos onde os dois principais mecanismos corrosivos podem atuar, a corrosão química e eletroquímica. Na corrosão química, comumente chamada de corrosão a seco, o ataque do agente químico ocorre sem o deslocamento de cargas, ou seja, o metal reage com um meio não iônico. Em uma corrosão eletroquímica há um fluxo de elétrons entre o metal e o agente corrosivo, envolvendo a presença de um eletrólito no sistema. Este mecanismo de corrosão é o mais comum, estando presente em corrosões resultante de soluções aquosas, corrosão atmosférica e corrosão em solos (GENTIL, 2011).

Figura 2 - Incidência do fenômeno corrosivo numa instalação industrial



Fonte: ATOM JR (2020)

2.3 O MECANISMO DA CORROSÃO

Grande parte dos materiais expostos a ambientes em condições naturais estabelecem uma relação termodinamicamente instável com o meio. Excetuando-se os metais nobres, que são encontrados em sua forma metálica na natureza e

apresentam reatividade menor que a do hidrogênio, tornando-os naturalmente resistentes à corrosão; os metais submetidos a processos metalúrgicos, quando em sua forma refinada, apresentam um nível energético superior ao material original extraído na natureza (GENTIL, 2011; PONTE, s.d.).

Quando submetido aos efeitos do fenômeno corrosivo, o substrato dos metais, é gradativamente transformado num composto semelhante ao primordial, indicando um processo inverso ao siderúrgico. Este fenômeno, conhecido como ciclo dos metais, exibe a razão termodinâmica da espontaneidade das reações de corrosão, onde a energia liberada no processo de deterioração do metal é cedida ao meio ambiente (GENTIL, 2011; PONTE, s.d.).

É de suma importância, numa etapa de projeto de seleção de materiais, que se considere a velocidade com que a reação de corrosão ocorrerá no cenário onde o material será inserido (GENTIL, 2011). Segundo Ponte (s.d.), apesar da termodinâmica indicar a possibilidade de uma reação, não estabelece a velocidade com que esta reação ocorre. Assim, diferentes metais e ligas metálicas apresentarão velocidades de corrosão únicas, fato que viabiliza sua aplicação no cotidiano e que pode ser ainda mais favorecido através do uso de inibidores de corrosão.

Durante o processo corrosivo ocorre uma perda ou ganho de elétrons, sendo a perda definida como oxidação, ou reação anódica; e o ganho de elétrons como redução, ou reação catódica. Este fenômeno faz com que as propriedades do material deteriorado sejam alteradas, como, por exemplo, o ferro que perde características relacionados ao seu caráter metálico (GENTIL, 2011).

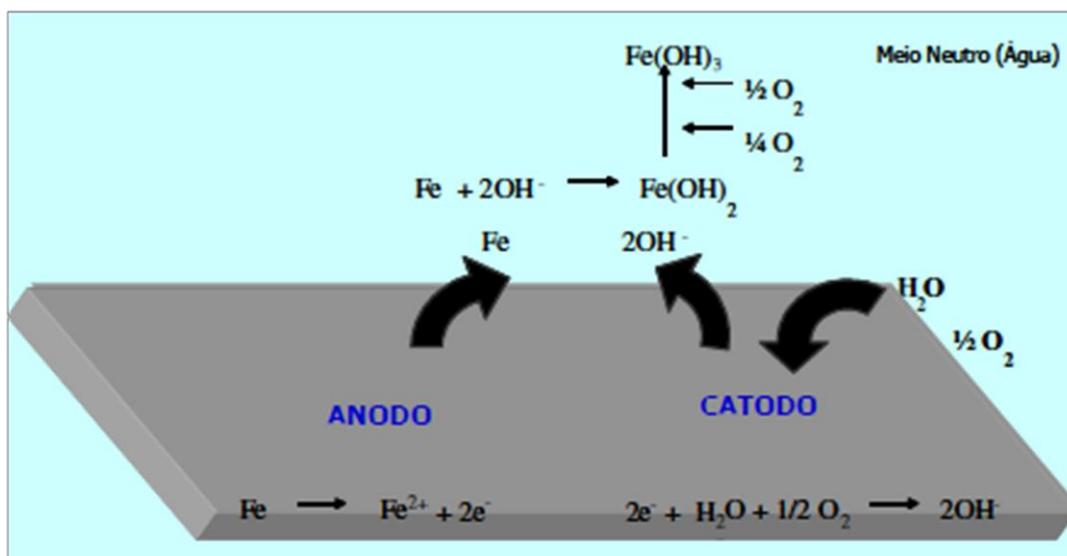
A partir do ponto de vista químico, e tomando o ferro como substrato sujeito a corrosão, as reações previamente citadas podem ser descritas pelas seguintes equações:



A imperfeição na superfície do aço carbono, ou seja, a não uniformidade superficial promove o aparecimento de pequenos pontos com diferentes potenciais eletroquímicos; característica do material que favorece a ação da corrosão

eletroquímica no meio atmosférico e aquoso (GROSSER, 2015). Na figura a seguir pode-se observar o fenômeno da corrosão eletroquímica num eletrodo de aço carbono.

Figura 3 - Mecanismo da corrosão eletroquímica num eletrodo aço carbono em meio aquoso



Fonte: MELLO (2008)

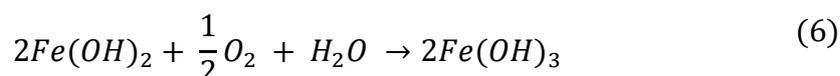
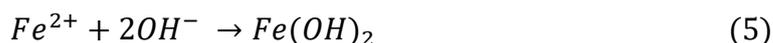
A diferença de potencial presente no eletrodo promove na região anódica uma transferência de elétrons do ferro metálico para a região catódica, viabilizando assim a solubilização do ferro presente no eletrodo, e causando o aparecimento de microcélulas eletroquímicas. As equações que descrevem as reações presentes no sistema (MELLO, 2008):



A corrente elétrica, resultado do fluxo de elétrons da região de dissolução do metal para a região de redução do oxigênio, possibilita a ocorrência simultânea dessas reações. Uma consequência conhecida das reações eletroquímicas é a formação da

ferrugem, resultante da oxidação do ferro durante o processo corrosivo (GROSSER, 2015, WOLYNEC, 2003).

A ferrugem é descrita como depósitos dos produtos de corrosão, ou seja, é constituída principalmente por hidróxido férrico e óxido férrico (GROSSER, 2015). A seguir são expostas as equações que representam este fenômeno:



Com a equação 7 representando a formação do depósito de cor marrom no substrato do eletrodo de aço carbono.

Compreende-se então que, para que se possa usufruir das propriedades dos aços carbonos, sem que as adversidades causadas pelos processos corrosivos se mostrem presentes, é necessário prover ao metal proteção satisfatória, protegendo-o contra agentes corrosivos presentes no ambiente onde o metal será empregue e da própria atmosfera.

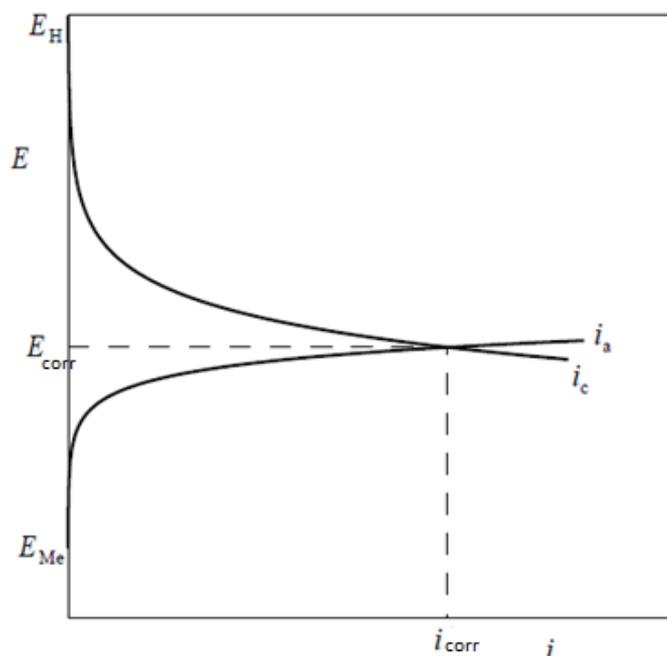
2.4 POTENCIAL DE CORROSÃO

O potencial de corrosão (E_{corr}) pode ser definido como um potencial assumido por um metal durante o fenômeno corrosivo, podendo ser obtido em ensaios simples e não destrutivos. Considerando um metal em processo de corrosão numa solução desareada qualquer, envolvendo apenas as reações catódica e anódica, ou seja, dissolução do metal e redução do hidrogênio respectivamente; considerando também que no processo a solução possui elevada condutividade (desprezando-se a queda ôhmica) e que cada uma das reações previamente mencionadas será caracterizada por seus respectivos potenciais de equilíbrio e curvas de polarização (WOLYNEC, 2003).

A corrosão ocorrerá quando o potencial de equilíbrio da reação anódica (dissolução do metal) for menor que o da reação catódica (redução de hidrogênio). O

processo pode ser melhor compreendido a partir da imagem abaixo: (WOLYNEC, 2003).

Figura 4 - Curvas de polarização representando de forma esquemática a corrosão de um metal em solução aquosa desaeerada



Fonte: WOLYNEC (2003)

O processo de dissolução do metal irá prosseguir quando a reação de hidrogênio aproveitar os elétrons produzidos pela reação de dissolução do metal, havendo então um fluxo de elétrons entre as reações, deste modo poderá ser observado tanto a corrosão do substrato metálico como a evolução do hidrogênio (WOLYNEC, 2003).

Isto irá ocorrer apenas quando a carga elétrica transferida na unidade de tempo da reação anódica seja igual à carga elétrica absorvida pela reação catódica. Circunstância resultante de quando a densidade de corrente anódica (i_a) é igual a densidade de corrente catódica (i_c) e para que isto seja possível o sistema deverá assumir um potencial de eletrodo intermediário, que corresponderá à interseção das curvas anódicas e catódicas no gráfico. Comumente o potencial intermediário é designado como potencial misto e no campo da corrosão ele é equivalente ao potencial de corrosão. Este potencial possui uma densidade de corrente

correspondente, sendo ela a densidade de corrente de corrosão (i_{corr}), também chamada taxa de corrosão ou velocidade de corrosão (WOLYNEC, 2003).

Logo, o potencial de corrosão pode ser descrito sucintamente como uma condição onde a cinética das reações anódica e catódica são iguais, condição necessária para se manter a eletroneutralidade.

2.5 VELOCIDADE DE CORROSÃO

O conceito de velocidade de corrosão é de fundamental importância no estudo dos processos corrosivos, e pode ser classificada em velocidade média de corrosão e velocidade instantânea de corrosão. Enquanto a velocidade média de corrosão pode ser obtida através de medições da diferença de peso apresentada pelo material metálico e também pela determinação da concentração de íons metálicos na solução durante um intervalo de tempo em exposição a um agente corrosivo; para se obter a velocidade instantânea de corrosão são necessários métodos para se determinar uma densidade de corrente de corrosão que possa ser correlacionada com a perda de massa do metal. (GENTIL, 2011).

Desta forma, a velocidade de corrosão pode ser expressa como corrente de corrosão, esta equivale à corrente anódica que circula pelo metal no potencial de corrosão. Como existe também uma corrente catódica circulando pelo metal, de valor igual e sentido oposto à corrente anódica, torna-se inviável a obtenção da corrente de corrosão através da medição direta (GENTIL, 2011).

Um dos métodos utilizados para se obter a velocidade de corrosão instantânea é através da extrapolação de *Tafel*, uma técnica baseada na teoria potencial mista que afirma que tanto a reação catódica como a reação anódica ocorrem simultaneamente, deste modo, a densidade de corrente total (i_{total}) é dado pela diferença das densidades de corrente das reações catódica e anódica, sendo expresso pela seguinte equação (ROJAS, 2014):

$$i_{total} = i_c - i_a \quad (8)$$

2.6 CURVAS DE POLARIZAÇÃO

As curvas de polarização podem ser obtidas com o auxílio de um potenciostato, um dispositivo eletrônico cuja função é impor ao eletrodo de trabalho um potencial de eletrodo específico (WOLYNEC, 2003).

Através da polarização de um eletrodo por meio de um potenciostato é gerada uma curva de polarização, esta curva relaciona o potencial imposto ao sistema com a corrente medida no dispositivo, representando, globalmente, os efeitos das reações que ocorrem simultaneamente na superfície do eletrodo (WOLYNEC, 2003).

Caso o potencial aplicado pelo dispositivo (E_a ou E_c) seja igual ao potencial de corrosão não será registrada corrente, uma vez que, a corrente anódica será completamente anulada pela corrente catódica. Então, para ser realizado o ensaio de polarização, é necessário que os potenciais apresentem valores distintos, assim, o valor da i_a irá divergir de i_c e o potenciostato registrará a diferença entre as reações presentes, gerando os seguintes casos: (WOLYNEC, 2003).

Para $E_a > E_{corr}$

$$\Delta i_a = i_a - |i_c| > 0 \quad (9.1)$$

Para $E_c < E_{corr}$

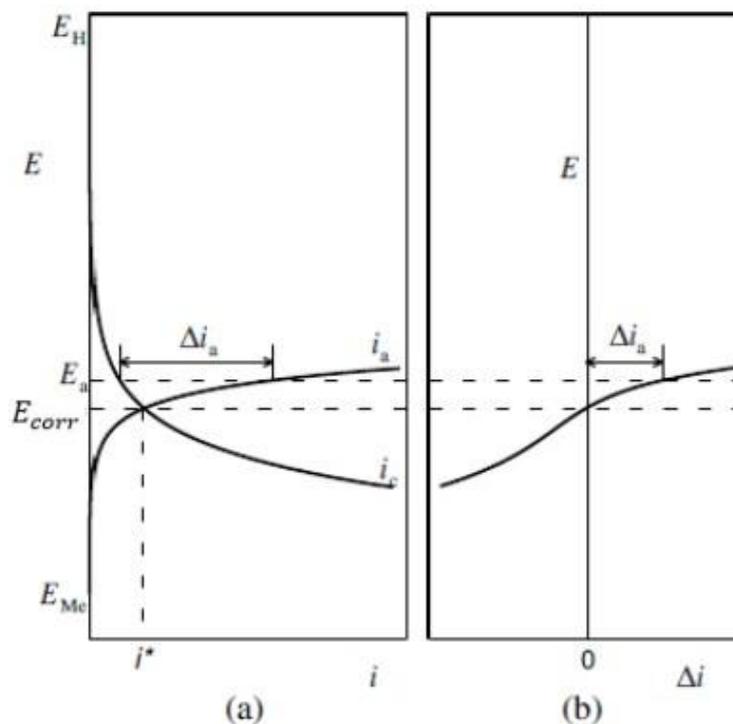
$$\Delta i_c = i_a - |i_c| < 0 \quad (9.2)$$

De forma geral temos:

$$\Delta i = i_a - |i_c| \quad (9.3)$$

Para um caso de corrosão simples o potenciostato irá prover a curva diferença ou resultante entre as curvas de polarização das reações anódica e catódica, como pode ser observado na figura a seguir (WOLYNEC, 2003):

Figura 5 - Curvas de polarização anódica e catódica (a) Curva de polarização resultante
(b)



Fonte: WOLYNEC, (2003)

A polarização pode prover múltiplos dados relevantes para o estudo de corrosão de materiais como por exemplo: a corrente de corrosão, a taxa de corrosão, coeficientes anódicos (b_a) e catódico (b_c) dos declives de *Tafel*, as resistências de polarização e a eficiência de inibidores de corrosão (WOLYNEC, 2003; MENNUCCI, 2006).

2.6.1 Equação de *Wagner-Traud* e Extrapolação de *Tafel*

Considerando que os coeficientes b_a e b_c sejam os declives de *Tafel* das retas de polarização das reações anódica e catódica respectivamente e para os potenciais considerarmos $\Delta E = E_a - E_{corr}$, teremos que (WOLYNEC, 2003):

$$\Delta E = b_a \log \frac{i_a}{i_{corr}} \quad (10)$$

$$\Delta E = -b_c \log \frac{i_{corr}}{|i_c|} = b_c \log \frac{|i_c|}{i_{corr}} \quad (11)$$

As equações anteriores resultam nas seguintes expressões para as correntes anódica e catódica respectivamente:

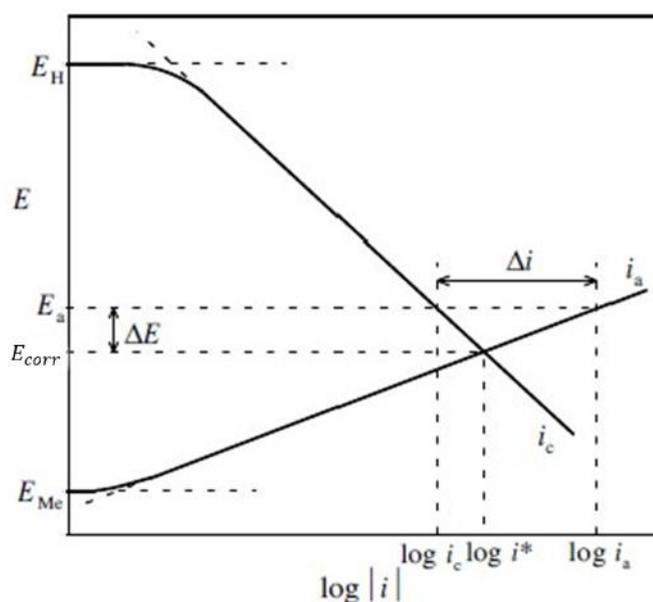
$$i_a = i_{corr} \exp\left(\frac{2.303\Delta E}{b_a}\right) \quad (12)$$

$$|i_c| = i_{corr} \exp\left(\frac{2.303\Delta E}{b_c}\right) \quad (13)$$

Ao aplicar as equações (12) e (13) em (9.3) obtêm-se a equação de *Wagner-Traud*:

$$\Delta i = i_{corr} \left[\exp\left(\frac{2.303\Delta E}{b_a}\right) - \exp\left(\frac{2.303\Delta E}{b_c}\right) \right] \quad (14)$$

Figura 6 - Esquema da corrosão de um metal cujas reações catódica e anódica são controladas pela polarização de ativação na região de *Tafel*



Fonte: WOLYNEC, (2003)

Para que a equação 14 seja válida, as seções que definem E_{corr} e i_{corr} no diagrama E vs. $\log|i|$ precisam ser retas. Então, em casos onde E_{corr} se aproxima

dos potenciais de equilíbrio E_{Me} e E_H , que ocorre geralmente a menos de 30mV desses potenciais, a equação de *Tafel*, mostrada a seguir, não será válida (WOLYNEC, 2003):

$$\eta = b \log \frac{|i|}{i_o} \quad (15)$$

Onde:

η = Sobreensão ou Sobrepotencial;

b = Coeficiente dos declives de *Tafel*;

i_o = Densidade de corrente de troca;

i = Densidade de corrente.

Teremos então que, para valores de $|\Delta E| \geq 30mV$, uma das exponenciais da equação 14 será desprezível em relação à outra, dando origem a equações semelhantes à de *Tafel*, como pode ser observado a seguir (WOLYNEC, 2003):

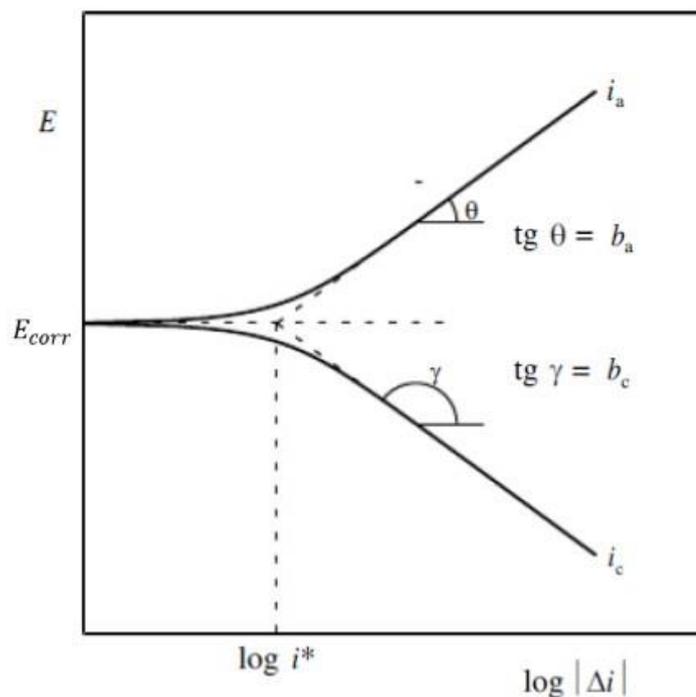
$$\Delta E_a = b_a \log \frac{\Delta i_a}{i_{corr}} \quad (16.1)$$

$$\Delta E_c = b_c \log \frac{|\Delta i_c|}{i_{corr}} \quad (16.2)$$

Ao representar equação de *Wagner-Traud* em um gráfico E vs. $\log|\Delta i|$ pode-se verificar que, a extrapolação das retas de *Tafel* para o potencial de corrosão E_{corr} irá determinar o valor da corrente de corrosão i_{corr} (WOLYNEC, 2003).

Considerada a base do método da Extrapolação de *Tafel*, esse método permite determinar, através das retas de *Tafel*, além de i_{corr} , os declives de *Tafel* b_a e b_c . Deste modo, para uma inclinação da reta anódica θ e γ para a catódica tem-se então que $b_a = tg \theta$ e $b_c = tg \gamma$, como pode-se visualizar na figura a seguir (WOLYNEC, 2003):

Figura 7 - Representação da equação de *Wagner-Traud* em um gráfico E vs. $\log|\Delta i|$



Fonte: WOLYNEC, (2003)

O método da Extrapolação de *Tafel* pode ser utilizado desde que o conjunto de valores obtidos através de ensaios experimentais, E e Δi , quando plotados no gráfico E vs. $\log|\Delta i|$, determine pelo menos uma reta bem definida, ou seja, a curva de *Tafel* terá de ser linear num intervalo de potenciais aproximados, esta é considerada uma condição essencial para sua aplicação (WOLYNEC, 2003; ROJAS, 2014).

2.8 MATERIAIS COMPÓSITOS E NANOMATERIAIS

Cada vez mais se intensificam as buscas por materiais compósitos com propriedades otimizadas e possíveis de serem produzidos em larga escala. A aplicação destes materiais tem sido indispensável para os mais diversos campos da ciência e indústria, podendo ser destinados para aplicações nos setores industriais como o automobilístico, aeroespacial e petroquímico (MARTINS et al., 2022).

Os materiais compósitos são formados através da combinação de uma fase matriz (contínua) com uma fase dispersa (não contínua), ambos representando materiais com propriedades geralmente distintas e complementares. Assim, são considerados uma classe de materiais multifásico e heterogêneo que possuem as

características intermediárias das fases que o integram. Por meio da combinação otimizada de propriedades obtêm-se um material com desempenho final superior em comparação as fases isoladas (MARTINS et al., 2022; EESCJR, 2022).

A fase matriz que irá compor o material pode ser polimérica, cerâmica ou metálica, com os compósitos baseados em matrizes poliméricas possuindo maior visibilidade e aplicabilidade devidas suas propriedades otimizadas. Polímeros orgânicos como a epoxy e o poliuretano, amplamente utilizados como revestimentos anticorrosivos, são exemplos de materiais que podem atuar como fase matriz para um compósito, que ao ser aditivado pela fase dispersa de interesse, resultará numa melhora significativa das propriedades do compósito resultante (MARTINS et al., 2022; OLAJIRE, 2018).

2.8.1 Nanocompósitos e suas particularidades

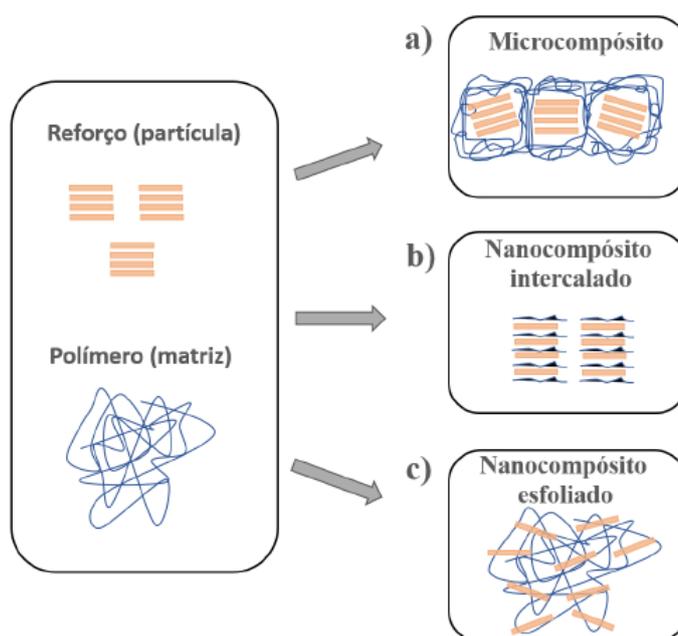
Quando a matriz polimérica é combinada com uma fase dispersa que possui pelo menos uma de suas dimensões em escala nanométrica, o material resultante será denominado um nanocompósito. De forma geral, estes materiais são constituídos de uma matriz polimérica contínua e um reforço (nanopartículas) de morfologia nanométrica, podendo ser classificados como zero-dimensional (nanopartículas), unidimensional (nanofibras) e bidimensional (nanoplaquetas). (MARTINS et al., 2022). A alta razão de área de superfície para volume presente nas nanopartículas faz com que sua área de interface com a matriz polimérica aumente consideravelmente, promovendo assim, uma sinergia entre fases que resulta em melhorias singulares na mobilidade, conformação e cristalinidade da cadeia polimérica, além de otimizar as propriedades elétricas, mecânicas, óticas e químicas do NC (MARTINS et al., 2022; OKPALA, 2014).

Baseando-se na estrutura da fase dispersa, um compósito pode ser classificado como microcompósito, nanocompósito esfoliado e intercalado; quando se considera a forma e o tamanho da fase dispersa, os compósitos podem ser classificados como particulados, fibrosos e lamelares (MARTINS et al., 2022; MIGNONI, 2022)

Num caso em que a fase dispersa presente no material é formada por estruturas de escala micrométrica têm-se um microcompósito. Os materiais onde a fase dispersa é constituída por aglomerados de nanomateriais que não intercalam com a matriz polimérica são chamados compósitos de fases distintas, estes tendem a

possuir propriedades semelhantes aos microcompósitos. Em NC esfoliados o reforço se encontra bem separado e uniformemente disperso na matriz polimérica, diferentemente dos NCs intercalados, onde a dispersão não é completa, apresentando regiões com concentrações desiguais de nanopartículas (MARTINS et al., 2022; MIGNONI, 2022). A figura abaixo ilustra de forma simplificada como as estruturas previamente mencionadas podem ser classificadas:

figura 8 - Classificações possíveis para os compósitos baseando-se na estrutura da fase dispersa



Fonte: MARTINS et al., (2022)

A melhor distribuição da fase dispersa na matriz polimérica dos nanocompósitos esfoliados propicia uma grande aplicabilidade, tornando-o então, a estrutura mais desejada (MARTINS et al., 2022).

Assim, para se atender parâmetros específicos de utilização, é de suma importância estudar minuciosamente a natureza dos componentes que se deseja combinar em um material. Como as diferentes estruturas possíveis para a fase dispersa, assim como seu tamanho e forma, influenciam diretamente nas propriedades do NC resultante, são considerados fatores que afetam diretamente a sua aplicabilidade (MARTINS et al., 2022).

O aumento da consciência e responsabilidade ecológica propicia a adoção de compostos orgânicos para atuarem como inibidores anticorrosivos. A resina epoxy,

um polímero termofixo, é um excelente candidato em aplicações de revestimentos anticorrosivos em superfícies metálicas. Contudo, a porosidade presente em sua estrutura torna-se um caminho ao substrato metálico, permitindo o ataque de um agente corrosivo. Através da incorporação de nanopartículas com propriedades anticorrosivas, como o grafeno e seus derivados, na matriz polimérica da epoxy, aumenta-se a performance anticorrosiva do material, tornando-o um nanocompósito polimérico biocompatível e com ótima eficiência na inibição do fenômeno corrosivo. (GENTIL, 2011; GROSSER, 2015; OTHMAN et al., 2019)

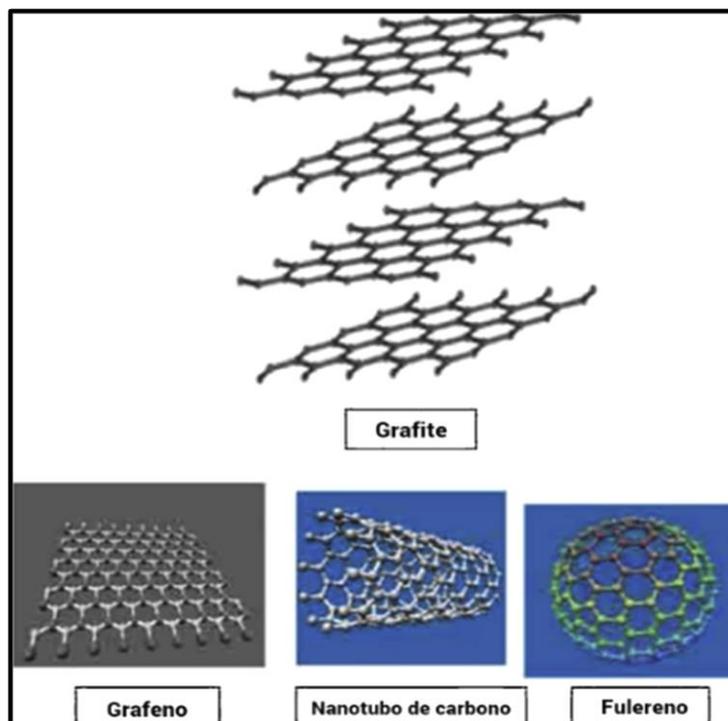
O grafeno, assim como seus derivados, são um exemplo de nanomateriais de reforço que despertaram grande interesse no meio acadêmico e industrial nas últimas décadas. Matrizes poliméricas podem ser incorporadas com o grafeno e seus derivados para produzir NCs de revestimento com propriedades excepcionais (OTHMAN et al., 2019).

2.8.2 Grafeno

A recente descoberta do grafeno, uma variação alotrópica do carbono formado por uma folha planar de espessura monoatômica, promoveu uma corrida científica e tecnológica no campo das ciências dos materiais. Cientistas buscam identificar rotas de produção otimizadas e aplicações as quais este nanomaterial pode se sobressair em comparação a materiais convencionais. O GF contém átomos de carbono em hibridização sp^2 compactados numa rede cristalina de alta qualidade com estrutura hexagonal, conhecida como *honeycomb* (favo de mel). Estas características o concedem propriedades notáveis como elevada resistência mecânica, altas condutividades térmica e elétrica, elevada área superficial específica além de ser quimicamente inerte e impermeável (NECOLAU; PANDELE, 2020; SINGH; KUMAR; SINGH, 2016).

Considerada um elemento primordial, a estrutura hexagonal do grafeno pode ser manipulada para se obter os alótropos do carbono, como os fulerenos, nanotubos de carbono e o grafite. A figura abaixo expõe como estes alótropos podem ser representados (OTHMAN et al., 2016):

Figura 9 - Representação geométrica das formas alotrópicas do carbono



Fonte: Adaptada de OTHMAN, et al (2019)

Os fulerenos, alótropos esferoidais e zero-dimensionais do carbono, fazem parte de uma classe de nanomoléculas constituída do embrulhamento de folhas de grafeno, tem-se como principal representante deste alótropo o buckminsterfulereno (C_{60}) (FILHO, 1996; OTHMAN et al., 2016).

Nanotubos de carbono (NTC), forma alotrópica unidimensional e de formato cilíndrico do carbono, são constituídos de uma ou mais folhas de grafeno enroladas de forma concêntrica e cilíndrica, podendo ser subdivididos em NTC de parede única e NTC de paredes múltiplas (OTHMAN et al., 2016; TRINDADE, 2011).

O grafite, forma alotrópica tridimensional e mais conhecida do carbono, é formado a partir do empilhamento planar de folhas de grafeno. Pode ser natural ou sintético, resultando em estruturas cristalinas (hexagonal ou romboédrica) semelhantes, mas propriedades físicas e químicas distintas (OTHMAN et al., 2016; SOUZA, 2014).

A produção de nanocompósitos incorporados com grafeno é feita a partir da mistura física ou química de um polímero com nanoplaquetas de GF nas concentrações desejadas (OTHMAN et al., 2016). A deposição do GF na matriz pode ser realizada por métodos como as técnicas baseadas em solventes (deposição de

camada-por-camada, revestimento por imersão, revestimento por rotação, etc.), deposição por vapor químico (CVD), deposição eletroforética (EPD) e o processamento térmico rápido (RTP) (OLLIK; LIEDER, 2020).

Quando feita a dispersão das nanopartículas de grafeno em NCs, cria-se um efeito barreira que resulta numa redução na permeabilidade da matriz polimérica. Isto se deve ao prolongamento da sinuosidade das rotas de deslocamento, diminuindo a mobilidade de íons e moléculas na estrutura do NC. Assim, pode-se obter NCs com excelentes propriedades, múltiplas funcionalidades e de ampla aplicabilidade. Quando empregues em revestimentos, os NCs baseados em GF podem ser utilizados como revestimentos inibidores de corrosão, retardantes de chamas, anti-incrustantes, adsorventes de partículas poluentes, entre outras aplicações (OLLIK; LIEDER, 2020; OTHMAN et al., 2016).

Necolau e Pandeale (2020) apontam que, apesar do grafeno dotar de propriedades notáveis, seu uso prático em larga escala enfrenta obstáculos como os altíssimos custos associados aos procedimentos de produção, sua baixíssima solubilidade e uma tendência em formar aglomerados moleculares quando aplicado em compósitos sem que haja um controle minucioso das condições de dispersão das nanopartículas. Assim, desperta-se o interesse pela busca de nanomaterias alternativos com propriedades e funcionalidades semelhantes ao do grafeno, sendo um deles a sua forma oxidada.

2.8.3 Óxido de grafeno

O óxido de grafeno é definido como um derivado do grafeno ligado a vários grupos funcionais contendo oxigênio, sendo os principais os grupamentos epóxi, hidroxila, carboxila, dentre outros. A presença e distribuição destes grupos dependem significativamente de fatores como o tipo de precursor utilizado (grafite, GF, ácidos carboxílicos, etc.), método de oxidação e procedimento de purificação. Estes grupos funcionais favorecem uma dispersão uniforme e esfoliada do GO na matriz polimérica, uma vez que promovem alterações nas interações de *Van der Waals* e aprimoram as ligações de interface da nanopartícula na matriz polimérica (OTHMAN et al., 2016; SINGH; KUMAR; SINGH, 2016).

Esta forma oxidada do GF, possui uma superior resistência mecânica e melhores estabilidades térmica e química, além de que, a presença dos grupos

funcionais altera drasticamente algumas propriedades do GO, comparado ao GF, como aumento da hidrofiliabilidade das estruturas oxidadas, interação com determinados constituintes químicos tais como espécies catiônicas, dentre outras propriedades advindas da oxidação do GF. Estas são umas das características as quais tornam o GO um excelente material para atuar na inibição de corrosão em ambientes com altas temperaturas e em meio aquoso (NECOLAU; PANDELE, 2020; SINGH; KUMAR; SINGH, 2016).

Para que os nanocompósitos baseados em GO apresentem propriedades e funcionalidades excepcionais, é fundamental que a incorporação das nanopartículas na matriz polimérica tenha uma distribuição uniforme e em concentrações adequadas (NECOLAU; PANDELE, 2020; OTHMAN et al., 2016).

Ao ser aplicado em nanocompósitos de revestimento, o GO normalmente liga-se em múltiplas camadas de folhas planares numa estrutura tridimensional, que mantém a integridade do revestimento ao dificultar a interação de moléculas voláteis de sua composição com o meio externo. (NECOLAU; PANDELE, 2020).

Assim, além de munir o NC com características hidrofóbicas e excelente estabilidade térmica e química, o GO também causa uma diminuição da adsorção e interação de agentes corrosivos com o substrato revestido, otimizando assim a resistência contra a corrosão e estendendo a vida útil do revestimento (NECOLAU; PANDELE, 2020)

Devido a imensa quantidade de componentes que podem atuar como matriz polimérica ou dispersa em um NC, é conveniente dispor de ferramentas que facilitem a análise e identificação das substâncias e concentrações ideais dos aditivos para a aplicação desejada.

Em artigos de referência no campo de desenvolvimento de NCs inibidores de corrosão, pode-se adquirir os valores das concentrações das nanopartículas presentes nos NCs além do potencial e corrente de corrosão obtidos através da extrapolação de *Tafel* nos gráficos $E \text{ vs. } \log|\Delta i|$, dados essenciais para avaliar a eficiência dos revestimentos anticorrosivos. Com a coleta de dados (referentes a atuação dos inibidores na superfície do aço carbono) e criação de uma base de dados estruturada, pode-se recorrer a ferramentas computacionais, como o *Machine Learning*, para realizar estudos de predição das propriedades anticorrosivas destes inibidores.

2.9 MACHINE LEARNING: FUNDAMENTOS

A obtenção e manipulação de dados por meio de métodos convencionais além de exigir uma mão de obra extremamente qualificada, tornou-se, nas últimas décadas, gradativamente inviável devido à complexidade, importância e volume das informações geradas pelos mais diversos setores. As limitações físicas e lógicas envolvendo a manipulação destes dados deu origem ao campo de estudo da ciência de dados, que, através de técnicas e ferramentas computacionais sofisticadas viabilizou o desenvolvimento de tecnologias que até pouco poderiam ser consideradas fictícias (FACELI et al., 2011; FARIA et al., 2021).

Neste cenário, pôde-se visualizar também, um grande aumento na disseminação de métodos computacionais integrados com inteligência artificial (IA), voltados para a solução de problemas de grande pluralidade (FACELI et al., 2011). Segundo Charniak e McDermott (1985), a partir do ponto de vista do pensamento racional, a IA pode ser definida como: O estudo das faculdades mentais por meio do uso de modelos computacionais.

As técnicas de IA envolvem o desenvolvimento de máquinas ou sistemas com a capacidade de pensar de forma inteligente, ou seja, simulando a capacidade humana de tomar decisões lógicas a partir da análise de dados. Para atuar desta forma, as técnicas devem ser capazes de criar por si próprias, baseadas em amostras ou experiências precedentes, hipóteses ou funções resposta para solucionar um problema. Este processo de indução de hipóteses baseado em experiências passadas ou dados de treinamento é chamado de *Machine Learning* (FACELI et al., 2011).

Um sistema baseado em *ML* é considerado inteligente quando ele é capaz de aprender continuamente, adaptando-se a novos cenários, desta forma, o desenvolvedor não terá de prever soluções para todos casos possíveis, tarefa muitas vezes considerada impraticável. A partir de um modelo matemático com parâmetros limitados, a aprendizagem de máquina será dada com a execução de um algoritmo computacional que realizará a otimização dos parâmetros deste modelo, que baseando-se no processo de indução, fará uso de uma base de dados de treinamento ou experiências precedentes (ALPAYDIN, 2020).

A utilização de modelos computacionais baseados em *ML* são excelentes ferramentas para otimização de processos, tornando-os mais rápidos e assertivos, como por exemplo na predição das propriedades de nanocompósitos, objeto de

estudo deste trabalho. Pode-se assim, avaliar a eficiência de um composto numa aplicação específica, extraíndo dados a partir de bibliotecas de quimioinformática, sem que seja necessário produzi-lo em laboratório.

Para se obter sucesso na resolução de problemas utilizando os modelos preditivos de *ML* é de suma importância que se utilizem técnicas de pré-processamento e estruturação do conjunto de dados, que juntamente as fontes com dados de qualidade resultarão num modelo dotado de um fluxo de processos bem estabelecido e com eficiência adequada (FACELI et al., 2011; GOMES, 2019)

De maneira genérica, o aprendizado de um sistema baseado em *ML* pode ser classificado, principalmente, de duas maneiras: supervisionado ou não supervisionado. A escolha do método será dada com base no tipo de problema tratado e de acordo com a categoria do conjunto de dados que será utilizado (PINHEIRO, 2020).

2.9.1 Aprendizado supervisionado

Nesta forma de aprendizado é preciso disponibilizar ao algoritmo de *ML* uma base de dados que servirá de conjunto de treino, onde as amostras “*x*” irão conter os valores de seus atributos e do rótulo “*y*”. O rótulo ou etiqueta é exatamente a variável de interesse que se pretende predizer, ou seja, a solução desejada, que pode assumir valores discretos ou contínuos num domínio conhecido (PINHEIRO, 2020; FACELI et al., 2011).

Assim, espera-se que o algoritmo aprenda uma função genérica que realizará o mapeamento das entradas e suas respectivas saídas para cada caso, $f: x \rightarrow y$. Nos algoritmos preditivos de *ML* constrói-se com base de um conjunto de exemplos rotulados, um estimador. Caso a função contenha rótulos com valores contínuos, ou seja, seu domínio for um conjunto infinito e ordenado de valores, ela será delimitada como um problema de regressão e o estimador gerado será um regressor; para rótulos cujos valores sejam discretos, isto é, seu domínio sendo um conjunto de valores nominais, ter-se-á um problema de classificação ou aprendizado de conceitos, e seu estimador será um classificador (PINHEIRO, 2020 apud BISHOP, 2006; FACELI et al., 2011).

Com o fim do treinamento do algoritmo, é importante validar se o modelo está realizando uma generalização adequada do problema em questão e dos dados não

vistos. Esta validação pode ser feita através do teste na predição de novos rótulos y , utilizando novos valores para as amostras x . Para avaliar a eficiência do modelo de predição pode-se criar, a partir do conjunto de dados, um conjunto de treino e um conjunto de teste. Esta divisão promove um refino na base de dados, ademais, proporciona o treinamento dos modelos de predição por meio dos dados presentes no conjunto de treino e sua avaliação através dos dados presentes no conjunto de teste (PINHEIRO, 2020; BIGOTO, 2020 apud GÉRON, 2019).

2.9.2 Aprendizado não supervisionado

Como o processo de coleta de dados (rotulagem) demanda tempo devido ao grande volume e complexidade das informações, além de requerer especialistas na área que ainda estarão sujeitos à erros humanos. Na aprendizagem não-supervisionada não haverá categorização ou rotulação dos dados, assim, o próprio modelo de *ML* deverá identificar padrões presentes na estrutura dos dados não rotulados, examinando conjuntos de dados tão extensos que inviabilizam a atuação direta de um profissional cientista de dados. Os métodos desta forma de aprendizado envolvem os algoritmos de agrupamento (*Clustering*), que consistem no agrupamento de subconjuntos criados pela máquina, baseando-se na conjectura de que dados que apresentem alguma semelhança devam pertencer ao mesmo subconjunto. Um ponto forte atrelado à esta forma de aprendizado é a redução de dimensionalidade, ação realizada pela máquina quando ela assume que existem dados redundantes no conjunto e realiza remoções de dimensões ou combinações dos dados, quando conveniente. Esta compressão promove um refino do conjunto de dados além de otimizar o desempenho do modelo de aprendizagem, assim, haverá a economia de tempo e poder computacional ao longo da execução dos modelos. (PINHEIRO, 2020 apud DUDA et al., 2012; TIBCO, 2023).

2.9.3 Abordagem do *Auto Machine Learning*

Técnicas envolvendo modelos baseados em *Machine Learning* evoluem e se expandem continuamente, e apesar desta disseminação de técnicas cada vez mais avançadas possibilite a solução de problemas ainda mais abrangentes nas mais

diversas áreas de aplicação, cria-se uma dependência de técnicas ainda mais avançadas no campo da ciência dos dados (TSIAKMAKI et al., 2019).

Ademais, ainda que os modelos computacionais agilizem os processos para os quais eles foram destinados, muitas vezes o desenvolvimento destes modelos requer tempo e profissionais especialistas na área que se deseja explorar. Este contratempo pode ser contornado a partir do uso de diferentes abordagens no desenvolvimento, implementação e validação de novos modelos. Assim, pode-se investigar a utilidade de novas técnicas e desenvolver soluções relacionadas a suas aplicações (MS., 2020).

Uma abordagem que tem se tornado cada vez mais visada neste contexto é o aprendizado automático de máquina ou *AutoML*. Através dele pode-se reduzir não só a necessidade da atuação humana no desenvolvimento de modelos complexos de *Machine Learning*, mas também, simplificar o seu entendimento para usuários ou desenvolvedores que não são especialistas na área (MS., 2020; TSIKMAKI et al., 2019).

Os sistemas de *AutoML* automatizam a exploração, desenvolvimento e otimização de modelos, permitindo a geração de processos, fluxos de tarefa e algoritmos de pesquisa de padrões ainda mais apurados. Além de diminuir a necessidade de cientistas de dados especialistas no desenvolvimento de modelos, esta abordagem traz diversas vantagens como: viabiliza a análise de diversos métodos para uma aplicação de interesse; melhora o desempenho, rastreabilidade e comparabilidade dos modelos desenvolvidos; simplifica a validação e otimização dos modelos; possibilita um maior aproveitamento na etapa de treinamento de modelos, dispondo as informações referentes aos modelos avaliados de forma ranqueada, servindo como um indicador de convergência para ajustes no algoritmo ou revisão do conjunto de dados (CLÉSIO, 2020; MS., 2020).

Deste modo, pode-se recorrer a técnicas de *AutoML* para avaliar que modelos preditivos podem atuar de eficientemente com os dados referentes às propriedades anticorrosivas de um nanocompósito. Provendo assim uma ferramenta de suporte a decisão que facilita a avaliação de quais matrizes poliméricas ou dispersas são ideais para integrar o NC, poupando assim, tempo e gastos em fases experimentais no desenvolvimento dos materiais.

2.9.4 Ferramentas de quimioinformática

A utilização de métodos computacionais se expandiu de tal forma que se tornaram indispensáveis para o estudo e desenvolvimento de tecnologias e produtos nos mais diversos campos da ciência. Quando a química é integrada com recursos computacionais e manipulação de dados, recebe o nome de quimioinformática, um campo de estudo de grande interdisciplinaridade que possibilita a resolução de problemas complexos na área da química (GASTEIGER; FUNATSU, 2006; ALVES et al., 2018).

A quimioinformática é uma ciência que se consolida cada vez mais com o passar dos anos, uma vez que, está pareada com métodos computacionais e de manipulação de dados, que também têm evoluído nas últimas décadas (ALVES et al., 2018). Este campo de pesquisa, também chamado “termo genérico” segundo GASTEIGER e FUNATSU (2006), abrange o desenvolvimento, criação, organização, administração, disseminação, visualização e utilização de dados químicos; além de englobar também a química computacional, onde se utiliza métodos matemáticos e computacionais para a resolução de cálculos e simulações envolvendo estruturas moleculares e compostos (OLIVEIRA, 2015).

Duas ferramentas da quimioinformática de extrema relevância em pesquisas de desenvolvimento atuais são a representação de estruturas químicas de forma estrutural e a criação de descritores moleculares, tendo sido ambas utilizadas no desenvolvimento do presente trabalho (ALVES et al., 2018).

2.9.4.1 Representação gráfica: *SMILES*

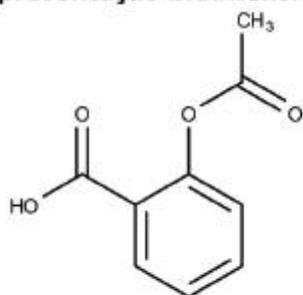
Por meio de uma representação gráfica pode-se expor a forma com que estão dispostos os átomos de um composto químico, que, apesar de serem úteis para a identificação visual através do ponto de vista humano, torna-se inviável para a análise e processamento de um computador. Sendo assim, métodos mais sofisticados são necessários para representar moléculas para facilitar a interpretação por meio da máquina (ALVES et al., 2018).

O *SMILES* (*Simplified Molecular-Input Line-Entry*), pode ser definido como um sistema na forma de notação que pode representar moléculas, estruturas químicas e reações de maneira linear em forma de grafos, método onde cada nó presente na

ramificação representará um átomo presente na estrutura molecular. Além de possuir uma excelente base de dados acessível em bibliotecas de quimioinformática, os *SMILES* são de fácil compreensão, possibilitando seu uso por leigos na simplificação da representação gráfica de estruturas moleculares de variada complexidade (ALVES et al., 2018; SILVA; NORILER, 2020). A imagem a seguir demonstra várias formas de representar o ácido acetilsalicílico (Aspirina), sendo eles sua representação bidimensional, tridimensional e a representação *SMILES* de sua estrutura molecular (ALVES et al., 2018).

Figura 10 - Representações para a estrutura molecular da Aspirina

A) Representação bidimensional



B) Representação tridimensional



C) Notações lineares

SMILES: CC(=O)Oc1ccccc1C(=O)O

Fonte: Adaptado de ALVES et al. (2018)

2.9.4.2 Criação de descritores: *mol2vec*

Como foi pontuado no tópico anterior, é de suma importância utilizar ferramentas de representação de estruturas moleculares de forma mais clara e compreensível, porém, também é preciso traduzir as informações representadas de forma gráfica ou em notações lineares em dados que podem ser lidos e processados por um computador. Assim, pode-se usar métodos lógicos e matemáticos para transformar as informações químicas codificadas numa representação simbólica de uma estrutura molecular, como o *SMILES*, em um valor numérico útil, como um vetor ou matriz (ALVES et al., 2018; OLIVEIRA, 2015).

Estes descritores químicos representam diferentes níveis de representação estrutural de moléculas, podendo ser classificados por sua dimensionalidade, fator que permite facilitar a interpretação das propriedades moleculares devido à amplitude de propriedades, características e fenômenos físico-químicos abrangidos pela

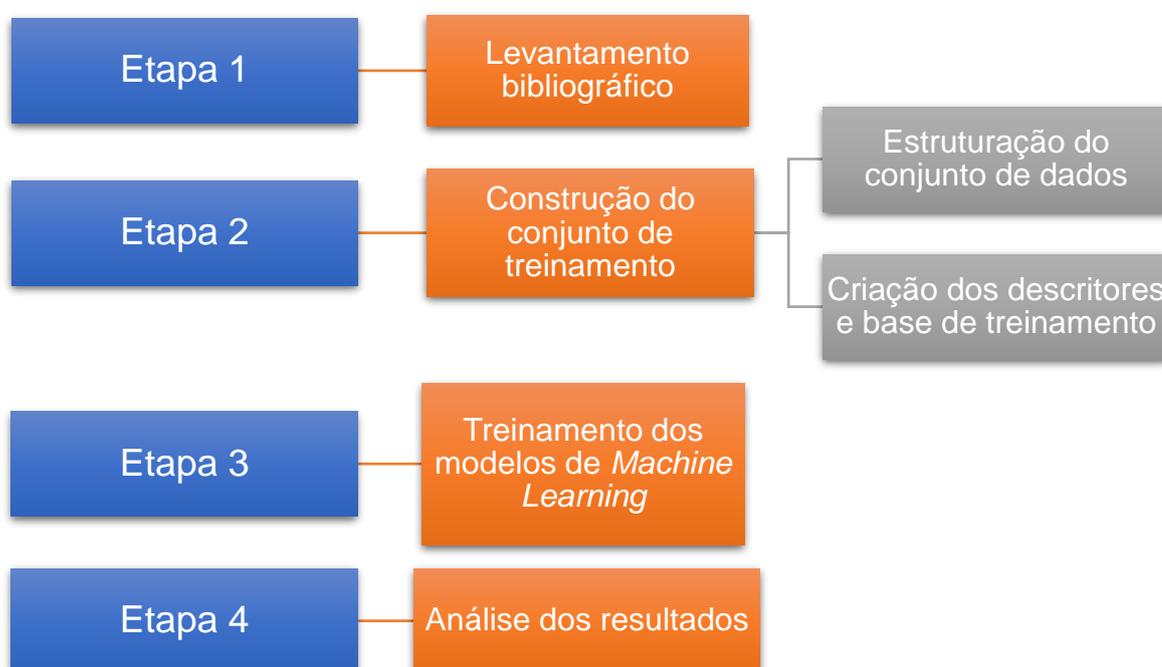
dimensionalidade do descritor, que pode ser classificado de uma até seis dimensões (ALVES et al., 2018).

Uma das formas de realizar a tradução das representações simbólicas de estruturas químicas é através de um método de aprendizagem de máquina não-supervisionado denominado *mol2vec*. Por meio deste método pode-se obter a codificação de compostos na forma de vetores que podem ser direcionados a técnicas de aprendizagem supervisionada para realizar a predição de propriedades destes compostos. A aplicação deste método consiste no treinamento de subestruturas de vetores sobrepostos e incorporados utilizando um método não-supervisionado chamado de *CORPUS* (JAEGER; FULLE; TURK, 2017), ferramenta da quimioinformática que detém as informações sobre as propriedades dos compostos presentes na base de dados de treinamento. A aplicação deste método resulta em agrupamento de dados de alta qualidade para as estruturas químicas que se deseja explorar, sendo assim, uma ferramenta essencial para o estudo dos descritores, que para o caso do presente trabalho são as propriedades anticorrosivas, de múltiplos compostos (JAEGER; FULLE; TURK, 2017).

3 METODOLOGIA

No desenvolvimento do trabalho foi seguida uma metodologia clássica para estudos baseados em ciências de dados, onde inicialmente prepara-se o material objeto de estudo para em seguida serem realizadas otimizações, filtragem de dado e adaptações necessárias para alcançar melhores resultados. A metodologia aplicada no desenvolvimento deste trabalho seguiu a execução das etapas mostradas na figura a seguir:

Figura 11 - Etapas realizadas durante o desenvolvimento do trabalho



Fonte: O autor (2022)

A seguir serão detalhadas as atividades realizadas em cada uma das etapas.

Etapa 1 - Levantamento Bibliográfico: levantamento do estado da arte sobre compósitos contendo GF/GO aplicados para revestimentos anticorrosivos, e aplicações de *ML* para a predição de propriedades de materiais. O levantamento bibliográfico será necessário para a criação de uma base de dados sobre nanocompósitos envolvendo GF, GO e derivados, com foco em características anticorrosivas, quando aplicados como revestimentos em superfícies metálicas, para

a criação de um conjunto de treinamento para os sistemas de *ML* a serem desenvolvidos;

Etapa 2 - Construção de Conjunto de Treinamento para *ML*: Com o levantamento realizado na etapa T1, será produzido um conjunto de dados para treinamento dos sistemas de *ML*, seguindo o seguinte fluxo de processos:

I – Estruturação dos dados: Extração dos dados pertinentes nos artigos sobre: o tipo de polímero, tipo de nanopartícula utilizada (GF, GO ou derivados), proporção entre os componentes e propriedades obtidas, especificamente os potenciais de corrosão obtidos através das curvas de *Tafel*. A seguir é apresentado o modelo de catalogação feito para os artigos extraídos;

Tabela 1 - Modelo de Catalogação dos dados referentes as propriedades anticorrosivas dos nanocompósitos

Artigo 54, DOI: 10.1016/j.surfcoat.2019.125227				
Nanocompósito	Matriz dispersa	Amostra (wt%)	E_{corr} (V)	i_{corr} (Acm ⁻²)
WBE	Blank	100	-0.38	9.94×10^{-8}
GO/WBE	GO	0.6	-0.38	1.83×10^{-8}
AGO/WBE	AGO	0.6	-0.39	4.94×10^{-10}
AGO/WBE	AGO	0.6	-0.37	6.25×10^{-9}

Fonte: Adaptado de Tian *et al.* (2022)

Neste artigo foi utilizado o polímero *Waterborne Epoxy* (WBE) como matriz polimérica. Como matriz dispersa foram utilizados o Graphene Oxide (GO) e o *Ammonium-Grafted Graphene Oxide* (AGO), em diferentes concentrações, resultando assim em valores específicos para as propriedades anticorrosivas.

II – Criação de descritores e da base de treinamento: Após o agrupamento dos polímeros obtidos na etapa I, serão implementadas rotinas em *PYTHON*, utilizando-se bibliotecas voltadas para quimioinformática, no intuito de se obter descritores que representem os polímeros e nanopartículas. Para este trabalho, optou-se pelo uso do método de *ML* não supervisionado denominado *mol2vec* (JAEGER; FULLE; TURK, 2017), no qual a partir da representação *SMILES* de moléculas orgânicas (WEININGER, 1988), um vetor numérico de dimensão

pré-definida é gerado a partir de um modelo treinado (*embedding*). Com os descritores gerados e com as respostas relacionadas as propriedades anticorrosivas obtidas experimentalmente pelos artigos, será criada a base de dados a ser utilizada para o treinamento dos sistemas de *ML*.

Etapa 3 – Treinamento para *ML*: Com a base de dados obtida na etapa **T2**, foram implementadas rotinas em *PYTHON* para o treinamento de sistemas inteligentes baseados em diferentes modelos de *ML*. Todos os treinamentos foram realizados utilizando-se o processo de validação cruzada e avaliando-se os resultados de predição obtidos através da comparação de figuras de mérito como R^2 , erro médio quadrático (*MSE*) e erro médio absoluto (*MAE*). Após avaliação dos resultados foi selecionado o modelo que melhor apresente resultados de predição. Como forma de automatizar e maximizar a qualidade de predição de modelos de *ML*, optou-se pelo uso da abordagem de *AutoML* para o treinamento automatizado e construção de ensembles de vários modelos de *ML* (FEURER, et al., 2020). Os modelos de regressão avaliados no processo de treinamento foram: *Adaboost*, *Ard Regression*, *Decision Tree*, *Extra Trees*, *Gaussian Process*, *Gradient Boosting*, *KNN*, *Libsvm SVR*, *MLP* e *Random Forest*.

Os treinamentos dos modelos seguiram a seguinte configuração geral:

```
automl = autosklearn.regression.AutoSklearnRegressor(  
    time_left_for_this_task=300,  
    per_run_time_limit=60,  
    ensemble_size = 10,  
    resampling_strategy='cv',  
    resampling_strategy_arguments={'folds': 10})
```

Etapa 4 – Análise dos resultados: Os resultados obtidos durante o desenvolvimento do trabalho foram avaliados levando em consideração a eficiência do modelo e das métricas de avaliação.

4 ANÁLISE DOS RESULTADOS

A partir das metodologias citadas e realizadas, obtivemos até o presente momento os seguintes resultados:

4.1 LEVANTAMENTO BIBLIOGRÁFICO

Nesta etapa, catalogou-se 100 artigos referentes a compósitos contendo GF ou GO como dopantes de matrizes poliméricas utilizadas como revestimento anticorrosivo para aço carbono. Como parâmetro de resposta da qualidade dos revestimentos desenvolvidos nos artigos catalogados, optou-se pelo valor do potencial de corrosão obtido a partir da geração das curvas de *Tafel*. Basicamente, quanto menos negativo o valor do potencial obtido, melhor a resistência a corrosão da superfície analisada.

4.2 CONSTRUÇÃO DE CONJUNTO DE TREINAMENTO PARA *ML*

Com o levantamento realizado na primeira etapa, foi produzido um conjunto de dados para treinamento dos sistemas de *ML*. Tendo em vista que o potencial de referência utilizado para avaliar o efeito de um revestimento contra corrosão baseia-se na superfície “nua”, isto é, sem o revestimento proposto, identificou-se alguns problemas no uso direto do potencial de corrosão: (i) o potencial da superfície “nua”, para o mesmo tipo de liga, variava de artigo para artigo, devido a questões experimentais de cada trabalho; (ii) como o objetivo do trabalho é avaliar o efeito de compósitos contendo GF ou GO, pôde-se constatar que alguns trabalhos apresentavam dois valores de potencial de referência: da superfície sem revestimento e da superfície recoberta com algum polímero sem a presença do GF ou GO.

Desta forma, optou-se por se avaliar duas variáveis de resposta aos modelos de *ML* a serem treinados: (i) potencial de corrosão dos compósitos, considerando-se as superfícies recobertas apenas com os respectivos polímeros como nosso potencial de referência; (ii) a diferença entre o potencial de corrosão obtido das superfícies recobertas com os compósitos contendo diferentes quantidades de GF ou GO, descontado o valor do potencial obtido das superfícies recobertas apenas com os respectivos polímeros. Logo, optou-se por selecionar apenas os dados de artigos

contendo polímeros e GO, o que resultou finalmente em uma base de dados com apenas 40 instâncias, o que já temos ciência que poderia diminuir a qualidade dos resultados obtidos dos modelos de *ML* para regressão, mas que permitiriam representações mais simples dos materiais presentes nos artigos, do ponto de vista de representação numérica para o bom entendimento dos modelos de *ML*. Lembrando-se que o valor de concentração do GO nos compósitos avaliados foi parametrizado para %*m/v* e o valor das variáveis de resposta (E_{corr} e ΔE_{corr}) estão representados em mV.

Outro ponto que dificultou a geração da base de dados foi a natureza dos compósitos desenvolvidos em cada artigo catalogado. Muitos revestimentos, além de um polímero e do uso do GF ou GO, possuíam nanopartículas de outras naturezas, como óxidos metálicos ou de metais puros. Além disso, a proporção entre artigos utilizando GF foi bem menor do que artigos utilizando GO, dentre os pesquisados.

4.3 CRIAÇÃO DE DESCRITORES

Cada polímero foi inicialmente representado pelo código *SMILES* das respectivas estruturas químicas de seus monômeros. Contudo, cada molécula possui uma quantidade específica de átomos, o que inviabiliza a construção de dados tabulares diretamente. Logo, optou-se pela transformação dos dados *SMILES* para uma codificação tabular de dimensão determinada, padronizando todas as moléculas estudadas. A transformação escolhida baseou-se no método de aprendizado de máquina não supervisionado, denominado de *mol2vec*. Através da utilização do modelo *mol2vec*, representou-se os polímeros em vetores numéricos de diferentes dimensões n , a saber: $n=50$ e $n=100$. Para tanto, necessitou-se previamente construir alguns *CORPUS* de *SMILES* para a construção dos modelos de codificação *mol2vec*. Neste caso, foram extraídos 7 milhões de *SMILES* de moléculas orgânicas diversas da base de dados do *PubChem*. Cada *CORPUS* foi construído com 1 milhão de *SMILES* como base de referência. Por fim, construiu-se um total de 8 modelos de codificação *mol2vec*, (4 *CORPUS* x 2 dimensões n de conversão). Ao final desta etapa, construiu-se um total de 8 conjuntos de treinamento/teste para serem utilizados para o treinamento dos modelos de regressão.

4.4 TREINAMENTO DE MODELOS DE REGRESSÃO

Como forma de automatizar e maximizar a qualidade de predição de modelos de *ML*, optou-se pelo uso da abordagem de *AutoML* para o treinamento automatizado e construção de ensembles de vários modelos de *ML*. Os modelos de regressão avaliados no processo de treinamento foram: *Adaboost*, *Ard Regression*, *Decision Tree*, *Extra Trees*, *Gaussian Process*, *Gradient Boosting*, *KNN*, *Libsvm SVR*, *MLP* e *Random Forest*.

Os resultados dos treinamentos podem ser vistos nas Tabelas 1 e 2, onde apresentam-se os quatro melhores regressores treinados para o potencial de corrosão e diferença entre potenciais de corrosão, respectivamente:

Tabela 2 - Valores das métricas de avaliação dos modelos de regressão obtidos, tendo como variável de resposta o valor de E_{corr} .

Corpus utilizado para Treinamento/ Teste	R^2 (Treino)	R^2 (Teste)	MSE (Treino)	MSE (Teste)
N = 50, CORPUS 1	0,918	0,128	0,003	0,034
N = 50, CORPUS 2	0,913	0,007	0,004	0,039
N = 50, CORPUS 3	0,913	0,140	0,004	0,034
N = 50, CORPUS 4	0,913	0,148	0,004	0,033

Fonte: O autor (2022)

Tabela 3 - Valores das métricas de avaliação dos modelos de regressão obtidos, tendo como variável de resposta o ΔE_{corr} .

Corpus utilizado para Treinamento/ Teste	R^2 (Treino)	R^2 (Teste)	MSE (Treino)	MSE (Teste)
N = 50, CORPUS 1	0,836	0,425	0,007	0,043
N = 50, CORPUS 2	0,824	0,429	0,008	0,043
N = 50, CORPUS 3	0,879	0,418	0,005	0,044
N = 50, CORPUS 4	0,881	0,414	0,005	0,044

Fonte: O autor (2022)

Observou-se que dos 8 treinamentos realizados, os melhores resultados foram obtidos com as representações de menor dimensão das conversões *mol2vec* ($n=50$). Contudo, devido ao tamanho reduzido da base de dados utilizada para treinamento/teste dos modelos treinados, observa-se que os valores das métricas obtidas, principalmente para o grupo de teste, indicam uma baixa capacidade de predição dos modelos de regressão. Porém, quando comparamos os resultados a partir do tipo de variável resposta, conclui-se que quando a variável de resposta está representada como diferença entre potenciais de corrosão das amostras com e sem a presença do GO nos compósitos, a capacidade de predição dos grupos de teste melhoram consideravelmente, quando se observa os valores de R^2 (TESTE). Por fim, para os melhores resultados obtidos nas duas situações de variável resposta discutidas acima, o melhor modelo de *ML* identificado foi baseado em *Gaussian Process*.

5 CONSIDERAÇÕES FINAIS

Foram levantados os artigos necessários para a geração da base de treinamento e teste dos modelos de *Machine Learning* a serem utilizados neste trabalho, contudo, devido à complexidade de representação dos dados, o conjunto de dados gerados apresentou um tamanho final bastante reduzido. Os dados foram validados antes do processo de treinamento dos modelos. Obteve-se como melhor regressor o modelo baseado em Processos Gaussianos (*Gaussian Process*). Uma vez treinado e com mais exemplos de dados para treinamento, o modelo poderá ser utilizado para a predição da capacidade de proteção à corrosão de revestimentos baseados em compósitos contendo GO.

Apesar de novas tentativas para refinar os resultados através do uso do modelo de Processos Gaussianos, aplicando o algoritmo diretamente na base de dados, sem utilizar a abordagem *AutoML*, tenham sido realizadas, não foram gerados resultados satisfatórios, sendo eles semelhantes ou as vezes incongruentes com a matemática do processo.

Para se obter melhores resultados para os valores da métrica de avaliação dos modelos sugere-se a adoção de um conjunto de dados de maior amplitude. Assim pode-se realizar uma varredura num maior número de dados de entrada, abrindo espaço para atuação dos modelos de predição na identificação de padrões em meio aos dados referentes as propriedades anticorrosivas dos nanocompósitos. Tomando como variável resposta para avaliação dos modelos os valores para as diferenças de potencial de corrosão (ΔE_{corr}), densidade de corrente de corrosão (i_{corr}) e também as concentrações (wt%) das nanopartículas presentes no nanocompósito.

REFERÊNCIAS

- ALPAYDIN, E. **Introduction to machine learning**. MIT press, 2020.
- ALVES, V. M. et al. Quimioinformática: uma introdução. **Química Nova**, v. 41, p. 202-212, 2018.
- BIGOTO, M. A. R. **Avaliação de modelos de machine learning para predição da temperatura crítica de supercondutores**. 2020. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Física) - Universidade de São Paulo, Lorena, 2020.
- CHARNIAK, E.; MCDERMOTT, D. **Introduction to artificial intelligence**. Boston, MA: Addison-Wesley Longman Publishing Co, 1985.
- CLÉSIO, F. Automated Machine Learning (AutoML): Aspectos práticos, teóricos, vantagens e limitações. **Medium**, 2020. Disponível em: <https://medium.com/data-hackers/automated-machine-learning-automl-parte-i-1d3219d57d31>. Acesso em: 16 de jan. de 2023.
- Corrosão: Entenda o que é e como ocorre. **Atom Jr.**, São Carlos, 1 de jul. de 2020. Disponível em: <https://www.atomjr.com.br/post/corrosao-entenda-o-que-e-e-como-ocorre>. Acesso em: 22 de jun. de 2023.
- FACELI, K. *et al.* **Inteligência artificial: Uma abordagem de aprendizado de máquina**. Rio de Janeiro, RJ: LTC, 2011.
- FARIA, L. *et al.* **Ciência de dados: algoritmos e aplicações**. Rio de Janeiro, RJ: IMPA, 2021.
- FERNANDES, C. M. **Avaliação da eficiência de inibidor de corrosão na presença de inibidores de incrustação e sequestrantes de H₂S**. 2016. Trabalho de Conclusão de Curso (Bacharelado em Química Industrial) - Universidade Federal Fluminense, Niterói, 2016.
- FEURER, M. et al. Auto-sklearn 2.0: Hands-free automl via meta-learning. **J Machine Learn Res**, v. 23, n. 261, p. 1-61, 2020.
- FRAUCHES-SANTOS, C. *et al.* A Corrosão e os Agentes Anticorrosivos. **Revista Virtual de Química**, v. 6, n. 2, p. 293–309, 2013.
- GASTEIGER, J.; FUNATSU, K. Chemoinformatics – An Important Scientific Discipline. **Journal of computer chemistry**, Japan, v. 5, n. 2, p. 53-58, 2006.
- GENTIL, V. **Corrosão**. 6. ed. Rio de Janeiro: LTC, 2011.
- GOMES, P. C. T. Conheça as Etapas do Pré-Processamento de dados. **DataGeeks**, 2019. Disponível em: <https://www.datageeks.com.br/pre-processamento-de-dados/>. Acesso em: 17 de jan. de 2023.

GROSSER, F. N. **Inibidores orgânicos de corrosão – Estudos com compostos naturais obtidos de diversas espécies de mentas**. 2015. Tese (Doutorado em Química) – Universidade Federal do Rio grande do Sul, Porto Alegre, 2015.

GU, G. X.; BUEHLER, M.J. Tunable mechanical properties through texture control of polycrystalline additively manufactured materials using adjoint-based gradient optimization. **Acta Mech**, vol. 229, p. 4033–4044, 2018.

HAMIDI, Y. K.; BERRADO, A.; ALTAN, M. C. Machine learning applications in polymer composites. **AIP Conference Proceedings**, vol. 2205, p. 020031-1 – 0230031-5, 2020.

JAEGER, S., FULLE, S., TURK, S. Mol2vec: Unsupervised Machine Learning Approach with Chemical Intuition, **Journal of Chemical Information and Modeling**, vol. 58, p. 27-35, 2017.

KUMAR, S. K., KRISHNAMOORTI, R. Nanocomposites: Structure, Phase Behavior, and Properties. **Annu. Rev. Chem. Biomol. Eng**, vol. 1, p. 37-58, 2010.

LOUIS, S. Y. et. al. Graph convolutional neural networks with global attention for improved materials property prediction. **Phys. Chem.**, vol. 22, p. 18141-18148, 2020.

MANAGEMENT SOLUTIONS. **Auto Machine Learning, rumo à automação dos modelos**. 2020. Disponível em: <https://www.managementsolutions.com/pt-br/publicacoes-e-eventos/informes-sectoriais/white-papers/auto-machine-learning-rumo-a-automacao-dos-modelos>. Acesso em: 16 jan. 2023.

MARTINS, D. *et al.* **Nanocompósitos poliméricos**. 2022. Disponível em: <https://www.alice.cnptia.embrapa.br/alice/handle/doc/1148354>. Acesso em: 20 jan de 2023.

Materiais Compósitos: O que são, propriedades e aplicações. **EESCJr.**, São Carlos, 22 de mar. de 2022. Disponível em: <https://eescjr.com.br/blog/materiais-compositos-o-que-sao-propriedades-e-aplicacoes/>. Acesso em: 15 de jan. de 2023.

MELLO, A. G. B. **INFLUÊNCIA DA DEFORMAÇÃO PLÁSTICA NA CORROSÃO DE AÇO CARBONO, AÇO INOXIDÁVEL E COBRE EM SISTEMAS DE RESFRIAMENTO**. 2008. Dissertação (Mestrado em Ciências) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

MENNUCCI, M. M. **Avaliação da potencialidade de aditivos como inibidores de corrosão do aço carbono CA-50 usado como armadura de estruturas de concreto**. 2006. Dissertação (Mestrado em Ciências na Área de Tecnologia Nuclear) - Instituto de pesquisas energéticas e nucleares, Universidade de São Paulo, São Paulo, 2006.

MIGNONI, M. L. **Síntese, caracterização e aplicações de nanocompósitos polímero-argila**. 2008. Dissertação (Mestrado em Ciências dos Materiais) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.

MOURÃO, Marcelo Breda, *et al.* **Introdução a Siderurgia**. São Paulo, SP: ABM, 2007.

NECOLAU, M-I.; PANDELE, A-M. Recent advances in graphene oxide-based anticorrosive coatings: an overview. **Coatings**, v. 10, n. 12, p. 1149, 2020.

OKPALA, C. C. The benefits and applications of nanocomposites. **Int. J. Adv. Eng. Tech**, v. 12, p. 18, 2014.

OLAJIRE, A. A. Recent advances on organic coating system technologies for corrosion protection of offshore metallic structures. **Journal of Molecular Liquids**, v. 269, p. 572-606, 2018.

OLIVEIRA, T. B. **Emprego de ferramentas de quimioinformática no estudo do perfil metabólico de plantas e na desreplicação de matrizes vegetais**. 2015. Tese (Doutorado em Ciências) – Universidade de São Paulo, Ribeirão Preto, 2015.

OLLIK, K.; LIEDER, M. Review of the Application of Graphene-Based Coatings as Anticorrosion Layers. **Coatings**, v. 10, n. 9, p. 883, 2020.

O que é aprendizagem supervisionada?. **Tibco**, 2023. Disponível em: <https://www.tibco.com/pt-br/reference-center/what-is-unsupervised-learning>. Acesso em: 17 jan. 2023.

OTHMAN, N. H. *et al.* Graphene-based polymer nanocomposites as barrier coatings for corrosion protection. **Progress in Organic Coatings**, v. 135, 82-99, 2019.

PAPAGEORGIOU, D. G.; KINLOCH, I. A.; YOUNG, R. J. Mechanical properties of graphene and graphene-based nanocomposites. **Progress in Mater. Sci**, v. 90, 75-127, 2017.

PINHEIRO, G. A. L. L. **Predição de propriedades moleculares via aprendizado de máquina e representação SMILES**. 2020. Dissertação (Mestrado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2020.

PONTE, H. A. **Fundamentos da Corrosão**. Paraná. Disponível em: <http://www.gea.ufpr.br/arquivos/lea/material/Fundamentos%20da%20Corrosao.pdf>. Acesso em: 28 dez. 2022.

ROCHA-FILHO, R. C. Os fulerenos e sua espantosa geometria molecular. **Química Nova na Escola**, v. 4, p. 7-11, 1996.

ROJAS, Y. A. H. **Estudo eletroquímico e caracterização das camadas de produtos de corrosão por CO₂ em aços API5L X80 e API 5CT P110**. 2014. Dissertação (Mestrado em Engenharia de Materiais, Processos Químicos e Metalúrgicos) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2014.

SILVA, L. F. Representação gráfica de moléculas na reconstrução molecular de misturas de hidrocarbonetos. **XXVIII Congresso de Iniciação Científica da Unicamp**. 2020.

SILVA, L F; NORILER, D. Representação gráfica de moléculas na reconstrução molecular de misturas de hidrocarbonetos. **Congresso {virtual} de Iniciação Científica da Unicamp**, Campinas, 2020. Disponível

em: <https://www.prp.unicamp.br/inscricao-congresso/resumos/2020P17567A35360O5643.pdf>. Acesso em: 21 mar. 2023.

SINGH, R. K.; KUMAR, R.; SINGH, D. P. Graphene oxide: strategies for synthesis, wreductionand frontier applications. **RSC Adv**, v. 6, p. 64993–65011, 2016.

SOUZA, N. S. **Síntese e Estudo das Propriedades Físicas de Grafite Ferromagnético Fluido Nanoestruturado**. 2014. Tese (Doutorado em Física) – Universidade Federal de São Carlos, São Carlos, 2014.

TIAN, Y. et al. Ammonium-grafted graphene oxide for enhanced corrosion resistance of waterborne epoxy coatings. **Surface and Coatings Technology**, v. 383, p. 125227, 2020.

TRINDADE, J. NANOTUBOS DE CARBONO: PROPRIEDADES E FUNCIONALIZAÇÃO. **Egitania Sciencia**, p. 5-17. 2011.

TSIAKMAKI, M. et al. Implementing AutoML in educational data mining for prediction tasks. **Applied Sciences**, v. 10, n. 1, p. 90-117, 2019.

WEININGER, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. **J. Chem. Inf. Comput. Sci.**, vol. 28, p. 31-36, 1988.

WOLYNEC, S. **Técnicas Eletroquímicas em Corrosão**. 1. Ed. São Paulo: EDUSP, 2003. Disponível em: https://edisciplinas.usp.br/pluginfile.php/6563355/mod_resource/content/1/libro%20de%20corrosao.pdf. Acesso em: 11 jan. 2023.

ZOLIN, I. **Materiais de Construção Mecânica**. 1. Ed. Santa Maria, RS: Colégio Técnico Industrial de Santa Maria, 2011. Disponível em: http://redeetec.mec.gov.br/images/stories/pdf/eixo_ctrl_proc_indust/tec_autom_ind/mat_constr_mec/161012_mat_const_mec.pdf. Acesso em: 24 jan. 2023.