



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
COMISSÃO NACIONAL DE ENERGIA NUCLEAR
CENTRO REGIONAL DE CIÊNCIAS NUCLEARES DO NORDESTE
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS ENERGÉTICAS E NUCLEARES

LUCAS EMANUEL ALMEIDA BARBOZA

**ELABORAÇÃO DE UM MODELO DE IDENTIFICAÇÃO DE FALHAS E ANÁLISE
DE CONFIABILIDADE PARA SISTEMAS DE RASTREAMENTO SOLAR EM
USINAS FOTOVOLTAICAS**

Recife

2023

LUCAS EMANUEL ALMEIDA BARBOZA

**ELABORAÇÃO DE UM MODELO DE IDENTIFICAÇÃO DE FALHAS E ANÁLISE
DE CONFIABILIDADE PARA SISTEMAS DE RASTREAMENTO SOLAR EM
USINAS FOTOVOLTAICAS**

Dissertação apresentada ao Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares da Universidade Federal de Pernambuco, como requisito para obtenção do título de mestre em Tecnologias Energéticas e Nucleares. Área de concentração: Fontes Renováveis de Energia.

Orientadora: Prof^a. Dr^a. Olga de Castro Vilela

Coorientador: Prof. Dr. Gustavo de Novaes Pires Leite

Recife

2023

Catálogo na fonte
Bibliotecário Gabriel Luz, CRB-4 / 2222

- B239e Barboza, Lucas Emanuel Almeida.
Elaboração de um modelo de identificação de falhas e análises de confiabilidade para o sistema de rastreamento solar em usina fotovoltaica / Lucas Emanuel Almeida Barboza, 2023.
168 f.: il.
- Orientador: Prof^a. Dr^a. Olga de Castro Vilela.
Coorientador: Prof. Dr. Gustavo de Novaes Pires Leite
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares. Recife, 2023.
Inclui referências.
1. Tecnologias Energéticas e Nucleares. 2. Confiabilidade. 3. Sistemas de energia fotovoltaica. 4. Rastreador solar. 5. Localização de falhas. 6. Métodos estatísticos. 7. Sistemas reparáveis. 8. Processo de renovação alternada. I. Vilela, Olga de Castro (orientadora). II. Leite, Gustavo de Novais (coorientador). III. Título.

621.042 CDD (22. ed.)

UFPE

BCTG / 2023 - 237

LUCAS EMANUEL ALMEIDA BARBOZA

**ELABORAÇÃO DE UM MODELO DE IDENTIFICAÇÃO DE FALHAS E ANÁLISE
DE CONFIABILIDADE PARA SISTEMAS DE RASTREAMENTO SOLAR EM
USINAS FOTOVOLTAICAS**

Dissertação apresentada ao Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares da Universidade Federal de Pernambuco, como requisito para obtenção do título de mestre em Tecnologias Energéticas e Nucleares. Área de concentração: Fontes Renováveis de Energia.

Aprovado em: 28/07/2023.

BANCA EXAMINADORA

Documento assinado digitalmente
 **ALEXANDRE CARLOS ARAUJO DA COSTA**
Data: 03/11/2023 00:17:19-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Alexandre Carlos Araújo da Costa (Examinador Interno)
Universidade Federal de Pernambuco - UFPE

Documento assinado digitalmente
 **MANOEL HENRIQUE DE OLIVEIRA PEDROSA FILHO**
Data: 30/10/2023 18:36:13-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Manoel Henrique de Oliveira Pedrosa Filho (Examinador Externo)
Instituto Federal de Pernambuco - IFPE

**JOSE ANGELO
PEIXOTO DA
COSTA**

Assinado digitalmente por JOSE ANGELO PEIXOTO DA COSTA
DN: cn=JCP, o=ICP-Brasil, ou=videoconferencia, ou=2088011000107, ou=Passos Faltas A1, ou=PROFESSOR, ou=Assinatura Certificadora, serial=1000000000, ou=JOSE ANGELO PEIXOTO DA COSTA
Razão: Eu sou o autor deste documento
Localização: sem localização de assinatura aqui
Data: 2023.10.30 17:14:35-0200
Foxit PDF Reader Versão: 11.1.0

Prof. Dr. Jose Ângelo Peixoto da Costa (Examinador Externo)
Instituto Federal de Pernambuco - IFPE

Dedico este trabalho à minha boa Mãe, Maria Santíssima, porque, como diria meu pai espiritual Dom Bosco, foi Ela quem tudo fez.

AGRADECIMENTOS

Venho por meio destas breves palavras fazer um singelo agradecimento, a pessoas que me auxiliaram e me ajudaram a portento. Primeiro, como todo bom jogador de futebol, agradecer a Deus, nossa força e sustento; e também à minha boa Mãe Maria que ouviu o meu lamento.

Agradeço à minha família e à Talita por toda acolhida e amor, isso sim é o verdadeiro energético e força para o meu labor.

A Professora Olga e Professor Gustavo, uma palavra de gratidão! Foram excelentes orientadores, que a esse trabalho deram cores; tiveram muita paciência com seu orientando, e nesse trabalho estiveram sempre guiando! Ao Professor Alexandre agradeço o auxílio, que sempre muito esguio, ajudou nos “perengues” teóricos, com comentários categóricos.

Não poderia me esquecer de agradecer a gente do CER, um pessoal muito bom e competente que não deixa cair a colher. A pesquisa lá é de alto nível, porque é feita de coração, é um povo muito bom que não deixa você na mão.

Agradeço a algumas empresas parceiras, que auxiliaram de forma certa este pesquisador que vos fala: a Companhia Hidro Elétrica do São Francisco – CHESF por meio do projeto Plataforma Solar de Petrolina CVI 80.2020.0010.00 (23076.009704/2020-56), a Multiner por meio do projeto de Acordo de Parceria Nº 16/20 FADE/UFPE/IFPE/RIO AMAZONAS (Convênio Nº 67/2020 UFPE) e a Ibitú por meio do projeto de Acordo de Parceria Nº 02/20 FADE/UFPE/QUEIROZ GALVÃO ENERGÉTICA/CESC/CEC (Convênio Nº 20/2020), todos Programas de P&D ANEEL. Agradeço também a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio aos bolsistas no âmbito do Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares – PROTEN.

Indo agora para os finalmentes, quero pedir de forma potente: muitas desculpas se esqueci de alguém, e aos erros durante a caminhada peço perdão também.

Sou apenas um marinheiro ajustando seus ponteiros, olhando para o céu a dizer: se navegar eu não fazer, de que vale a pena o viver?

Sub tuum praesidium confugimus, Sancta Dei Genetrix...

RESUMO

As usinas fotovoltaicas têm presenciado um crescente desenvolvimento com aumento dos seus fatores de capacidade devido a avanços tecnológicos. Dentre esses, os rastreadores solares são peças chave, pois, rastreando o movimento solar, aumentam a coleta de energia, levando a um aumento na produção energética entre 30% e 40%. Contudo, o emprego dos rastreadores também acarreta um acréscimo nos custos operacionais e de manutenção, estimado entre 10% e 15%, os quais são compensados pelo aumento no fator de capacidade proporcionado. Assim, este trabalho tem como objetivo desenvolver uma metodologia para inferir índices relacionados à confiabilidade dos rastreadores solares. Com eles, é possível compreender a distribuição temporal das falhas e desenvolver estratégias mais eficientes de manutenção, além de obter estimativas mais precisas dos custos associados. A inferência dos índices de confiabilidade consiste em utilizar os dados de ocorrência de falhas dos sistemas de rastreamento para determinar as distribuições estatísticas associadas à confiabilidade do equipamento. Entre os principais índices de confiabilidade inferidos estão a Curva de Manutenção, a Função Taxa de Reparo, a Curva de Confiabilidade e a Função Taxa de Falhas. Entretanto, uma limitação presente no contexto desse estudo é a indisponibilidade de registros de falhas para os sistemas de rastreamento. Para contornar esse desafio, foi desenvolvido um método de detecção de falhas que utiliza a comparação entre a potência medida e a potência simulada a partir da radiação solar. Este método de detecção representa um avanço na literatura, pois é ajustável às características locais e infere os limiares de classificação de forma automática. A metodologia foi aplicada em dois sistemas de rastreamento, um de eixo duplo e outro de eixo único, disponibilizados pelo *Desert Solar Centre* de Alice Springs, na Austrália. Os resultados obtidos demonstraram a capacidade do método em descrever adequadamente as distribuições dos períodos de operação e reparo do sistema. Por fim, a aplicação dos dados de ocorrência de falhas na inferência dos índices probabilísticos de confiabilidade preenche uma lacuna na literatura relacionada à aplicação da teoria da confiabilidade aos rastreadores solares. A metodologia proposta pode ser empregada por usinas para a construção de calendários de manutenção mais eficientes, impactando positivamente o setor fotovoltaico.

Palavras-chave: Confiabilidade; Sistemas de energia fotovoltaica; Rastreador Solar; Localização de Falhas; Métodos Estatísticos; Sistemas Reparáveis; Processo de Renovação Alternada.

ABSTRACT

Photovoltaic plants have increasing development with increased capacity factors due to technological advances. Among these, solar trackers are essential, as by aligning the panels with the Sun, they increase energy collection, leading to an increase in energy production of between 30% and 40%. However, using trackers also leads to an increase in operational and maintenance costs, estimated between 10% and 15%, which are offset by the increase in the capacity factor provided. Therefore, this work aims to develop a methodology to infer indices related to the reliability of solar trackers. With them, it is possible to understand the temporal distribution of failures, develop more efficient maintenance strategies, and obtain more accurate estimates of the associated costs. The inference of reliability indices consists of using failure occurrence data from tracking systems to determine the statistical distributions associated with equipment reliability. The leading reliability indices inferred are the Maintenance Curve, the Repair Rate Function, the Reliability Curve, and the Failure Rate Function. However, a limitation present in the context of this study is the absence of failure records for the tracking systems. For this purpose, a fault detection method was developed, comparing measured power and simulated power from solar radiation. This detection method represents an advance in the literature, adjusting to local characteristics and automatically inferring classification thresholds. The methodology was applied to two tracking systems, one with a dual axis and the other with a single axis, provided by the Desert Solar Center in Alice Springs, Australia. The results obtained demonstrated the method's ability to adequately describe the distributions of the system's operating and repair periods. Finally, applying failure occurrence data to infer probabilistic reliability indices fills a gap in the literature on using reliability theory in solar trackers. The proposed methodology can be used by plants to create more efficient maintenance schedules, positively impacting the photovoltaic sector.

Keywords: Reliability; Solar Power Systems; Solar Trackers; Fault Detection; Statistical Methods; Repairable Systems; Alternating Renewal Process.

LISTA DE ILUSTRAÇÕES

Figura 1 – Eixo de rotação unitário $\mathbf{e}$ e o eixo $\mathbf{n}$ normal a um rastreador de eixo único.....	26
Figura 2 – Rastreadores de Inclinação de Ponta e de Azimute-Altura, respectivamente, e seus eixos de rotação.	28
Figura 3 – Exemplo de Rastreador de Inclinação de Ponta.....	28
Figura 4 – Exemplo de Rastreador Azimute-Altura.	29
Figura 5 – Resumo do Esquema de Classificação dos Rastreadores Solares apresentado.	31
Figura 6 – Desenho esquemático dos principais componentes de um rastreador de eixo único, com duas fileiras de módulos na posição vertical	33
Figura 7 – Desenho esquemático dos principais componentes de um rastreador de eixo duplo	33
Figura 8 – Sistema de coordenadas $x' - y' - z'$ localizado na superfície horizontal do local analisado, vetor solar $\hat{n}_s$ e ângulos das coordenadas polares	34
Figura 9 – Vistas superior e lateral do coletor (módulo fotovoltaico) posicionado na superfície com ângulo de inclinação β , sendo também representado o sistema de coordenadas $x' - y' - z'$ localizado na superfície horizontal do local analisado.....	38
Figura 10 – Sistema de coordenadas $x' - y' - z'$ localizado na superfície horizontal do local analisado com representação dos versores e ângulos relativos à posição do sol e à do coletor	38
Figura 11 - Vista lateral do coletor e as relações existentes entre o ângulo β , o vetor $\mathbf{n}_C$ (perpendicular ao coletor) e o eixo x'	39
Figura 12 – Representação do vetor $\hat{n}_C$ para um rastreador horizontal com eixo de rotação norte-sul.....	41
Figura 13 – Representação do vetor $\hat{n}_s$ e suas projeções no plano $x'-y'$ para dois instantes: anterior $\hat{n}_{s-0}$ e posterior $\hat{n}_{s-1}$	42
Figura 14 – Representação do vetor da posição do módulo $\hat{n}_C$ para o caso de rastreamento segmentado de meio de arco e a projeções de $\hat{n}_s$ no plano $x'-y'$ para os instantes de início $\hat{n}_{s-0}$ e final $\hat{n}_{s-1}$ do passo considerado.	44
Figura 15 – Representação genérica de uma Curva Banheira, ou Função Taxa de Falhas (taxa de falhas λ em função do tempo t)	46
Figura 16 – Curva Banheira para Sistemas Eletrônicos e Mecânicos	47
Figura 17 –Taxa de Falha λ com distribuição exponencial (constante), juntamente com a PDF e a Curva de Confiabilidade ($R(t)$)	47
Figura 18 - Metodologia de Análise de Vida Útil para estudo da variável aleatória T a partir do conjunto de tempos t_i de ocorrência de falha para um conjunto de sistemas	50
Figura 19- Metodologia de Análise de Vida Útil para estudo da variável aleatória T	50
Figura 20 – Curvas de potência gerada de um sistema fotovoltaico com rastreamento de eixo único (funcionado e quebrado) comparadas com curvas de irradiância no plano para coletores com rastreamento e afixados	58
Figura 21 - Etapas do Método de Detecção de Falhas nos Sistema de Rastreamento.....	63
Figura 22 -Etapas para inferência dos índices probabilísticos de confiabilidade dos rastreadores	64
Figura 23 – Curvas de potências medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no local 2 do <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs, Austrália Central, em 6 de abril de 2011.....	65

Figura 24 – Curvas de potências medida e simuladas para um inversor com 5 sistemas fotovoltaicos com rastreadores de eixo único, no local 5 do <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs, Austrália Central, em 13 de fevereiro de 2012.....	66
Figura 25 – Curvas de potências medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no local 2 do <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs, Austrália Central, em 10 de março de 2012.....	67
Figura 26 – Curvas de Erro (CE) das potências medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no local 2 do <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs, Austrália Central, em 10 de março de 2012	68
Figura 27 - Esquema do procedimento de seleção de dias utilizados para o ajuste dos coeficientes do modelo de potência	73
Figura 28 – Radiações Global Horizontal medida e de Céu Claro para um dia cuja correlação entre essas curvas é maior que 0,99	74
Figura 29 – Radiações Global Horizontal medida e de Céu Claro para um dia cuja correlação entre essas curvas está entre 0,95 e 0,97	75
Figura 30 – Valores normalizados de Potência medida e de Radiação no plano do coletor com rastreamento para um dia cuja correlação entre essas curvas é maior que 0,98 e para o qual a correlação entre as radiações globais horizontal medida e de céu claro	75
Figura 31 – Valores normalizados de Potência medida e de Radiação no plano do coletor com rastreamento para um dia cuja correlação entre essas curvas é maior que 0,98 e para o qual a correlação entre as radiações globais horizontal medida e de céu claro está entre 0,95 e 0,9976	
Figura 32 – Curvas de Potência medida para os dias 15/10/2010 e 24/07/2018, os quais a correlação entre as radiações globais horizontal medida e de céu claro está entre 0,97 e 0,9976	
Figura 33 - Exemplo de Diagrama de Taylor.....	79
Figura 34 – Curvas de potência medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no local 2 no <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs, Austrália Central, em 24 de agosto de 2010	80
Figura 35 – Correlações de Funcionamento e de Falha para dias rotulados visualmente como dias indefinidos, no local 2 do <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs	86
Figura 36 – Correlações de Funcionamento e de Falha para dias rotulados visualmente como dias indefinidos, no local 5 do <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs	86
Figura 37 – Representação esquemática dos 3 grupos de atributos extraídos das curvas analisadas.....	87
Figura 38 - Exemplo Genérico de Método de Classificação por Limiar, cujos valores acima do limiar são classificados com pertencentes à classe positiva.....	89
Figura 39 - Representação Esquemática da construção da Curva ROC a partir da variação do limiar de classificação do modelo.....	91
Figura 40 - Representação de como se comporta a curva ROC de um modelo ideal de classificação (verde).....	92
Figura 41 – Diagrama representando o conjunto das classificações dos dias em relação ao sistema de rastreamento, excluindo os dias de dados ausentes e de dias em que o rastreadores esteve na posição de <i>stow</i>	93
Figura 42 – Representação esquemática do método de identificação de limiares a partir dos valores dos atributos	94
Figura 43 – Representação esquemática com histogramas contínuos do método de identificação de limiares	94
Figura 44 - Representação Esquemática do Método de Classificação a partir do par de atributos	95
Figura 45 - Pressupostos sobre em relação a ocorrência de falha, tempo de reparo e tempo entre falhas a partir da classificação dos dias.....	98

Figura 46 - Representação esquemática dos procedimentos de Ajuste de Curva para determinação dos coeficientes da CDF de Weibull de dois parâmetros a fim de verificar a eficácia no uso dos dados de falhas do modelo de detecção	103
Figura 47 – Localização geográfica do Desert Knowledge Australia Solar Centre, DKASC	105
Figura 48 – Esquema em 3D do Sistemas Fotovoltaicos presentes no DKASC.....	106
Figura 49 - Sistema Fotovoltaico do sítio 2 do DKASC, com rastreador duplo de azimute-altura, modelo ADES 5F-27M.	107
Figura 50 - Sistema Fotovoltaico do sítio 5 do DKASC, composto por cinco rastreadores de eixo simples segmentados, modelo BW Solar F2 5 Star Tracker, com abertura de 90°, andando em passos de 15°.....	108
Figura 51 - Potência medida e Potência simulada pelos modelos linear, Q3 quadrático e Pv Watt para o sistema FV do sítio 2 do DKASC, no dia 11 de julho de 2012, cuja correlação entre I_h e $I_{h,CC}$ está entre 0,95 e 0,97.....	115
Figura 52 - Potência medida e Potência simulada pelos modelos linear, Q3 quadrático e Pv Watt para o sistema FV do sítio 5 do DKASC, no dia 22 de setembro de 2009, cuja correlação entre I_h e $I_{h,CC}$ é maior que 0,99.....	116
Figura 53 - Diagrama de Taylor dos modelos de potência para o sistema fotovoltaico do sítio 2 do DKASC	116
Figura 54 - Diagrama de Taylor dos modelos de potência para o sistema fotovoltaico do sítio 5 do DKASC	117
Figura 55 – Potências medida e simuladas para um inversor com 5 sistemas fotovoltaicos com rastreadores de eixo único, no local 5 no <i>Desert Knowledge Australia Solar Centre</i> , em Alice Springs, Austrália Central, em 19 de março de 2011	119
Figura 56 – Curva ROC dos três atributos com melhor desempenho na classificação dos dias indefinidos (bipartição do conjunto), para o sistema do Sítio 2	121
Figura 57 – Curva ROC dos três atributos com melhor desempenho na classificação dos dias de falha a partir do subconjunto de dias definidos identificados na rotulação visual, para o sistema do Sítio 2.....	122
Figura 58 – Gráfico dos valores de AC razão entre as Áreas entre Curvas de funcionamento e de falha para o subconjunto de dias indefinidos, para o sítio 2, durante o período de calibração	122
Figura 59 – Curva ROC dos três atributos com melhor desempenho na classificação dos dias indefinidos (bipartição do conjunto), para o sistema do Sítio 5	123
Figura 60 Curva ROC dos três atributos com melhor desempenho na classificação dos dias de falha a partir do subconjunto de dias definidos identificados na rotulação visual, para o sistema do Sítio 5	124
Figura 61 - Matrizes de Confusão para os períodos de calibração e validação do Par 1*1 (maior MRPA) para o sítio 2	126
Figura 62 - Matrizes de Confusão para os períodos de calibração e validação do Par 1*1 (maior MRPA) para o sítio 5	128
Figura 63 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas da Rotulação Visual para o Sítio 2, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente	132
Figura 64 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas da Método do Limiar Móvel para o Sítio 2, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente	133
Figura 65 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas da Rotulação Visual para o	

Sítio 5, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente	133
Figura 66 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas do Método do Limiar Móvel para o Sítio 5, com os eixos y e x correspondendo aos valores de $\ln \ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente	134
Figura 67 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 5 oriundos dos dados da Rotulação Visual, e representação dos ajustes locais a partir da partição das amostras, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_O))] \}$ e $\ln(t_O)$, respectivamente	137
Figura 68 – Função Taxa de Falhas para o Tempo de Operação para o Sítio 5 obtida do ajuste da distribuição de Weibull de 4 parâmetros modificada com os dados originados da Rotulação Visual sem tratamento	138
Figura 69 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 5 oriundos dos dados da Rotulação Visual, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_O))] \}$ e $\ln(t_O)$, respectivamente	139
Figura 70 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 5 oriundos dos dados do Método do Limiar Móvel, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_O))] \}$ e $\ln(t_O)$, respectivamente	141
Figura 71 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 2 oriundos dos dados da Rotulação Visual, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_O))] \}$ e $\ln(t_O)$, respectivamente	142
Figura 72 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 2 oriundos dos dados do Método do Limiar Móvel, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_O))] \}$ e $\ln(t_O)$, respectivamente	142
Figura 73 – Curva de Manutenção do sistema de rastreamento dos Sítios 2 e 5	144
Figura 74 - Função Taxa de Reparo do sistema de rastreamento dos Sítios 2 e 5	145
Figura 75 - Curva de Confiabilidade do sistema de rastreamento do Sítio 2	146
Figura 76 - Função Taxa de Falhas do sistema de rastreamento do Sítio 2	146

LISTA DE TABELAS

Tabela 1 – Tipos Mais Comuns De Rastreadores de Eixo Único	27
Tabela 2 – Valores aproximados com duas casas decimais da declinação solar (δ), do versor nS no sistema horizontal $x' - y' - z'$ e do azimute solar ΦS para os dias julianos 172 (21/06) e 355 (21/12), em uma latitude Ω de 0°	37
Tabela 3 – Análise Simplificada de Modo de Falhas e Efeitos para componentes de um rastreador de eixo duplo para sistema HCPV	57
Tabela 4 - Limites impostos aos coeficientes durante o ajuste dos modelos de potência.....	77
Tabela 5 - Curvas comparadas para extração de Atributos de Similaridade.....	83
Tabela 6 – Exemplo de rótulos na biblioteca python TSFRESH e atributos correspondentes	85
Tabela 7 - Exemplo genérico de matriz de confusão para um conjunto com três classes	90
Tabela 8 - Exemplo Genérico de Cálculo da Métrica de Ranqueamento de Par de Atributo (MRPA) a partir das métricas de desempenho levando em conta os períodos de calibração e validação.....	97
Tabela 9 - Exemplo Genérico de estimativa da CDF pela aproximação de Bernard do ranqueamento da mediana.....	99
Tabela 10 – Dados Geográficos DKASC.....	106
Tabela 11 – Dados do Sistema Fotovoltaico do Sítio 2 do DKASC	107
Tabela 12 – Dados do Sistema Fotovoltaico do Sítio 5 do DKASC	108
Tabela 13 - Resultados dos Testes Locais para as séries das variáveis de interesse dos sistemas FV analisados do DKASC.....	111
Tabela 14 - Dias selecionados do sistema FV do sítio 2 para ajuste dos coeficientes do modelo de potência.....	113
Tabela 15 - Dias selecionados do sistema FV do sítio 5 para ajuste dos coeficientes do modelo de potência.....	113
Tabela 16 - Valores dos ajustes dos coeficientes por modelo e classe de dia.....	114
Tabela 17 – Parâmetros estatísticos de avaliação dos modelos de potência	117
Tabela 18 – Quantitativo de dias por classe para o Sítio 2 do DKASC.....	118
Tabela 19 – Quantitativo de dias por classe para o Sítio 5 do DKASC.....	118
Tabela 20 – Os três melhores Pares de Atributos ordenados pelo MRPA para o Sítio 2	125
Tabela 21 - Os três melhores Pares de Atributos ordenados pelo MRPA para o Sítio 5.....	127
Tabela 22 - Ordenamento dos Tempos de Reparo e inferência dos valores da CDF estimada pela aproximação de Bernard do ranqueamento da mediada, por sítio e por conjunto de origem dos dados de falha (Rotulação Visual e Método Limiar Móvel).....	131
Tabela 23 – Comparação das CDFs ajustadas aos Tempos de Reparo oriundos dos dados da Rotulação Visual e do Método Limiar Móvel.....	134
Tabela 24 - Ordenamento dos Tempos de Operação e inferência dos valores da CDF estimada pela aproximação de Bernard do ranqueamento da mediada, por sítio e por conjunto de origem dos dados de falha (Rotulação Visual e Método Limiar Móvel).....	135
Tabela 25 - Ordenamento dos Tempos de Operação e inferência dos valores da CDF estimada pela aproximação de Bernard do ranqueamento da mediada, por sítio e por conjunto de origem dos dados de falha (Rotulação Visual e Método Limiar Móvel).....	136
Tabela 26 – Valores do ajuste ascendente da primeira partição do conjunto de tempos de operação oriundos da Rotulação Visual para o Sítio 5	139
Tabela 27 - Análise das partições dos conjuntos de Tempos de Operação e aplicação do procedimento de tratamento dos dados.....	140
Tabela 28 – Comparação das CDFs ajustadas aos Tempos de Operação oriundos dos dados da Rotulação Visual e do Método Limiar Móvel.....	143

LISTA DE ABREVIATURAS E SIGLAS

A: Disponibilidade Estacionária

A e B: Coeficientes do modelo de Leng e Ackerman de céu claro

AC: Atributos de comparação

AG: Atributos de caráter geral

ARP: Processo de Renovação Alternada, *Alternating Renewal Process*

AS: Atributos de similaridade

$A(t)$: Disponibilidade não estacionária

AUC: Área sobre a Curva ROC, *Area Under Curve*

C95: Dias cuja correlação entre I_h e $I_{h,CC}$ esteve entre 0.95 e 0.97

C97: Dias cuja correlação entre I_h e $I_{h,CC}$ esteve entre 0.97 e 0.99

C99: Dias cuja correlação entre I_h e $I_{h,CC}$ foi maior que 0.99

C_0 , C_1 e C_2 : Coeficientes do modelo Q3 Quadrático

CA: Corrente Alternada

CAL: Período de Calibração

CC: Corrente Contínua

CDF: Função de Distribuição Acumulada, *Cumulative Distribution Function*

CE: Curva de Erro

$C_{PEARSON}$: Correlação de Pearson

DJ: Dia Juliano

DKASC: *Desert Knowledge Australia Solar Centre*

DP: Desvio padrão

DP_{OBS} : Desvio padrão da observação

DTW: Atributo Distorção Dinâmica do Domínio Temporal, *Dynamic Time Warping*

f_1 : Parâmetro de avaliação da classificação dos modelos

FN: Falsos negativos

FP: Falsos positivos

$\hat{F}_{R,i}$: Estimativa da CDF na i -ésima posição a partir da aproximação de Bernard para o ranqueamento da mediana

$f(t)$: Função Densidade de Probabilidade de ocorrência de falhas

$F(t)$: Função de Distribuição Acumulada de ocorrência de falhas

FV: Fotovoltaica

$I_{0,eff}$: Irradiância Solar Extraterrestre efetiva (W/m^2)

I_h : Irradiância global no plano horizontal (W/m^2)

$I_{h,CC}$: Irradiância Global Horizontal de Céu Claro (W/m^2)

$I_{h,col}$: Irradiância global no plano do coletor (W/m^2)

I_d : Irradiância difusa no plano horizontal (W/m^2)

K : Coeficiente de conversão de radiação em temperatura, conhecido como K de Ross ($^{\circ}C\ m^2/W$)

K_T : Coeficiente térmico de perda de potência ($W/^{\circ}C$)

L_{LOC} : Longitude local

LM: Método de ajuste não linear de Levenberg-Marquardt

L_{STC} : Longitude padrão da localidade

$M(t)$: Curva de Manutenção

MAE : Erro médio absoluto, *Mean Average Error*

MBE : Erro Médio, *Mean Bias Error*

MSE : Erro quadrático médio, *Mean Square Error*

$MTTF$: Tempo médio para a falha, *Mean time to failure*

MRPA: Métrica de Ranqueamento e Par de Atributo

N : Número total de amostras de um conjunto

$\hat{n}_C$: Versor de posição do coletor

$nI_{h,rast}$: Irradiância global no plano de um coletor com rastreamento normalizada, sem unidade

$nPot$: Potência normalizada, sem unidade

$\hat{n}_S$: Versor da posição do Sol

ORP: Processo Ordinário de Renovação, *Ordinary Renewal Process*

P_{CA} : Potência de saída de corrente alternada do inversor (W)

P_{CC} : Potência de corrente contínua (W)

PCM: Atributo Mapeamento Parcial de Curvas *Partial Curve Mapping*

PLC: Controlador Lógico Programável, *Programmable Logic Controller*

P_{MIN} : potência mínima necessária para o funcionamento do inversor (W)

P_{nom} : Potência fotovoltaica nominal do sistema (W)

POA: Plano do módulo, *Plane-of-array*

$posição_{CALIBRAÇÃO}$: posição de ordenamento do par na calibração em relação aos demais, levando em conta o $f1$ médio: as primeiras posições são dos pares com maiores $f1$

$posição_{VALIDAÇÃO}$: posição de ordenamento do par na validação em relação aos demais, levando em conta o $f1$ médio: as primeiras posições são dos pares com maiores $f1$

Pot : Potência (W)

Qtd: Quantidade de vezes que determinada variável se repetiu

$R(t)$: Curva de Confiabilidade

$RMSD$: Raiz da diferença centrada quadrática

$RMSE$: Raiz do erro médio quadrático, *Root Mean Square Error*

$SS4$: Métrica de avaliação de Taylor

$STIR$: *Subspace Trust-Region Interior-Reflective*, método de ajuste não linear

$Stow$: Posição fixa em que o rastreador se coloca em situações ambientais extremas

T_{AMB} : Temperatura ambiente ($^{\circ}C$)

TFP: Taxa de Falsos Positivos

t_o : Tempo de operação

t_R : Tempo de reparo

T_{ref} : Temperatura de referência dos módulos

TVP: Taxa de Verdadeiros Positivos

VAL: Períodos de Validação

VN: Verdadeiros negativos

VP: Verdadeiros positivos

$x' - y' - z'$: Sistema horizontal de coordenadas

β : Ângulo de inclinação do coletor ($^{\circ}$)

δ : Declinação Solar ($^{\circ}$)

ε : Fator de forma da distribuição de Weibull

ξ_{dia} : Duração em horas do dia

η : eficiência da conversão CC para CA de um inversor

η_{FV} : Coeficiente de representação do comportamento médio das perdas relacionadas ao gerador fotovoltaico

$\eta_{NOMINAL}$: rendimento nominal do inversor, razão entre potências nominais de corrente alternada e de corrente contínua apresentadas no catálogo de informações do inversor

$\eta_{REFERÊNCIA}$: rendimento de referência do inversor

θ : Ângulo de abertura do coletor (°)

θ_Z : Ângulo zenital (°)

$\lambda(t)$: Função Taxa e Falhas

$\lambda_R(t)$: Função Taxa de Reparo

μ_O : Média dos tempos de operação o sistema

μ_R : Média dos tempos de reparo do sistema

ρ_{ground} : Refletividade do solo

€: razão que tem por numerador $[\eta_{NOMINAL} \times (\sum_{i=1}^N P_{CC,i})]$ e por denominador a potência nominal de corrente alternada apresentada no catálogo do inversor

Φ : Azimute do coletor (°)

Φ_S : Azimute solar (°)

Ψ : Ângulo entre o versor $\hat{n}_C$ e o eixo x' (°)

Ω : Latitude local (°)

v : Velocidade do vento (m/s)

x : Tempo entre falhas

τ : Fator de escala da distribuição de Weibull

ω : Ângulo horário (°)

SUMÁRIO

1. INTRODUÇÃO	20
2. CONCEITOS PRELIMINARES	25
2.1 Introdução aos Rastreadores Solares.....	25
2.1.1 Mecanismo.....	25
2.1.2 Sistema de Controle	29
2.1.3 Sistema de Apontamento	30
2.2 Principais Componentes dos Rastreadores Solares.....	31
2.3 Cinemática dos Rastreadores Solares.....	34
2.3.1 Coletor com Rastreamento Solar de Eixo Duplo.....	40
2.3.2 Coletor com Rastreamento Solar de Eixo Único	41
2.4 Conceitos da Teoria da Confiabilidade.....	44
2.4.1 Índices Probabilísticos de Confiabilidade	45
2.4.2 Função Taxa de Falhas e Seus Modelos.....	46
2.4.3 Modelagem da Confiabilidade de Sistemas.....	48
3. REVISÃO DE LITERATURA	52
3.1 Análise de Confiabilidade no Setor de Energia Solar	52
3.2 Análise de Confiabilidade para Rastreadores Solares	56
4. METODOLOGIA	62
4.1 Método de Detecção de Falhas	64
4.1.1 Modelo de Potência.....	68
4.1.2 Rotulação Visual e Classificação dos Dias.....	80
4.1.3 Extração de atributos.....	81
4.1.4 Avaliação do Método de Classificação	88
4.1.5 Método de Classificação.....	93
4.2 Ajuste dos Dados de Ocorrência de Falha para Inferência dos Índices Probabilísticos de Confiabilidade.....	97
4.3 Dados Empregados.....	105
5. RESULTADOS E DISCUSSÕES	112
5.1 Método de Detecção de Falhas	112
5.1.1 Modelo de Potência.....	112
5.1.2 Rotulação Visual e Classificação dos Dias.....	118
5.1.3 Método de Classificação.....	119
5.2 Ajuste dos Dados de Ocorrência de Falha para Inferência dos Índices Probabilísticos de Confiabilidade.....	129
5.2.1 Ajuste dos Coeficientes de Weibull e Comparação dos Conjuntos de Dados.....	130
5.2.2 Índices Probabilísticos de Confiabilidade dos Sistemas	143
6. CONCLUSÕES	148
REFERÊNCIAS	152

1. INTRODUÇÃO

A busca por construir uma matriz energética mundial mais limpa e renovável tem sido uma constante atual. Dentre as fontes renováveis, a energia solar fotovoltaica vem se destacando ao longo do tempo. No final de 2022, a energia Fotovoltaica (FV), juntamente com a eólica, representava 12,1% da produção de energia elétrica mundial, sendo a energia solar desde 2016 a fonte renovável com a maior capacidade de potência anual instalada (REN21, 2020, 2023). Segundo dados da Agência Internacional de Energia (IEA, 2022), a energia solar FV terá uma expansão esperada de 30% de sua capacidade instalada entre 2022 e 2030.

O crescimento do mercado de energia fotovoltaica está diretamente relacionado à redução de custos e ao aumento do fator de capacidade dos sistemas. De acordo com relatórios anuais da NREL, nos Estados Unidos, os custos totais para instalação e gerenciamento de usinas fotovoltaicas de grande escala diminuíram de aproximadamente US\$5/Wp em 2007 para US\$1,3/Wp em 2022 (NREL, 2023). Além disso, o fator de capacidade aumentou para 30% em 2022. No Brasil, a EPE destaca em seu relatório que, entre 2014 e 2018, o fator de capacidade das usinas fotovoltaicas aumentou de 20% para cerca de 30% (EPE, 2018). Em suas análises, tanto o NREL quanto a EPE atribuem esse fortalecimento do mercado FV ao desenvolvimento tecnológico, principalmente dos módulos fotovoltaicos e seus processos de fabricação, dos inversores e do emprego dos rastreadores solares — equipamento estudado neste trabalho.

Na geração centralizada, os módulos fotovoltaicos eram tradicionalmente instalados de forma fixa, gerando menos energia no início da manhã e fim do dia devido à inclinação dos raios do Sol com relação ao plano de abertura. Assim, o uso do rastreador solar altera essa perspectiva ao permitir que o módulo siga a trajetória solar, aumentando a radiação direta recebida ao longo do dia (RECA-CARDEÑA e LÓPEZ-LUQUE, 2018). Dessa forma, é possível reduzir a área de módulos necessária para geração energética, aumentando o fator de capacidade do sistema (MORADI; ABTAHI; MESSENGER, 2016). Os ganhos na produtividade pelo uso de rastreadores podem ser, em média, entre 30 a 40% por ano (MOUSAZADEH *et al.*, 2009).

No mercado fotovoltaico, o rastreamento solar é um dos setores que mais cresce, passando de um valor de mercado de US\$2,92 bilhões em 2018 para US\$4,41 bilhões em 2022, com expectativas de triplicação do valor atual em 2030 (GRAND VIEW RESEARCH, 2019, 2023; VANTAGE MARKET RESEARCH, 2023). Sua presença em usinas centralizadas de 20% em 2016 aumentou para aproximadamente 30% em 2020 (FORTUNE BUSINESS

INSIGHTS, 2020, 2021; JÄGER-WALDAU, 2018). Na Espanha, o rastreamento solar insignificamente presente em 2006 esteve em 25% a 30% das novas instalações já em 2008 (RENEWABLE ENERGY WORLD, 2009). Na geração solar centralizada dos Estados Unidos, o uso de rastreadores, incipiente até 2010, aumentou consideravelmente, fazendo com que a porcentagem total de usinas com a tecnologia fosse maior que 60% em 2021 (BOLINGER; SEEL; ROBSON, 2019; MORDOR INTELLIGENCE, 2022). No Brasil, no Ambiente de Contratação Regulada, dos 31 projetos do primeiro leilão em 2014, 10 apresentaram rastreamento (EPE, 2014); pouco depois, em 2017, a tecnologia era empregada em todos (EPE, 2018).

Segundo Bolinger, Seel e Robson (2019), a recente e crescente presença do rastreamento nas usinas fotovoltaicas se deve ao barateamento e ao aumento da confiabilidade da tecnologia. Por exemplo, o preço médio de uma usina estadunidense com rastreamento diminuiu quase pela metade entre 2007 e 2012 (BARBOSE *et al.*, 2015). Com respeito à confiabilidade, os fabricantes de rastreadores, como a Nextracker (2021), Politec (2021) e a STI Norland (2021), informam disponibilidade do equipamento próxima de 100%, com garantias em torno de 5 anos para os sistemas de direção e controle e 10 anos para a estrutura.

No entanto, a literatura carece de estudos independentes sobre a confiabilidade do equipamento (ANDERSON *et al.*, 2022; RUTH; MULLER, 2018). Além disso, garantia e disponibilidade não correspondem à totalidade de informações relevantes numa análise de confiabilidade (LEWIS *et al.*, 1994). Confiabilidade é aqui definida, conforme a área da Teoria da Confiabilidade, como a capacidade de um equipamento ou sistema não apresentar falha para determinado período de operação, dentro de condições operacionais pré-estabelecidas (BAZOVSKY, 2004). Nesse sentido, a avaliação da capacidade de operação do sistema é realizada a partir dos índices de confiabilidade, sendo os mais conhecidos os índices probabilísticos de confiabilidade, como a Curva de Confiabilidade e a Função Taxa de Falhas (SINGH; JIRUTITIJAROEN; MITRA, 2019).

García *et al.* (2009) analisaram os registros de falhas fornecidos pela equipe de manutenção de um conjunto de usinas fotovoltaicas presentes na região de Navarra, na Espanha. Na localidade, os rastreadores solares apresentaram a maior quantidade de eventos de mal funcionamento em relação ao demais equipamentos. No entanto, o autor afirma que, embora as falhas no sistema de rastreamento representem cerca de 45% das ocorrências de eventos indesejados, elas têm um impacto menor na produção por apenas reduzir parcialmente a energia produzida. Mesmo assim, do ponto de vista de perda energética, ficaram atrás apenas das falhas

nos inversores e nos sistemas de proteção. No relatório apresentado por Klise (2015), do total de falhas registradas no banco de dados da Sandia, 18% estavam relacionadas a rastreadores.

Por sua vez, Jordan *et al.* (2020) realizaram a análise dos registros de falhas de mais de 100 000 sistemas FV nos Estados Unidos, entre usinas residenciais, comerciais e de grande escala, comentando que os registros de falhas no sistema de rastreamento são bastante subestimados. Nos sistemas nos quais foram registrados mal funcionamento dos rastreadores, as falhas ocorreram em aproximadamente 0,14% do tempo total de funcionamento das usinas, representando uma perda média na potência esperada de 15,73%. Spertino e Corona (2013) também realizaram um estudo semelhante para 13 sistemas FV na Europa, concluindo que as falhas devido a desalinhamentos dos rastreadores são um dos principais fatores que afetam a desempenho das usinas. Os autores sugerem uma manutenção anual para prevenção de falhas nos rastreadores.

Com base nas informações apresentadas anteriormente, pode-se concluir que a utilização de rastreadores solares aumenta os custos anuais de Operação e Manutenção (O&M) de sistemas fotovoltaicos (COSSU; BACCOLI; GHIANI, 2021). Segundo Ramasamy *et al.* (2022), nos Estados Unidos, esse aumento é estimado em aproximadamente US\$2/KWp por ano, comparando-se usinas de grande escala com instalação fixa e com rastreadores de eixo único. Cossu; Baccoli; Ghiani (2021) afirmam que, de uma forma geral, é uma boa estimativa um aumento na ordem de 10% a 15% nos custos de O&M. No entanto, apesar do aumento nos custos de manutenção, a utilização do sistema de rastreamento é compensada pelo aumento no fator de capacidade proporcionado. Pois, para aumentar o fator de capacidade de um sistema fixo na mesma proporção que os rastreadores aumentam, o custo seria muito maior (CAMPBELL *et al.*, 2009).

Nesse contexto, analisar a confiabilidade dos rastreadores fotovoltaicos é importante para entender como as falhas podem ocorrer e qual é o tempo esperado de operação sem falhas. Isso permite projetar as necessidades de intervenções no sistema e incluir esses custos na elaboração dos projetos fotovoltaicos. Dessa forma, é possível planejar a manutenção do sistema de maneira mais eficiente e prever os custos associados (SCHENKELBERG, 2015; THERISTIS; PAPAZOGLU, 2014). De forma geral, as usinas FV estão sujeitas a diversos tipos de falhas, como por exemplo incertezas associadas aos equipamentos empregados, condições ambientais adversas e problemas sistêmicos na rede de distribuição. Analisando a confiabilidade, é possível quantificar os riscos de perdas na produção energética esperada e minimizá-los (ZHANG *et al.*, 2013).

A confiabilidade de empreendimentos fotovoltaicos é uma área de pesquisa extremamente relevante no contexto atual (SINGH; SAKET; KHAN, 2023). Diversos artigos tratam do tema para sistemas autônomos (ACUÑA *et al.*, 2018; CARRASCO *et al.*, 2013; GEORGILAKIS e KATSIGIANNIS, 2009; THERISTIS e PAPAZOGLU, 2014). Outros desenvolvem modelos que calculam a Curva de Confiabilidade e a Função Taxa de Falhas para equipamentos específicos, como módulos fotovoltaicos (GAUTAM e KAUSHIKA, 2002; VÁZQUEZ e REY-STOLLE, 2008) e inversores (MORRIS; FIFE, 2009; SCHENKELBERG, 2015); bem como para a usina como um todo (FRAAS e PARTAIN, 2010; SAYED *et al.*, 2019). No entanto, poucos trabalhos pesquisados mencionam a análise de confiabilidade para os rastreadores solares, mesmo aqueles que analisam a usina em sua totalidade. Dessa forma, não se sabe ao certo como a inserção de partes móveis em usinas fotovoltaicas afeta a confiabilidade global, tendo em vista que um ponto positivo de sistemas FV é justamente sua alta confiabilidade, devido ao baixo índice de falhas e à baixa necessidade de manutenção.

Os índices de confiabilidade de muitos equipamentos presentes em usinas fotovoltaicas são conhecidos, entretanto essa informação para rastreadores solares são escassas (SAYED *et al.*, 2019). Dos poucos trabalhos presentes na literatura, existe uma metodologia para determinação da disponibilidade do equipamento a partir de dados de potência (RUTH e MULLER, 2018), um trabalho de identificação de falhas a partir de dados de corrente (ANDERSON *et al.*, 2022), um estudo da curva de disponibilidade do equipamento a partir de testes de aceleração e modelos teóricos (ELERATH, 2010, 2011, 2017, 2018; ELERATH; SPENCER; HORNE, 2011) e existem modelos relacionados à análise do tema para heliostatos — rastreadores utilizados em usinas de solar térmica, similares aos rastreadores fotovoltaicos — (BENAMMAR e TEE, 2019).

Diante de tudo que foi abordado nos parágrafos anteriores, o presente estudo tem por objetivo principal o desenvolvimento de uma metodologia para inferência da Curva de Confiabilidade e da Função Taxa de Falhas (índices probabilísticos de confiabilidade) para rastreadores solares empregados em usinas fotovoltaicas. Assim, para alcançar este objetivo, os objetivos específicos serão dois. Primeiro, a aplicação de um método de detecção de dias nos quais os rastreadores falharam a partir da comparação entre a potência medida e a potência simulada. Segundo, com os dias de falha fornecido pelo método anterior, aplicar um modelo de ajuste de curva para inferência dos índices de confiabilidade.

Para o desenvolvimento do trabalho, duas premissas principais serão consideradas. A primeira é que o formato da curva diária de potência medida carrega informação sobre a funcionalidade dos rastreadores solares. Se o formato da curva de potência medida for

semelhante a curva de potência simulada de um sistema fixo, é o indicativo de que naquele dia o rastreador falhou. Essa é a mesma premissa apresentada no trabalho de Ruth e Muller (2018) e também utilizada no trabalho de Anderson *et al.* (2022). A segunda premissa é justamente supor que a confiabilidade dos sistemas de rastreamento possa ser modelada a partir de um modelo de Processo de Renovação (*Renewal Process*, em inglês), o qual pode ser um Processo Ordinário de Renovação ou Processo de Renovação Alternada.

Nesse sentido, a principal contribuição inovadora do trabalho é se utilizar de informações operacionais para o desenvolvimento de uma metodologia de inferência de índices probabilísticos de confiabilidade de rastreadores solares fotovoltaicos. Em outras palavras, os índices de confiabilidade serão inferidos a partir de dados de operação do equipamento, ao invés de modelos teóricos e de testes de laboratório. Para obter dados de falhas operacionais, o trabalho aplicará um método de detecção de falhas. Esse método de detecção apresenta melhorias em relação ao apresentado por Ruth e Muller (2018) e Anderson *et al.* (2022), pois emprega uma abordagem não subjetiva de determinação de limiares e utiliza uma quantidade maior de atributos de comparação entre as potências medida e simulada. Inclusive, o presente trabalho determina em quais classes de dias o método de detecção de falhas não funciona, denominando tais dias como indefinidos.

Este trabalho é dividido em seis seções, começando com esta Introdução. A segunda seção, Conceitos Preliminares, apresenta informações importantes para o desenvolvimento da pesquisa, incluindo a cinemática dos rastreadores e conceitos da teoria da confiabilidade. Na Revisão de Literatura, é apresentada uma análise de como a literatura aborda o tema da confiabilidade para sistemas de potência, para usinas fotovoltaicas e para rastreadores solares. A seção de Metodologia descreve o método de detecção dos dias de falha e a metodologia para inferir os índices probabilísticos de confiabilidade. Na seção de Resultados, são apresentados e discutidos os resultados da aplicação da metodologia em duas usinas fotovoltaicas. Por fim, a seção de Conclusão traz os apontamentos finais e perspectivas futuras.

2. CONCEITOS PRELIMINARES

2.1 Introdução aos Rastreadores Solares

O rastreador solar é um sistema utilizado em tecnologias solares com a finalidade de acompanhar o movimento relativo do Sol em relação à Terra, utilizando tanto um único eixo de movimento, como dois. Com o eixo duplo, obtêm-se uma melhor precisão no rastreio, em contrapartida, torna-se uma tecnologia mais cara devido aos elevados custos de operação e manutenção (COSSU; BACCOLI; GHIANI, 2021). O sistema de rastreamento solar não é usado exclusivamente para painéis fotovoltaicos, mas também para outros equipamentos como refletores, lentes, entre outros serviços ópticos (RHIF, 2011). A sua principal finalidade é diminuir o ângulo de incidência entre a reta normal ao plano do equipamento e reta base da radiação direta provinda do Sol, aumentando assim a energia captada ao longo do dia (RECA-CARDEÑA; LÓPEZ-LUQUE, 2018).

Na prática, o aumento da produtividade energética pelo uso de rastreadores podem ser, em média, entre 20 a 40% por ano, a depender da localidade (BOLINGER; SEEL; ROBSON, 2019; RACHARLA; RAJAN, 2017). Levando em consideração a área, a partir de 1/GCR (inverso da taxa de ocupação de terreno, *Ground Coverage Ratio* — em inglês) maiores que 2, o uso de rastreadores aumentam de 10% a 30% a energia produzida em relação ao sistema fixo (NARVARTE; LORENZO, 2008).

Os rastreadores solares podem ser classificados a partir de três partes empregadas em suas construções: os mecanismos, os sistemas de controle de rastreamento e os sistemas de apontamento (RECA-CARDEÑA; LÓPEZ-LUQUE, 2018).

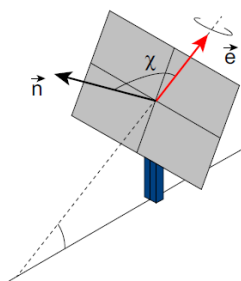
2.1.1 Mecanismo

É a parte responsável por movimentar o sistema, seguindo a posição do Sol. Em relação aos mecanismos, os rastreadores solares podem ser classificados em eixo único ou eixo duplo, dependendo dos graus de liberdade de movimento.

- Eixo Único

Os sistemas de rastreamento solar de eixo único apresentam, como o próprio nome sugere, apenas um grau de liberdade de movimento, isto é, rotacionam apenas ao longo de um eixo. O eixo de rotação, em relação ao módulo fotovoltaico, pode ser horizontal, vertical ou inclinado. A Figura 1 mostra um esquema geral de um sistema de rastreamento de eixo único, com o eixo de rotação unitário $\vec{e}$ e o eixo $\vec{n}$ normal ao plano coletor. O ângulo χ entre esses vetores unitários se mantém constante (RECA-CARDEÑA; LÓPEZ-LUQUE, 2018). São vários os tipos de sistemas desenvolvidos com eixo único. Na Tabela 1 são apresentados os principais

Figura 1 – Eixo de rotação unitário $\vec{e}$ e o eixo $\vec{n}$ normal a um rastreador de eixo único.

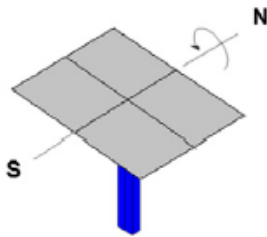
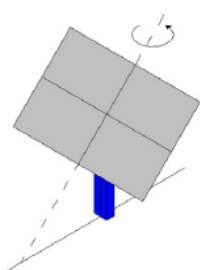
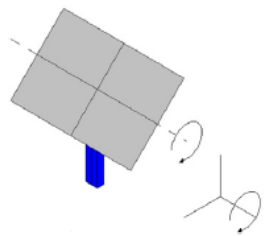
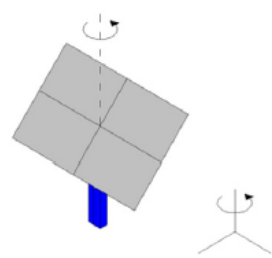


Fonte: Reza-Cardeña; López-Luque (2018).

As vantagens de um sistema de rastreamento com eixo único são, dentre outras, sua baixa complexidade tecnológica, baixa manutenção e seu baixo consumo de energia para funcionamento; no entanto, possui ganhos energéticos menores em comparação aos sistemas de eixo duplo pelo fato de seguir o Sol em apenas um eixo (DGS, 2008).

- Eixo Duplo

Os rastreadores solares de eixo duplo apresentam dois graus de liberdade, o que os permite acompanhar diretamente o vetor solar. Esse tipo de rastreador é utilizado principalmente em sistemas de concentração de potência, em especial em solar térmica e sistemas de concentração fotovoltaica, nos quais o grau de aproveitamento da radiação direta afeta consideravelmente a produtividade (HAFEZ; YOUSEF; HARAG, 2018).

Tabela 1 – Tipos Mais Comuns De Rastreadores de Eixo Único	Imagem
<i>Rastreador de eixo horizontal Norte-Sul</i> O eixo de rotação é horizontal e se encontra na direção Norte-Sul, com $\chi = 90^\circ$, movimentando-se de leste a oeste.	
<i>Rastreador Polar: eixo polar Norte-Sul inclinado com ângulo igual a latitude local</i> O eixo de rotação fica inclinado em relação à reta Norte-Sul de um ângulo igual à latitude local.	
<i>Rastreador de eixo horizontal Leste-Oeste</i> Nesta configuração, o eixo de rotação está na direção Leste-Oeste, e o sistema se movimenta de norte a sul.	
<i>Rastreador Azimutal</i> A inclinação do ângulo do módulo é constante e igual à latitude local, rastreando o ângulo de azimute do Sol.	

Segundo Racharla e Rajan (2017), os dois principais tipos de rastreadores de eixo duplo são os de inclinação de ponta e os de azimute-altura (Figura 2). Os rastreadores de inclinação de ponta (Figura 3) se movimentam com eixos de rotação Norte-Sul e Leste-Oeste. Já nos rastreadores azimute-altura (Figura 4), o equipamento funciona a partir das coordenadas angulares azimute e inclinação. Os módulos fotovoltaicos, ao rotacionar em torno de uma estrutura em formato de anel, variam o azimute; e, rotacionando no outro eixo, que fica no plano do módulo, variam a inclinação.

A depender do preço da energia, para sistemas fotovoltaicos para os quais a radiação direta não é primordial (a exemplo dos sistemas de concentração), o rastreamento de eixo duplo não é muito empregado. O eixo duplo aumenta em cerca de 10% a geração em comparação ao eixo único, o que muitas vezes não compensa o capital investido devido ao aumento nos custos de operação e manutenção, visto que rastreadores de eixo duplo apresentam maior risco de falha (BOLINGER; SEEL e ROBSON, 2019). Por isso, atualmente, as principais pesquisas na área focam no barateamento da tecnologia, aumento da confiabilidade e diminuição de sua complexidade (HAFEZ; YOUSEF e HARAG, 2018).

Figura 2 – Rastreadores de Inclinação de Ponta e de Azimute-Altura, respectivamente, e seus eixos de rotação.



Fonte: Hong et al. (2016).

Figura 3 – Exemplo de Rastreador de Inclinação de Ponta.



Fonte: SEI (2016).

Figura 4 – Exemplo de Rastreador Azimute-Altura.



Fonte: SEI (2016).

2.1.2 Sistema de Controle

O rastreamento solar pode ser classificado, de acordo com o sistema de controle, em dois tipos: malha aberta, ou malha fechada (também chamado de controle baseado em resposta).

Os de malha fechada são aqueles que utilizam repostas diretas de fotossensores (como os resistores dependentes de luz) para seguir a posição do Sol (CHONG e WONG, 2009). Neste tipo de rastreamento um par de sensores elétricos ópticos (fotoresistores ou células fotovoltaicas) é instalado em regiões antiparalelas para captação da radiação solar (MOUSAZADEH *et al.*, 2009). Ao perceber diferença na medição entre os sensores do par, o controlador ativa o motor de rotação para mover o sistema fotovoltaico até a equalização do sinal. Em dias com muitas nuvens ou com pouca radiação solar, ocorre uma desorientação nesse tipo de rastreador (RECA-CARDEÑA e LÓPEZ-LUQUE, 2018). A fim de evitar essas falhas, tem sido instalado um sistema auxiliar que controla o rastreio até a estabilização da radiação e dos fatores climáticos locais (CHONG e WONG, 2009).

Por sua vez, os rastreadores solares de malha aberta utilizam microprocessadores para realizarem cálculos astronômicos para determinação da posição do Sol (CHONG e WONG, 2009). Sua principal vantagem é a independência de condições externas favoráveis para seu funcionamento. Suas principais desvantagens são: a baixa adaptabilidade frente a passagem de nuvens; e o fato de que, caso os valores de tempo e localidade não estejam corretos, haverá um desalinhamento em relação à trajetória solar, ocasionando perda de energia.

Mousazadeh *et al.* (2009) relata que alguns rastreadores utilizam sensores junto ao controle de malha aberta para auxiliar na identificação do posicionamento do Sol. São os chamados sistemas mistos, que mesclam componentes dos sistemas de malha aberta e fechada.

2.1.3 Sistema de Apontamento

Os rastreadores solares são classificados em ativos e passivos a depender dos equipamentos responsáveis por sua movimentação mecânica, isto é, seus sistemas de apontamento. Os ativos utilizam, de forma geral, motores elétricos de corrente contínua ou de corrente alternada para rotação, enquanto os passivos o fazem utilizando princípios termo-hidráulicos (RECA-CARDEÑA e LÓPEZ-LUQUE, 2018).

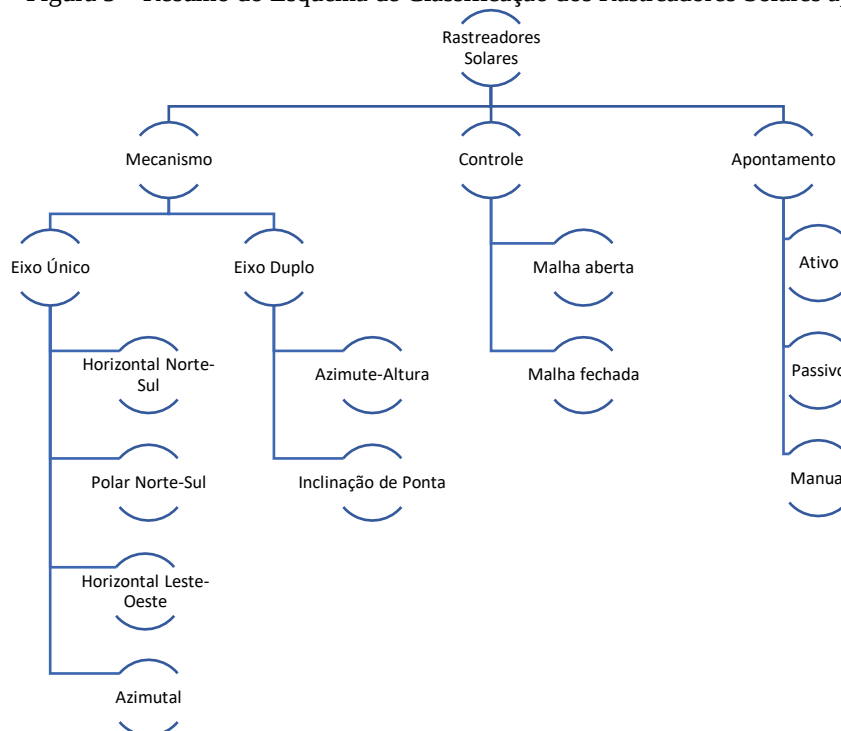
O rastreamento solar ativo utiliza motores elétricos para direcionar as placas solares. Para isso, o sistema consome energia, o que requer a instalação de baterias. Surge, por conseguinte, uma das suas principais desvantagens: os gastos energéticos necessários à movimentação do sistema. Mesmo assim, devido a ser um sistema muito confiável, é o mais empregado no mercado fotovoltaico. O principal tipo de rastreador ativo é aquele que utiliza microprocessadores e sensores elétricos para a movimentação do sistema (HAFEZ; YOUSEF; HARAG, 2018).

Por sua vez, os rastreadores solares passivos são sistemas em equilíbrio estável, como fluidos voláteis, molas ou ligas com efeito de forma, que, ao serem retirados de sua posição natural, tendem a retornar a ela (HAFEZ; YOUSEF; HARAG, 2018; MOUSAZADEH *et al.*, 2009). Nesse tipo de sistema, quando a radiação recebida não é perpendicular aos painéis fotovoltaicos, um dos agentes recebe mais incidência que outro, fazendo o sistema ficar desbalanceado e tender a movimentar os módulos em busca de um ponto de equilíbrio no recebimento da radiação solar. Possuem baixa complexidade, mas apresentam baixa eficiência a baixas temperaturas ambiente (MOUSAZADEH *et al.*, 2009).

Por último, vale mencionar os sistemas manuais, que, como o próprio nome sugere, utilizam o trabalho braçal para a mudança periódica dos ângulos de inclinação dos sistemas. Em algumas situações no rastreamento de eixo duplo, um dos eixos é construído para ser movido manualmente. Apesar de sua deficiência no seguimento do Sol, são sistemas com baixa complexidade tecnológica e mais baratos. Sua principal aplicação é na área de secadores solares e na de sistemas fotovoltaicos flutuantes (HAFEZ; YOUSEF; HARAG, 2018).

A Figura 5 apresenta um esquema resumindo a classificação dos rastreadores solares a partir de suas três partes principais.

Figura 5 – Resumo do Esquema de Classificação dos Rastreadores Solares apresentado.



Fonte: Elaboração Própria.

2.2 Principais Componentes dos Rastreadores Solares

Papathanasiou (2021) em um relatório organizado pela *Taiyang News*, um veículo de comunicação especializado no mercado solar, afirma que o mercado de energia fotovoltaica utiliza predominantemente rastreadores de eixo único com rotação de leste a oeste, com apontamento ativo e malha aberta. O relatório destaca que os principais componentes de um rastreador podem ser divididos em duas partes: os componentes mecânicos e os componentes elétricos.

Os componentes mecânicos incluem o eixo de torque, ao qual os módulos estão conectados, sendo a estrutura responsável pela rotação do sistema; o suporte do rastreador, que é um esqueleto metálico que sustenta os módulos; e os grampos e fixadores (PAPATHANASIOU, 2021). Segundo o relatório, os fabricantes de rastreadores solares oferecem uma ampla variedade de opções de montagem para os módulos fotovoltaicos, permitindo que eles sejam dispostos em diferentes configurações. Isso inclui a possibilidade de montar os módulos na orientação vertical, com uma ou duas fileiras por rastreador, ou na

orientação horizontal, com duas ou quatro fileiras por rastreador. Essa flexibilidade permite que os sistemas de rastreamento solar sejam adaptados às necessidades específicas de cada projeto.

Já os componentes elétricos são compostos por motores, caixas de engrenagens e atuadores. De acordo com Papathanasiou (2021), os motores utilizados são principalmente motores de corrente contínua sem escovas (*brushless DC motors*, em inglês), compostos por ímãs permanentes e um estator feito de bobinas. Junto ao motor, há um tipo de caixa de engrenagem conhecida como *slew drive*, que pode suportar com segurança cargas radiais e axiais sem freio, além de transmitir torque. Para alimentação dos componentes elétricos, a maioria dos fabricantes de rastreadores de um eixo oferece sistemas independentes formados por módulos fotovoltaicos e baterias (PAPATHANASIOU, 2021).

Por sua vez, os sistemas de apontamento possuem um sistema de controle e comunicação, geralmente um Controlador Lógico Programável (*Programmable Logic Controller* — PLC), responsável por indicar a posição ótima do rastreador. O PLC dos rastreadores pode utilizar diversos sensores, como GPS, inclinômetros, relógios a hora de Greenwich, piranômetros, anemômetros e até informações dos inversores.

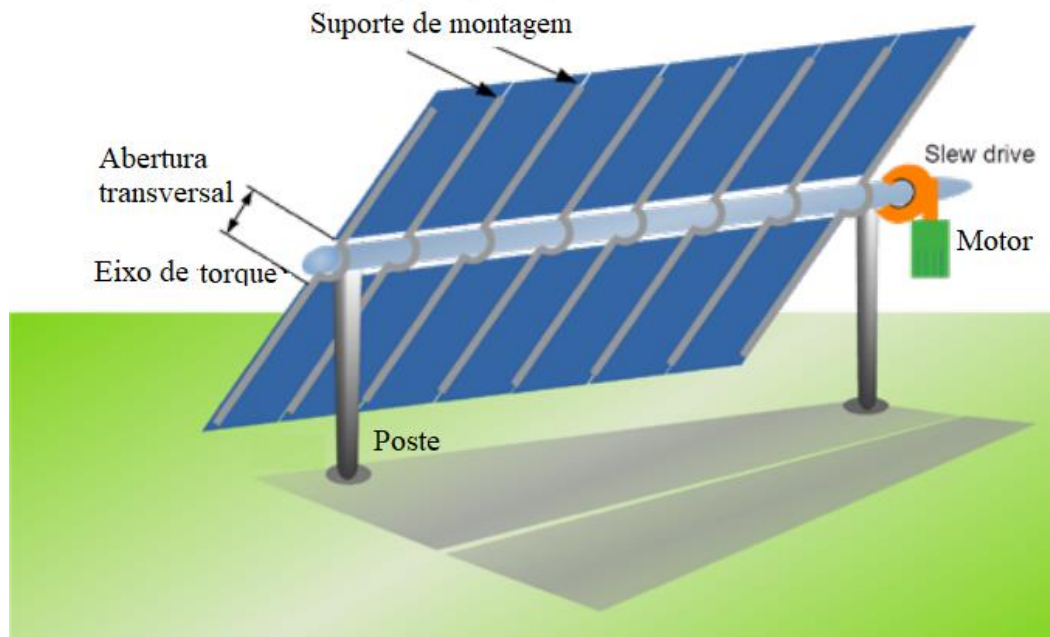
Por fim, os sistemas de rastreamento podem ser de linha independente ou multilinhas conjugadas. Nos de linha independente, cada eixo de rotação possui um atuador específico, enquanto os de multilinhas possuem duas ou mais fileiras de módulos manejadas pelo mesmo atuador.

A Figura 6 apresenta um desenho esquemático dos principais componentes de um rastreador de eixo único, com duas fileiras de módulos na posição vertical.

Em relação ao rastreadores de eixo duplo, os principais componentes são basicamente iguais aos apresentados para o sistema de rastreamento de eixo único. A principal diferença consiste no fato de que o conjunto motor e *slew drive* é responsável pelo movimento azimutal do sistema, enquanto um atuador linear (equipamento que converte energia em movimento linear, geralmente um *screw jack*) é responsável pelo movimento de elevação (BSQ SOLAR, 2023; HORNE, 2012; SEME *et al.*, 2020). A Figura 7 apresenta um esquema com os principais componentes de um rastreador de eixo duplo.

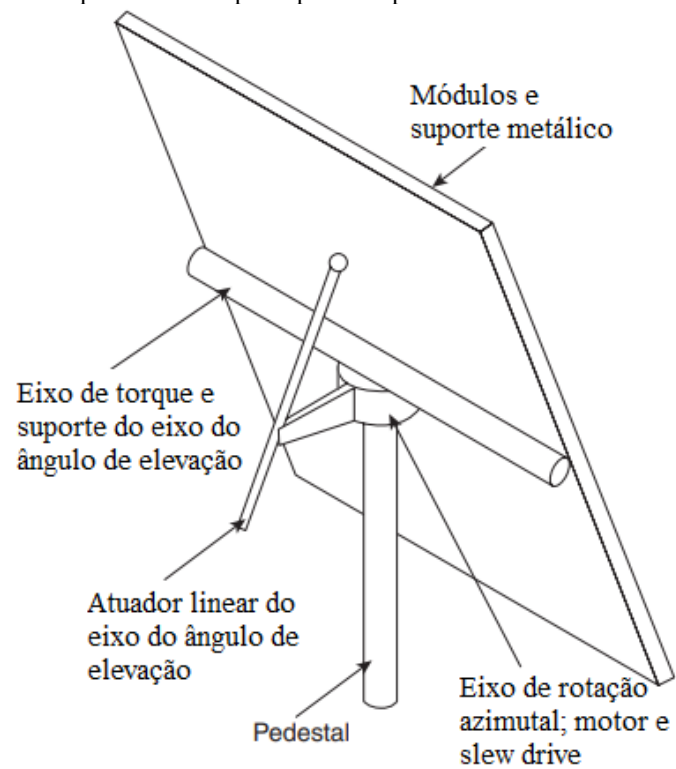
Para finalizar, é importante abordar que atualmente os rastreadores solares possuem algoritmos de otimização para maximizar a energia produzida ao evitar sombreamento entre fileiras próximas, conhecido como *backtracking* (PAPATHANASIOU, 2021). Além disso, os sistemas possuem mecanismos para reduzir os impactos relacionados a ventos fortes, neve e tempestades, colocando os módulos em uma posição fixa conhecida como *stow* que minimize os danos.

Figura 6 – Desenho esquemático dos principais componentes de um rastreador de eixo único, com duas fileiras de módulos na posição vertical



Fonte: Papathanasiou (2021).

Figura 7 – Desenho esquemático dos principais componentes de um rastreador de eixo duplo



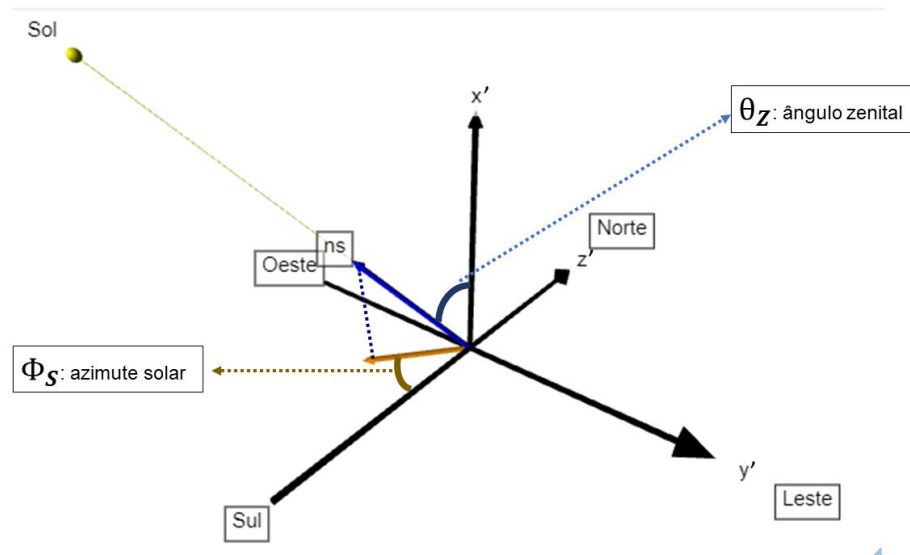
Fonte: Horne (2012).

2.3 Cinemática dos Rastreadores Solares

A cinemática dos rastreadores solares se refere aos cálculos necessários para a determinação da posição do sistema em determinado instante de tempo. Inicialmente, é importante ter em mente o sistema de coordenadas $x' - y' - z'$ mostrado na Figura 9 e os ângulos zenital e de azimute, que determinam a posição do vetor da posição do Sol $\hat{n}_s$.

O sistema $x' - y' - z'$ de coordenadas é denominado sistema horizontal, pois coloca como referência o plano da superfície local. Definindo, pois, como eixo x' a reta perpendicular à superfície apontando para o zênite, como eixo z' o eixo norte-sul apontando para o norte e como eixo y' o eixo leste-oeste apontando para leste.

Figura 8 – Sistema de coordenadas $x' - y' - z'$ localizado na superfície horizontal do local analisado, vetor solar $\hat{n}_s$ e ângulos das coordenadas polares



Fonte: Elaboração Própria.

Em azul está representado o vetor da posição solar $\hat{n}_s$, e em amarelo escuro sua projeção no plano $y' - z'$. A posição do Sol, neste caso, é determinada ao se conhecer os valores do ângulo zenital θ_z e do azimute solar Φ_S .

θ_z é o ângulo entre $\hat{n}_s$ e $\hat{x}'$; e Φ_S é o ângulo entre a projeção de $\hat{n}_s$ no plano $y' - z'$ e $-\hat{z}'$, sendo positivo no sentido horário.

O vetor $\hat{n}_s$, em $x' - y' - z'$, pode ser calculado pela Equação 1 (RABL, 1985).

$$\hat{n}_s = \begin{pmatrix} \cos \delta \cos \omega \cos \Omega + \sin \delta \sin \Omega ; \\ - \cos \delta \sin \omega ; \\ - \cos \delta \cos \omega \sin \Omega + \sin \delta \cos \Omega \end{pmatrix}_{x'-y'-z'} \quad (1)$$

Em que:

- Ω : latitude local ($^\circ$);
 - δ : declinação solar ($^\circ$), dada por
 - $\sin \delta = -\sin 23,45^\circ \cos\left[\frac{360^\circ}{365,25} (DJ + 10)\right]$ sendo DJ o dia juliano;
 - ω : ângulo horário ($^\circ$), dado por:
 - $\omega = \frac{\Delta t}{\xi_{dia}} 360^\circ$, com Δt sendo a diferença em horas da hora solar atual para o meio-dia solar, negativo para antes do meio-dia solar; e ξ_{dia} sendo a duração em horas do dia, 24h;
 - Devido à hora marcada no relógio não corresponder à hora solar, é necessário aplicar a seguinte correção:
 - $Hora_{solar} = Hora_{local} + 4(L_{LOC} - L_{STC}) + E$, sendo L_{STC} : longitude padrão da localidade em graus, é igual ao fuso-horário local em horas multiplicado por 15° (a oeste de Greenwich positivo e a leste negativo); L_{LOC} : longitude local em graus (mesma referência de sinais anterior); $E = 9,87 \sin 2B - 7,53 \cos B - 1,5 \sin B$, em que $B = \frac{365^\circ(DJ - 81)}{364}$, com DJ igual ao dia juliano.
- As parcelas $4(L_{LOC} - L_{STC}) + E$ estão em minutos. A unidade da constante 4 são $min/^\circ$, já que 15° equivaleria a $60 \text{ min} = 1h$.

No sistema de coordenadas $x' - y' - z'$, os ângulos necessários para a determinação da posição do Sol são o azimute solar Φ_S e o ângulo zenital θ_Z . θ_Z é o ângulo entre $\hat{n}_S$ e $\hat{x}'$; e Φ_S é o ângulo entre a projeção de $\hat{n}_S$ no plano $y' - z'$ e $-\hat{z}'$, sendo positivo no sentido horário.

O ângulo zenital pode ser calculado como o produto vetorial entre $\hat{n}_S$ e $\hat{x}'$, resultando na Equação 2. Como $0^\circ \leq \theta_Z \leq 180^\circ$ e a função arco cosseno está definida para o intervalo $[0^\circ; 180^\circ]$, o arco cosseno da Equação 2 retornará o valor correto para a declinação.

$$\cos \theta_Z = \cos \delta \cos \omega \cos \Omega + \sin \delta \sin \Omega \quad (2)$$

Em que:

- Ω : latitude local ($^\circ$);
- δ : declinação solar ($^\circ$),
- ω : ângulo horário ($^\circ$).

Tomando a projeção do versor $\hat{n}_S$ sobre o plano $y' - z'$, a tangente do azimuth solar será a razão entre as componentes y' e z' de $\hat{n}_S$, conforme mostra a Equação 3. No entanto, é preciso prestar atenção ao fato de que $-180^\circ \leq \Phi_S \leq 180^\circ$, enquanto a função arcotangente está definida para o intervalo $[-90^\circ; 90^\circ]$. Nesse sentido, para o semiplano $z' \leq 0$, o arco tangente da razão entre as componentes y' e z' de $\hat{n}_S$ retornará o valor correto para o azimuth solar. Já para o semiplano $z' > 0$, existe a seguinte particularidade: para valores com $y' \geq 0$, o resultado da função arco tangente deve ser acrescido de -180° ; e para $y' < 0$, deve ser acrescido de 180° .

$$\Phi_S = \begin{cases} \tan^{-1} \frac{\cos \delta \sin \omega}{\cos \delta \cos \omega \sin \Omega - \sin \delta \cos \Omega}, \hat{n}_{S-z'} \leq 0 \\ -180^\circ + \tan^{-1} \frac{\cos \delta \sin \omega}{\cos \delta \cos \omega \sin \Omega - \sin \delta \cos \Omega}, \hat{n}_{S-z'} > 0 \text{ e } \hat{n}_{S-y'} \geq 0 \\ 180^\circ + \tan^{-1} \frac{\cos \delta \sin \omega}{\cos \delta \cos \omega \sin \Omega - \sin \delta \cos \Omega}, \hat{n}_{S-z'} > 0 \text{ e } \hat{n}_{S-y'} < 0 \end{cases} \quad (3)$$

Em que:

- Ω : latitude local ($^\circ$);
- δ : declinação solar ($^\circ$),
- ω : ângulo horário ($^\circ$).

Para exemplificar as correções nas equações do azimuth solar aplicadas na Equação 3, é possível tomar como exemplo a Tabela 2, para os dias 21/06 e 21/12 para uma latitude Ω de 0° . Esses dias possuem uma declinação solar de aproximadamente 23.45° e -23.45° , respectivamente. Isso significa que, para a latitude mencionada, o Sol estará sempre ao norte do ponto de vista da Terra no dia 21/06 e ao sul no dia 21/12, com o mesmo ângulo em módulo.

No dia 21/06, de acordo com a Tabela 2, o azimuth solar Φ_S , sem as correções, indicaria que o Sol nasceu a oeste e se pôs a leste, além de estar no hemisfério sul. Essa informação contradiz a componente z' do versor $\hat{n}_S$ no sistema horizontal, que é positiva, indicando que o Sol está ao norte. Além disso, o fato de caminhar de oeste para leste contradiz a componente y' do versor $\hat{n}_S$ no sistema horizontal, que varia de um valor positivo para negativo ao longo das horas do dia, indicando que o Sol caminhou de leste a oeste. Nesse contexto, com as correções

propostas, pode-se observar na tabela que o ângulo de azimute passa a indicar que o Sol está no hemisfério norte e percorre de leste a oeste. Além disso, é importante destacar que, quando o Sol se encontra ao sul da localidade, como nas informações apresentadas para o dia 21/12, as correções no ângulo de azimute solar não são necessárias.

Tabela 2 – Valores aproximados com duas casas decimais da declinação solar (δ), do versor $\hat{n}_S$ no sistema horizontal $x' - y' - z'$ e do azimute solar Φ_S para os dias julianos 172 (21/06) e 355 (21/12), em uma latitude Ω de 0°

Hora Solar, ω ($^\circ$)		-30°	-15°	0°	15°	30°
DJ = 172 (21/06)	δ ($^\circ$)	23,45°	23,45°	23,45°	23,45°	23,45°
	$\hat{n}_{S_{x'-y'-z'}}$	(0,79; 0,46; 0,40)	(0,89; 0,24; 0,40)	(0,92; 0,00; 0,40)	(0,89; -0,24; 0,40)	(0,79; -0,46; 0,40)
	Φ_S ($^\circ$) – sem correções	49,06°	30,82°	0°	-30,82°	-49,06°
	Φ_S ($^\circ$) – com correções	-130,94°	-149,18°	180,00°	149,18°	130,94°
	δ ($^\circ$)	-23,45°	-23,45°	-23,45°	-23,45°	-23,45°
	$\hat{n}_{S_{x'-y'-z'}}$	(0,79; 0,46; -0,40)	(0,89; 0,24; -0,40)	(0,92; 0,00; -0,40)	(0,89; -0,24; -0,40)	(0,79; -0,46; -0,40)
DJ = 355 (21/12)	Φ_S ($^\circ$) – sem correções	-49,06°	-30,82°	0°	30,82°	49,06°
	Φ_S ($^\circ$) – com correções	-49,06°	-30,82°	0°	30,82°	49,06°

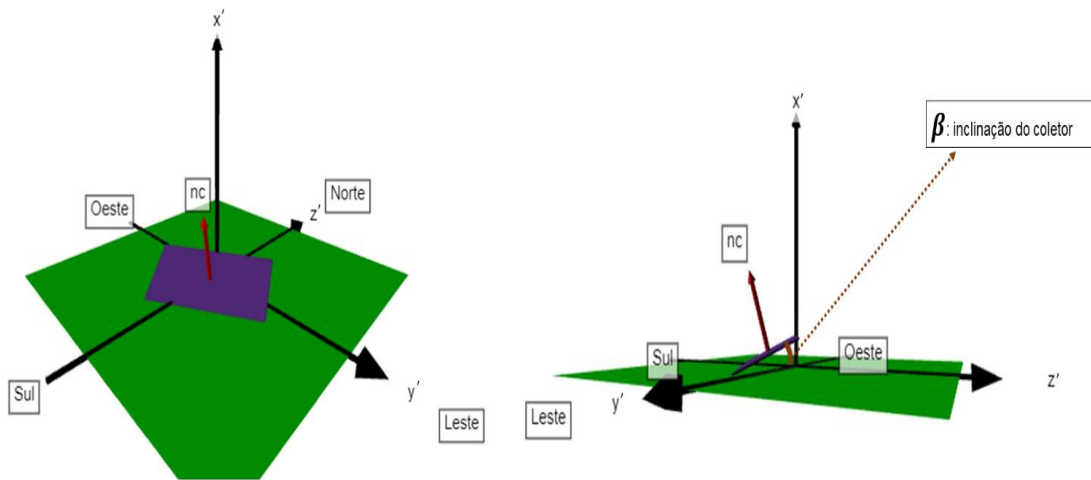
Fonte: Elaboração Própria.

Tendo os valores do azimute solar e do ângulo zenital, também é possível definir $\hat{n}_S$ para o sistema de coordenadas $x' - y' - z'$ conforme a Equação 4.

$$\hat{n}_S = (\cos \theta_Z ; -\sin \theta_Z \sin \Phi_S ; -\sin \theta_Z \cos \Phi_S)_{x'-y'-z'} \quad (4)$$

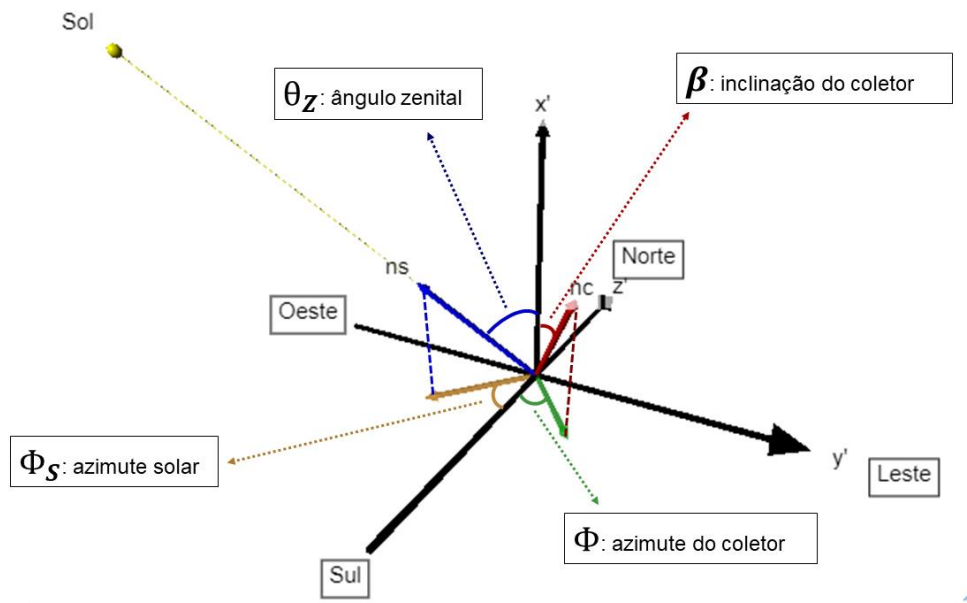
É importante também adicionar à análise o coletor de energia solar, neste caso, o módulo fotovoltaico. Como pode ser visto na Figura 9, o versor que identifica o coletor com inclinação β em relação à horizontal é aquele da reta normal à superfície do módulo, denominado $\hat{n}_C$. Assim sendo, dois ângulos são necessários para identificar a posição de $\hat{n}_C$ no sistema $x' - y' - z'$: o azimute Φ e a inclinação β . Na Figura 10, há a representação gráfica de todos os ângulos de interesse.

Figura 9 – Vistas superior e lateral do coletor (módulo fotovoltaico) posicionado na superfície com ângulo de inclinação β , sendo também representado o sistema de coordenadas $x' - y' - z'$ localizado na



Fonte: Própria

Figura 10 – Sistema de coordenadas $x' - y' - z'$ localizado na superfície horizontal do local analisado com representação dos vetores e ângulos relativos à posição do sol e à do coletor



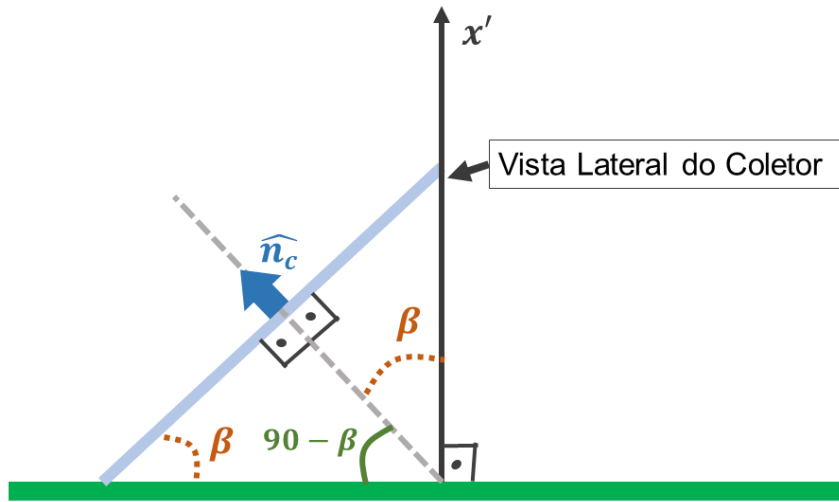
Fonte: Elaboração Própria.

Em azul está representado o vetor da posição do Sol $\hat{n}_s$, e em amarelo escuro sua projeção no plano $y' - z'$. Em vermelho está representado o vetor $\hat{n}_c$ da reta normal à superfície do coletor, e em verde sua projeção no plano $y' - z'$. A posição do coletor, neste caso, é determinada ao se conhecer os valores da sua inclinação β e do azimute Φ . β é o ângulo entre $\hat{n}_c$ e $\hat{x}'$; e Φ é o ângulo entre a projeção de $\hat{n}_c$ no plano $y' - z'$ e $\hat{z}'$, sendo positivo no sentido horário (portanto, no caso desta figura, Φ possui um valor negativo).

Por conseguinte, $\hat{n}_c$ pode ser definido no sistema $x' - y' - z'$ conforme Equação 5. Perceba que a convenção de sinais adotada para Φ é igual a adotada para o azimuth solar, e, por isso, o azimuth do coletor representado na Figura 10 possui um valor negativo. Vale destacar também que β , além de ser o ângulo entre o plano horizontal e os módulos, também é igual ao ângulo entre $\hat{n}_c$ e $\hat{x}'$, conforme pode ser verificado na **Erro! Autoreferência de indicador não válida..**

$$\hat{n}_c = (\cos \beta; -\sin \beta \sin \Phi; -\sin \beta \cos \Phi)_{x'-y'-z'} \quad (5)$$

Figura 11 - Vista lateral do coletor e as relações existentes entre o ângulo β , o versor $\hat{n}_c$ (perpendicular ao coletor) e o eixo x'



Fonte: Elaboração Própria.

Por último, determinado as coordenadas de $\hat{n}_c$ e $\hat{n}_s$, é possível calcular o ângulo entre eles, denominado ângulo de abertura θ , já que o produto vetorial $\hat{n}_s \cdot \hat{n}_c = \cos \theta$. Sabendo que $0 \leq \theta \leq 180^\circ$ e a função arco cosseno está definida para o intervalo $[0^\circ; 180^\circ]$, o arco cosseno, neste caso, retornará o valor correto para o ângulo de abertura. θ é de suma importância para, dadas as irradiâncias global e difusa no plano horizontal, inferir-se a irradiância total que chega no plano do coletor (*plane of array*, POA), conforme pode ser visto na Equação 6.

$$I_{h,col} = (I_h - I_d) \frac{\cos \theta}{\cos \theta_z} + I_d \frac{(1 + \cos \beta)}{2} + I_h \rho_{ground} \frac{(1 - \cos \beta)}{2} \quad (6)$$

Em que:

- $I_{h,col}$: irradiância global no plano do coletor (W/m^2)
- I_h : irradiância global no plano horizontal (W/m^2);
- I_d : irradiância difusa no plano horizontal (W/m^2);
- θ : ângulo de abertura do coletor;
- θ_z : ângulo zenital;
- β : inclinação do coletor;
- ρ_{ground} : refletividade do solo; 0,7 para solos com neve e 0,2 para os demais solos (RABL, 1985).

A posição do Sol, que varia no céu ao longo do tempo, é calculada a partir do versor $\hat{n}_s$. Quando os módulos são instalados de forma fixa, enquanto $\hat{n}_s$ está variando, a posição de $\hat{n}_c$ permanece a mesma. Isso muda com a utilização dos rastreadores solares, fazendo $\hat{n}_c$ também variar posicionalmente no decorrer do tempo. Nesse sentido, modelar o movimento de um rastreador é entender a movimentação temporal de $\hat{n}_c$, visando calcular o ângulo de abertura para um determinado instante a fim de obter a irradiância no plano do coletor.

2.3.1 Coletor com Rastreamento Solar de Eixo Duplo

O rastreador solar de eixo duplo é construído com dois graus de liberdade para movimento. Por isso, ele consegue acompanhar completamente o movimento do Sol ao continuamente fazer $\hat{n}_c$ estar na mesma posição que $\hat{n}_s$. No entanto, devido às limitações tecnológicas, sempre existe um erro associado ao seguimento, medido pelo valor do ângulo de abertura, já que, para um sistema de rastreamento contínuo de eixo duplo, θ deveria ser sempre zero. Na prática, um rastreador contínuo é aquele no qual o ângulo de abertura é pequeno o suficiente para considerar o seguimento contínuo, enquanto o rastreador segmentado é aquele que possui um erro considerável de rastreo solar.

De uma forma geral, os rastreadores de dois eixos são construídos apresentando uma limitação quanto aos ângulos mínimo e máximo de azimuth e de inclinação. Por exemplo, um sistema de rastreamento com 240° de variação no azimuth, caso esteja virado para o sul, só se movimenta em relação ao azimuth de -120° a $+120^\circ$ (ou -60° a $+60^\circ$, se virado para o norte). Para o caso de um rastreador com inclinação máxima de 60° , por exemplo, sua variação será

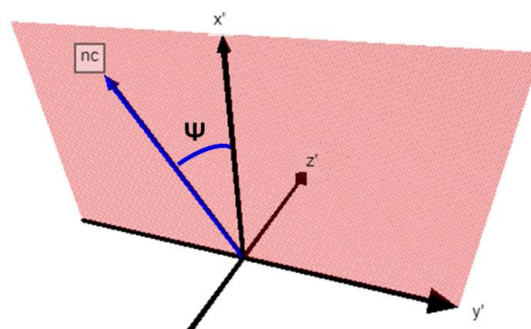
de 0° a 60° em relação à superfície horizontal. É importante destacar que tais limitações podem fazer com que, para determinada localidade, o erro indicado pelo fabricante do rastreador seja extrapolado para certas faixas do dia no ano, tendo em vista que a variação do azimuth solar e do ângulo zenital podem estar fora da faixa de abrangência do sistema de rastreo. E isso é particularmente relevante para sistemas fotovoltaicos de alta concentração, nos quais o ângulo de abertura máximo tem de estar bem controlado para que a produção energética seja efetiva.

Por último, vale abordar a existência da diferença entre as limitações mecânicas e as limitações relacionadas ao algoritmo. Todo algoritmo de rastreamento tem seu erro em relação à real posição do Sol, geralmente na casa de milésimos. No entanto, há também limites maiores, associados à mecânica do equipamento: engrenagens, sistema de transmissão, entre outros. Nesse sentido, à medida que são exigidos sistemas de rastreamento mais precisos, mais as limitações mecânicas surgem como uma barreira impeditiva.

2.3.2 Coletor com Rastreamento Solar de Eixo Único

Os rastreadores de eixo único não acompanham completamente o movimento do Sol, mas apenas parte dele, em outras palavras, o versor do módulo $\hat{n}_c$ só segue determinada projeção de $\hat{n}_s$ a depender do tipo de sistema (eixo Norte-Sul, Leste-Oeste, azimutal ou polar). Aqui o sistema de rastreamento analisado será horizontal com eixo de rotação na direção norte-sul rotacionando de leste a oeste, pois é amplamente empregado nas usinas fotovoltaicas (FORTUNE BUSINESS INSIGHTS, 2021; ONS, 2021; PAPATHANASIOU, 2021). Ele apresenta a mesma configuração da Figura 10, porém tendo $\beta=0$. Nesse caso, o vetor $\hat{n}_c$ estará sempre localizado no plano $x' - y'$ (Figura 12), sendo definido conforme a Equação 7.

Figura 12 – Representação do vetor $\hat{n}_c$ para um rastreador horizontal com eixo de rotação norte-sul



Fonte: Elaboração Própria.

$$\hat{n}_C = (\cos \Psi; -\sin \Psi; 0) \quad (7)$$

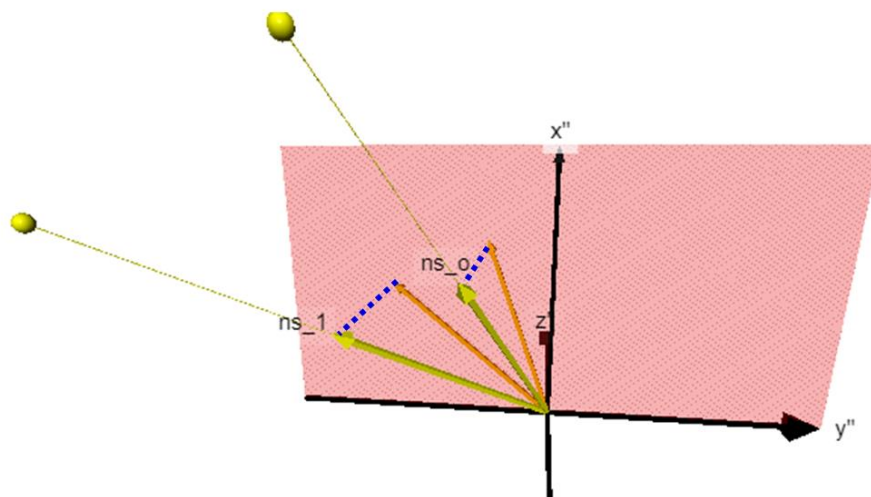
Em que:

- Ψ : ângulo entre o versor $\hat{n}_C$ e o eixo x' , positivo no sentido negativo do eixo y' (a oeste).

O ângulo Ψ é aquele entre o versor $\hat{n}_C$ e o eixo x' , sendo definido como positivo no sentido negativo do eixo y' (a oeste). É também o ângulo que caracteriza o movimento do rastreador horizontal com eixo de rotação Norte-Sul. Assim, ao se determinar o valor desse ângulo, a posição do sistema de rastreamento estará necessariamente definida. Vale destacar que, quando se estabelece que um rastreador de eixo único possui uma abertura com valor de, por exemplo, 120° , é normalmente assumido que o ângulo de rotação Ψ varia entre -60° a 60° .

Ao se olhar mais atentamente para a Figura 12, percebe-se que o rastreador de eixo único com eixo de rotação norte-sul na realidade não seguirá o versor $\hat{n}_S$ propriamente dito, mas sim a sua projeção no plano $x' - y'$ (Figura 13). Assim, Ψ deverá ser igual a Ψ_S , o ângulo entre a projeção de $\hat{n}_S$ no plano $x' - y'$ (denominada $\hat{n}_{S-X'-Y'}$) e o eixo x' .

Figura 13 – Representação do vetor $\hat{n}_S$ e suas projeções no plano $x' - y'$ para dois instantes: anterior $\hat{n}_{S-0}$ e posterior $\hat{n}_{S-1}$



Fonte: Elaboração Própria.

Sendo:

$$\hat{n}_{s-x'-y'} = (\cos \theta_Z; -\sin \theta_Z \sin \Phi_S; 0) \quad (8)$$

Logo:

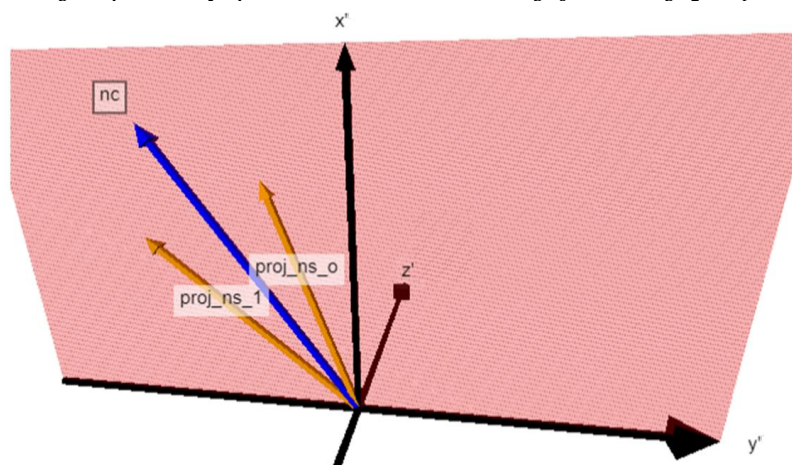
$$\tan \psi_S = -\frac{\sin \theta_Z \sin \Phi_S}{\cos \theta_Z} \quad (9)$$

É importante diferenciar os sistemas de rastreamento contínuo e segmentado. No segmentado ou discreto, o rastreador não estará continuamente posicionando o vetor $\hat{n}_C$ na mesma direção e sentido da projeção $\hat{n}_{s-x'-y'}$, mas realizará o rastreio de forma discretizada, a partir de determinados passos. Esses passos podem ser de início, de meio e fim de arco. O procedimento ficará mais evidente a partir do exemplo a seguir.

Para o caso da Figura 14, tome por exemplo que, para o primeiro instante, a projeção $\hat{n}_{s-x'-y'-0}$ está localizada a 20° do eixo x' e, para o último instante considerado, $\hat{n}_{s-x'-y'-1}$ estará localizado a 30° . Durante todo o trajeto no qual a projeção do versor solar no plano $x' - y'$ estará variando de 20° a 30° , o versor $\hat{n}_C$ estará parado e posicionado a 20° do eixo x' no rastreamento discreto de início de arco, 25° no do meio de arco e 30° no de fim de arco. Em seguida, quando começar o próximo trajeto do Sol indo de 30° a 40° no plano $x' - y'$, $\hat{n}_C$ estará localizado a 30° , 35° ou 40° do eixo x' a depender se o rastreio é de início, meio e fim de arco, respectivamente. Esse processo irá continuamente acontecer até que seja atingido o limite de abertura do rastreador em questão. Conclui-se que quem determina o tamanho do arco para o qual o rastreador estará parado é o passo considerado. No exemplo anterior, esse passo foi de 10° .

Por último, deve-se ter em conta que, na prática, devido às limitações tecnológicas, não existe rastreador contínuo, mas sim sistemas nos quais o passo é tão pequeno que podemos modelá-lo assertivamente como tal.

Figura 14 – Representação do vetor da posição do módulo $\hat{n}_C$ para o caso de rastreamento segmentado de meio de arco e a projeções de $\hat{n}_S$ no plano $x'-y'$ para os instantes de início $\hat{n}_{S-0}$ e final $\hat{n}_{S-1}$ do passo considerado.



Fonte: Elaboração Própria.

2.4 Conceitos da Teoria da Confiabilidade

Todos os agentes envolvidos, incluindo empresários, fabricantes, reguladores, operadores de serviços públicos e consumidores, necessitam de um modelo preditivo para o funcionamento de determinado equipamento ou sistema, a fim de auxiliar nas decisões de custo-benefício, indicar melhorias no desenho e fornecer estimativas do desempenho ao longo da vida útil (COLLINS *et al.*, 2009). Com essa finalidade, surge a teoria da confiabilidade.

Nascida durante o período da revolução industrial, a teoria da confiabilidade evoluiu suas abordagens à medida que novos desafios surgiam para setores como o elétrico, aeroespacial e eletrônico (BAZOVSKY, 2004). Deixando de ser algo mais intuitivo, a teoria da confiabilidade foi adquirindo a robustez teórica necessária para se transformar em área de conhecimento, utilizando conceitos de *design* e de estatística em suas análises de engenharia.

Nesse contexto, a Teoria da Confiabilidade é a área da engenharia responsável pelo estudo do comportamento das falhas durante a operação de determinado equipamento ou serviço. A principal definição relacionada à área é a do próprio conceito de confiabilidade: a capacidade de um sistema funcionar satisfatoriamente, dentro de condições operacionais estipuladas, durante determinado período de tempo (BAZOVSKY, 2004).

A avaliação da confiabilidade de um sistema, ou de sua capacidade de operacionalidade é realizada de uma forma geral a partir dos denominados índices de confiabilidade (SINGH; JIRUTITIJAROEN; MITRA, 2019). Uma classe de índices de confiabilidade muito conhecida são os referentes ao atendimento de cargas energéticas, tais como o *loss of load expectation*

(LOLE) e o *loss of energy expectation* (LOEE). Neste trabalho, porém, o principal foco está na classe dos índices probabilísticos de confiabilidade, como a Curva de Confiabilidade e a Função Taxa de Falhas, principalmente. Os índices probabilísticos permitem inferir a distribuição probabilística das falhas do sistema e fornecem informações valiosas para o gerenciamento dos equipamentos, principalmente na fase de projeto do sistema.

2.4.1 Índices Probabilísticos de Confiabilidade

Definamos a variável aleatória $\mathbf{T}$ como o período de tempo para que ocorra uma falha num determinado sistema ou equipamento. Sendo $F(t) = P\{\mathbf{T} \leq t\}$ a função de distribuição acumulada, a curva de confiabilidade é definida como:

$$R(t) = 1 - F(t) \quad (10)$$

Assim, a função densidade de probabilidade para a variável aleatória $\mathbf{T}$ é:

$$f(t) = -\frac{d}{dt}R(t) \quad (11)$$

Por sua vez, a taxa de falha (também chamada de taxa de risco ou de mortalidade) pode ser definida como (LEWIS *et al.*, 1994):

$$\lambda(t) = \frac{f(t)}{R(t)} \quad (12)$$

Utilizando a Equação 12, a confiabilidade também pode, após alguns cálculos, ser definida como:

$$R(t) = \exp \left[- \int_0^t \lambda(t') dt' \right] \quad (13)$$

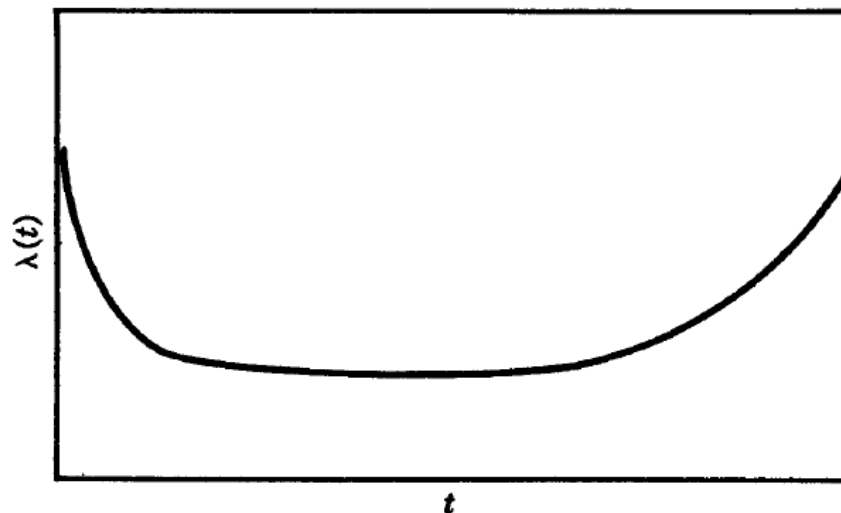
O tempo médio para a falha (*mean time to failure*, MTTF) é o valor esperado (esperança) da variável aleatória $\mathbf{T}$, que, após alguns cálculos algébricos, pode ser calculado como:

$$MTTF = \int_0^{\infty} R(t) dt \quad (14)$$

2.4.2 Função Taxa de Falhas e Seus Modelos

A função taxa de falhas também é conhecida como curva banheira, um exemplo genérico pode ser visto na Figura 15. Essa curva apresenta três momentos bastante característicos: o início, com as falhas conhecidas como de nascimento; o meio, com falhas conhecidas como aleatórias; e o fim, com falhas relacionadas à degradação e ao fim da vida útil (envelhecimento) (LEWIS *et al.*, 1994). A depender do sistema analisado, uma dessas falhas pode ser preponderante em relação às demais, como no caso de sistema mecânicos com as falhas de envelhecimento, e os componentes eletrônicos com as falhas aleatórias (Figura 16).

Figura 15 – Representação genérica de uma Curva Banheira, ou Função Taxa de Falhas (taxa de falhas λ em função do tempo t)



Fonte: Lewis et al. (1994).

As falhas de nascimento são aquelas que ocorrem principalmente durante o início da vida do sistema e estão relacionadas a problemas de qualidade na fabricação. As falhas aleatórias acontecem no decorrer da vida útil e estão relacionadas a fatores não controláveis, como por exemplo fenômenos naturais. Por último, as falhas de envelhecimento estão relacionadas aos efeitos de desgaste que a natureza proporciona ao sistema durante o tempo, um exemplo é a corrosão.

Para equipamentos para os quais as falhas aleatórias são preponderantes, uma boa aproximação é modelar a função taxa de falhas $\lambda(t)$ a partir de uma taxa constante, o caso denominado de distribuição exponencial. Neste caso:

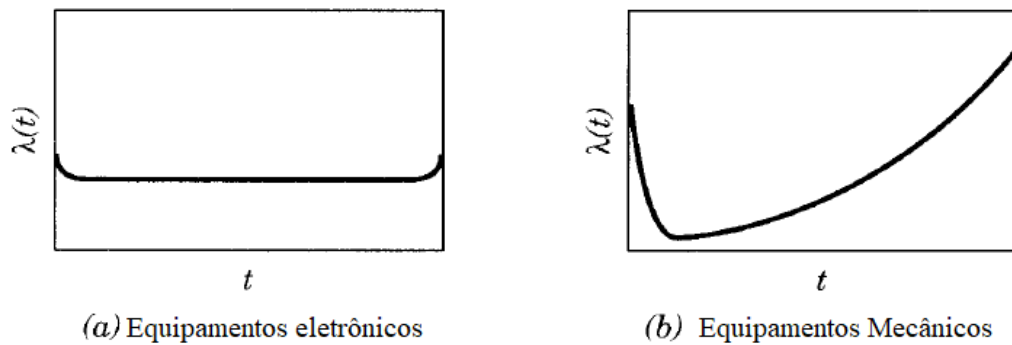
$$\lambda(t) = \lambda \quad (15)$$

$$R(t) = e^{-\lambda t} \quad (16)$$

$$f(t) = \lambda e^{-\lambda t} \quad (17)$$

$$MTTF = 1/\lambda \quad (18)$$

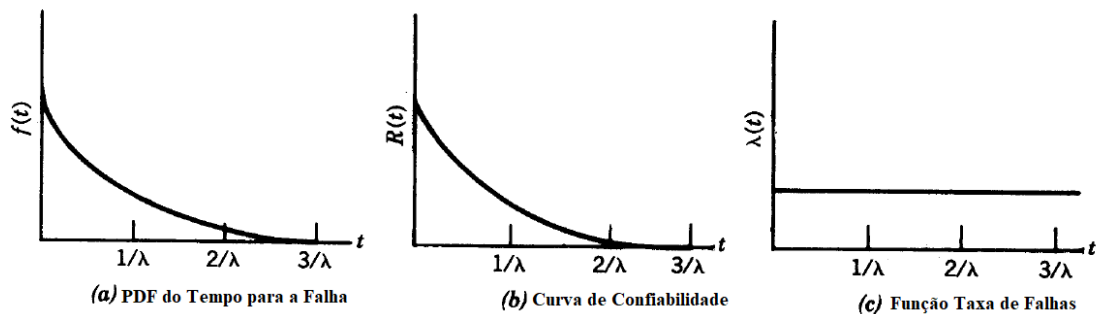
Figura 16 – Curva Banheira para Sistemas Eletrônicos e Mecânicos



Fonte: Lewis et al. (1994).

A Figura 17 traz a curva de confiabilidade, a função taxa de falhas e a função de probabilidade para a distribuição exponencial.

Figura 17 –Taxa de Falha λ com distribuição exponencial (constante), juntamente com a PDF e a Curva de Confiabilidade ($R(t)$)



Fonte: Lewis et al. (1994).

No entanto, em muitos casos a taxa de falhas de um equipamento ou sistema varia ao longo do tempo. Para sistemas nos quais essa variação é monotônica, a variável aleatória T pode ser

modelada por distribuições como a normal, a log-normal e a de Weibull. Quando, porém, a variação da função não é monotônica, apresentando um comportamento semelhante ao da Figura 16, faz-se necessário utilizar distribuições específicas – como a Weibull modificada ou a exponencial modificada (SARHAN; APALOO, 2013) –, ou particionar o intervalo para ajustes locais que permitam a representação não monotônica do conjunto. Nesse sentido, para cada caso, faz-se necessário analisar quais são as características do sistema a fim de escolher a melhor metodologia de ajuste que represente a distribuição de probabilidade com as propriedades mais adequadas para a modelagem da confiabilidade.

2.4.3 Modelagem da Confiabilidade de Sistemas

As principais metodologias para estudar a confiabilidade de um sistema são três: os testes de aceleração, as abordagens teóricas e o ajuste de curvas.

Os testes de aceleração são procedimentos de simulação acelerada das condições operacionais, usados para identificar o tempo para falha de um sistema ou equipamento quando não há registros suficientes de falhas de operação disponíveis. Esses testes podem ser realizados através de taxas de aceleração, nos quais o sistema opera em velocidade maior que a normal, ou através de processos de estresse, nos quais o sistema é exposto a variações bruscas de variáveis físicas, como temperatura e pressão (FARA *et al.*, 2017). Os principais objetivos são identificar: o tempo até a falha; as relações (matemáticas ou físico-químicas) entre o tempo para a falha e o estresse aplicado; e evidências sobre se a simulação de falha ocorrerá a nível operacional (COLLINS *et al.*, 2009). No entanto, é importante ter cuidado com os procedimentos empregados nos testes de aceleração, pois eles podem identificar falhas que não ocorreriam na prática e deixar de identificar aquelas que realmente acontecem. (SCHENKELBERG, 2015).

A confiabilidade de um sistema também pode ser analisada através de modelos teóricos qualitativos, como o modelo de árvore de falhas (em inglês, *Fault Tree Analysis* — FTA), a análise de modo de falhas e efeitos (em inglês, *Failure Modes and Effects Analysis* — FMEA) e o diagrama de blocos de confiabilidade (em inglês, *Reliability Block Diagram* — RBD). O FTA é um diagrama de eventos lógicos que relaciona eventos de falha com componentes ou subsistemas através de operadores lógicos (GARRO e BARRARA, 2011; STEMBER, 1981). O FMEA lista as consequências das falhas ou funcionamento de cada subsistema e suas possíveis causas (STEMBER, 1981). O RBD é usado para entender conceitualmente as conexões entre os subcomponentes do sistema. Quando há dados disponíveis sobre a ocorrência

de falhas em cada subsistema, o FMEA e o RBD podem ser usados para calcular a probabilidade total de falha do sistema em um determinado instante de tempo (LEWIS *et al.*, 1994; STEMBER, 1981).

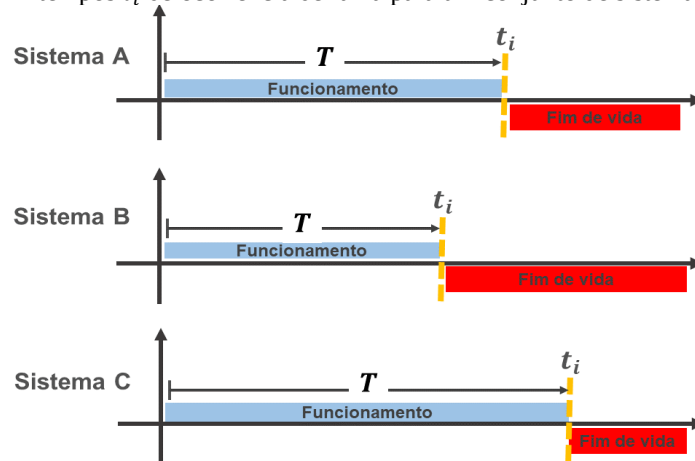
Outra abordagem teórica que merece destaque são as análises de Estados de Markov. Essa análise permite determinar a probabilidade de o sistema estar em um determinado estado em um intervalo de tempo específico, usando equações diferenciais de transição de estado (LEWIS *et al.*, 1994). As taxas de transição são compostas pelas taxas de falha dos componentes, taxas de reparo ou ambos. Essa abordagem pode ser usada em conjunto com outras técnicas, como a FTA, a FMEA e o RBD, para fornecer uma análise abrangente da confiabilidade do sistema.

Apesar da relevância dos testes de aceleração e das abordagens teóricas, a principal metodologia para análise de confiabilidade é a de ajuste de curvas, na qual $F(t)$ ou $R(t)$ é inferida a partir de dados reais de operação do sistema (KALBFLEISCH; PRENTICE, 2002). Para isso, são utilizados modelos paramétricos e não paramétricos (NELSON, 1982; O'CONNOR; KLEYNER, 2012). Nesse contexto, os conjuntos de dados são classificados em dois grupos: dados de vida útil e dados de falhas recorrentes (COLLINS *et al.*, 2009; KALBFLEISCH; PRENTICE, 2002; O'CONNOR; KLEYNER, 2012; RELIASOFT, 2015b).

No primeiro grupo, a análise foca no tempo para ocorrência da primeira falha destrutiva de determinado sistema ou equipamento, isto é, no período total de vida útil (NELSON, 1982). A análise de dados de vida é utilizada principalmente em sistemas nos quais as falhas resultam em fim total de operação, como satélites em órbita ou quedas de avião (CASTET e SALEH, 2009). Assim, são necessários registros de um conjunto significativo de sistemas ou equipamentos iguais, em operação, tendo em vista que cada um deles apresenta apenas uma falha de fim de sua vida útil. Do ponto de vista da variável aleatória T , o ajuste de curvas na Análise de Vida Útil é realizado a partir do conjunto de registro dos tempos t_i de ocorrência das falhas dos diferentes sistemas, conforme pode ser visto na Figura 18.

O segundo grupo analisa um conjunto de dados de falhas recorrentes, em outras palavras, analisa aqueles sistemas que, após a falha, voltam a funcionar, seja por uma particularidade sistêmica, seja por uma manutenção pontual (COOK e LAWLESS, 2007; NELSON, 2003). Um exemplo desse grupo são os carros com suas manutenções e o tema de estudo deste trabalho, os rastreadores solares. Nesse sentido, diferentemente da análise de dados de vida, na análise de falhas recorrentes, um mesmo sistema ou equipamento pode ter registros de falhas de tamanho significativo para gerar o ajuste das curvas.

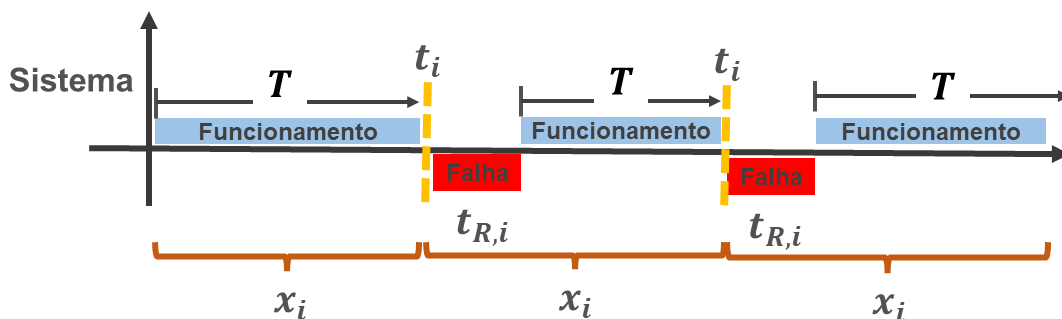
Figura 18 - Metodologia de Análise de Vida Útil para estudo da variável aleatória T a partir do conjunto de tempos t_i de ocorrência de falha para um conjunto de sistemas



Fonte: Elaboração Própria.

Do ponto de vista da variável aleatória T , o ajuste de curvas na Análise de Falhas Recorrentes é realizado a partir do conjunto de registro dos denominados ciclos de renovação, isto é, dos intervalos de tempo relacionados a falha e funcionamento do sistema, conforme esquematizado na Figura 19. Um modelo muito comum desta metodologia é o de Processo Ordinário de Renovação (*Ordinary Renewal Process*, em inglês), que considera o tempo de reparo $t_{R,i}$ desprezível e, portanto, analisa o conjunto de intervalos entre os registros de ocorrência de falhas x_i . Contudo, quando o tempo de reparo não é desprezível, abordagens que levem em conta o tempo de reparo devem ser utilizadas, uma delas por exemplo é o Processo de Renovação Alternada (*Alternating Renewal Process*, em inglês).

Figura 19- Metodologia de Análise de Vida Útil para estudo da variável aleatória T



Fonte: Elaboração Própria.

O tempo de ocorrência da falha são representados pelas variáveis t_i ; os períodos de reparo são representados por $t_{R,i}$; e x_i é o intervalo entre as ocorrências de falhas.

Compreender o conceito de reparo é essencial na Análise de Falhas Recorrentes. Uma das suas definições relaciona-se ao tempo que o sistema volta a operar normalmente após uma ocorrência de falha, podendo envolver intervenção humana ou não. Dessa forma, considera-se não apenas o processo em si, mas também o tempo necessário para que o sistema retorne ao funcionamento adequado.

3. REVISÃO DE LITERATURA

O setor elétrico teve uma importância crucial para avanços na teoria da confiabilidade. Ao longo do século XIX, as análises de confiabilidade para sistemas de potência utilizavam duas principais abordagens: simulações Monte Carlo e enumeração sucessiva de prováveis estados graves do sistema (BILLINTON e WENYUAN, 1991; EL-SAYED, 1996). Ambas aplicam conceitos derivados da teoria dos jogos e da teoria dos números randômicos. Atualmente, métodos determinísticos, probabilísticos e de aprendizagem de máquina foram integrados na estrutura de avaliação da confiabilidade dos sistemas de potências, com vistas a enfrentar os novos desafios do setor, principalmente relacionados a integração da rede e inserção de renováveis (BIGLARBEGLIAN *et al.*, 2018; FERNANDEZ SAVARI *et al.*, 2021; PEYGHAMI; PALENSKY; BLAABJERG, 2020; SINGH; JIRUTITIJAROEN; MITRA, 2019).

A seguir será apresentada a evolução da aplicação da Teoria de Confiabilidade no setor de energia solar, culminando com seu uso para o caso dos rastreadores solar.

3.1 Análise de Confiabilidade no Setor de Energia Solar

No setor elétrico, a partir da construção das primeiras usinas solares em meados do século XIX, viu-se a necessidade da aplicação da teoria da confiabilidade ao segmento. Em um primeiro momento, foi aplicada principalmente a sistemas heliotérmicos (CHOPRA; WOLOSEWICZ, 1980; EL-SAYED, 1986; WAITE; MAVEC; WOLOSEWICZ, 1982). Com o avanço e consolidação do mercado, a análise de confiabilidade também começou a ser feita para o setor fotovoltaico. Entretanto, no início, eram escassos os dados de falhas operacionais para validação dos modelos. Assim, de uma forma geral, eram aplicadas metodologias teóricas da teoria da confiabilidade, tais como o FTA, FMEA e RBD, para obter informações sobre os sistemas fotovoltaicos ou equipamentos fotovoltaicos (HAMDY; BESHIR; ELMASRY, 1989; STEMBER, 1981, 1982).

Dentro do setor fotovoltaico, os sistemas autônomos, isto é, aqueles não conectados à rede e que possuem um banco de baterias, intrinsicamente são os que mais necessitam de análises de confiabilidade, tendo em vista sua finalidade de ser a única fonte energética para abastecimento de determinada carga. Sistemas autônomos já levam em conta na fase de projeto índices de confiabilidade relacionados à sua capacidade de atendimento, como *loss of load*

probability (LOLP), *loss of load expectation* (LOLE), *loss of energy expectation* (LOEE), *loss of load health* (LOLH) e *expected energy not supplied* (ACUÑA *et al.*, 2018; CANALES *et al.*, 2021; GEORGILAKIS; KATSIGIANNIS, 2009). Raghuwanshi e Arya (2020) apresentam um método de decisão entre as possíveis configurações de sistemas híbridos autônomos a partir da avaliação dos índices de confiabilidade de atendimento de carga, utilizando a análise dos estados de Markov associada a técnicas de frequência-duração. No âmbito atual, a tendência é maximizar a confiabilidade dos sistemas autônomos a partir da utilização de algoritmos de otimização para desenho e gerenciamento, tais como *Teaching-Learning-Based Optimization* (TLBO), *Particle Swarm Optimization* (PSO), *Coyote Optimization Algorithm* (COA) e *Grey Wolf Optimizer* (GWO) (HADIDIAN MOGHADDAM *et al.*, 2019; JAHANNOOSH *et al.*, 2021; JANAMALA; SREENIVASULU REDDY, 2021). Por sua vez, utilizando a análise de vida útil, o trabalho de Carrasco *et al.* (2013) calculou a curva de confiabilidade e a função taxa de falhas a partir de registros de falhas ao longo de 5 anos de um conjunto de 7500 sistemas autônomos da zona rural do Marrocos.

Em relação aos sistemas fotovoltaicos conectados à rede, a revisão de literatura mostrou um amplo emprego da teoria da confiabilidade para análise tanto das usinas como dos equipamentos e subsistemas. De uma forma geral, podem ser encontrados trabalhos nos três tipos de metodologia apresentadas na Seção 2.4.3: testes acelerados; abordagens teóricas; e utilização dos registros de falhas para inferência de índices de confiabilidade.

No setor fotovoltaico, os testes acelerados são usados para avaliar a confiabilidade dos equipamentos, principalmente dos módulos fotovoltaicos, que têm uma vida útil de cerca de 30 anos. No caso dos módulos, esperar o fim de sua vida útil para avaliar a confiabilidade seria muito custoso, então os testes acelerados são usados para simular as condições operacionais em um período mais curto de tempo. Um trabalho de referência nessa área é o de Wohlgemuth *et al.* (2006), no qual foram aplicados testes de estresse utilizando a temperatura e a radiação em módulos fotovoltaicos, visando encontrar os mecanismos responsáveis pelas falhas e assim calcular sua vida útil. Mais recentemente, Kim *et al.* (2021) discutem os tipos de degradação dos módulos e sugerem testes acelerados para cada um deles. Gambogi *et al.* (2018) descrevem uma sequência de testes de aceleração para os módulos levando em consideração as características ambientais e comparando os resultados obtidos com os registros de falhas disponíveis. Owen-Bellini *et al.* (2021) introduzem um novo método denominado combinação de testes de estresse acelerado para descrever melhor determinadas falhas operacionais que vinham ocorrendo nos módulos FV e que não eram detectadas pelos testes acelerados anteriores.

Partindo para as abordagens teóricas, a literatura é vasta em relação a sua utilização para estudo da confiabilidade de sistemas fotovoltaicos. No âmbito dos processos markovianos, Cristaldi, Khalil e Faifer (2015) os empregaram para estimar a curva de confiabilidade e a probabilidade de falhas ao longo do tempo para módulos fotovoltaicos, utilizando como estados do sistema os fatores causadores de perda de potência como corrosão, ranhuras, entre outros. Já Shivarama e Kumar (2017) associaram lógica difusa aos Modelos de Markov justamente para estimar de forma mais assertiva a quantidade de energia gerada por sistemas fotovoltaicos. Por sua vez, Oprea *et al.* (2019) analisou uma planta fotovoltaica de 0,5 MW para desenvolver um modelo probabilístico baseado na análise de estados de Markov com a finalidade de calcular índices de confiabilidade, como tempo médio entre falhas, tempo médio para falha e tempo máximo das interrupções na produção energética. Maihulla e Yusuf (2021) desenvolveram um modelo teórico para sistemas fotovoltaicos conectados à rede utilizando uma técnica estatística de análise da dependência entre variáveis aleatórias denominada método Gumbel Hugaard Copula. A partir desse modelo, realizaram uma análise de confiabilidade, inferindo as curvas de confiabilidade e de disponibilidade. Sánchez-Barroso *et al.* (2021) utilizou os Estados de Markov para otimizar a periodicidade de limpeza dos módulos fotovoltaicos em regiões áridas da Europa.

Além da análise dos estados de Markov, constata-se o emprego de outras abordagens teóricas, como o diagrama de blocos de confiabilidade (RBD) e a análise de árvore de falhas (FTA). Bousshoua e Elmaouhab (2019) utilizou RBD para desenvolver uma avaliação da confiabilidade de sistemas híbridos com eólica e fotovoltaica na Argélia. Sonawane *et al.* (2023) utilizou uma FTA para avaliar a confiabilidade de sistemas fotovoltaicos de grande escala a fim de identificar as falhas mais críticas. De acordo com os resultados, as situações mais graves que afetam significativamente o desempenho dos sistemas fotovoltaicos estão relacionadas a: problemas na solda; células danificadas; quebra dos interconectores; problemas na estrutura do suporte; problemas no sistema de proteção contra descargas elétricas; delaminação; descoloração dos módulos; e sombreamento do sistema. Nesse mesmo sentido, Syamsuddin (2021) utilizou a abordagem do diagrama de blocos de confiabilidade aplicado a sistemas fotovoltaicos em Gili Trawangan, na Indonésia, para identificar os componentes mais críticos dos sistemas.

Ainda nesse contexto, alguns trabalhos modelam as falhas dos equipamentos a partir de variáveis físicas para a partir disso inferir os índices de confiabilidade. Yan; Liu e Kong (2021) analisaram a confiabilidade a partir de modelos de dispersão exponencial de degradação de potência dos módulos, determinando os parâmetros de seu modelo de degradação a partir de

um conjunto de dados de degradação de módulos. Morris e Fife (2009) aplicaram um modelo de falha de inversores a partir da temperatura de operação para análise da disponibilidade do equipamento. Depois de uma revisão de literatura sobre degradação da potência de módulos fotovoltaicos, Vázquez e Rey-Stolle (2008) propuseram que a degradação segue uma distribuição normal, utilizando-a para calcular parâmetros de confiabilidade como curva de confiabilidade, MTTF, taxa de falhas, entre outros.

De certa forma, o método de detecção de falhas em rastreadores com vistas à inferência dos índices probabilísticos de confiabilidade desenvolvido neste texto é semelhante ao apresentado pelos trabalhos do parágrafo anterior. Em ambos os casos, existe uma indisponibilidade de registros de falhas que permitam a inferência dos índices de confiabilidade a partir de dados operacionais. Então, utiliza-se um modelo de detecção da falha a partir de uma variável operacional disponível, como a temperatura ou irradiância. Nesse sentido, na metodologia desenvolvida neste trabalho, a variável operacional de entrada do modelo de detecção da falha é justamente a comparação entre os formatos das curvas de potência medida e simulada.

Por sua vez, quando se trata do uso de registros de falhas para análise de confiabilidade, há menos trabalhos disponíveis em comparação com as duas abordagens anteriores. Isso pode ser atribuído, em parte, à dificuldade de acessar esses registros e à sua indisponibilidade em muitos casos. Collins *et al.* (2009) ajustaram os parâmetros de um modelo de diagrama de blocos de confiabilidade para centrais fotovoltaicas utilizando 5 anos de registros de falhas de um sistema localizado em Springerville, Arizona (EUA). Fizanti, Musyafa' e Muhammad (2020) determinaram a taxa de falhas de uma usina de energia renovável híbrida localizada na Indonésia a partir das falhas registradas ao longo do tempo, a fim de fazer uma análise de confiabilidade, disponibilidade e de manutenção do sistema. Kuitche, Tamizh-Mani e Pan (2011) inspecionaram 46 módulos com câmera infravermelha, encontrando falhas como delaminação, pontos quentes, entre outros. De posse das frequências de ocorrência, eles as utilizaram em uma análise FMEA.

No entanto, existem trabalhos que analisam os registros de ocorrências de falhas de maneira quantitativa, não necessariamente uma análise de confiabilidade. Por exemplo, García *et al.* (2009) analisaram os registros de falhas de 6 plantas fotovoltaicas em Navarra na Espanha, dividindo-os por componentes. Lillo-Bravo *et al.* (2018) fizeram uma análise parecida, mas para 15 sistemas espalhados na Espanha e na Itália, contabilizando também falhas relacionadas a eventos externos como chuvas. Jordan *et al.* (2020) realizaram a análise dos registros de falhas de mais de 100 000 sistemas FV nos Estados Unidos, entre usinas residenciais, comerciais e de

grande escala, subdividindo-os por equipamento e contabilizando a perda de energia associada a cada um. Spertino e Corona (2013) também realizaram um estudo semelhante para 13 sistemas FV na Europa, visando o desenvolvimento de um conjunto de boas práticas de manutenção de sistemas FV.

3.2 Análise de Confiabilidade para Rastreadores Solares

Percebe-se que, no setor fotovoltaico, a teoria da confiabilidade já foi aplicada tanto a usinas como um todo, como a equipamentos específicos, como módulos fotovoltaicos e inversores. No entanto, como dito na Introdução, os rastreadores fotovoltaicos têm ocupado um papel de relevância no setor, sendo muito difícil encontrar atualmente usinas de grande porte sendo instaladas sem a tecnologia. Por consequência, as falhas de rastreamento também ocasionam perdas na produção de energia.

No relatório apresentado por Klise (2015), das falhas registradas no banco de dados da Sandia, 18% estavam relacionadas a rastreadores; número semelhante foi encontrado por García *et al.* (2009), em sistemas fotovoltaicos em Navarra (Espanha), que encontrou os sistemas de rastreamento como sendo responsáveis por 11% do total das perdas de geração. Jordan *et al.* (2020) na sua análise dos registros de falhas de mais de 100 000 sistemas FV nos Estados Unidos, entre usinas residenciais, comerciais e de grande escala, comenta que os registros de falhas no sistema de rastreamento são bastante subestimados e onde foram registrados representaram uma perda média na potência esperada para todo período de análise de 15,73%. Nesse sentido, alguns autores que desenvolvem análises de confiabilidade para usinas fotovoltaicas como um todo já vêm inserindo os rastreadores em seus modelos (GRANATA *et al.*, 2012; ROTAR *et al.*, 2021; SAYED *et al.*, 2019).

No entanto, é importante destacar que existem poucos registros disponíveis de falhas da operação dos rastreadores solares em usinas fotovoltaicas. Dessa forma, os autores que se debruçam sobre o tema da confiabilidade para o equipamento têm usado outras abordagens. Elerath, Spencer e Horne (2011) realizaram testes de aceleração em um sistema de rastreamento de eixo duplo para um sistema fotovoltaico de alta concentração (*high concentrator photovoltaics*, HCPV), com sensores realizando medições em pontos estratégicos. Os autores concluíram que variáveis como corrente do motor e vibração trazem informações relevantes sobre as falhas no rastreamento. Mais uma vez, Elerath (2011) realiza testes de aceleração em um rastreador HCPV, com medições em pontos estratégicos. Ele afirma que, fora a quebra total,

o desgaste mecânico (atrito, fadigas, deformações de superfícies entre outros) é o principal causador das imprecisões no seguimento do Sol pelos rastreadores, apresentando uma tabela (Tabela 3) com uma análise do modo de falhas e seus efeitos para o equipamento.

Tabela 3 – Análise Simplificada de Modo de Falhas e Efeitos para componentes de um rastreador de eixo duplo para sistema HCPV

Componente	Subcomponentes	Modos	Causa
Controle do Rastreador	- Transistores de efeito de campo de potência	- Falha de ligamento	- Alta temperatura ou corrente - Queda de tensão de corrente contínua durante a operação do motor
	- CPU	- Intermitência	
	- Memória	- Perda de dados	
Motores (azimute e inclinação)	- Rolamentos e engrenagens	- Ruído, vibração, travamento	- Atrito, brinelamento, fadiga - Altas temperatura e corrente
	- Enrolamentos	- Curto-circuito	
	- Escovas	- Arco elétrico	
Engrenagens (<i>slew drive</i> e <i>screw jack</i>) e cursor do atuador linear	- Rolamentos	- Ruído, vibração, travamento	- Atrito, brinelamento, fadiga - Desalinhamento, falta de lubrificante
	- Engrenagens	- Rachado, desgastado	
Rolamentos: elevação do painel e extremidade do <i>screw jack</i>	- Rolamentos	- Ruído, vibração, travamento	- Atrito, brinelamento, fadiga

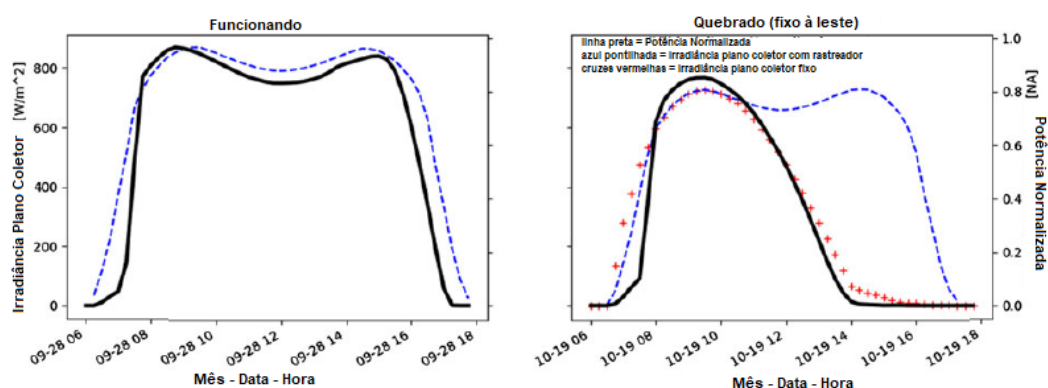
Fonte: Elerath (2011).

Por outro caminho, Benammar e Tee (2019) criaram um modelo que relaciona a confiabilidade de um heliostato (um rastreador solar de eixo duplo utilizado em sistema heliotérmicos) com a velocidade do vento, baseando-se no diagrama de tensões oriundas da força gravitacional e das forças resultantes da interação com o vento. Elerath (2017) aborda que a efetividade de um rastreador é medida pela disponibilidade ao invés da confiabilidade, tendo em vista que a disponibilidade também leva em conta o tempo para reparo e manutenção. Nesse contexto, ele analisa dois tipos de rastreadores de eixo único, um com fileiras independentes e outro multilinhas. A caixa de engrenagem, o motor e o controlador são os subsistemas comuns aos dois tipos. O autor utilizou simulações Monte Carlo para estimar o tempo médio de parada por ano de cada sistema, visando o gráfico de probabilidade em função do percentual de

disponibilidade. Por fim, chegou à conclusão de que rastreadores com fileiras independentes apresentam uma disponibilidade muito maior.

Por último, é importante abordar a metodologia desenvolvida por Ruth e Muller (2018) para identificação de falhas no rastreadores de eixo único, a partir dos dados de irradiância e de potência. O princípio básico por trás da metodologia é o seguinte: quando o rastreamento estiver funcionando corretamente ao longo do dia, as curvas de potência e de irradiância no plano de um coletor com seguimento apresentarão formas muito semelhantes. Por outro lado, se o rastreador passou grande parte do dia parado por um problema de desalinhamento, a curva de potência gerada possuirá maior similaridade com o gráfico de irradiância no plano de um coletor fixo. Isso pode ser visto na Figura 20.

Figura 20 – Curvas de potência gerada de um sistema fotovoltaico com rastreamento de eixo único (funcionando e quebrado) comparadas com curvas de irradiância no plano para coletores com rastreamento e afixados



Fonte: Ruth e Muller (2018).

Os gráficos mostram os dados de potência medidos para um rastreador de eixo único (linha sólida preta) para 2 dias e a irradiância no plano do coletor. No gráfico à esquerda, potência gerada e a irradiância no plano de um coletor com rastreamento de eixo único têm um formato semelhante (rastreador em funcionamento), mas, no gráfico à direita, os formatos diferem (rastreador preso ou quebrado). Alternativamente, a curva de irradiância para um coletor fixo com inclinação de 45° voltada para o leste (sinais de mais vermelhos) mostra uma boa correspondência com a curva de potência gerada, sugerindo a posição onde o rastreador ficou parado.

O trabalho de Ruth e Muller (2018) visa determinar a disponibilidade (razão entre dias funcionamento sobre dias totais de operação) dos sistemas de rastreamento empregados em usinas fotovoltaicas. Os autores analisam apenas o subconjunto de dias de céu claro, considerando como premissa que a proporção de dias de funcionamento nesse subconjunto de dias é a mesma que no conjunto total. Nesse sentido, a aplicação do método se dá apenas em dias de céu claro.

O passo a passo da metodologia de Ruth e Muller (2018) consiste em primeiro selecionar os dias de céu claro, para em seguida calcular a correlação de Pearson entre a curva de potência gerada e a curva de irradiância no plano de um coletor (*plane of array*, POA) com rastreamento. Caso essa correlação seja maior que determinado limite estipulado pelo pesquisador, indica-se que o rastreador funcionou durante aquele dia. Caso contrário, verifica-se se a correlação é alta para com a irradiância POA de um coletor fixo. Quer dizer, calculam-se vários gráficos de irradiância no plano de um coletor fixo variando o ângulo, comparando-os com a curva de potência medida e calculando-se a correlação entre elas. Se a máxima correlação das amostras de coletores fixos for maior que determinado limite estipulado pelo pesquisador, o rastreador falhou durante aquele dia. Não acontecendo nenhuma dessas duas situações, fica-se indeterminada a classificação. É importante ressaltar que a metodologia só classifica como falha caso ocorra uma alta correlação entre as curvas de potência e de irradiância no plano de um coletor fixo.

A correlação de Pearson é calculada de acordo com a Equação 19, em que as variáveis x e y representam as curvas que se deseja comparar. A correlação pode variar de -1 a 1, sendo 1 uma correlação positiva perfeita, -1 uma correlação negativa perfeita e 0 uma ausência de correlação.

$$C_{PEARSON} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (19)$$

Em que:

- $C_{PEARSON}$: correlação de Pearson;
- x_i : valor da variável x na i -ésima posição;
- $\bar{x}$: média do conjunto de valores da variável x ;
- y_i : valor da variável y na i -ésima posição;
- $\bar{y}$: média do conjunto de valores da variável y ;
- N : número total de elementos do conjunto das variáveis.

Na esteira do trabalho de Ruth e Muller (2018), Anderson *et al.* (2022) exploraram a aplicação da metodologia na identificação de falhas para além dos dias de céu claro. A novidade é justamente a utilização da potência simulada ao invés da irradiância no plano do coletor, inclusive inserindo no modelo de potência as perdas de geração fotovoltaica, como

sombreamento e ângulo de incidência. Os resultados se mostraram satisfatórios até determinados índices de céu claro (razão entre a irradiância global medida e a irradiância global estimada pelo modelo de céu claro). No entanto, não fica claro como os autores sabiam quais dias efetivamente os rastreadores falharam, seja por terem tido informações do sistema de controle dos rastreadores, seja por terem feito uma inspeção visual.

Alguns destaques devem ser feitos em relação aos trabalhos de Ruth e Muller (2018) e Anderson *et al.* (2022). O primeiro se dá em relação ao uso da correlação de Pearson como atributo de determinação. Como poderá ser visto na Seção 4.1, a correlação mostra uma deficiência em atestar falha para dias cujas curvas de potência medida e potência simulada para um coletor com rastreamento e coletor fixo apresentam as mesmas tendências. Em outras palavras, possíveis dias de falha poderão ser classificados como dias de funcionamento. Outro destaque é que a subjetividade da estipulação do limite de classificação dos dias impede a automatização do processo. Nesse sentido, a metodologia presente neste trabalho apresenta como contribuição inovadora a utilização de uma vasta gama de atributos de determinação além da correlação e a determinação de um limiar de classificação de forma automatizada a partir da maximização de um parâmetro de desempenho.

Em relação ao trabalho de Anderson *et al.* (2022), pode-se afirmar que sim é possível aplicar a metodologia a dias com baixo índice de céu claro. No entanto, isso é verdade até determinado ponto, visto que existe uma classe de dias para os quais é impossível se distinguir o funcionamento ou falha do rastreador. Nesse contexto, outra contribuição inovadora do presente trabalho é justamente a determinação dos dias para os quais o método de detecção de falhas não pode ser aplicado, os denominados dias indefinidos apresentados na Seção 4.3. Em outras palavras, é estabelecido o escopo de aplicação da metodologia. Por isso, é sempre importante se ter clara a definição adotada para falha, conforme apresentado no início da Seção 4.

Finalmente, conclui-se que, no âmbito do setor fotovoltaico, muito tem se estudado sobre a confiabilidade dos sistemas. Assim, com a utilização crescente dos rastreadores solares, a avaliação da confiabilidade deste equipamento é uma lacuna presente nos trabalhos científicos. No entanto, apesar dos artigos apresentados acima que se debruçam sobre o tema, percebe-se um problema semelhante ao encontrado pelos sistemas fotovoltaicos no início de sua história: a ausência de registros de falhas que permitam melhores análises de confiabilidade. Nesse sentido, o presente trabalho se propõe de maneira inovadora a inferir os índices probabilísticos de confiabilidade dos rastreadores solares fotovoltaico a partir da utilização dos

dias de falha operacional detectados por um método que se baseia na premissa utilizada no trabalho de Ruth e Muller (2018).

4. METODOLOGIA

Como dito anteriormente, existe uma indisponibilidade de registros de falhas de rastreadores solares em usinas fotovoltaicas, ou pelo menos pode ser dito que esses dados não são de fácil acesso. Nesse sentido, o processo proposto neste trabalho é dividido em duas etapas: num primeiro momento será desenvolvida uma metodologia para detecção de dias nos quais o sistema de rastreamento falhou; em seguida, esses dados serão utilizados para inferência dos índices probabilísticos de confiabilidade.

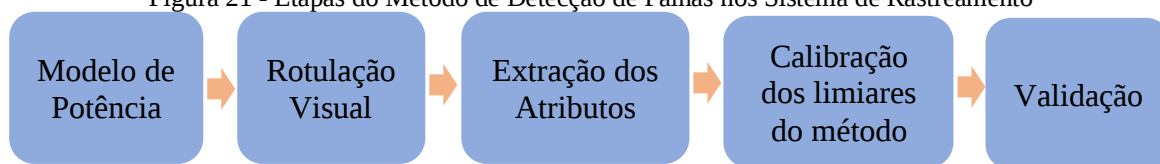
Segundo Bazovsky (2004), uma falha ocorre quando um equipamento ou serviço não atinge um desempenho adequada para sua finalidade durante determinado período de tempo, dentro de determinadas condições operacionais. Nesse contexto, como o principal objetivo de um rastreador solar é o aumento na produção de energia, a falha de um sistema de rastreamento será definida neste trabalho a partir da queda de potência por ela ocasionada. Portanto, apenas quando uma perda na produção de energia de um sistema FV puder ser diretamente relacionada a um desalinhamento no rastreador é que será considerado uma falha. Ainda de acordo com a definição acima, outro ponto a se definir é a escala de tempo. Mais uma vez, tendo em vista que a finalidade de um sistema de rastreamento é o aumento na produção energética, a escala diária foi a escolhida neste trabalho, pois maus funcionamentos no rastreador em intervalos de tempo curtos não tendem a gerar perdas significativas na geração de potência. Em resumo, serão consideradas falhas no sistema de rastreamento aquelas na qual a diminuição na geração diária de potência for decorrente do desalinhamento do rastreador.

O processo de detecção de dias de falhas está baseado na premissa adotada por Ruth e Muller (2018) de que a relação entre a curva de potência e a curva de irradiância no plano do coletor (*plane of array*, POA; o $I_{h,col}$ da Equação 6) traz implicitamente informações sobre se o seguimento do Sol ocorreu da forma esperada ou não. No entanto, ao invés da irradiância POA será utilizada a Potência Simulada, tanto para um coletor em posição fixa quanto para um coletor com rastreamento. A Potência Simulada é aquela obtida a partir de modelos que visam descrever a geração de potência de um sistema fotovoltaico.

As etapas do método de detecção de falhas no sistema de rastreamento estão resumidas na Figura 21. O primeiro passo é ajustar os parâmetros do modelo de potência à realidade do sistema fotovoltaico analisado. Esse ajuste será feito a partir de uma técnica de ajuste de curvas utilizando os dados de potência medida. Em seguida, para cada dia, serão analisados visualmente as curvas normalizadas de potência medida, potência simulada com o rastreador

funcionando e potência simulada com o coletor na posição fixa de máxima correlação. Com isso, será realizada uma rotulação visual para classificação de cada dia em relação ao funcionamento ou falha do rastreador, ou a indefinição quanto ao estado do equipamento. A rotulação visual servirá de referência para verificação da eficácia do método de classificação. O terceiro passo, por sua vez, consiste em extrair um conjunto de atributos das curvas de potência medida e de potências simuladas (rastreamento e fixo). De posse dos atributos, o período de análise será dividido em 2/3 para calibração e 1/3 para validação. Durante a calibração será verificado a eficácia de cada atributo em classificar corretamente os dias, sendo determinado o limiar de classificação que maximize o parâmetro de desempenho escolhido. No caso deste trabalho, o parâmetro selecionado foi o $f1$ (que será melhor explicado mais à frente). Por último, os limiares de classificação determinados durante o período de calibração serão utilizados no período de validação para analisar a desempenho de cada atributo. Nesse sentido, serão escolhidos os atributos que obtiverem um desempenho bom e consistente durante os períodos de calibração e validação.

Figura 21 - Etapas do Método de Detecção de Falhas nos Sistema de Rastreamento



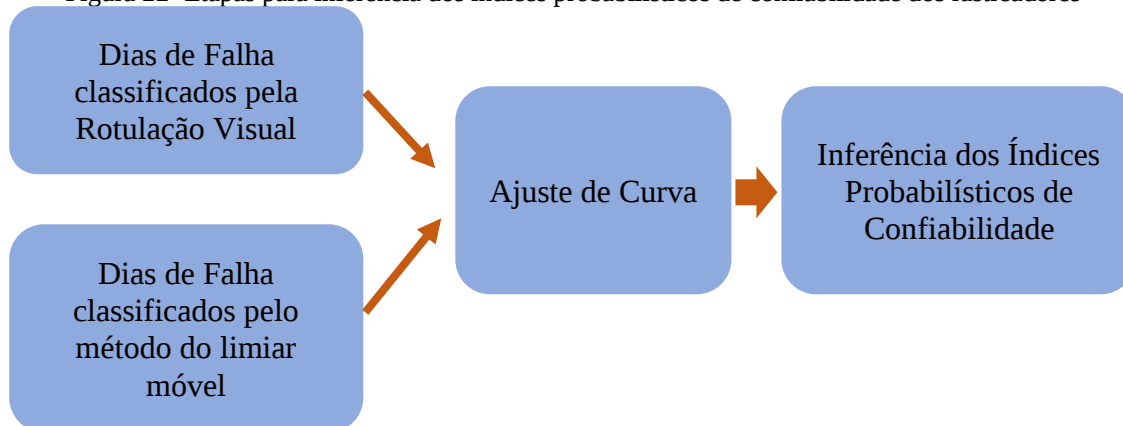
Fonte: Elaboração Própria.

Posteriormente, com as amostras de dias de falhas é realizada a inferência dos índices probabilísticos de confiabilidade a partir de um modelo de reparo geral. Como explicado na Seção 2.4.3, tal modelo faz parte da análise de falhas recorrentes, isto é, falhas que ocorrem mais de uma vez para um mesmo sistema ou equipamento. Isso é particularmente afim com os desalinhamentos ocorridos em rastreadores tendo em vista que desalinhamentos e funcionamentos podem ocorrer de forma recorrente.

As etapas para inferência dos índices probabilísticos de confiabilidade dos rastreadores estão esquematizadas na Figura 22. O primeiro passo é fazer o ajuste de curvas utilizando o modelo de reparo geral determinando a função de distribuição acumulada (CDF) $F(t)$ descrita na seção 2.4.1. Serão determinadas as CDFs a partir dos dias de falha classificados pela rotulação visual e também dos dias classificados pelo método do limiar móvel. Assim, poderá ser verificado se a CDF obtida com os indicativos de falha apontados pelo método de classificação diferem significativamente da referência classificatória que é a rotulação visual.

Feita essa primeira análise, os demais índices probabilísticos de confiabilidade serão inferidos: a curva de confiabilidade, a função taxa de falhas e tempo médio entre falhas.

Figura 22 -Etapas para inferência dos índices probabilísticos de confiabilidade dos rastreadores



Fonte: Elaboração Própria.

De uma forma geral, este trabalho apresenta uma metodologia de detecção de falhas no rastreamento com vistas à inferência dos índices probabilísticos de confiabilidade. Essa metodologia pode ser aplicada aos mais diversos contextos das usinas fotovoltaicas, desde que sejam realizadas as devidas adaptações. Em especial, a principal adaptação é referente ao modelo para potência simulada, pois deve levar em conta as especificidades da usina analisada, como o acesso a dados de potência de corrente contínua, o *backtracking* do sistema de rastreamento, sombreamento, entre outros.

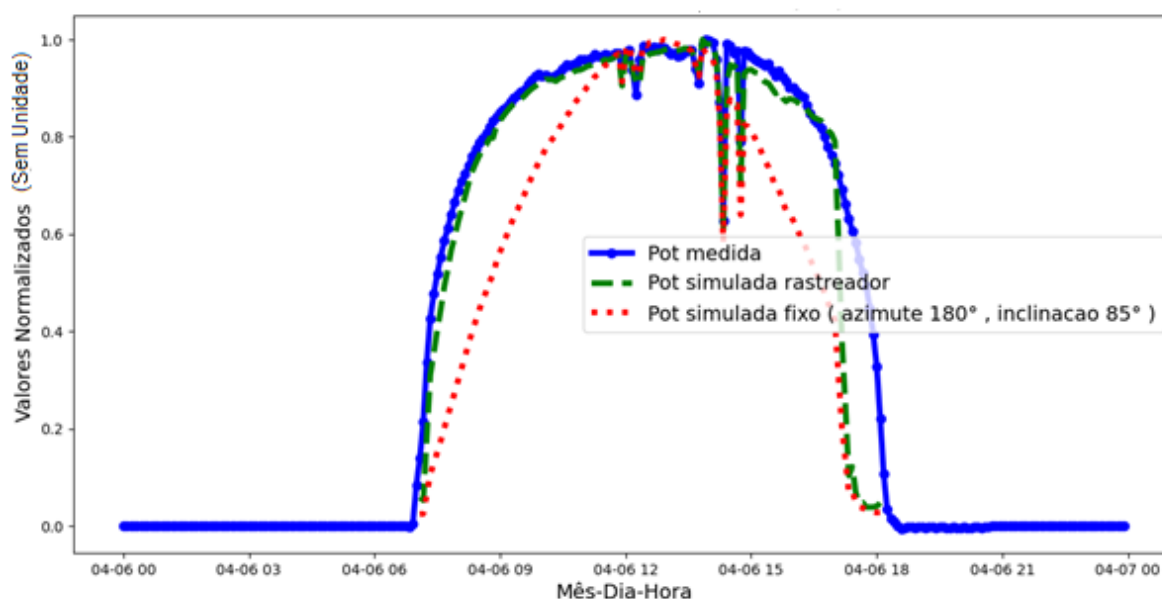
4.1 Método de Detecção de Falhas

O método de detecção de falhas baseia-se na premissa utilizada por Ruth e Muller (2018) de que, caso durante uma quantidade considerável do dia o rastreador fique parado, a curva de potência terá um formato similar à curva de irradiância no plano de um coletor fixo. Neste trabalho, porém, essa ideia será adaptada substituindo-se a irradiância pela potência simulada, tendo em vista que com isso será possível analisar sistemas FV compostos por mais de um rastreador, os quais podem estar desalinhados em posições diferentes entre si.

Como pode ser visto na Figura 23, quando o sistema de rastreamento está funcionando, a curva de potência do sistema fotovoltaico possui uma forma semelhante à curva de potência simulada para um sistema com rastreamento. Por outro lado, se o rastreador falhar passando a maior parte do dia travado, a curva de potência do sistema FV terá um formato semelhante a

uma curva de potência simulada para um coletor em posição fixa, como mostrado na Figura 24. A comparação entre as curvas de potência acontece com os valores normalizados de potência, sendo essa normalização realizada com respeito aos valores máximos de potência apresentados em cada caso (medida, simulada rastreador e simulada fixo).

Figura 23 – Curvas de potências medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no sítio 2 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 6 de abril de 2011.

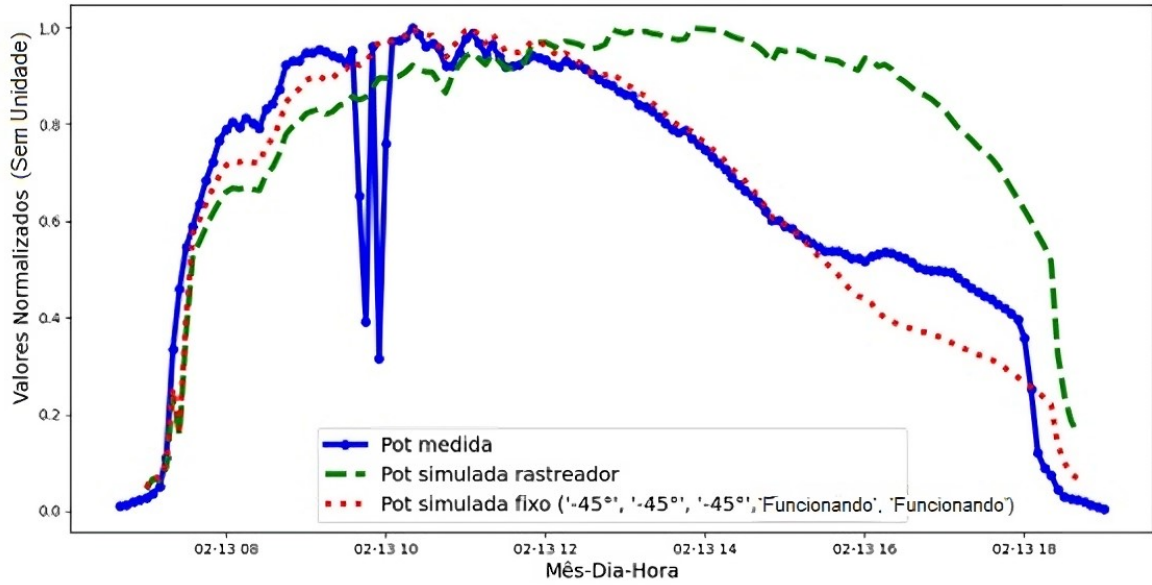


Fonte: Elaboração Própria.

A linha azul sólida representa a potência medida; a linha verde tracejada é a potência simulada para um coletor com rastreamento; a linha pontilhada vermelha a potência simulada de um coletor em posição fixa. Observe que quando o rastreador está funcionando, a curva de potência é semelhante à curva de potência simulada para um coletor com rastreamento.

Além das potências medida e simulada, um comparativo de curvas que também carrega informação sobre o funcionamento ou falha do rastreador é a comparação entre as denominadas Curvas de Erro (CE), cuja fórmula é apresentada na Equação 20. Nela, a irradiância no coletor com rastreamento é o referencial, atentando ao fato de que só é possível a realização da diferença entre potência e irradiância devido à normalização dos valores, pois essas são variáveis de natureza diferente. Seguindo a mesma lógica do comparativo das curvas de potência, quando o rastreador falhar passando a maior parte do dia em posição fixa, a CE da potência medida tem um formato similar à CE da potência simulada com o coletor em posição fixa; enquanto, se o rastreador estiver funcionando, a CE da potência medida é mais parecida com a CE da potência simulada de um coletor com rastreamento.

Figura 24 – Curvas de potências medida e simuladas para um inversor com 5 sistemas fotovoltaicos com rastreadores de eixo único, no sítio 5 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 13 de fevereiro de 2012.



Fonte: Elaboração Própria.

A linha azul sólida representa a potência medida; a linha verde tracejada é a potência simulada para o sistema com funcionamento dos rastreadores; a linha pontilhada vermelha o sistema com 3 rastreadores quebrados na posição de -45° , com os outros dois funcionando. Observe que quando o rastreador está quebrado, a curva de potência medida é semelhante à curva de potência simulada para uma configuração com os coletores em posição fixa.

$$CE_{k,i} = nPot_{k,i} - nI_{h,rast,k,i} \quad (20)$$

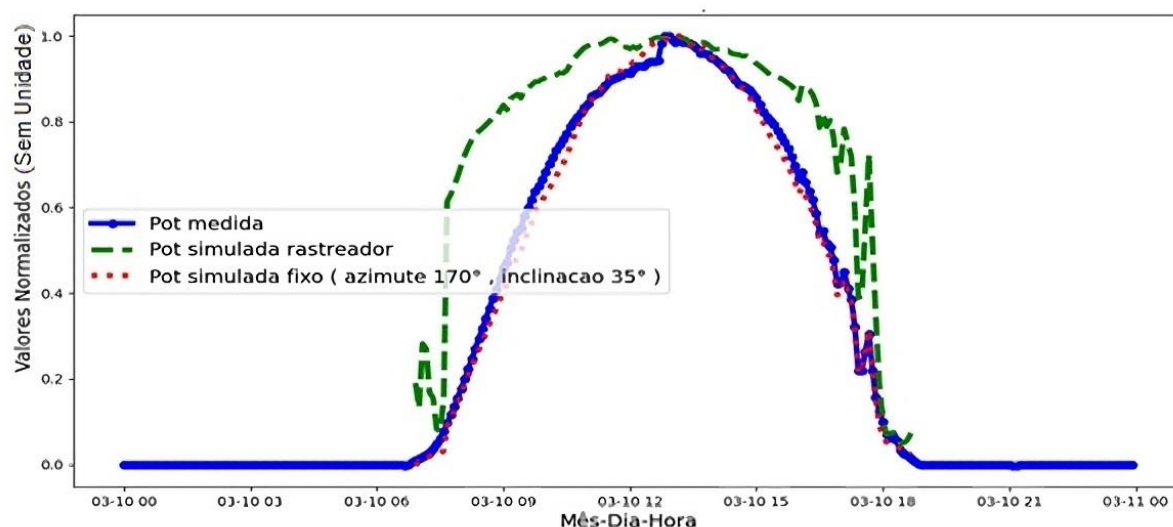
Em que:

- $CE_{k,i}$: valor da Curva de Erro na posição i para o dia k ;
- $nPot_{k,i}$: valor na posição i da Potência, normalizada pelo valor máximo da mesma para o dia k (essa potência pode ser tanto a potência medida como a simulada);
- $nI_{h,rast,k,i}$: valor na posição i da Irradiância no plano de um coletor com rastreamento, normalizada pelo valor máximo da mesma para o dia k .

A curva de erro se mostrou particularmente relevante em assegurar os dias de falha. Tendo em vista que, em dias como o da Figura 25, embora seja visualmente claro que o sistema FV esteve travado, a metodologia de classificação proposta mais à frente, ao analisar apenas as curvas de potência a partir da correlação entre elas, tende a fazer uma falsa afirmação concluindo que o rastreador funcionou. Para dias semelhantes a esse, existe em certo sentido

uma similaridade entre a potência medida e a potência simulada com rastreamento, visto que as tendências de crescimento e descida são iguais, mudando apenas as larguras das curvas. De forma especial, isso acontece bastante para rastreadores de eixo duplo. No entanto, com a introdução da comparação entre curvas de erro, esse equívoco é mitigado, conforme pode ser visto na Figura 26.

Figura 25 – Curvas de potências medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no local 2 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 10 de março de 2012.

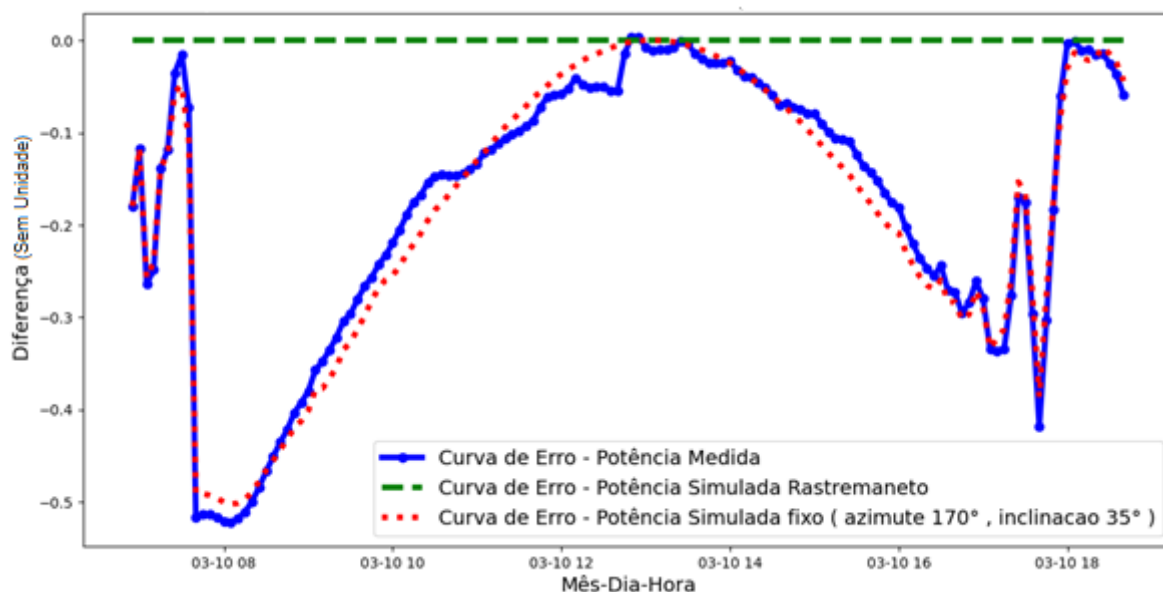


Fonte: Elaboração Própria.

A linha azul sólida representa a potência medida; a linha verde tracejada é a potência simulada para um coletor com rastreamento; a linha pontilhada vermelha a potência simulada de um coletor em posição fixa. Observe que é visualmente claro que o rastreador esteve parado ao longo deste dia. Entretanto, as tendências de crescimento e descida são as mesmas entre as três curvas, o que pode ocasionar um possível equívoco no método de classificação.

É importante se destacar que a comparação entre curvas se dá estritamente em relação aos seus formatos, já que a alteração na forma esperada está diretamente vinculada a desalinhamentos no rastreador. Os demais tipos de falhas, como sujeira, módulos quebrados, falhas no inversor, entre outros, geram diminuição na potência gerada. Nesse sentido, se a metodologia proposta levasse em conta as perdas nas magnitudes, seria bastante difícil vinculá-las a falhas no sistema de rastreamento. Assim sendo, a metodologia deste trabalho se propõe a analisar apenas as falhas relacionadas ao mau funcionamento do sistema de rastreamento a partir da análise de similaridade de curvas, por isso a comparação das curvas é sempre feita a partir de seus valores normalizados pelo máximo do dia analisado.

Figura 26 – Curvas de Erro (CE) das potências medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no local 2 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 10 de março de 2012



Fonte: Elaboração Própria.

A linha azul sólida representa a CE da potência medida; a linha verde tracejada é a CE da potência simulada para um coletor com rastreamento; a linha pontilhada vermelha a CE da potência simulada de um coletor em posição fixa. Observe que a comparação entre curvas de erro é um distintivo que permite mitigar equívocos na classificação dos dias de falha.

Por último, vale ressaltar também que a posição de desalinhamento indicada pela metodologia é a mais provável, podendo não estar totalmente correta. Entretanto, o indicativo de ocorrência de falha se mostra bastante assertiva.

Os próximos tópicos abordam: o modelo de potência adotado neste trabalho e utilizado no cálculo das potências simuladas; o procedimento de rotulação visual e de classificação dos dias; os atributos extraídos das curvas; e o método automático de classificação que utiliza esses atributos.

4.1.1 Modelo de Potência

É importante destacar que esta etapa é crucial na metodologia de detecção de falhas no sistema de rastreamento, pois é necessário que o modelo de potência adotado consiga simular bem o comportamento do sistema fotovoltaico analisado. No caso dos sistemas analisados neste trabalho, apenas os dados de potência de saída do inversor são fornecidos e, por isso, foi necessária a utilização dos modelos de potência de corrente alternada (CA). No entanto, se

informações de geração de corrente contínua estiverem disponíveis, a utilização dos modelos de potência CA não se faz necessária. Outro ponto, caso os sistemas FV analisados possuam rastreadores com *backtracking*, devem ser estudadas as alterações na cinemática dos rastreadores decorrentes do *backtracking* e, por consequência, como a irradiância no plano do coletor é alterada por conta disso. Da mesma forma, deve ser analisado como o sombreamento de objetos e até o sombreamento mútuo entre fileiras afeta a irradiância que chega no plano do coletor.

Nesse contexto, os sistemas analisados neste trabalho possuem rastreadores que não utilizam a técnica de *backtracking* e para os quais o sombreamento não interfere de forma significativa. Por isso, foram empregadas as equações apresentadas nesta seção.

De uma forma geral, a literatura utiliza para simular a potência gerada por um sistema fotovoltaico a Equação 21 (NARVARTE e LORENZO, 2008). Essa potência é chamada também de potência de entrada ou potência de corrente contínua (CC), pois é a potência produzida pelos módulos fotovoltaicos e que entrará no inversor para ser convertida em potência de corrente alternada. Aqui, a equação foi acrescida de um coeficiente denominado η_{FV} , que pretende representar o comportamento médio do conjunto de perdas relacionadas ao gerador fotovoltaico, como sujeira, perdas ôhmicas, perdas por descasamento, ângulo de inclinação modificado (IAM), entre outras.

$$P_{CC} = \eta_{FV} \times P_{nom} \times \frac{I_{h,col}}{I_{ref}} \times [1 - K_T(T_{AMB} + K \times I_{h,col} - T_{ref})] \quad (21)$$

Em que:

- P_{CC} : potência de corrente contínua (W);
- η_{FV} : coeficiente de representação do comportamento médio das perdas relacionadas ao gerador fotovoltaico, variando de 0 a 1;
- P_{nom} : potência fotovoltaica nominal do sistema (Wp);
- $I_{h,col}$: irradiância no plano do coletor (W/m^2);
- I_{ref} : irradiância de referência, $1000 W/m^2$;
- K_T : coeficiente térmico de perda de potência ($W/^\circ C$);
- K : coeficiente de conversão de radiação em temperatura, conhecido como K de Ross ($^\circ C m^2/W$);
- T_{AMB} : temperatura ambiente ($^\circ C$);

- T_{ref} : temperatura de referência dos módulos ($^{\circ}C$), $25^{\circ}C$.

Perceba que as principais variáveis de entrada do modelo são a temperatura ambiente e a irradiância POA. Portanto, se o modelo receber como entrada uma irradiância para um coletor com rastreamento, a potência modelada será a de um sistema FV com rastreador; se receber uma $I_{h,col}$ de um coletor fixo, o modelo será de um sistema FV fixo.

Por sua vez, ainda existe uma outra classe de modelos para simular a conversão da potência de corrente contínua em potência de corrente alternada (CA): os modelos de potência de inversor. Neste trabalho, foram testados os modelos de potência CA das Equações 22, 23 e 24, mais conhecidos como modelo linear, modelo quadrático Q3 e modelo PV Watt, respectivamente (ABD EL-AAL *et al.*, 2006; DOBOS, 2014). Os três modelos serão analisados e comparados com base em parâmetros estatísticos de avaliação dos resultados. O modelo que apresentar o melhor ajuste aos dados será escolhido como o mais adequado para simular a conversão da potência de corrente contínua em potência de corrente alternada.

$$P_{CA} = \eta \times \left[\left(\sum_{i=1}^N P_{CC,i} \right) - P_{MIN} \right] \quad (22)$$

$$P_{CA} = \begin{cases} 0, \sum_{i=1}^N P_{CC,i} < P_{MIN} \\ C_0 + C_1 \left(\sum_{i=1}^N P_{CC,i} \right) + C_2 \left(\sum_{i=1}^N P_{CC,i} \right)^2, \sum_{i=1}^N P_{CC,i} \geq P_{MIN} \end{cases} \quad (23)$$

$$P_{CA} = \left(\sum_{i=1}^N P_{CC,i} \right) \left[\frac{\eta_{NOMINAL}}{\eta_{REFERÊNCIA}} \left(-0.0162\epsilon - \frac{0.059}{\epsilon} + 0.9858 \right) \right] \quad (24)$$

Em que:

- P_{CA} : potência de saída de corrente alternada do inversor (W);
- η : eficiência da conversão CC para CA;
- $P_{CC,i}$: potência fotovoltaica gerada pelo i -ésimo sistema FV conectado ao inversor (W);
- N : número de sistemas FV ligados às entradas MPPT's dos inversores;

- P_{MIN} : potência mínima necessária para o funcionamento do inversor, esse valor não ultrapassa 1% da potência nominal do inversor (ABD EL-AAL *et al.*, 2006);
- C_0, C_1 e C_2 : coeficientes do modelo;
- $\eta_{NOMINAL}$: rendimento nominal do inversor, razão entre potências nominais de corrente alternada e de corrente contínua apresentadas no catálogo de informações do inversor;
- $\eta_{REFERÊNCIA}$: rendimento de referência do inversor, coeficiente do modelo;
- ϵ : razão que tem por numerador $[\eta_{NOMINAL} \times (\sum_{i=1}^N P_{CC,i})]$ e por denominador a potência nominal de corrente alternada apresentada no catálogo do inversor.

Pode ser percebido pelas equações que o coeficiente térmico de perda de potência (K_T), o coeficiente de conversão de radiação em temperatura (K), o coeficiente de representação do comportamento médio das perdas relacionadas ao gerador fotovoltaico (η_{FV}), a eficiência de conversão CC para CA (η), a potência mínima necessária para o funcionamento do inversor (P_{MIN}), os parâmetros C_0, C_1 e C_2 do modelo Q3 e o rendimento de referência do inversor ($\eta_{REFERÊNCIA}$) são coeficientes necessários aos modelos apresentado. Neste trabalho, optou-se por estimar esses coeficientes de forma estatística a partir de um de ajuste de curva utilizando os dados de potência medida dos geradores fotovoltaicos analisados. Um destaque sobre o ajuste do coeficiente relativo à Equação 24: para trabalhos futuros seria interessante analisar se ocorre um aumento na eficácia caso o ajuste fosse realizado em relação à razão $\eta_{NOMINAL}/\eta_{REFERÊNCIA}$ ao invés de $\eta_{REFERÊNCIA}$.

Os ajustes de curva acontecem quando se tem uma função paramétrica e se deseja ajustar seus coeficientes para melhor representar um conjunto de dados observacionais. Uma maneira muito comum de fazê-lo é minimizando o erro quadrático médio (somatório do quadrado da diferença entre o valor real e o estimado pela função). Um caso muito conhecido é justamente o ajuste linear, no qual uma função linear é ajustada aos dados. Inclusive nesse caso, a minimização do erro quadrático médio pode ser feita de forma analítica.

No entanto, para os denominados ajustes não lineares, nem sempre se é possível minimizar o erro quadrático médio analiticamente, sendo necessário a utilização de técnicas de otimização visando encontrar o mínimo global da função objetivo, no caso, a função do erro quadrático médio. As duas técnicas de otimização empregadas neste trabalho foram Levenberg-Marquardt (LM) e a *Subspace Trust-Region Interior-Reflective* (STIR) (BRANCH; COLEMAN e LI, 1999; MORE, 1977).

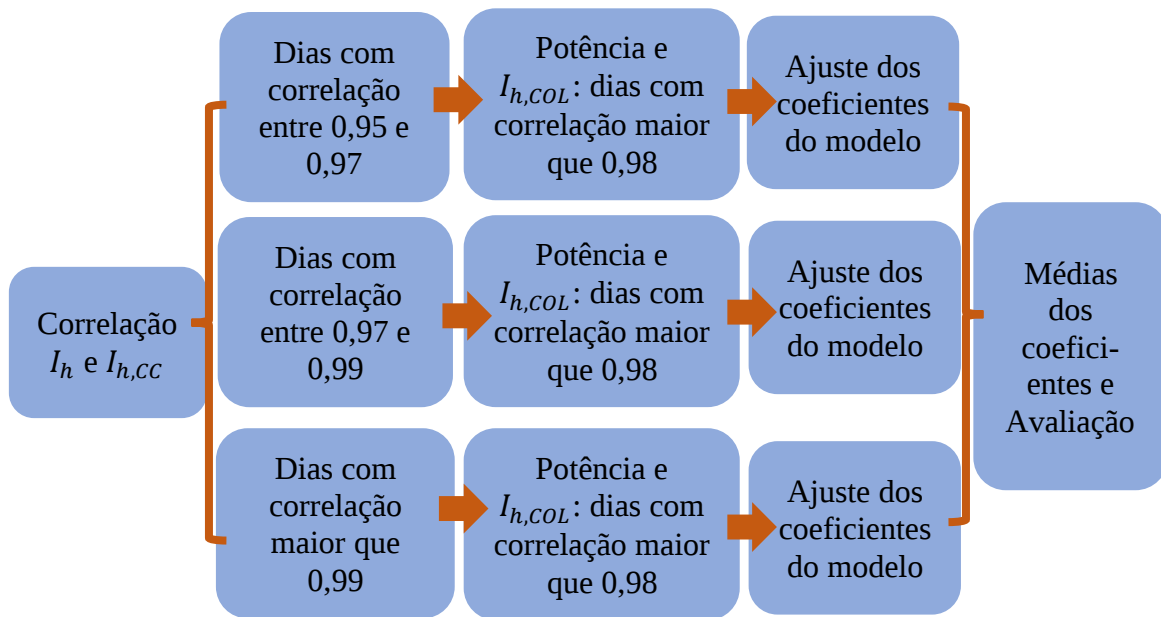
A explicação mais aprofundada dos procedimentos de cada uma delas foge do escopo deste trabalho. No entanto, de uma forma geral, o algoritmo STIR utiliza uma estratégia de aproximação em espaço reduzido para resolver problemas de otimização não linear com restrições de igualdade e desigualdade. A ideia é aproximar a solução do problema de otimização por uma solução em um subespaço de dimensão reduzida. O STIR é a técnica de otimização recomendada quando se possuem limites para os coeficientes das funções. Por sua vez, o LM é um algoritmo iterativo que combina o método do gradiente descendente e o método de Gauss-Newton, ajustando um parâmetro chamado parâmetro de Levenberg-Marquardt para equilibrar as propriedades de convergência dos dois métodos. O LM é conhecido por sua robustez e eficiência na resolução de problemas de mínimos quadrados não lineares.

O primeiro passo para a determinação dos coeficientes a partir de um ajuste não linear foi a seleção de dias. O procedimento está esquematizado na Figura 27. No processo de seleção, primeiramente são calculados os valores diários da correlação entre a irradiância global horizontal medida (I_h) e a de céu claro ($I_{h,CC}$), selecionando os dias que obtiveram correlação entre 0,95 e 0,97, 0,97 e 0,99 e maior que 0,99. Essa primeira etapa tem por finalidade selecionar dias que se aproximam minimamente do céu claro, e por isso que o menor valor da correlação foi escolhido como sendo 0,95. Por outro lado, dias para os quais a correlação entre I_h e $I_{h,CC}$ é maior que 0,99 apresentam um alto grau de claridade e se aproximam dos dias que são selecionados na metodologia de Ruth e Muller (2018). Nesse sentido, dias cuja correlação entre I_h e $I_{h,CC}$ está entre 0,97 e 0,99 representam uma classe intermediária.

A partir desses três grupos, são selecionados os dias com correlação maior que 0,98 entre a potência medida e a irradiância no plano do coletor ($I_{h,COL}$) com rastreamento. Vale salientar que a primeira etapa é referente a seleção tendo por base o modelo de céu claro; enquanto na segunda etapa, o embasamento é a comparação entre a potência medida e a $I_{h,COL}$ com rastreamento.

Em seguida, os dias resultantes de cada grupo são utilizados para ajuste dos coeficientes do modelo, gerando valores de coeficientes para cada um dos três grupos de dias. Ao final, é calculada a média dos coeficientes e avaliada a eficácia do modelo em simular a observação para o conjunto total dos dias selecionados. A correlação de Pearson apresentada na Equação 19 da Seção 3.2 é o principal parâmetro utilizado nas etapas do processo de seleção. A seguir serão melhor explanadas cada uma das etapas da seleção dos dias.

Figura 27 - Esquema do procedimento de seleção de dias utilizados para o ajuste dos coeficientes do modelo de potência



Fonte: Elaboração Própria.

A primeira seleção é referente aos dias de céu claro. Para isso, para cada dia foi calculada a correlação de Pearson entre a irradiância global horizontal e a irradiância global horizontal de céu claro obtida pelo ajuste dos coeficientes A e B da Equação 25. Os coeficientes A e B foram ajustados de maneira diária utilizando o algoritmo LM. A Equação 25 é conhecida como modelo de céu claro de Long e Ackerman (2000), sendo uma lei de potência cuja variável independente é o cosseno do ângulo zenital.

$$I_{h,cc} = A(\cos \theta_Z)^B \quad (25)$$

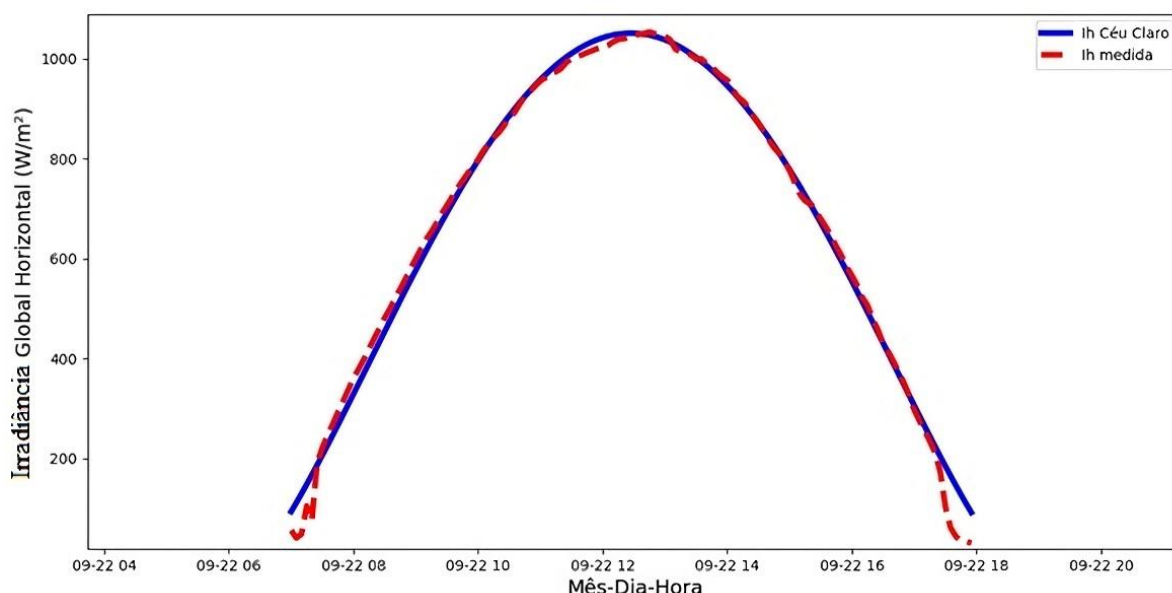
Em que:

- $I_{h,cc}$: irradiância global horizontal estimada pelo modelo de céu claro (W/m^2);
- A e B : coeficientes do modelo;
- θ_Z : ângulo zenital.

Assim, foram determinadas três grupos de dias a partir dos valores da correlação entre a irradiância global e a irradiância global de céu claro: aqueles com valores superiores a 0,99; com valores entre 0,97 e 0,99; e aqueles com valores entre 0,95 e 0,97. As Figuras 28 e 29 trazem exemplos de dias para as classes de correlação de céu claro maior que 0,99 e entre 0,95

e 0,97. É possível notar a ligeira diferença existente entre cada uma delas, sendo os dias com valores maiores que 0,99 os que mais se assemelham ao modelo de céu claro adotado.

Figura 28 – Irradiâncias Global Horizontal medida e de Céu Claro para um dia cuja correlação entre essas curvas é maior que 0,99



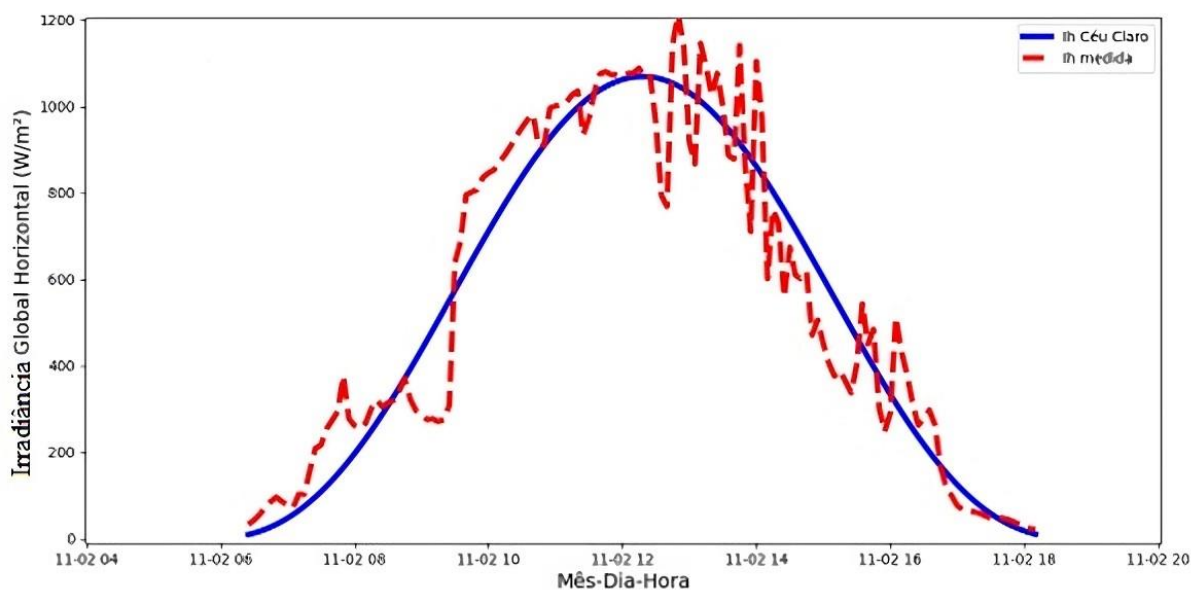
Fonte: Elaboração Própria.

Sistema fotovoltaico com rastreadores de eixo único, no local 5 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 22 de setembro de 2009. A linha azul sólida representa a irradiância global horizontal medida e a linha vermelha tracejada é a irradiância global de céu claro.

Em seguida, para cada um dos três grupos, foram selecionados os dias em que as curvas de potência medida e irradiância no plano do coletor com rastreamento apresentassem uma correlação maior que 0,98. Essa etapa foi realizada para garantir que os coeficientes determinados pelo ajuste fossem referentes a dias nos quais o sistema fotovoltaico esteve em funcionamento normal, com indicativos de funcionamento do rastreador. As Figuras 28 e 29 exemplificam isso utilizando os mesmos dias apresentados nas Figuras 30 e 31. A partir delas, é possível notar a semelhança entre os formatos da curva de potência medida e irradiância no plano de um coletor com rastreamento, um indicativo de funcionamento normal.

Para finalizar o processo de seleção, a partir de uma inspeção visual dos dias resultantes, foram eliminados aqueles cuja potência gerada média comparativamente teve um elevado decaimento, o que seria um indicativo de algum tipo de regulação na produção do sistema FV que viesaria o ajuste dos coeficientes. A Figura 32 traz um exemplo de dia no qual pode ser percebido uma diminuição da potência gerada em comparação com outro dia do mesmo grupo.

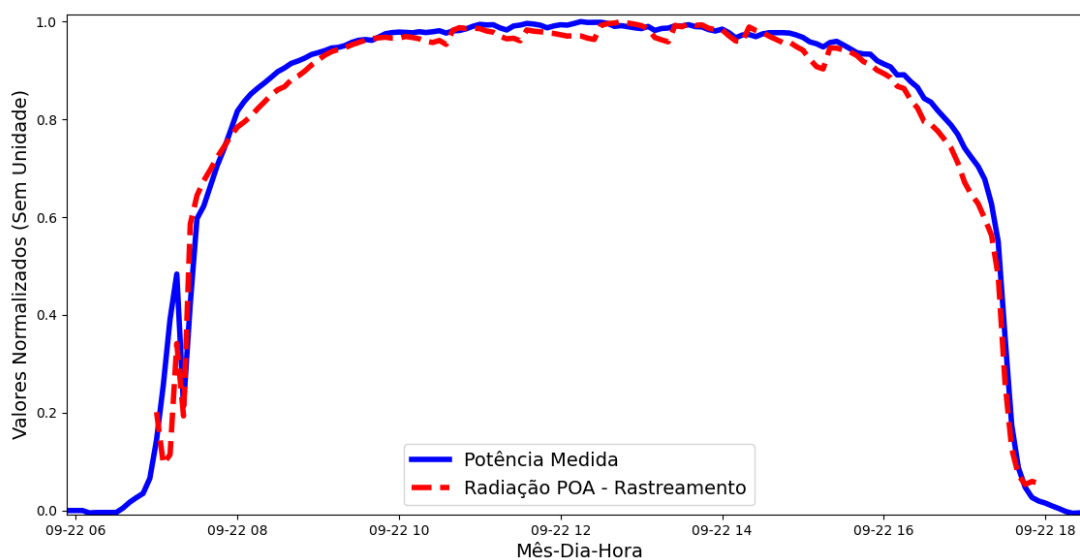
Figura 29 – Irradiâncias Global Horizontal medida e de Céu Claro para um dia cuja correlação entre essas curvas está entre 0,95 e 0,97



Fonte: Elaboração Própria.

Sistema fotovoltaico com rastreadores de eixo duplo, no local 2 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 02 de novembro de 2011. A linha azul sólida representa a irradiância global horizontal medida e a linha vermelha tracejada é a irradiância global de céu claro.

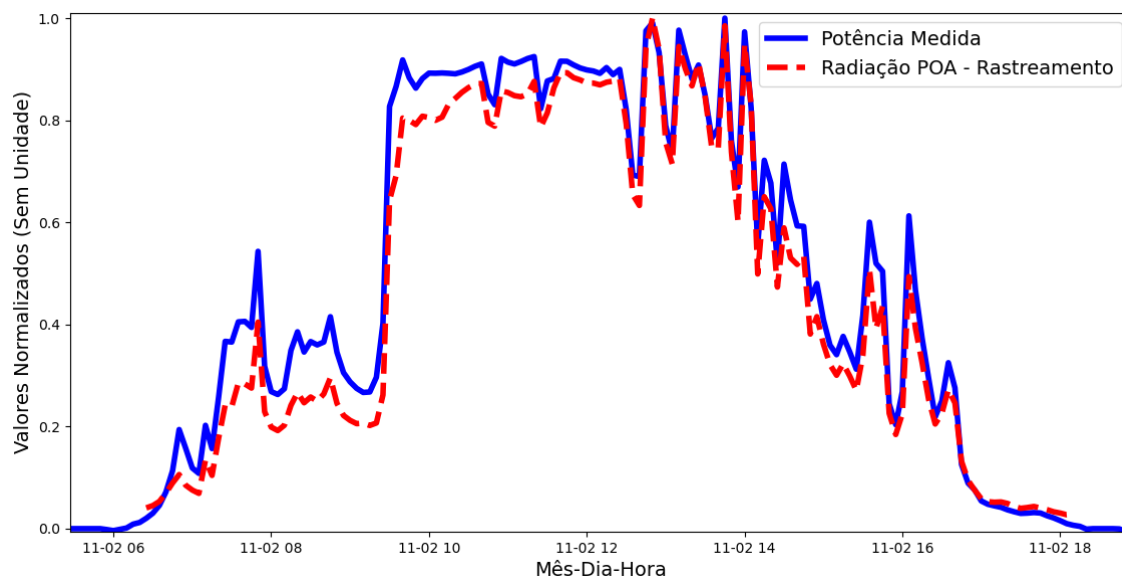
Figura 30 – Valores normalizados de Potência medida e de Radiação no plano do coletor com rastreamento para um dia cuja correlação entre essas curvas é maior que 0,98 e para o qual a correlação entre as radiações globais horizontal medida e de céu claro



Fonte: Elaboração Própria.

Sistema fotovoltaico com rastreadores de eixo único, no local 5 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 22 de setembro de 2009. A linha azul sólida representa a potência medida e a linha vermelha tracejada é a radiação global no plano do coletor com rastreamento.

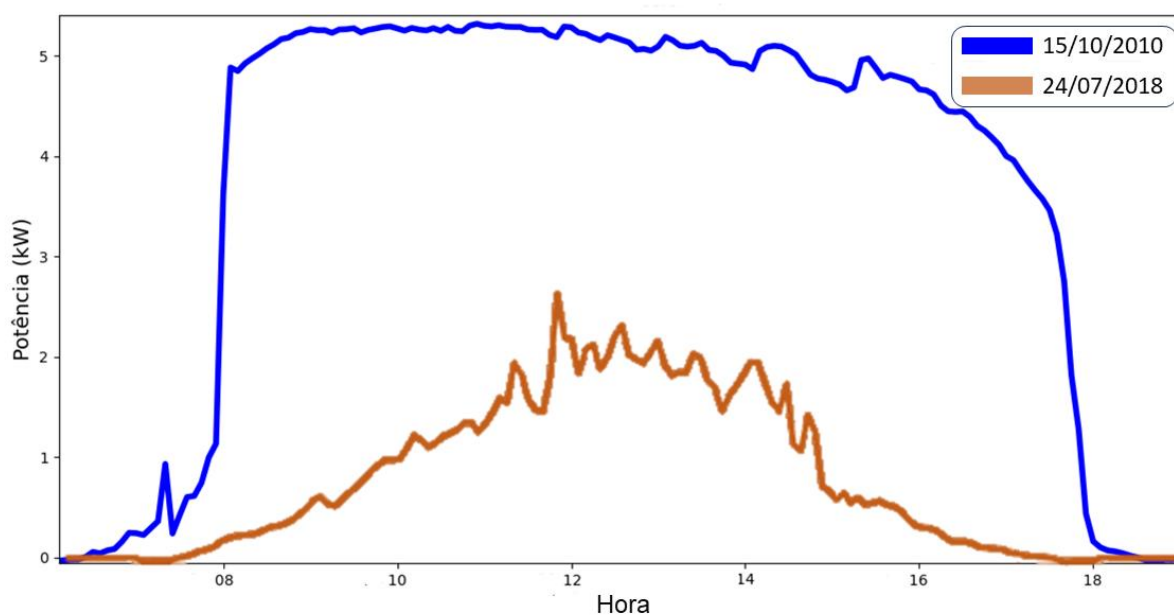
Figura 31 – Valores normalizados de Potência medida e de Radiação no plano do coletor com rastreamento para um dia cuja correlação entre essas curvas é maior que 0,98 e para o qual a correlação entre as radiações globais horizontal medida e de céu claro está entre 0,95 e 0,99



Fonte: Elaboração Própria.

Sistema fotovoltaico com rastreadores de eixo duplo, no local 2 do *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 02 de novembro de 2011. A linha azul sólida representa a potência medida e a linha vermelha tracejada é a radiação global no plano do coletor com rastreamento.

Figura 32 – Curvas de Potência medida para os dias 15/10/2010 e 24/07/2018, os quais a correlação entre as radiações globais horizontal medida e de céu claro está entre 0,97 e 0,99



Fonte: Elaboração Própria.

Por fim, para cada grupo, os dias selecionados foram utilizados para inferir os coeficientes dos modelos analisados através do algoritmo STIR. Restrições foram impostas aos valores aceitáveis de alguns coeficientes, com base em análises físicas do comportamento dos módulos e inversores em usinas fotovoltaicas (FUENTES *et al.*, 2007; NARVARTE; LORENZO, 2008). A Tabela 4 apresenta os limites estabelecidos para os coeficientes que foram restringidos durante o ajuste.

Tabela 4 - Limites impostos aos coeficientes durante o ajuste dos modelos de potência

Coeficiente	Limite Inferior	Limite Superior
η_{FV}	0.7	1
η	0.9	1
K_T	$-0.005 \text{ } ^\circ\text{C}^{-1}$	$-0.003 \text{ } ^\circ\text{C}^{-1}$
K	$0.02 \text{ } ^\circ\text{C}^{-1} \text{ m}^2/\text{W}$	$0.04 \text{ } ^\circ\text{C}^{-1} \text{ m}^2/\text{W}$
P_{MIN}	0	1% da Potência Nominal do Inversor

Fonte: Elaboração Própria.

Utilizando os dias selecionados de cada grupo, os coeficientes dos modelos são ajustados, resultando em 3 valores diferentes por coeficiente, um para cada grupo. O objetivo do procedimento é calcular a média desses três valores para utilizá-la nos modelos de simulação de potência do sistema fotovoltaico.

Com os coeficientes médios, a avaliação da eficácia de cada modelo é realizada verificando seu comportamento durante os dias selecionados. Essa avaliação é realizada a partir de parâmetros estatísticos de avaliação do resultado, que são a raiz do erro médio quadrático (em inglês, *root mean square* — RMSE), o erro médio (em inglês, *mean bias error* — MBE) e uma métrica de avaliação proposta no trabalho de Taylor (2001) denominada SS4.

As Equações 26 e 27 apresentam as fórmulas para o RMSE e o MBE, respectivamente. Ambas as equações necessitam dos dados observacionais (o) e dos dados gerados pelo modelo (m). Por sua vez, o SS4 é apresentado na Equação 30 (MIRANDA, 2023). O SS4 é uma métrica de avaliação que combina as informações da correlação de Pearson (Equação 19, na Seção 3.2) e da razão entre os desvios padrão dos dados observacionais e das estimativas do modelo (Equações 28 e 29). Ela é comumente usada para ranquear modelos, pois quanto mais próximo da unidade melhor será o mesmo.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (o_i - m_i)^2} \quad (26)$$

$$MBE = \bar{m} - \bar{o} \quad (27)$$

$$DP = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (28)$$

$$DP_{RAZÃO} = DP_{MODELO} / DP_{OBSERVAÇÃO} \quad (29)$$

$$SS4 = \frac{(1 + C_{PEARSON})^4}{4(DP_{RAZÃO} + 1/DP_{RAZÃO})^2} \quad (30)$$

Em que:

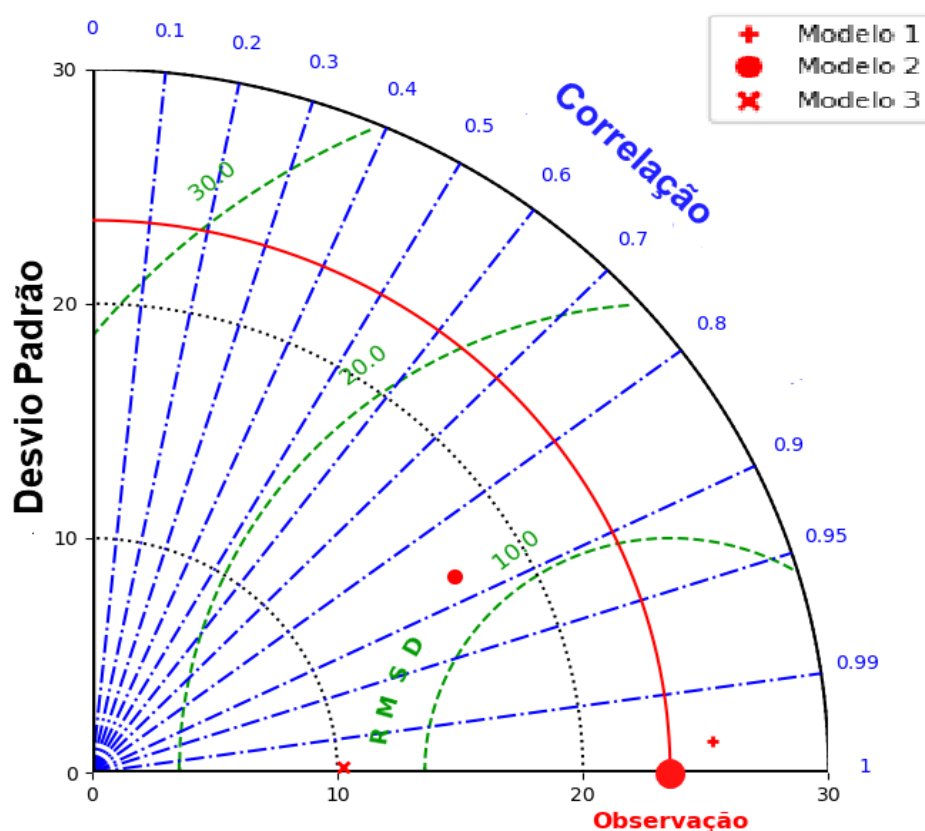
- $RMSE$, MBE , DP e $SS4$: respectivamente, raiz do erro médio quadrático, erro médio, desvio padrão e métrica de avaliação de Taylor;
- o_i e $\bar{o}$: i-ésimo valor dos dados observacional e média do conjunto;
- m_i e $\bar{m}$: i-ésimo valor da estimativa do modelo e média do conjunto;
- x_i e $\bar{x}$: i-ésimo valor da estimativa de uma variável analisada e média do conjunto;
- N : quantidade total de valores do conjunto;
- $C_{PEARSON}$: correlação de Pearson apresentada na Equação 19 da Seção 3.2.

Uma forma visual de fazer a análise do ajuste é a partir do Diagrama de Taylor (TAYLOR, 2001), que é uma representação que permite comparar modelos visualmente, a partir de três estatísticos: a correlação de Pearson, o desvio padrão e o RMSD, a raiz da diferença centrada quadrática — uma variação do RMSE, apresentada na Equação 31.

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N [(o_i - \bar{o}) - (m_i - \bar{m})]^2} \quad (31)$$

A Figura 33 apresenta um exemplo do Diagrama de Taylor contendo a comparação entre 3 modelos genéricos. No diagrama, a correlação é representada pela posição azimutal do ponto, isto é, o comprimento de arco partindo o eixo das abscissas é proporcional a magnitude da correlação dos dados estimados, quanto menor esse arco maior a correlação. Isso está representado pela cor azul da figura, permitindo a constatação de que o modelo 3 apresenta a maior correlação. Por sua vez, o desvio padrão é proporcional ao tamanho do raio das circunferências concêntricas pontilhadas em preto, de maneira que quanto maior o raio maior o desvio padrão. No exemplo da figura, o modelo 3, representado por um x, possui um desvio padrão muito próximo de 10, tendo em vista que está na vizinhança da circunferência cujo raio é proporcional a um desvio padrão de 10. Por último, a ideia por trás da representação do RMSD é igual à do desvio padrão, contudo as circunferências concêntricas neste caso possuem centro no valor da observação, o que está representado em verde. No exemplo da Figura 33, pode ser percebido, então, que o modelo 3 apresentou maior correlação com os dados observacionais, enquanto o modelo 1 possui um desvio padrão e um RMSD mais afim com a observação.

Figura 33 - Exemplo de Diagrama de Taylor



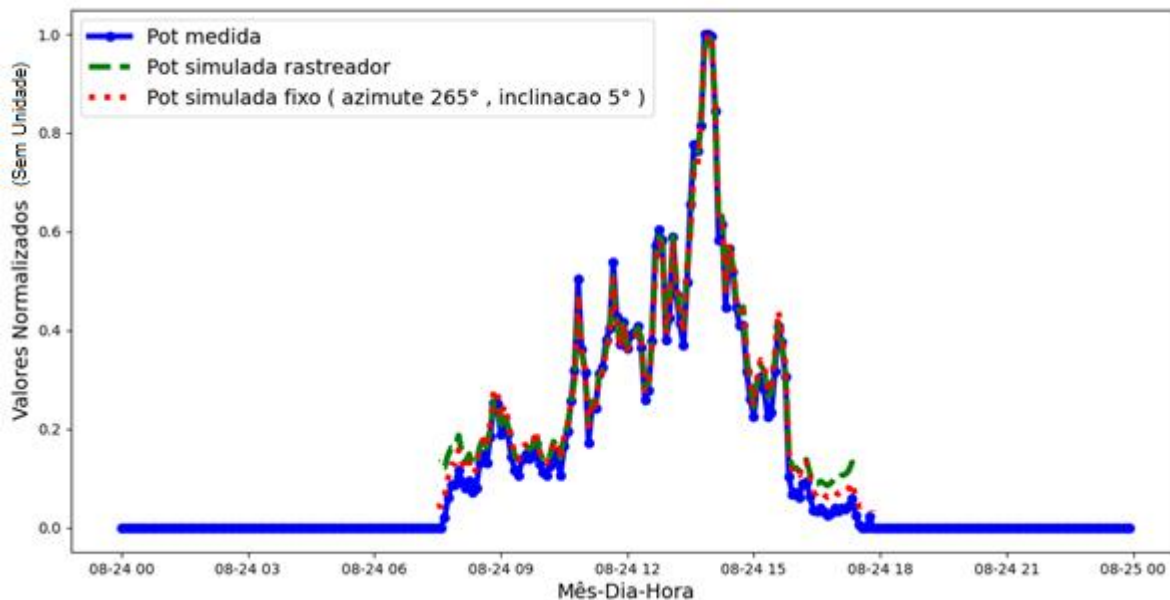
Fonte: Elaboração Própria.

4.1.2 Rotulação Visual e Classificação dos Dias

A fim de construir um conjunto de controle que permitisse a verificação da eficácia dos modelos de classificação, foram inspecionadas visualmente as curvas de potências medida e simulada para uma quantidade significativa de dias. Para cada dia, foram inseridos em um gráfico os valores normalizados da potência medida e das potências simuladas para o rastreador funcionando e para a configuração de coletor fixo com maior correlação.

A partir dessa inspeção visual, constatou-se a existência de três classes: os dias de funcionamento normal do rastreador, os dias de ocorrência de falhas e os dias indefinidos. Os dias de falha e funcionamento seguem a premissa apresentada no início da Seção 4.1. Já os dias indefinidos são aqueles nos quais visualmente é impossível diferenciar se o rastreador esteve funcionando ou falhou, pois as duas possibilidades são válidas e não excludentes, conforme pode ser visto na Figura 34.

Figura 34 – Curvas de potência medida e simuladas para um sistema fotovoltaico com um rastreador de eixo duplo, no sítio 2 no *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 24 de agosto de 2010



Fonte: Elaboração Própria.

A linha azul sólida representa a potência medida; a linha verde tracejada é a potência simulada para um coletor com rastreamento; a linha pontilhada vermelha a potência simulada de um coletor em posição fixa. Observe que este é um dia indefinido, pois a curva de potência possui um formato semelhante tanto à curva de potência simulada de um sistema com rastreador, como a à curva de potência simulada de um sistema fixo. Assim, não é possível afirmar se o rastreador esteve funcionando ou quebrado.

Os dias indefinidos representam a limitação do método de detecção de falhas, pois para essa classe não é possível saber o estado do sistema de rastreamento, se ele funcionou ou falhou. Esse ponto é muito relevante porque é uma contribuição inovadora do presente trabalho ao determinar para quais tipos de dia a metodologia de detecção de falhas nos rastreadores pode atuar, em outras palavras, determina o escopo de aplicação do método.

Por outro lado, neste trabalho, apenas dias com pelo menos 78 amostras de irradiância e potência entre o nascer e o pôr-do-sol (equivalente a 6,5 horas de dados com passos de 5 minutos) foram analisados. Dias com menos de 6,5 horas de medições foram considerados como tendo dados insuficientes.

Em relação à posição de *stow* (rever final da Seção 2.2), foi atribuído um valor limite de 16 *m/s* de operacionalidade ao rastreador de um eixo e de 12 *m/s* para dois eixos, ou seja, acima dessas velocidades o sistema será considerado na posição de *stow*. Para o rastreador de um eixo, a velocidade foi determinada com base em informações fornecidas por um relatório da Archtech Solar (uma das grandes empresas de rastreadores solares do mundo) que indica que 16 *m/s* é o limite geralmente empregado no mercado (GUO e MAGALHÃES, 2022). Já para o rastreador de dois eixos, foram utilizadas as informações do catálogo da empresa ADES, fabricante do sistema de rastreamento de dois eixos analisado (ADES, 2023). Portanto, dias em que mais de 39 amostras de velocidade do vento (equivalente a 3,25 horas de dados com passos de 5 minutos, metade do mínimo para potência e radiação) forem maiores que as velocidades de *stow* serão considerados dias em que o rastreador estará em posição fixa de *stow* e não serão analisados.

4.1.3 Extração de atributos

Um conjunto de atributos (ou estatísticos) foram extraídos das curvas com o objetivo de permitir a criação de um método automático de classificação dos dias. Os atributos utilizados podem ser divididos em 3 grupos: de similaridade; gerais; e de comparação.

A premissa de identificação de falha no rastreamento se baseia na similaridade que a curva de potência medida tem com a curva de potência simulada para um sistema com rastreamento ou para um sistema fixo. Essa mesma lógica é aplicada à curva de erro, conforme explicado anteriormente. Portanto, a primeira ideia que surge é justamente extrair atributos que meçam o grau de similaridade entre curvas.

Nesse sentido, os primeiros estatísticos de similaridade extraídos serão os mais usuais: a correlação de Pearson, o erro médio absoluto (MAE), o erro quadrático médio (MSE), a raiz

do erro quadrático médio (RMSE) e a área entre curvas (CHICCO; WARRENS; JURMAN, 2021). A correlação de Pearson e o RMSE já foram apresentados nas Equações 19 e 26; a área entre curvas é a diferença entre as áreas da primeira e segunda curva; e o MSE e o MAE são apresentados a seguir nas Equações 32 e 33, respectivamente.

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad (32)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - y_i| \quad (33)$$

Em que:

- MSE e MAE : respectivamente, erro médio quadrático e erro médio absoluto;
- x_i : i-ésimo valor dos dados da primeira curva;
- y_i : i-ésimo valor dos dados da segunda curva;
- N : quantidade total de valores do conjunto.

Em relação à similaridade das cruvas, também foram extraídos outros atributos de similaridade descritos na literatura:

- A Distância Hausdorff, que é a o valor máximo presente no conjunto de distâncias entre os pontos da primeira curva em relação à segunda (BUCHIN, 2007);
- A Distância Fréchet, a qual é semelhante à distância Hausdorff, mas que aplica técnicas de parametrização para calcular a máxima distância existente levando em consideração o formato das curvas e a localização de seus pontos (BUCHIN, 2007);
- O Mapeamento Parcial de Curvas (em inglês, *Partial Curve Mapping* – PCM), uma combinação entre os comprimentos de arco e as áreas das curvas (WITOWSKI; STANDER, 2012);
- O Comprimento de Curva, o qual aplica determinadas ponderações sobre funções objetivas (CAO; LIN, 2008); e
- A Distorção Dinâmica do Domínio Temporal (em inglês, *Dynamic Time Warping* – DTW), que é uma técnica específica que mensura a distância entre

as curvas a partir de alinhamentos específicos entre ambas (PETITJEAN; KETTERLIN; GANÇARSKI, 2011).

Nesse contexto, todos os atributos de similaridade serão extraídos fazendo-se a comparação entre as curvas apresentadas na Tabela 5. É importante destacar que a comparação diária entre potências é sempre feita normalizando-se os valores pelo valor máximo apresentado naquele dia, tendo em vista que se busca comparar o formato e não a magnitude. Ressalte-se também que a inserção da comparação entre irradiância global e irradiância global de céu claro foi inserida por existir a possibilidade de apresentar informações relevantes na classificação dos dias indefinidos.

Tabela 5 - Curvas comparadas para extração de Atributos de Similaridade

Curva 1	Curva 2
Potência Medida	Potência Simulada (rastreamento e coletor fixo)
Potência Simulada com Rastreamento	Potência Simulada, coletor fixo
Curva de Erro da Potência Medida	Curva de Erro da Potência Simulada (rastreamento e coletor fixo)
Curva de Erro da Potência Simulada com rastreamento	Curva de Erro da Potência Simulada com coletor fixo
Radiação Global	Radiação Global de Céu Claro

Fonte: Elaboração Própria.

Por sua vez, os atributos de caráter geral são formados por uma variedade de tipos, tais como média, mediana, desvio padrão, entre outros, extraídos justamente para verificar se outras informações além da similaridade são relevantes para a classificação dos dias. Para a extração desses estatísticos foi utilizada a biblioteca python Tsfresh, a qual conta com um conjunto de mais de 1000 tipos de atributos (CHRIST *et al.*, 2018). Para extração, foram utilizadas as curvas de potência medida e de potência simulada (ambas normalizadas), e as curvas de irradiância global e irradiância global de céu claro.

Os atributos extraídos pela biblioteca TSFRESH consistem em medidas estatísticas, transformações e características que são amplamente utilizadas na análise de séries temporais. Esses atributos podem ser categorizados em diferentes grupos, sendo os principais mencionados a seguir.

Os atributos básicos englobam medidas simples e comuns, como soma, média, desvio padrão, mínimo, máximo, mediana, curtose, assimetria e entropia. Essas medidas fornecem informações essenciais sobre a distribuição e o comportamento da série temporal. Os atributos

relacionados à autocorrelação são utilizados para medir a semelhança entre uma série temporal e suas versões defasadas em alguns passos. Eles incluem o coeficiente de autocorrelação, valor-p, intercepto, inclinação e erro padrão de uma regressão linear entre a série e sua versão defasada. Esses atributos auxiliam na identificação de padrões de dependência temporal na série.

A transformada de Fourier é utilizada para decompor a série em componentes harmônicas, e os atributos correspondentes incluem coeficientes da transformada, frequências, amplitudes e fases das componentes. Esses atributos permitem detectar padrões de frequência presentes na série temporal. A análise *wavelet* também desempenha um papel importante, fornecendo informações sobre diferentes escalas e frequências presentes na série. Os atributos *wavelet* incluem coeficientes da transformada, energias e entropias das sub-bandas resultantes. Esses atributos são úteis para identificar mudanças de padrão em diferentes escalas de tempo.

Os atributos de ordem exploram as propriedades dos valores ordenados da série, como quantis, percentis, valores máximos e mínimos locais, comprimentos de sequências acima ou abaixo da média e porcentagens de valores repetidos. Esses atributos fornecem informações sobre a distribuição e o comportamento da série em relação a diferentes pontos de referência. Por fim, os atributos de complexidade medem a irregularidade, aleatoriedade ou imprevisibilidade da série temporal. Eles incluem a entropia aproximada, entropia de amostra, correlação de Benford, número de máximos locais e expoente de Lyapunov. Esses atributos ajudam a caracterizar a complexidade e a natureza dinâmica da série.

Assim, os atributos gerais proporcionam uma ampla gama de informações sobre as curvas analisadas, trazendo informações complementares para o método de classificação dos dias. A Tabela 6 apresenta um recorte com alguns estatísticos extraídos pela biblioteca TSFRESH e seus rótulos.

Por fim, o conjunto de atributos denominados de comparação são aqueles que comparam os valores de estatísticos conceitualmente opostos. Para entender melhor essa ideia, use-se a correlação como exemplo. Ao extrair a correlação entre as curvas normalizadas de potência medida e potência simulada com rastreamento, obtêm-se um valor que traz em si uma informação relacionada ao funcionamento do rastreador, ou seja, é um atributo que mede se o sistema de rastreamento funcionou. Em contrapartida, a correlação entre as curvas normalizadas de potência medida e potência simulada para um coletor fixo armazena informações relacionadas à falha. Em outras palavras, apesar de serem do mesmo tipo (correlação de Pearson), esses dois estatísticos trazem como finalidade principal informações opostas. É claro que com a correlação de funcionamento o modelo também pode identificar falhas, assim como

com a correlação de falha o sistema pode identificar funcionamentos. Porém, o ponto é que a principal finalidade de um é apontar funcionamento, e do outro, a falha; e, nesse sentido, são opostos.

Tabela 6 – Exemplo de rótulos na biblioteca python TSFRESH e atributos correspondentes

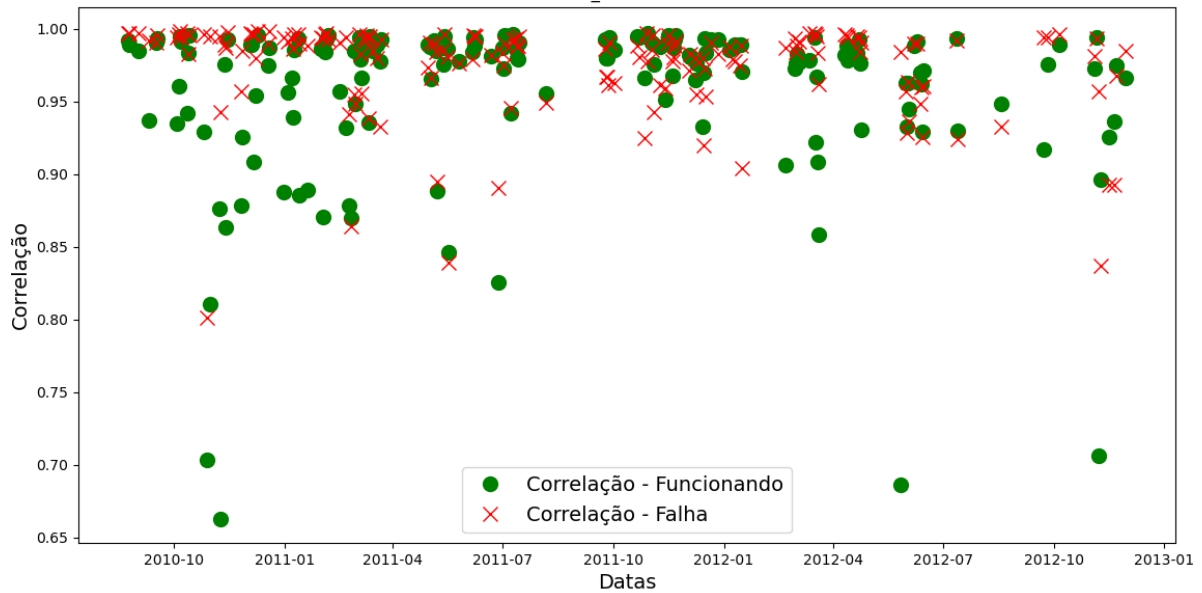
Rótulo na biblioteca python TSFRESH	Atributo
sum_values	Soma dos valores da série temporal
mean_abs_change	Média das mudanças absolutas entre dois valores consecutivos da série temporal
skewness	Medida da assimetria da distribuição dos valores da curva
autocorrelation_5	Coefficiente de autocorrelação da série temporal com uma defasagem de 5 passos
fft_coefficient_3	Terceiro coeficiente da transformada de Fourier da série temporal
wavelet_entropy_2	Entropia da segunda sub-banda da transformada wavelet da série temporal
quantile_0.9	Valor que separa os 90% menores valores dos 10% maiores valores da série temporal
sample_entropy	Medida da complexidade ou regularidade da série temporal baseada na frequência dos padrões repetidos

Fonte: Christ et al. (2023).

O desenvolvimento dos atributos de comparação surgiu da observação de que nos dias classificados como indefinidos, os estatísticos opostos apresentam magnitudes semelhantes. Isso pode ser visualizado nas Figuras 35 e 36, que mostram os valores das correlações opostas e dos RMSEs opostos, respectivamente, para dias rotulados visualmente como indefinidos. De uma forma geral, é possível notar que os valores são muito próximos, chegando até mesmo a se sobrepor visualmente em alguns casos.

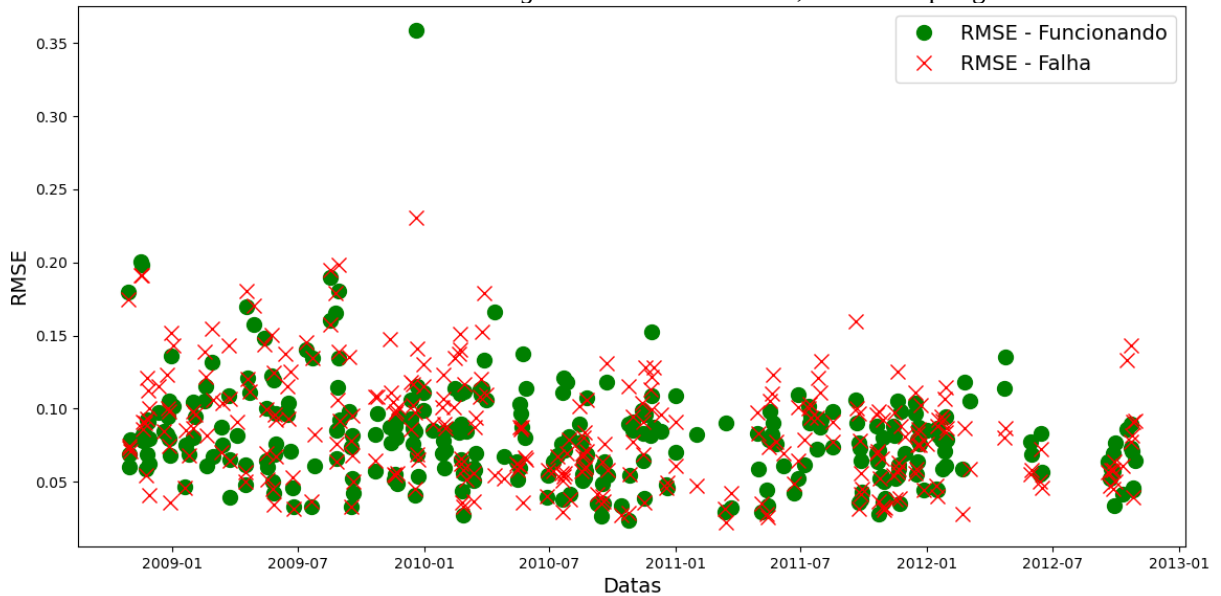
Com base nessa semelhança, foram calculadas três medidas para avaliar a proximidade dos valores de atributos opostos: diferença absoluta, média e razão. Essas medidas fornecem informações sobre a magnitude da diferença entre os valores opostos.

Figura 35 – Correlações de Funcionamento e de Falha para dias rotulados visualmente como dias indefinidos, no local 2 do *Desert Knowledge Australia Solar Centre*, em Alice Springs



Fonte: Elaboração Própria.

Figura 36 – Correlações de Funcionamento e de Falha para dias rotulados visualmente como dias indefinidos, no local 5 do *Desert Knowledge Australia Solar Centre*, em Alice Springs

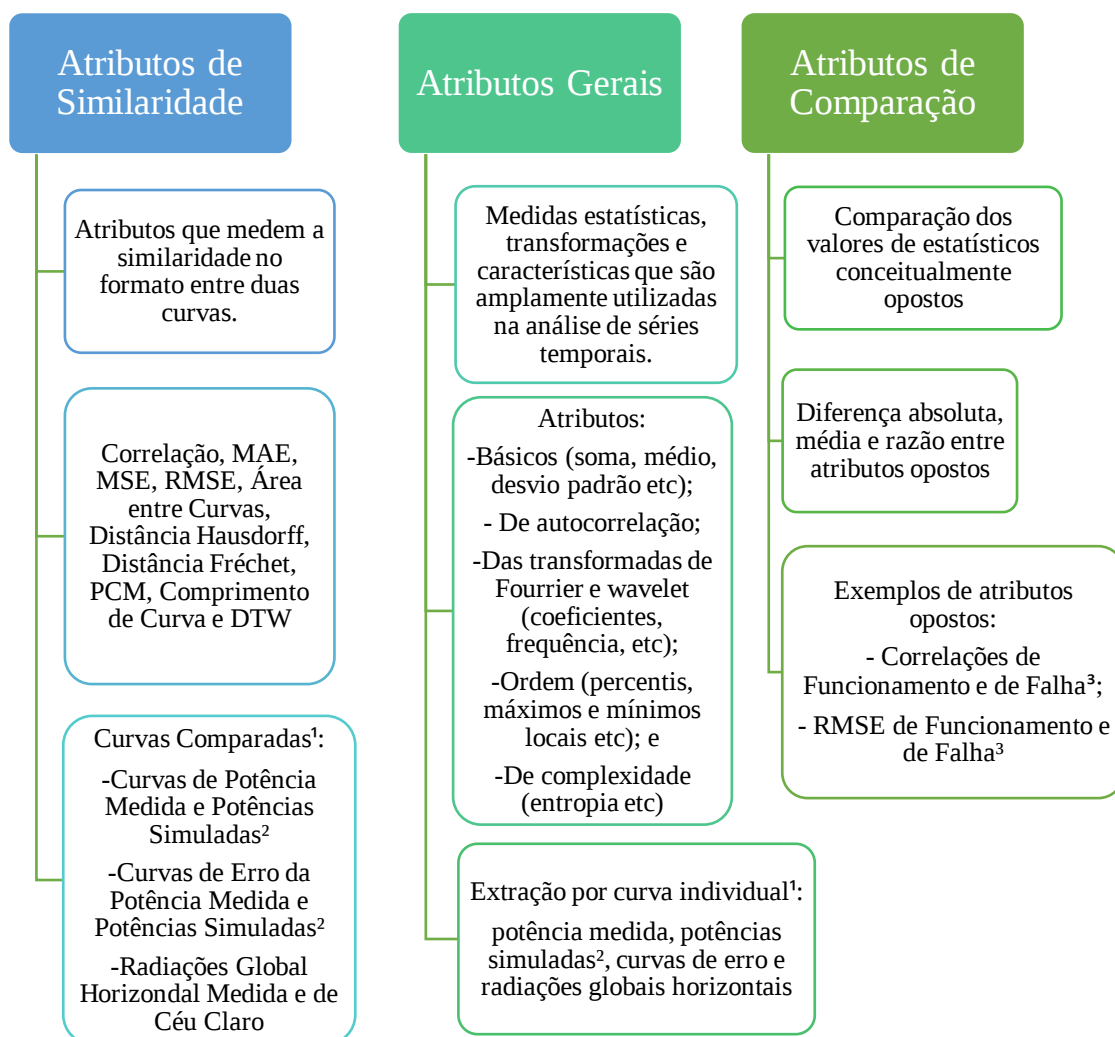


Fonte: Elaboração Própria.

De uma forma geral, nota-se que a oposição que se fala no conceito de atributos de comparação é essencialmente relacionada aos atributos de similaridade, como correlação, RMSE, DTW, PCM e os demais. No entanto, para trazer essa comparação também para o grupo dos atributos gerais, foi calculada a diferença absoluta entre os estatísticos extraídos das curvas de potência simuladas com rastreamento e com coletor fixo, já que elas seriam conceitualmente opostas no âmbito desse grupo.

A Figura 37 apresenta de forma esquemática os grupos de atributos explanados acima.

Figura 37 – Representação esquemática dos 3 grupos de atributos extraídos das curvas analisadas



Fonte: Elaboração Própria.

¹As curvas diárias de potência sempre são normalizadas pelos seus máximos diários.

²As curvas de potência simulada correspondem a potência simulada para o coletor com sistema de rastreamento e para o coletor em posição fixa.

³A correlação e o RMSE de funcionamento do rastreador são a correlação e o RMSE entre as curvas de potência medida e potência simulada com rastreamento, enquanto a correlação e o RMSE de falha são a correlação e o RMSE entre as curvas de potência medida e potência simulada com coletor fixo.

Por último, é importante ressaltar que não se conhece à priori a posição na qual o rastreador possivelmente falhou ao longo de determinado dia. Portanto, é necessário comparar a curva de potência medida com várias curvas de potência simulada para um coletor fixo, em que cada curva representa uma possível posição em que o sistema FV pode ter travado. Nesse

sentido, a metodologia proposta busca encontrar a posição fixa que maximiza a correlação entre a potência medida e a simulação de potência para um coletor fixo. Essa configuração de máxima correlação é então utilizada como referência para a extração dos demais atributos associados à curva de potência simulada para um coletor fixo.

À medida que múltiplos sistemas fotovoltaicos com rastreamento são conectados a um mesmo inversor e se possui apenas as medições da potência de saída do inversor, a quantidade de possíveis configurações formadas pela combinação das diversas posições em que cada rastreador pode ter falhado aumenta consideravelmente. Nesse cenário, o uso de técnicas de otimização, como a seleção de variáveis, torna-se necessário, levando em consideração sempre a avaliação do impacto na precisão resultante ao remover uma determinada configuração. Além disso, é importante considerar que a perda de precisão associada à remoção de uma possível configuração deve ser criteriosamente avaliada. A seleção de variáveis eficiente é essencial para equilibrar a simplificação do modelo com a manutenção de um nível adequado de precisão e acurácia nas estimativas.

4.1.4 Avaliação do Método de Classificação

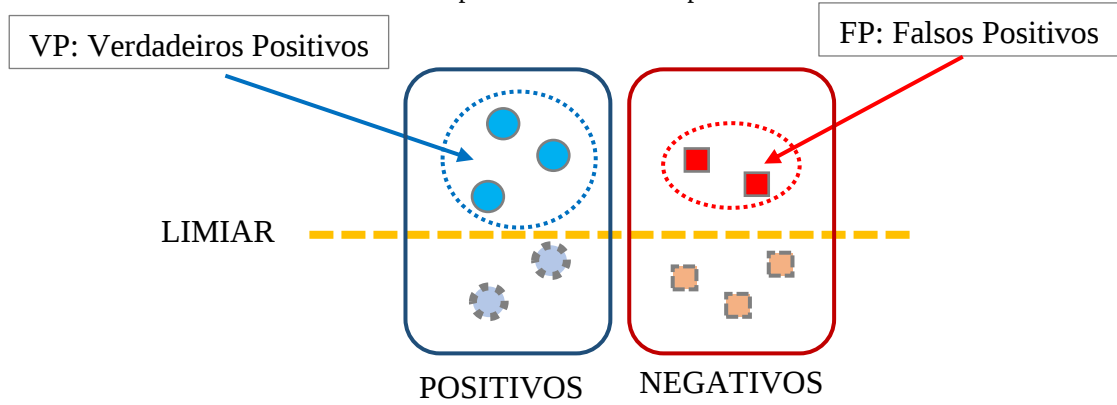
A seguir serão apresentadas as formas de avaliação do modelo de classificação utilizado na detecção de falhas no sistema de rastreamento.

Considere-se o exemplo genérico ilustrado na Figura 38. Nesse exemplo, os círculos representam os valores da classe positiva e os quadrados representam a classe negativa. O modelo desse exemplo estabelece um limiar para o qual acima dele os valores são classificados como positivos. Com base nisso, é possível haver valores corretamente classificados como positivos (verdadeiros positivos, VP) e valores erroneamente classificados como positivos (falsos positivos, FP), assim como valores erroneamente classificados como negativos (falsos negativos, FN) e valores corretamente classificados como negativos (verdadeiros negativos, VN).

Nesse contexto, a acurácia é a proporção dos valores verdadeiros em relação a quantidade total do conjunto. A Equação 34 apresenta a fórmula da acurácia para o caso de um conjunto com duas classes. Já o $f1$ (Equação 36) é uma média harmônica entre dois parâmetros conhecidos como *precision* e *recall* (COSTA *et al.*, 2007). O *precision* representa a proporção de verdadeiros positivos em relação a todos os valores classificados como positivos (Equação

35), enquanto o *recall* representa a proporção de verdadeiros positivos em relação a todos os valores da classe positiva (Equação 36).

Figura 38 - Exemplo Genérico de Método de Classificação por Limiar, cujos valores acima do limiar são classificados com pertencentes à classe positiva



Fonte: Elaboração Própria.

$$ACURÁCIA = \frac{VP + VN}{VP + FP + VN + FN} \quad (34)$$

$$precision = \frac{VP}{VP + FP} \quad (35)$$

$$recall = \frac{VP}{VP + FN} \quad (36)$$

$$f1 = 2 \frac{precision \times recall}{precision + recall} \quad (37)$$

Em que:

- *VP* e *FP*: respectivamente, verdadeiros positivos e falsos positivos;
- *VN* e *FN*: respectivamente, verdadeiros negativos e falsos negativos.

Em situações em que o conjunto de dados apresenta desbalanceamento entre as classes, a métrica da acurácia pode ser enganosa, uma vez que a classe predominante pode influenciar significativamente seu valor. Isso pode acarretar uma interpretação equivocada do desempenho do modelo, pois ele pode estar classificando corretamente apenas a classe predominante,

enquanto apresenta baixo desempenho nas demais classes. Nesse contexto, o f1 se mostra uma métrica relativamente melhor, pois leva em consideração informações de ambas as classes.

Além disso, em relação aos conjuntos com mais de duas classes, a acurácia pode ser calculada diretamente, porém o f1, por definição, não. Para contornar essa questão, é possível calcular o f1 para cada classe individualmente, considerando a classe em questão como positiva e as demais como negativas. Em seguida, pode ser realizada a média dos f1 por classe, fornecendo uma medida geral do desempenho do modelo em relação a todas as classes.

Uma ferramenta de visualização de desempenho do modelo é a matriz de confusão (COSTA *et al.*, 2007). Um exemplo genérico de matriz de confusão é apresentado na Tabela 7, na qual pode ser visto que a matriz é, simplificada, a relação entre as previsões feitas pelo modelo e as classes verdadeiras dos dados. Ao analisar a primeira linha da matriz de confusão na Tabela 7, que corresponde às amostras classificadas pelo modelo como Classe 0, 25 dos valores previstos são realmente da Classe 0, enquanto 20 pertencem à Classe 1 e 10 à Classe 2. Do ponto de vista percentual, o total de dados classificados pelo modelo como classe 0 corresponde a 55% do total, dos quais apenas 25% foram classificados corretamente, e 30% classificados de forma incorreta.

Tabela 7 - Exemplo genérico de matriz de confusão para um conjunto com três classes				
		Verdadeiros		
		Classe 0	Classe 1	Classe 2
Previstos	Classe 0	25 (25%)	20 (20%)	10 (10%)
	Classe 1	5 (5%)	20 (20%)	8 (8%)
	Classe 2	7 (7%)	10 (10%)	5 (5%)

Fonte: Elaboração Própria.

Uma outra forma de se visualizar o desempenho de um modelo de classificação é a partir da Curva ROC (FAWCETT, 2004). Enquanto as formas apresentadas anteriormente analisavam o modelo para um valor específico de limiar, a curva ROC enxerga como a variação do limiar interfere na classificação obtida. Assim, para cada possível limiar do conjunto de amostras, é registrado os valores da taxa de verdadeiros positivos (TVP) e da taxa de falsos positivos (TFP). As taxas, apresentadas nas Equações 38 e 39, são colocadas em um gráfico, no qual o eixo das abscissas é a TFP e o eixo das ordenadas é a TVP.

$$TVP = \frac{VP}{VP + FN} \quad (38)$$

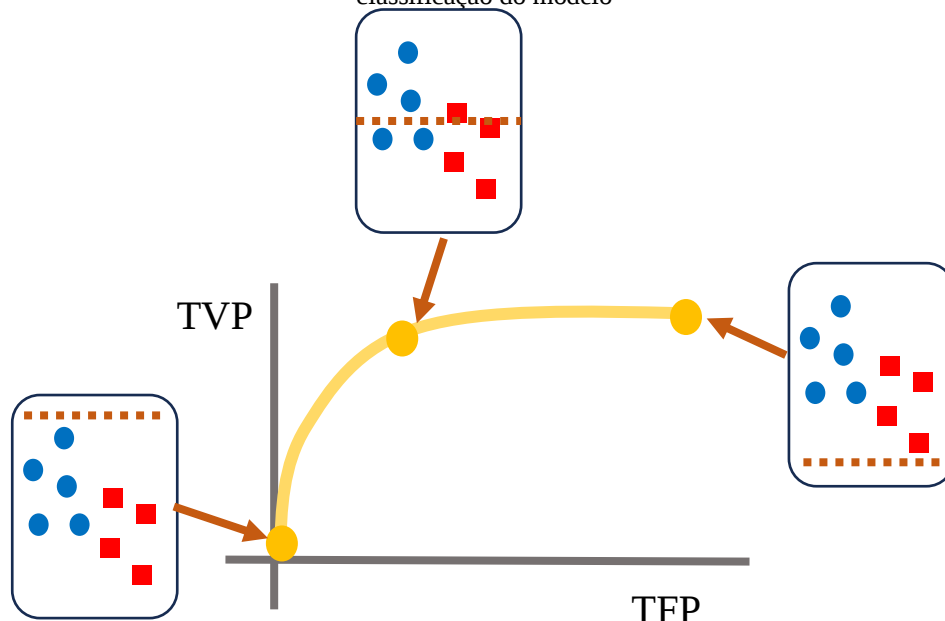
$$TFP = \frac{FP}{FP + VN} \quad (39)$$

Em que:

- *TVP* e *TFP*: respectivamente, taxa de verdadeiros positivos e taxa falsos positivos;
- *VP* e *FP*: respectivamente, verdadeiros positivos e falsos positivos;
- *VN* e *FN*: respectivamente, verdadeiros negativos e falsos negativos.

A representação esquemática da construção da curva ROC é apresentada na Figura 39. Ao observar a figura, pode-se perceber que, à medida que o limiar varia do valor máximo para o valor mínimo, os valores de TFP e TVP também variam, percorrendo o espaço entre as coordenadas (0;0) e (1;1). É importante ressaltar que o modelo representado na figura considera a classificação das amostras como positivas acima do limiar. No entanto, a mesma lógica se aplica caso a classificação dos valores positivos seja inferior ao limiar. Vale ressaltar também que a metodologia da curva ROC é aplicável apenas a um conjunto com apenas duas classes, a positiva e a negativa (FAWCETT, 2004).

Figura 39 - Representação Esquemática da construção da Curva ROC a partir da variação do limiar de classificação do modelo

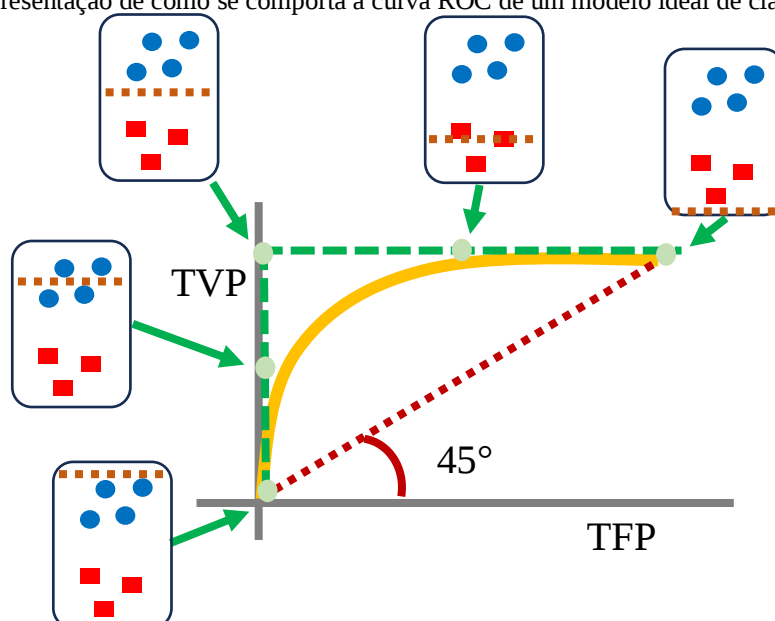


Fonte: Elaboração Própria.

Os retângulos vermelhos representam a classe negativa; os círculos azuis, a classe positiva. Por sua vez, TVP representa a taxa de verdadeiros positivos e TFP a taxa de falsos positivos.

O modelo de classificação ideal é caracterizado pela clara separação das classes, resultando em uma significativa margem entre elas. Nesse modelo, à medida que o limiar de classificação avança do valor máximo para o mínimo, a taxa de verdadeiros positivos aumenta gradualmente até atingir o valor máximo de 1, ao passo que a taxa de falsos positivos permanece em zero. Somente após o limiar ultrapassar o ponto de separação entre as classes é que a TFP começa a aumentar em direção ao valor máximo de 1, enquanto a TVP permanece constante e igual a 1. Essa dinâmica de variação da TVP e TFP para uma curva ROC ideal está representada pela curva verde tracejada na Figura 40.

Figura 40 - Representação de como se comporta a curva ROC de um modelo ideal de classificação (verde)



Fonte: Elaboração Própria.

Os retângulos vermelhos representam a classe negativa; os círculos azuis, a classe positiva. Por sua vez, TVP representa a taxa de verdadeiros positivos e TFP a taxa de falsos positivos.

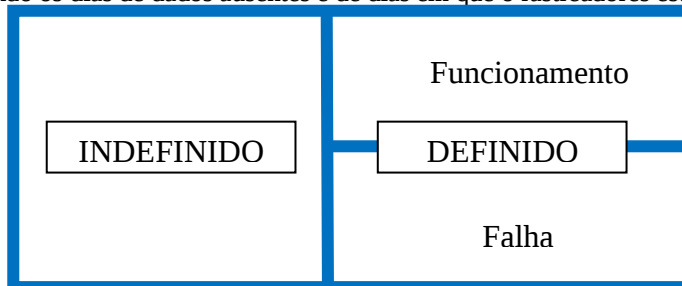
Nesse sentido, quanto mais próximo a curva ROC do modelo estiver da curva ideal, aqui denominada de quadrado, melhor será a qualidade do modelo de classificação. A área sob a curva ROC, conhecida como AUC (em inglês, *area under curve*), é uma métrica amplamente utilizada para avaliar a eficácia do modelo. Valores próximos a 1 indicam um desempenho ideal, enquanto valores abaixo de 0.5 sugerem que o modelo tem desempenho inferior a um classificador aleatório. Visualmente, um AUC inferior a 0.5 indica que a curva ROC do modelo se encontra inferior a reta de inclinação de 45°, a curva em vinho pontilhada da Figura 40 (FAWCETT, 2004).

4.1.5 Método de Classificação

Primeiramente, o período é dividido em calibração e validação sendo respectivamente $\frac{2}{3}$ e $\frac{1}{3}$ do total. A classificação dos dias pela rotulação visual (Seção 4.1.2) é usada como referência para verificação do desempenho do modelo de classificação. O método desenvolvido neste trabalho é denominado método do limiar móvel. Nele, dois procedimentos são realizados durante o período de calibração: o de identificação de limiares; e o de seleção de pares de atributos.

Removendo os dias de dados ausentes e os dias em que o rastreador esteve em posição de *stow*, é possível constatar que o conjunto de dias pode ser biparticionado em dias indefinidos e definidos. O subconjunto dos dias definidos, por sua vez, pode ser particionado em dias de falha e dia de funcionamento. Essa ideia está representada na Figura 41. Nesse contexto, o procedimento de identificação de limiares é desenvolvido para analisar a capacidade de um atributo em duas frentes: primeiro, na bipartição do conjunto de dias em indefinidos e definidos; e segundo, na identificação se um dia pertence à categoria de falha ou funcionamento, considerando apenas o subconjunto dos dias definidos.

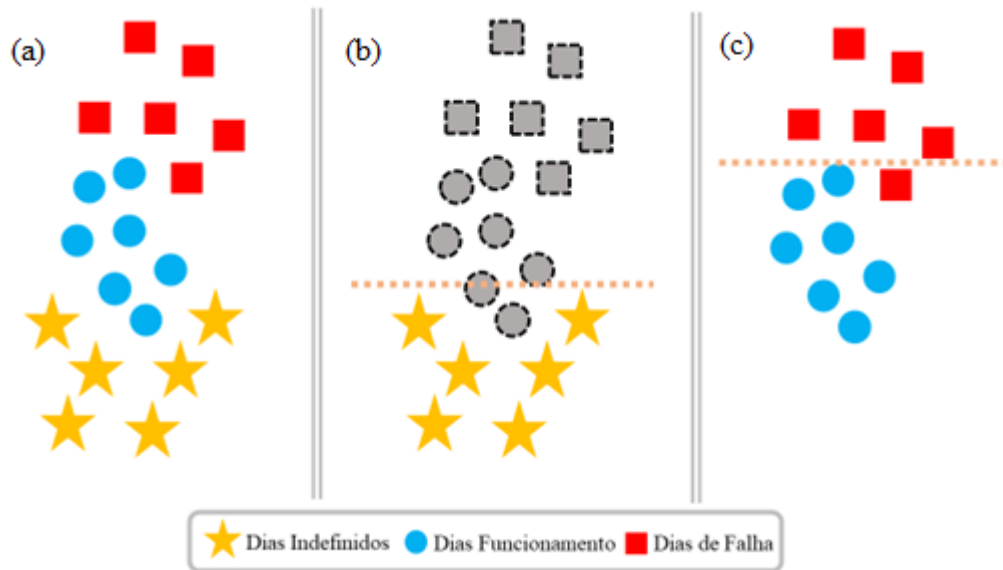
Figura 41 – Diagrama representando o conjunto das classificações dos dias em relação ao sistema de rastreamento, excluindo os dias de dados ausentes e de dias em que o rastreadores esteve na posição de *stow*



Fonte: Elaboração Própria.

A Figura 42 esquematiza o processo de identificação de limiares. Primeiramente, dado a série de valores de atributo (Figura 42a), um limiar móvel varia de forma que acima ou abaixo dele (ambas as possibilidades são testadas) os dias sejam classificados como indefinidos (Figura 42b). Em seguida, utilizando o mesmo método do limiar móvel, os dias de falha são classificados no subconjunto de dias definidos determinados pela rotulação visual (Figura 42c). O ajuste do limiar em ambas as etapas é realizado para maximizar um parâmetro de desempenho, neste caso, o $f1$. Uma representação alternativa desse procedimento pode ser visualizada na Figura 43, que utiliza histogramas para ilustrar a distribuição dos valores.

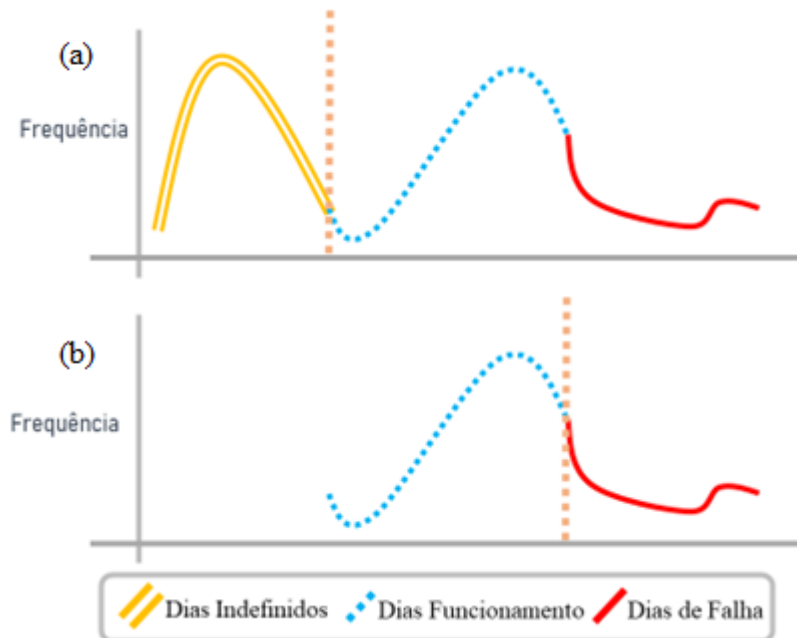
Figura 42 – Representação esquemática do método de identificação de limiares a partir dos valores dos atributos



Fonte: Elaboração Própria.

Em (b) é mostrado o primeiro passo do método, em que um limiar móvel é ajustado para avaliar a capacidade do atributo em distinguir os dias indefinidos e definidos (funcionamento e falha). No segundo passo, representado por (c), os dias indefinidos classificados na rotulação visual são removidos do conjunto, e o limiar móvel é aplicado para analisar a capacidade do atributo em identificar falhas no subconjunto de dias definidos.

Figura 43 – Representação esquemática com histogramas contínuos do método de identificação de limiares



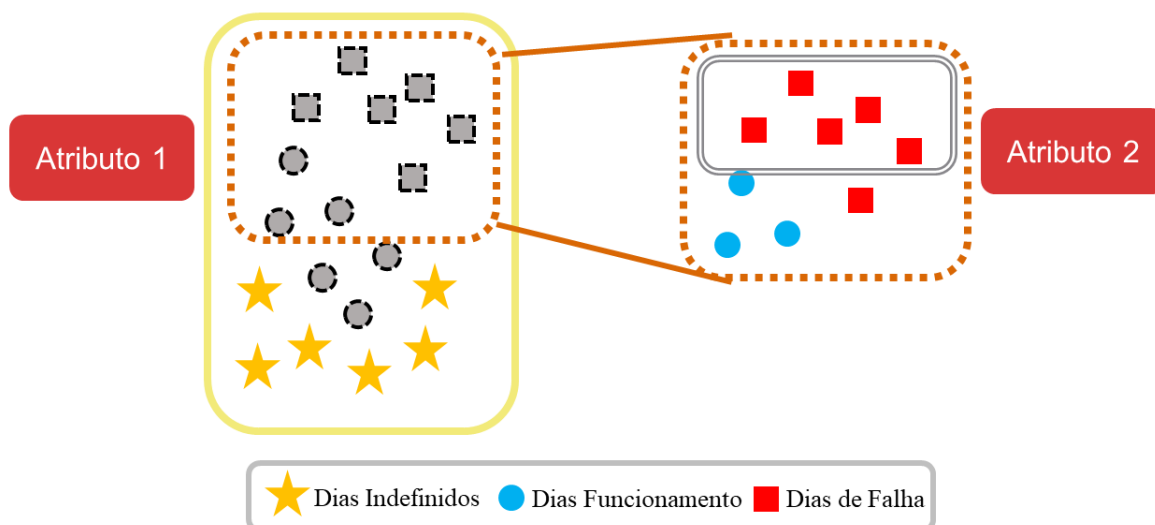
Fonte: Elaboração Própria.

Em (a) e (b) são representados os passos 1 e 2 do método do limiar móvel, utilizado para definir os limiares de identificação dos dias indefinidos e dos dias de falha, a partir do conjunto completo e do subconjunto de dias definidos, respectivamente.

Durante todo o processo de identificação de limiares, são calculados a Curva ROC, o f1 máximo, a acurácia associada ao limiar do f1 máximo e a matriz de confusão associada a esse mesmo limiar (COSTA *et al.*, 2007).

Durante a calibração, após a identificação de limiares, a próxima etapa é a seleção de pares de atributos, realizada a partir das possíveis combinações de estatísticos dois a dois. Para isso, os limiares de uma dupla de atributos são utilizados da seguinte maneira: do primeiro atributo é empregado o limiar para biparticionar o conjunto em dias indefinidos e definidos; e do segundo atributo o limiar para classificar os dias como falha ou funcionamento a partir do subconjunto de dias definidos identificados pelo primeiro atributo do par. Assim, ao final, obtêm-se a classificação dos dias em três classes: indefinidos, falha e funcionamento. A Figura 44 esquematiza esse processo. Durante a seleção de pares de atributos, são obtidos a matriz de confusão, a acurácia e o f1 médio das três classes. Os pares de atributos são ordenados de acordo com o valor do f1, com os pares que possuem maiores valores ocupando as primeiras posições.

Figura 44 - Representação Esquemática do Método de Classificação a partir do par de atributos



Fonte: Elaboração Própria.

Com o primeiro atributo do par, o conjunto de dias é biparticionado em dias indefinidos e dias definidos. A partir do subconjunto de dias definidos classificados pelo primeiro atributo do par, o segundo atributo classifica os dias em funcionamento e falha. Ao final, o conjunto de dias fica classificado em dias indefinidos, de funcionamento e de falha.

Na seleção de pares de atributos, é feita a classificação dos três tipos de dias para o período de calibração, por isso a necessidade de se utilizar o f1 médio. Em contrapartida, durante o processo de identificação de limiares, o contexto já é diferente, tendo em vista que

naquele momento em cada um dos dois passos existe apenas duas classes, sendo os dias indefinidos a classe positiva no passo 1 e os dias de falha no passo 2.

Em seguida, utilizando os limiares calculados durante o período de calibração, é feita uma análise de desempenho dos pares de atributos para o período de validação. São calculados a matriz de confusão, a acurácia e o f1 médio. Na validação, os pares também são ordenados a partir dos valores de f1.

Por fim, o principal objetivo do método de classificação é encontrar os pares de atributos que tiveram melhor desempenho durante os períodos de calibração e validação, ao mesmo tempo em que apresentam a menor variação de desempenho ao passarem do período de calibração para o período de validação. Isso é alcançado por meio da Métrica de Ranqueamento de Par de Atributo (MRPA) apresentada na Equação 40. A MRPA leva em consideração o f1 e a acurácia do par de atributos no período de validação, bem como a variação desses parâmetros entre os períodos de calibração e validação. Além disso, a posição ocupada pelo par durante os processos de calibração e validação é considerada na métrica, favorecendo pares que apresentaram menor variação e ocuparam posições superiores no processo de ordenamento. Nesse sentido, quanto maior o valor do MRPA, melhor é o par em termos de acurácia, f1 e consistência entre os períodos de calibração e validação.

$$MRPA = \frac{f1_{VALIDAÇÃO} \times ACURÁCIA_{VALIDAÇÃO}}{\Delta f1 \times \Delta ACURÁCIA \times posição_{CALIBRAÇÃO} \times posição_{VALIDAÇÃO}} \quad (40)$$

Em que:

- MRPA: métrica de ranqueamento de par de atributo;
- $f1_{VALIDAÇÃO}$: valor do f1 médio do par de atributos, durante o período de validação;
- $ACURÁCIA_{VALIDAÇÃO}$: acurácia do par de atributos, durante o período de validação;
- $\Delta f1$: diferença absoluta entre os valores de f1 médio do par durante a calibração e a validação;
- $\Delta ACURÁCIA$: diferença absoluta entre os valores de acurácia do par durante a calibração e a validação;

- $posição_{CALIBRAÇÃO}$: posição de ordenamento do par na calibração em relação aos demais, levando em conta o f1 médio: as primeiras posições são dos pares com maiores f1;
- $posição_{VALIDAÇÃO}$: posição de ordenamento do par na validação em relação aos demais, levando em conta o f1 médio: as primeiras posições são dos pares com maiores f1.

A Tabela 8 apresenta um exemplo ilustrativo do MRPA aplicado aos pares de atributos. Nela, observa-se que o par 2 obteve o maior valor de MRPA devido ao seu desempenho durante o período de calibração e validação, mesmo apresentando variações de f1 e de acurácia semelhantes ao par 3 e valores de f1 e de acurácia durante a validação iguais ao par 1. Além disso, o par 2 manteve posições consistentes e elevadas no processo de ordenamento durante a calibração e validação, contribuindo para o seu destaque.

Tabela 8 - Exemplo Genérico de Cálculo da Métrica de Ranqueamento de Par de Atributo (MRPA) a partir das métricas de desempenho levando em conta os períodos de calibração e validação

Par de Atributo	f1 Validação	$\Delta f1$	Acurácia Validação	$\Delta ACURÁCIA$	Posição Calibração	Posição Validação	MRPA
Par 1	0,91	0,05	0,93	0,04	1	6	70,52
Par 2	0,91	0,01	0,93	0,02	4	3	352,63
Par 3	0,85	0,01	0,90	0,02	2	8	239,06

Fonte: Elaboração Própria.

4.2 Ajuste dos Dados de Ocorrência de Falha para Inferência dos Índices Probabilísticos de Confiabilidade

Terminado o processo de detecção de falhas, dá-se lugar à inferência dos índices probabilísticos de confiabilidade, utilizando os indicativos de falhas do modelo de detecção da Seção 4.1.

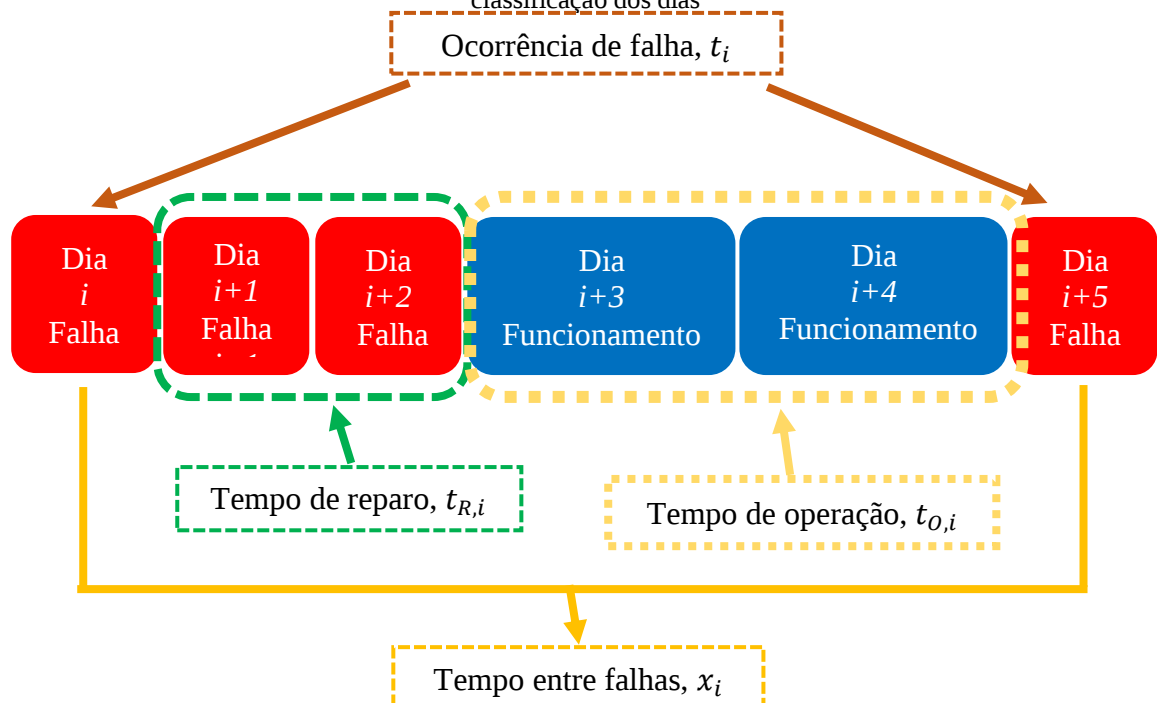
É importante discutir alguns pontos sobre as falhas por desalinhamento dos rastreadores analisados neste trabalho. O primeiro ponto é que o desalinhamento pode acontecer num determinado dia e já no dia seguinte não ocorrer, o que pode ser explicado pelo desgaste mecânico ou ausência de manutenção (ELERATH, 2011; SPERTINO; CORONA, 2013), assim

como por falhas no sistema de controle (KLISE, 2015). Mas também é importante destacar que a existência dessas falhas pontuais não exclui a ocorrência da possibilidade de falhas por grandes períodos consecutivos.

Não faz sentido, portanto, a aplicação da metodologia de Análise de Dados de Vida às falhas dos rastreadores, pois não são acontecimentos que acarretam fim da vida útil. Por isso, a Análise de Falhas Recorrentes é a mais recomendada tendo em vista a recorrência dos dias de desalinhamento (COLLINS *et al.*, 2009).

Nesse contexto, o pressuposto assumido é que, em dias consecutivamente classificados como falha, o primeiro dia é considerado o tempo t_i de ocorrência da falha, enquanto os dias seguintes são considerados como tempo de reparo, $t_{R,i}$. Essa escolha é fundamentada no fato de que uma das definições de reparo está relacionada ao tempo que o sistema volta a operar normalmente. Por outro lado, o tempo de operação, $t_{O,i}$, representa os dias de funcionamento. Esses pressupostos estão representados na Figura 45. Por último, os dias de dados ausentes e indefinidos também foram enquadrados em operação por não representarem falhas no sistema de rastreamento, de acordo com a definição adotada.

Figura 45 - Pressupostos sobre em relação a ocorrência de falha, tempo de reparo e tempo entre falhas a partir da classificação dos dias



Fonte: Elaboração Própria.

O ajuste de curva para inferência dos índices probabilísticos de confiabilidade começa pela escolha de qual índice será ajustado (a CDF $F(t)$, a curva de confiabilidade ou a função taxa de falha) e qual função paramétrica será empregada.

Uma prática muito comum é ajustar a $F(t)$ a partir do uso do ranqueamento da mediana (em inglês, *median rank*) (NELSON, 1982; O’CONNOR; KLEYNER, 2012). Esse ranqueamento ordena os dados do conjunto de amostras e estima a probabilidade acumulada por uma relação que utiliza a função beta (JACQUELIN, 1993). A fórmula do ranqueamento da mediana em si é um tanto quanto complexa e por isso são utilizadas aproximações, uma delas apresentada na Equação 41, chamada de aproximação de Bernard (CARRASCO *et al.*, 2013; JACQUELIN, 1993; O’CONNOR; KLEYNER, 2012).

$$\hat{F}_{R,i} = \frac{o_i - 0,3}{n + 0,4} \quad (41)$$

Em que:

- $\hat{F}_{R,i}$: estimativa da CDF na i -ésima posição a partir da aproximação de Bernard para o ranqueamento da mediana;
- o_i : índice de ordenamento da amostra;
- n : número total de elementos do conjunto de amostras.

Um exemplo genérico de uso do ranqueamento da mediana para estimar a função densidade de probabilidade de um conjunto de amostras de tempo entre falhas x é apresentado na Tabela 9.

Tabela 9 - Exemplo Genérico de estimativa da CDF pela aproximação de Bernard do ranqueamento da mediana

Tempo entre falhas x (ordenado)	Índice de ordenamento, o	$\hat{F}_{R,i}$ estimativa da CDF
100	1	0,09
200	2	0,23
700	3	0,36
700	4	0,50
700	5	0,64
1500	6	0,77
2000	7	0,91

Fonte: Elaboração Própria.

Em seguida, faz-se necessário a definição de qual distribuição paramétrica adotar. Nesse sentido, dentro da Análise de Falhas Recorrentes, a classe de modelos escolhida para o ajuste de curvas foi a de Processo de Renovação (PR). Um PR analisa as renovações de um sistema, isto é, os ciclos de falha e funcionamento do equipamento. Essa classe tem por premissa algumas simplificações que fazem com que a variável de interesse durante a inferência dos índices probabilísticos de confiabilidade seja o tempo entre falhas x_i (Figura 45). A primeira premissa é que o tempo de reparo pode ser considerado desprezível; e a segunda é que a distribuição de probabilidade é a mesma para todos os períodos entre falhas (RELIASOFT, 2015b).

Ainda na classe de modelos PR, existem os modelos de Processo Ordinário de Renovação (em inglês, *Ordinary Renewal Process* — ORP), que utilizam o tempo entre falhas x em funções conhecidas como a de Weibull (WANG; YANG, 2012). A função de probabilidade acumulada $F(x)$ de Weibull de dois parâmetros, assim como as respectivas curva de confiabilidade $R(x)$, função densidade de probabilidade $f(x)$ e função taxa de falha $\lambda(x)$ para o ORP são apresentadas nas Equações 42, 43, 44 e 45, respectivamente.

$$F(x) = 1 - \exp \left[-\left(\frac{x}{\varepsilon}\right)^\tau \right] \quad (42)$$

$$R(x) = \exp \left[-\left(\frac{x}{\varepsilon}\right)^\tau \right] \quad (43)$$

$$f(x) = \frac{\tau}{\varepsilon} \left(\frac{x}{\varepsilon}\right)^{\tau-1} \exp \left[-\left(\frac{x}{\varepsilon}\right)^\tau \right] \quad (44)$$

$$\lambda(x) = \frac{\tau}{\varepsilon} \left(\frac{x}{\varepsilon}\right)^{\tau-1} \quad (45)$$

Em que:

- x : tempo entre falhas;
- ε, τ : coeficientes do modelo: fatores de forma e de escala, respectivamente.

No entanto, nem sempre na prática é possível assumir que o tempo de reparo pode ser considerado desprezível. Uma forma de se verificar é justamente calculando a média aritmética entre as razões do tempo de reparo sobre o tempo entre falhas, isto é, $t_{R,i}$ sobre x_i . Para essas

situações, Azevedo (2013) e Cha e Finkelstein (2018) afirmam que um procedimento simples amplamente empregado é assumir que o sistema segue um Processo de Renovação Alternada (em inglês, *Alternating Renewal Processes* — ARP).

O ARP considera justamente o sistema como alternando entre dois estados, no caso deste trabalho, alternando entre operação e reparo. Nesse sentido, considerando o conjunto $\{t_{O,i}\}$, independente e igualmente distribuído (iid) e com distribuição $F(\cdot)$, e o conjunto $\{t_{R,i}\}$, também iid e com distribuição $M(\cdot)$, o processo $\{t_{O,i} + t_{R,i} \stackrel{\text{def}}{=} Z_i\}$ é um processo ordinário de renovação, cuja CDF $C(\cdot)$ é descrita pela convolução de $F(\cdot)$ e $M(\cdot)$, representada na Equação 46 (CHA; FINKELSTEIN, 2018).

$$C(v) = P(Z_i \leq v) = \int_0^v F(v-u)m(u)du = \int_0^v M(v-u)f(u)du \quad (46)$$

Em um sistema modelado pelo ARP, a principal característica analisada é a disponibilidade do sistema, $A(t)$, que é definida como a probabilidade do sistema se encontrar em operação em um determinado tempo t . A Equação 47 traz a denominada equação dependente do tempo para a disponibilidade de um processo ARP, também chamada de disponibilidade não estacionária (RELIASOFT, 2015a). Nela, o integrando representa a probabilidade de o sistema passar por uma renovação entre os tempos $(u, u + du)$, o que é representado pela função de probabilidade de renovação $h_Z(u)$ do ORP. Essa probabilidade leva em consideração que a última renovação ocorreu em algum momento entre 0 e t e que o sistema está operacional no tempo $t - u$, representado pelo termo de confiabilidade $R(t - u)$ (CHA; FINKELSTEIN, 2018).

$$A(t) = R(t) + \int_0^t h_Z(u)R(t-u)du \quad (47)$$

Na prática, a disponibilidade não estacionária é obtida numericamente. Contudo, a disponibilidade estacionária A , isto é, o $\lim_{t \rightarrow \infty} A(t)$, pode ser obtida pela Equação 48, a partir das médias μ_O e μ_R dos processos $\{t_{O,i}\}$ e $\{t_{R,i}\}$, respectivamente (CHA; FINKELSTEIN, 2018).

$$A = \lim_{t \rightarrow \infty} A(t) = \frac{\mu_O}{\mu_O + \mu_R} \quad (48)$$

Pode ser percebido, portanto, que, ao modelar um sistema como ARP, é importante inferir a distribuição dos tempos de operação e de reparo, a fim de se obter a disponibilidade e

também se analisar os processos associados a esses tempos. Nesse sentido, este trabalho vai utilizar a distribuição de Weibull de dois parâmetros para os ajustes, podendo ser o ajuste do tempo entre falha x_i caso o sistema possa ser modelado por um ORP ou o ajuste dos processos $\{t_{O,i}\}$ e $\{t_{R,i}\}$ se o sistema for modelado por um ARP. Para o segundo caso, as Equações 42 a 45 podem ser empregadas, substituindo a variável x_i por $t_{O,i}$ e $t_{R,i}$. Um detalhe é que a CDF do tempo de reparo é comumente denominada Curva de Manutenção, $M(t)$, e representa a probabilidade de que o componente esteja reparado até o tempo t . Vale ressaltar também que $1 - M(t)$ não representa a curva de confiabilidade da manutenção. Por fim, baseando-se no resultado da Equação 14 da Seção 2.4.1, as Equações 49 e 50 representam as médias μ_O e μ_R dos processos $\{t_{O,i}\}$ e $\{t_{R,i}\}$, respectivamente.

$$\mu_O = \int_0^{\infty} R(t)dt \quad (49)$$

$$\mu_R = \int_0^{\infty} [1 - M(t)]dt \quad (50)$$

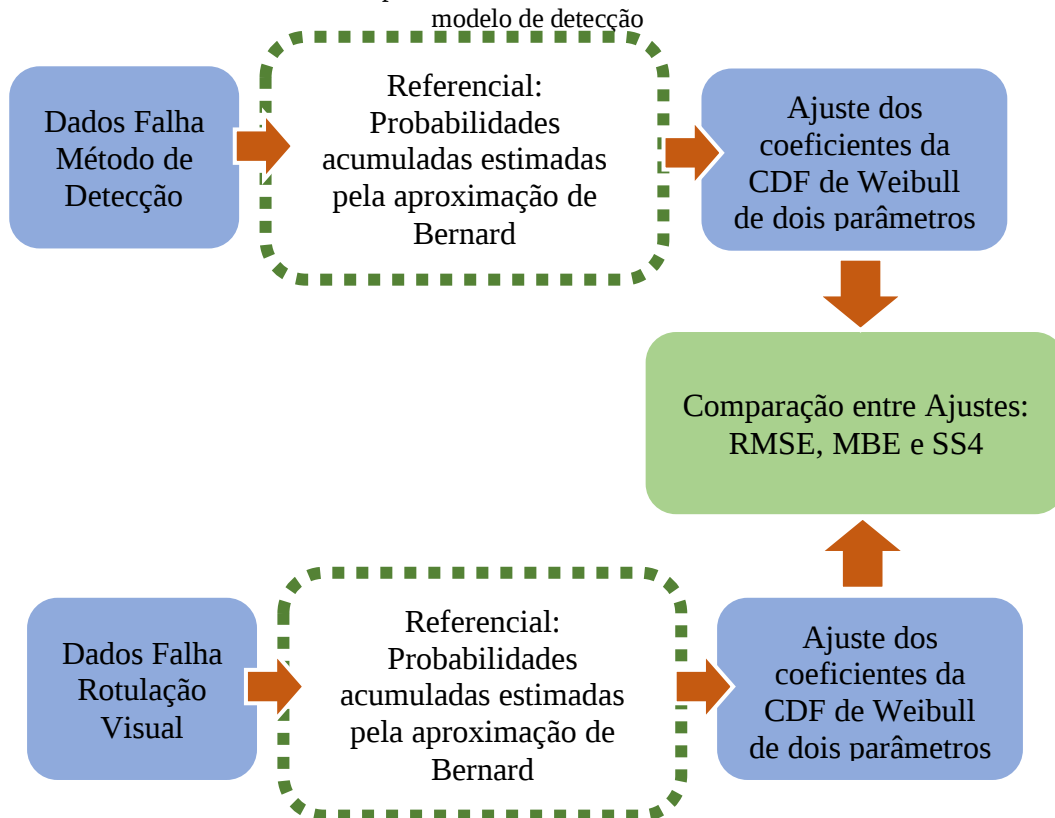
Assim sendo, o processo de ajuste de curvas empregado neste trabalho consiste em ajustar os coeficientes da CDF de Weibull de dois parâmetros. Para isso, o referencial utilizado será as estimativas de probabilidade acumulada calculadas pela aproximação de Bernard do ranqueamento da mediana. Se o sistema puder ser modelado por um processo ordinário de renovação, a variável de interesse será o tempo entre falhas. Por sua vez, caso seja modelado por um processo de renovação alternado, as variáveis de interesse serão os tempos de reparo e de operação.

O objetivo é ajustar os coeficientes da CDF de Weibull utilizando os dados de detecção de falhas obtidos pelo método do limiar móvel e comparar a curva resultante com a curva obtida a partir dos dados de falha da rotulação visual. Essa comparação permite avaliar a eficácia do método de detecção em inferir os índices probabilísticos de confiabilidade, verificando se as curvas geradas com os dados do método são consistentes com as curvas obtidas com os dados da rotulação visual. A comparação entre os dois tipos de ajustes se dará a partir do RMSE, MBE e SS4 (Equações 26, 27 e 30, respectivamente). Na Figura 46, esse processo está esquematizado.

É importante observar, contudo, que a função taxa de falhas de Weibull de dois parâmetros é uma função monotônica, podendo ser constante a depender dos coeficientes.

Nesse sentido, durante o ajuste, o conjunto de amostras será analisado quanto às suas tendências de crescimento e decrescimento a fim de que, caso haja um indicativo de não monotonicidade, o conjunto seja particionado para realização do ajuste dos coeficientes localmente.

Figura 46 - Representação esquemática dos procedimentos de Ajuste de Curva para determinação dos coeficientes da CDF de Weibull de dois parâmetros a fim de verificar a eficácia no uso dos dados de falhas do modelo de detecção



Fonte: Elaboração Própria.

O ajuste dos dois coeficientes da CDF de Weibull às amostras de probabilidades acumuladas estimadas pela aproximação de Bernard será feito minimizando o erro quadrático médio, semelhante ao que foi realizado na Seção 4.1.1 para o modelo de potência. Contudo, neste caso, é possível resolver analiticamente o ajuste ao linearizar a CDF de Weibull de dois parâmetros.

A linearização ocorre ao rearranjar a Equação 42 na forma da Equação 51. Em seguida, tomando duas vezes o logaritmo neperiano da equação, obtêm-se a Equação 52, na qual a variável dependente (Y) seria o termo $\ln \left[\ln \left(\frac{1}{1 - F(v)} \right) \right]$, a variável independente (X) seria o termo $\ln v$, o coeficiente angular (a) seria o fator de escala τ e o coeficiente linear (b) seria o termo $-\tau \ln \varepsilon$ (KHAN; PASHA, 2009; O'CONNOR; KLEYNER, 2012).

$$\frac{1}{1 - F(v)} = \exp \left[\left(\frac{v}{\varepsilon} \right)^\tau \right] \quad (51)$$

$$\ln \left[\ln \left(\frac{1}{1 - F(v)} \right) \right] = \tau \ln v - \tau \ln \varepsilon \quad (52)$$

Como o referencial para o ajuste é a estimativa de probabilidade pela aproximação de Bernard, isto é, o $\hat{F}_{R,i}$, o erro médio quadrático para a linearização da CDF de Weibull de dois parâmetros será justamente o somatório da diferença apresentada na Equação 57, levando em conta as variáveis das Equações 53, 54, 55 e 56. A minimização do erro quadrático médio é calculada derivando a Equação 57 em relação aos coeficientes a e b e igualando a zero, obtendo um sistema de equações lineares (CUSTODIO; ANDRADE; AUGUSTO, 1997). A Equações 58 e 59 para τ e ε são obtidas analiticamente a partir da resolução do sistema de equações lineares.

$$\hat{Y}_{R,i} = \ln \left[\ln \left(\frac{1}{1 - \hat{F}_{R,i}} \right) \right] \quad (53)$$

$$X_i = \ln v \quad (54)$$

$$a = \tau \quad (55)$$

$$b = -\tau \ln \varepsilon \quad (56)$$

$$\text{Erro Quadrático Médio} = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n [\hat{Y}_{R,i} - (a X_i + b)] \quad (57)$$

$$\tau = \frac{n \sum_{i=1}^n (\hat{Y}_{R,i} X_i) - \sum_{i=1}^n (\hat{Y}_{R,i}) \sum_{i=1}^n (X_i)}{n \sum_{i=1}^n (X_i) - (\sum_{i=1}^n (X_i))^2} \quad (58)$$

$$\varepsilon = \exp \left(- \frac{\sum_{i=1}^n (\hat{Y}_{R,i}) - \tau \sum_{i=1}^n (X_i)}{\tau n} \right) \quad (59)$$

Uma última observação relevante em relação ao procedimento de ajuste diz respeito à possibilidade de existirem mais de um valor de $\hat{F}_{R,i}$ para um mesmo período de tempo, como ilustrado na Tabela 9, na qual o tempo entre falhas de 700 dias se repete três vezes. Nesse sentido, o ajuste dos coeficientes utiliza uma função paramétrica de forma que a cada variável independente estará associado apenas um único valor de variável dependente e, portanto, o subconjunto de $\hat{F}_{R,i}$ com o mesmo período de tempo irá ser representado por apenas um ponto. Diante disso, ao invés de utilizar durante o ajuste os diferentes $\hat{F}_{R,i}$ associados a um mesmo período de falha, optou-se por se empregar a mediana desse subconjunto. Tome-se o exemplo da Tabela 9, em que os valores de $\hat{F}_{R,i}$ 0,36, 0,50 e 0,64 estão associados ao mesmo tempo entre falhas x_i de 700 dias. Neste caso, optou-se por utilizar a mediana, que seria 0,50.

Essa abordagem é justificada pelo fato de que o ajuste linear é suscetível a dados anômalos, o que poderia levar a imprecisões. Em outras palavras, a presença de dados muito discrepantes em relação ao subconjunto de valores de $\hat{F}_{R,i}$ relacionados a um mesmo período de tempo poderia influenciar negativamente o processo de ajuste. Portanto, ao utilizar a mediana como estatístico de centralidade, o ajuste se torna mais robusto, uma vez que assim o método é menos suscetível à influência de dados anômalos. Como resultado, a presença de valores discrepantes durante o ajuste é significativamente reduzida, promovendo resultados mais consistentes para o ajuste de coeficientes.

4.3 Dados Empregados

Os dados utilizados neste trabalho foram provenientes do complexo de Alice Springs, pertencente ao Centro Solar de Conhecimento do Deserto da Austrália (*Desert Knowledge Australia Solar Centre*, DKASC), localizado no centro do país (DKA, 2023). O mapa da Figura 47 mostra sua localização, e a Tabela 10 seus dados geográficos.

Figura 47 – Localização geográfica do Desert Knowledge Australia Solar Centre, DKASC



Fonte: Google Maps (adaptada).

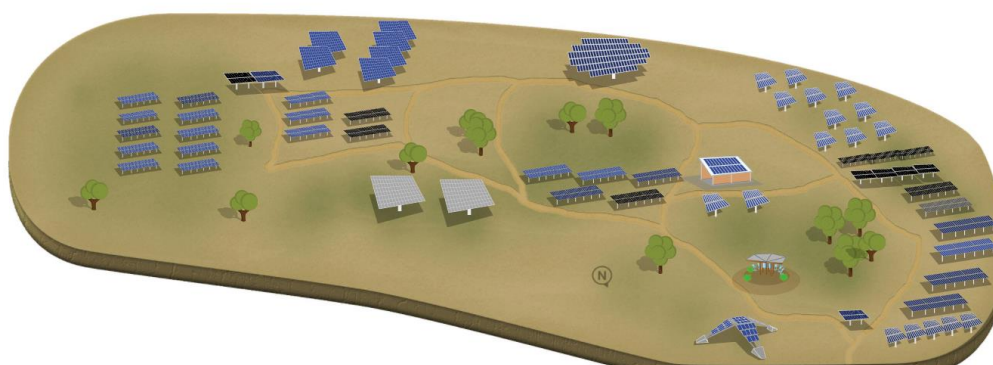
Tabela 10 – Dados Geográficos DKASC

Variável	Valor
Latitude	-23,76 °
Longitude	133,87°
Elevação Local	541 m

Fonte: Remund et al. (2003).

O DKASC possui ao todo um banco de dados com 46 sistemas fotovoltaicos, entre ativos e inativos, cada um dos quais com diferentes tipos de módulos, inversores e sistemas de rastreamento. Desses, 9 possuem rastreadores, sendo 8 de eixo duplo e 1 de eixo único. A Figura 48 mostra esquematicamente o complexo em 3D.

Figura 48 – Esquema em 3D do Sistemas Fotovoltaicos presentes no DKASC



Fonte: DKA (2023).

Dos sistemas do complexo, dois foram escolhidos para a análise. O primeiro foi o do sítio 2, composto por 156 painéis fotovoltaicos de 170 Wp acoplados a um sistema de rastreamento contínuo de eixo duplo azimute-altura ADES 5F-27M (Figura 49) (ADES, 2023), ligados em dois inversores de 15 kW. Suas características são apresentadas na Tabela 11.

A vantagem do sistema do sítio 2 está no fato de que, como os painéis estão todos inseridos em um mesmo rastreador, os dados de potência dos inversores estarão exclusivamente associados àquele rastreador. Isso não ocorre em outros sistemas, nos quais mais de um equipamento de rastreamento está conectado ao mesmo inversor.

Figura 49 - Sistema Fotovoltaico do sítio 2 do DKASC, com rastreador duplo de azimute-altura, modelo ADES 5F-27M.



Fonte: DKA (2023).

Tabela 11 – Dados do Sistema Fotovoltaico do Sítio 2 do DKASC

Variável	Descrição
Potência Nominal	26,52 kWp
Quantidade de rastreadores	1
Modelo do rastreador	ADES 5F-27M, eixo duplo
Número de módulos por rastreador x Potência por módulo	156 x 170 Wp
Modelo dos módulos	eco-Kinetics ECOKES 170M
Quantidade x Potência Nominal por Inversores	2 x 15 kW
Modelo dos Inversores	Fronius Symo 15 M X2

Fonte: DKA (2023).

Os dados de potência de saída do inversor para o sítio 2 começam às 00h00 do dia 24 de agosto de 2010 e são fornecidos em um intervalo de 5 em 5 minutos. Nas anotações relacionadas ao sistema feitas pela equipe de manutenção, duas se mostram importantes para este estudo:

- 22 de março de 2012: o motor do rastreador foi removido para manutenção até o dia 30 de maio de 2012, pois foi percebida uma falha no seguimento do Sol. Durante esse período, o sistema foi posto de forma fixa.
- Setembro de 2012: notou-se diversos problemas de desalinhamento no rastreador e, como o suporte técnico não conseguiu resolvê-los, o sistema foi definitivamente posto

de forma fixa no dia 01 de dezembro do mesmo ano, com azimuth -180° e inclinação de 20° .

Por conta deste último fato, o método de detecção de falhas por desalinhamento para o sítio 2 se deu com os dados de 24 de agosto de 2010 a 01 de dezembro de 2012.

Em relação ao rastreador de eixo único, o sistema fotovoltaico analisado no âmbito deste estudo foi o do sítio 5 do DKASC. O sistema é composto por 40 módulos fotovoltaicos de 135 Wp, os quais são conectados a cinco rastreadores segmentados de eixo único horizontal, modelo BW Solar F2 5 Star Tracker (Figura 50). Os rastreadores possuem uma abertura total de 90° , movendo-se em incrementos de 15° .

Figura 50 - Sistema Fotovoltaico do sítio 5 do DKASC, composto por cinco rastreadores de eixo simples segmentados, modelo BW Solar F2 5 Star Tracker, com abertura de 90° , andando em passos de 15°



Fonte: DKA (2023).

As características detalhadas dos rastreadores são apresentadas na Tabela 12. Neste sistema, há mais de um rastreador conectado a um mesmo inversor, o que requer uma análise das possíveis combinações de falhas entre eles durante a metodologia de detecção de falhas.

Tabela 12 – Dados do Sistema Fotovoltaico do Sítio 5 do DKASC	
Variável	Descrição
Potência Nominal	5,4 kWp
Quantidade de rastreadores	5
Modelo do rastreador	BW Solar F2 5 Star Tracker
Número de módulos por rastreador x Potência por módulo	8 x 135 Wp
Modelo dos módulos	Kyocera KD135GX-LP
Quantidade x Potência Nominal por Inversores	1 x 5 kW
Modelo dos Inversores	SMA SMC 5000A

Fonte: DKA (2023).

Os dados de potência de saída do inversor para o sítio 5 começam às 00h00 do dia 30 de setembro de 2008 e são analisados até 13 de julho de 2022, fornecidos em um intervalo de 5 em 5 minutos. Nas anotações relacionadas ao sistema feitas pela equipe de manutenção, três se destacam:

- 17 de fevereiro de 2012: detecção de que dois dos cinco rastreadores estavam travados, um requereu a troca do eixo principal de rotação e o outro a troca de bateria do controlador. Os sistemas voltaram a funcionar devidamente em 30 de abril de 2012.
- 20 de abril de 2016: notou-se a ocorrência de diversas falhas no sistema de rastreamento. Desde então, tem sido feitos esforços para a solução da problemática, o que é agravado pela ausência de fornecedores.
- 5 de fevereiro de 2019: a equipe constatou a necessidade de trocar o quadro do sistema de controle de um rastreador e de substituir a bateria em outro. O quadro foi prontamente trocado, porém a bateria foi substituída apenas no dia 8 de fevereiro de 2019.

As séries de dados utilizadas neste estudo foram fornecidas pelo DKASC: irradiância direta, irradiância difusa, temperatura ambiente, velocidade do vento a 10 metros de altura do solo e potência CA de saída do inversor. Essas séries foram registradas em intervalos de 5 minutos.

As séries de variáveis foram submetidas a testes globais de qualidade, abordando os seguintes aspectos: a ordenação cronológica dos dados, a detecção de duplicidade nos horários das medições, a identificação de lacunas e a uniformização dos intervalos de tempo (PETRIBÚ *et al.*, 2017). No caso da duplicidade, quando foram encontrados horários repetidos com amostras repetidas, apenas a primeira ocorrência foi mantida, preservando o valor correspondente. Entretanto, se as amostras duplicadas apresentavam valores diferentes, todas as duplicatas foram removidas e o valor da primeira amostra foi atribuído como vazio.

Por sua vez, as lacunas correspondem a registros de medições que deveriam estar presentes na série devido ao intervalo regular de medição, mas que estão ausentes. Quando uma lacuna é detectada, os tempos de ocorrência que deveriam constar na sequência são inseridos e atribuído valores vazios às amostras correspondentes. Isso garante que a série de dados seja preenchida adequadamente, mesmo nos intervalos em que não houve registros disponíveis.

Por fim, a uniformização dos intervalos de tempo foi realizada para garantir que as medições seguissem uma sequência regular. Por exemplo, se a sequência de tempos de registro de medição deveria ser 2023-06-15 14:10, 2023-06-15 14:15 e 2023-06-15 14:20, mas os tempos registrados foram 2023-06-15 14:10, 2023-06-15 14:17 e 2023-06-15 14:20, o valor

2023-06-15 14:17 foi ajustado para 2023-06-15 14:15 com base no intervalo de medição conhecido de 5 minutos.

Os resultados dos testes globais para as séries de dados dos sistemas dos sítios 2 e 5 acusaram apenas a existência de lacunas, sendo necessário a inserção de 233 tempos de ocorrência no sítio 2 e 9885 no sítio 5. Esses valores representaram aproximadamente 0,1% e 0,7% da série, respectivamente.

Agora, em relação aos testes locais de qualidade de cada variável, foram aplicadas testes físicos e estatísticos. Vale ressaltar que apenas os dados correspondentes a ângulos zenitais θ_z entre 0° e 87° foram considerados para análise dos testes locais, ou seja, foram selecionados apenas os períodos compreendidos aproximadamente entre o nascer e o pôr do sol (BRENNAND; PAULA, 2018). Isso foi feito justamente porque serão apenas os dados diários que serão empregados na metodologia de detecção.

Para a série de potência de saída do inversor, foi verificado se havia valores negativos ou 50% superiores a potência nominal do inversor. Caso algum valor se enquadrasse nessas condições, ele era substituído por um valor vazio.

No caso da série de velocidade do vento, foram aplicados dois testes (BRENNAND e PAULA, 2018). Primeiro, um teste de limites, em que valores abaixo de 0,1 m/s e acima de 96 m/s foram considerados anômalos e substituídos por valores vazios. Além disso, foi realizado um teste de persistência, no qual amostras repetidas por oito vezes consecutivas (equivalente a 40 minutos de dados a intervalo de 5 minutos) receberam também valores vazios.

Por fim, para as séries de irradiância, foram aplicados limites físicos sugeridos pela BRSN, conforme apresentado nas Equações 61 e 62, de forma que valores fora desses limites foram considerados anômalos e substituídos por valores vazios (LONG e DUTTON, 2010).

$$I_{0,eff} = 1367 \left[1 + 0,033 \cos\left(\frac{360^\circ}{365,25} DJ\right) \right] \quad (60)$$

$$0 \leq I_h \leq I_{0,eff} \times 1,5 \cos^{1,2} \theta_z + 100 \quad (61)$$

$$0 \leq I_d \leq I_{0,eff} \times 0,95 \cos^{1,2} \theta_z + 50 \quad (62)$$

Em que:

- DJ : dia juliano;
- θ_z : ângulo zenital, Equação 2 da Seção 2.3.

A Tabela 13 apresenta os resultados dos testes locais aplicados às séries de potência, velocidade do vento, irradiância global horizontal e irradiância difusa dos sistemas fotovoltaicos dos sítios 2 e 5 do DKASC. Ao analisar os resultados, é possível observar que a série de velocidade do vento se destaca em ambos os sítios, apresentando comparativamente um maior número de dados anômalos. No sítio 2, o percentual em relação ao total de dados que foram acusados como anômalos pelos testes locais na série de velocidade do vento totalizou 1,40%, enquanto no sítio 5 foi 0,64%. No Sítio 5, destaca-se também a série de potência com um percentual de anômalos de 1,12%, número expressivamente maior que o do sítio 2. Quanto às séries de irradiância global e irradiância difusa, os testes locais verificaram que praticamente não existem dados anômalos.

Tabela 13 - Resultados dos Testes Locais para as séries das variáveis de interesse dos sistemas FV analisados do DKASC

Testes Locais		Sítio 2	Sítio 5
Potência (Pot_{AC})	$Pot_{AC} < 0$ ou $Pot_{AC} > 1,5Pot_{AC,Nom}$	2 (0,00%)	7826 (1,12%)
Velocidade do Vento (v)	$v < 0,1 \text{ m/s}$ ou $v > 96 \text{ m/s}$	1552 (1,34%)	4279 (0,61%)
	Persistência em mais de 8 amostras	70 (0,06%)	175 (0,03%)
Radiação Global (I_h)	$I_h < 0$ ou $I_h > I_{0,eff} \times 1,5 \cos^{1.2} \theta_z + 100$	1 (0,00%)	25 (0,00%)
Radiação Difusa (I_d)	$I_d < 0$ ou $I_d > I_{0,eff} \times 0,95 \cos^{1.2} \theta_z + 50$	0 (0%)	2 (0,00%)
Total de Amostras por Série		115474	697680

Fonte: Elaboração Própria.

5. RESULTADOS E DISCUSSÕES

5.1 Método de Detecção de Falhas

Como explicado na Seção 4.1.5, o conjunto de dias analisados são divididos em calibração e validação.

Para o sítio 2, todos os 830 dias nos quais o sistema de rastreamento esteve funcionando foram considerados para a metodologia de detecção de falhas. Desse total, 2/3 foram para calibração, de 24 de agosto de 2010 a 27 de fevereiro de 2012; e 1/3 para validação, de 28 de fevereiro de 2012 a 30 de novembro de 2012.

Para o sítio 5, foram considerados 1501 dias. Da quantidade considerada, 2/3 foram para calibração, de 30 de novembro de 2008 a 26 de junho de 2011; e 1/3 para validação, de 27 de junho de 2011 a 8 de novembro de 2012.

5.1.1 Modelo de Potência

Conforme explicado na Seção 4.1.1, a primeira parte do ajuste dos coeficientes do modelo de potência é a seleção dos dias com alta correlação entre a potência medida e a irradiância no plano do coletor com rastreamento, a partir de três classes que levam em conta a correlação entre as irradiâncias globais horizontais medida (I_h) e de céu claro ($I_{h,CC}$). Após isso, ainda é realizada uma inspeção visual para verificar regulações na potência medida, isto é, dias nos quais a magnitude da potência medida é significativamente menor que a nominal.

Os dias selecionados para os sistemas FV dos sítios 2 e 5 do DKASC são apresentados nas Tabelas 14 e 15, respectivamente. Para o sítio 2 a inspeção visual não acarretou na exclusão de nenhum dia e o final da seleção resultou em um total de 10 dias, a maior parte deles nas classes de dias cuja correlação entre I_h e $I_{h,CC}$ esteve entre 0,95 e 0,97 e entre 0,97 e 0,99. Já para o sítio 5, os seguintes dias foram removidos a partir da inspeção visual: 24/07/2018, 19/05/2014, 01/11/2016, 09/03/2016, 04/07/2020 e 09/03/2017. Assim, o conjunto final de dias utilizados para o ajuste dos coeficientes do modelo de potência do sítio 5 totalizou 18, maior parte deles na classe de dias cuja correlação entre I_h e $I_{h,CC}$ esteve entre 0,97 e 0,99. Vale destacar que, para o sítio 5, a classe de dias com correlação entre I_h e $I_{h,CC}$ foi maior que 0,99 foi maior que para o sítio 2.

Tabela 14 - Dias selecionados do sistema FV do sítio 2 para ajuste dos coeficientes do modelo de potência

Correlação $C_{PEARSON}$ entre I_h com $I_{h,CC}$	Dias Selecionado
$0,95 \leq C_{PEARSON} < 0,97$	04/04/2011,
	27/06/2011
	02/11/2011
	11/07/2012
$0,97 \leq C_{PEARSON} < 0,99$	07/04/2011
	16/08/2011
	05/09/2011
	12/10/2011
	02/01/2012
$C_{PEARSON} \geq 0,99$	11/11/2011

Fonte: Elaboração Própria.

Tabela 15 - Dias selecionados do sistema FV do sítio 5 para ajuste dos coeficientes do modelo de potência

Correlação $C_{PEARSON}$ entre I_h com $I_{h,CC}$	Dias Selecionado
$0,95 \leq C_{PEARSON} < 0,97$	01/11/2008
	13/11/2008
	10/03/2009
	27/06/2011
	02/11/2011
	10/10/2012
$0,97 \leq C_{PEARSON} < 0,99$	19/02/2009
	03/03/2009
	04/01/2010
	08/10/2010
	15/10/2010
	10/10/2011
	12/10/2011
	14/10/2011
$C_{PEARSON} \geq 0,99$	22/09/2009
	16/10/2011
	21/10/2012
	31/10/2012

Fonte: Elaboração Própria.

Em seguida, após a seleção de dias, foi realizado o ajuste dos modelos de potência das Equações 22, 23 e 24 da Seção 4.1.1 para cada classe, resultando nos valores da Tabela 16.

Tabela 16 - Valores dos ajustes dos coeficientes por modelo e classe de dia

		Sítio 2				Sítio 5			
Modelo	Coeficiente	Classe de dias				Classe de dias			
		C 95*1	C 97*1	C 99*1	Média	C 95*1	C 97*1	C 99*1	Média
Linear	η_{FV}	0,78	0,79	0,84	0,80	0,80	0,83	0,86	0,83
	K_T	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003
	K	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
	η	0,93	0,93	0,94	0,93	0,93	0,94	0,95	0,94
	P_{MIN}	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Quadrático Q3	η_{FV}	0,71	0,70	0,70	0,71	0,82	0,73	0,75	0,76
	K_T	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003
	K	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
	C_0	0,44	1,05	1,70	1,06	0,04	0,18	0,13	0,12
	C_1	1,00	0,99	1,02	1,00	0,90	1,03	1,06	0,99
	C_2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	P_{MIN}	0,15	0,15	0,15	0,15	0,03	0,03	0,03	0,03
PV Watt	η_{FV}	0,77	0,78	0,83	0,79	0,80	0,84	0,89	0,84
	K_T	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003	-0,003
	K	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
	$\eta_{REFERÊNCIA}$	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

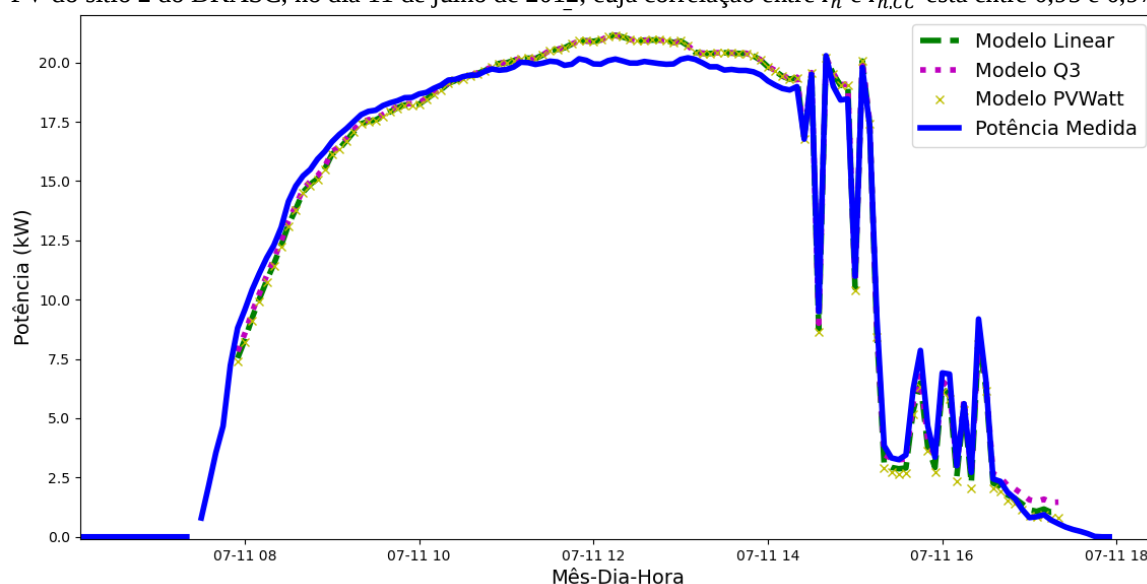
Fonte: Elaboração Própria.

*1 As classes C95, C97 e C99 correspondem, respectivamente, aos dias cuja correlação entre I_h e $I_{h,CC}$ esteve entre 0,95 e 0,97, entre 0,97 e 0,99 e maiores que 0,99.

Na Tabela 16, alguns resultados são interessantes. Primeiramente, o coeficiente térmico de potência K_T e o coeficiente de conversão de radiação em temperatura K (K de Ross) possuem os mesmos valores para os três modelos nas três classes de dias, $-0,003 \text{ W}/^\circ\text{C}$ e $0,02 \text{ }^\circ\text{C m}^2/\text{W}$, respectivamente. Isso indica uma consistência nas estimativas desses parâmetros independentemente das variações nos modelos e nas condições dos dias. Por sua vez, o coeficiente η_{FV} de representação do comportamento médio das perdas do gerador fotovoltaico possui valores muito próximos entre as classes de dias de um mesmo modelo. Outro ponto é que o mesmo η_{FV} dos modelos Linear e do Pv Watt possuem valores bastante similares.

Também pode ser percebido que a potência mínima de acionamento do inversor P_{MIN} possui valores pequenos, muito inferiores a 1% da potência nominal do inversor. Por último, no modelo quadrático Q3, C_2 assume um valor aproximadamente nulo, tornando o modelo quadrático em um modelo linear. Assim, os três modelos descrevem a potência do inversor de maneira linear. As Figuras 51 e 52 apresentam o ajuste dos modelos de potência para dois dias específicos nos sistemas fotovoltaicos dos sítios 2 e 5. Pode ser observado que nesses dias os modelos simularam a potência de forma bastante adequada, demonstrando uma boa concordância com os dados observados. Além disso, visualmente, é perceptível que os três modelos exibem resultados bastante similares, sugerindo uma consistência no desempenho deles na representação da potência do inversor.

Figura 51 - Potência medida e Potência simulada pelos modelos linear, Q3 quadrático e Pv Watt para o sistema FV do sítio 2 do DKASC, no dia 11 de julho de 2012, cuja correlação entre I_h e $I_{h,CC}$ está entre 0,95 e 0,97



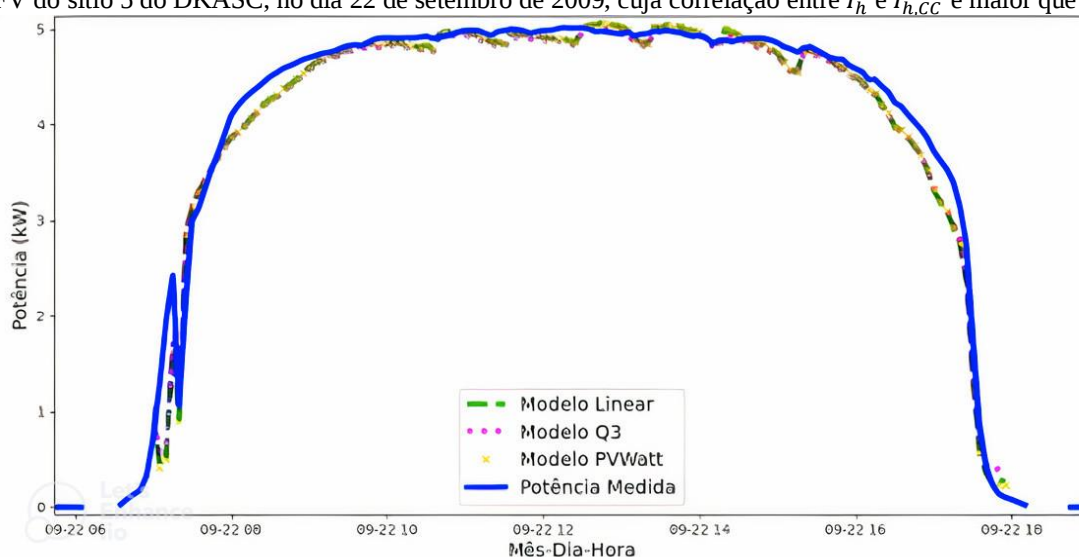
Fonte: Elaboração Própria.

Em seguida, tomando as médias dos coeficientes, é verificada a eficácia dos modelos em descrever o comportamento da potência no conjunto total de dias selecionados, isto é, o conjunto que aglomera todos os dias selecionados presentes nas três classes (correlação entre I_h e $I_{h,CC}$ entre 0,95 e 0,97, entre 0,97 e 0,99 e maiores que 0,99).

As Figuras 53 e 54 ilustram os diagramas de Taylor referentes aos modelos de potência aplicados nos sistemas fotovoltaicos dos sítios 2 e 5, respectivamente. Ao analisar o sítio 2, pode ser observado que os três modelos exibiram correlações bastante próximas, em torno de 0,95. No entanto, tanto o modelo Linear quanto o modelo Pv Watt apresentaram desvios padrão mais próximos dos valores observados, sendo que o modelo Pv Watt demonstrou uma maior

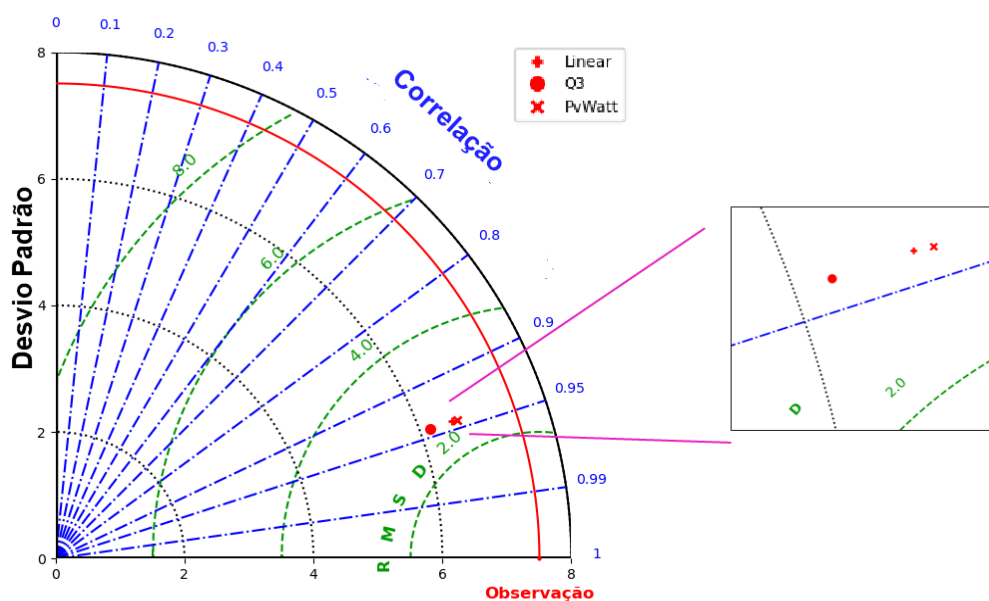
proximidade, como indicado pela sua posição mais próxima do arco em vermelho. No caso do sítio 5, a situação é semelhante, com as correlações dos três modelos também próximas entre si, porém com um valor ligeiramente superior a 0,95. Novamente, os modelos Linear e Pv Watt exibiram desvios padrão mais próximos dos valores observados, sendo que o modelo Pv Watt se mostrou levemente mais próximo. De uma forma geral, é perceptível, pelo Diagrama de Taylor, que os modelos têm resultados muito próximos, demonstrando que eles conseguem representar bem a potência de saída do inversor.

Figura 52 - Potência medida e Potência simulada pelos modelos linear, Q3 quadrático e Pv Watt para o sistema FV do sítio 5 do DKASC, no dia 22 de setembro de 2009, cuja correlação entre I_h e $I_{h,CC}$ é maior que 0,99



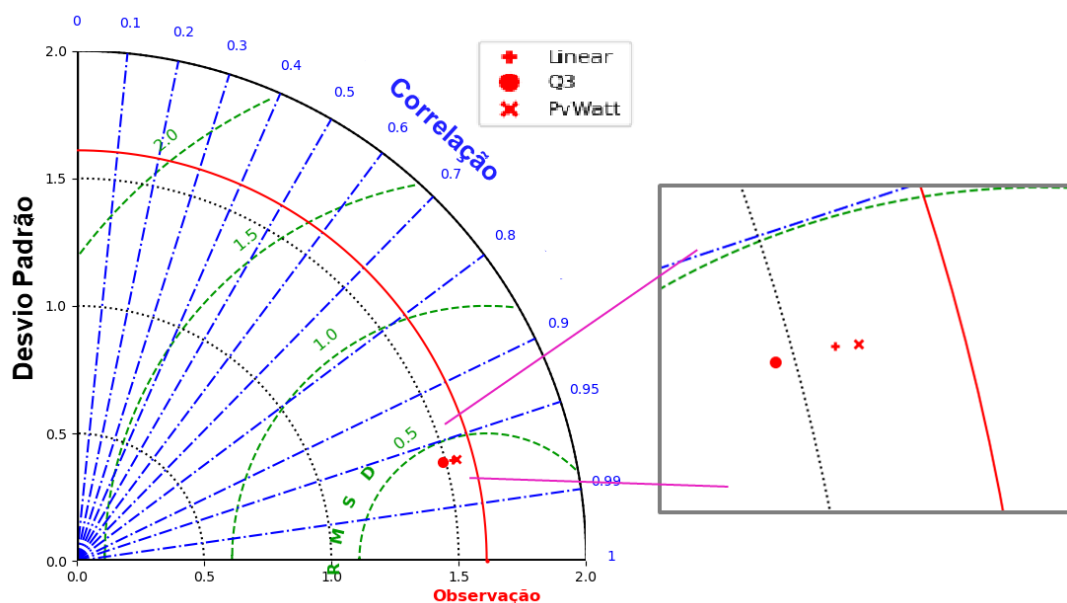
Fonte: Elaboração Própria.

Figura 53 - Diagrama de Taylor dos modelos de potência para o sistema fotovoltaico do sítio 2 do DKASC



Fonte: Elaboração Própria.

Figura 54 - Diagrama de Taylor dos modelos de potência para o sistema fotovoltaico do sítio 5 do DKASC



Fonte: Elaboração Própria.

A Tabela 17 fornece os valores dos parâmetros estatísticos utilizados na avaliação dos modelos de potência. Conforme observado nos diagramas de Taylor, os modelos apresentam correlações aproximadamente iguais, sendo que os modelos Linear e Pv Watt exibem desvios padrão mais próximos dos valores observados, com uma ligeira vantagem para o modelo Pv Watt. Isso implica em valores de SS4 similares e maiores para esses dois modelos. No que diz respeito ao RMSE e ao MBE, o modelo Q3 Quadrático obtém melhores resultados, seguido pelo modelo Linear. Com base nesses resultados, optou-se por adotar o modelo Linear, que apresenta um valor de SS4 aproximadamente igual ao do modelo Pv Watt, além de ter um MBE e um RMSE menores, perdendo nesses parâmetros apenas para o modelo Q3 Quadrático.

Tabela 17 – Parâmetros estatísticos de avaliação dos modelos de potência

		MBE	RMSE	Desvio Padrão Observação	Desvio Padrão	$C_{PEARSON}$	SS4
Sítio 2	Linear	-3,23	4,12		6,52	0,94	0,87
	Q3 Quadrático	-2,85	3,90	7,51	6,16	0,94	0,85
	Pv Watt	-3,30	4,16		6,61	0,94	0,87
Sítio 5	Linear	-0,17	0,45		1,52	0,97	0,93
	Q3 Quadrático	-0,13	0,44	1,61	1,48	0,97	0,92
	Pv Watt	-0,18	0,45		1,54	0,97	0,93

Fonte: Própria

5.1.2 Rotulação Visual e Classificação dos Dias

Como explicado na Seção 4.1.2, uma rotulação visual foi realizada para classificar os dias analisados pelo método de classificação. A rotulação visual será utilizada como referência para análise do desempenho do método de classificação. Os dias foram classificados em dias de falha, funcionamento, indefinidos, dias com dados ausentes e *stow*. Não foram verificadas nos bancos de dados velocidade acima de 12 *m/s* para o rastreador de eixo duplo, nem acima de 16 *m/s* para o rastreador de eixo único, o que caracterizaria uma situação em que rastreador estaria em posição de *stow*.

Ao todo foram rotulados, 830 dias para o sítio 2 e 1501 para o sítio 5. As Tabelas 18 e 19 apresentam os quantitativos das classes para os sítios 2 e 5 do DKASC, respectivamente.

Tabela 18 – Quantitativo de dias por classe para o Sítio 2 do DKASC – Rotulação Visual

Classe	Quantidade de Dias
Indefinido	188
Funcionamento	341
Falha	294
Dados Ausentes	7
<i>Stow</i>	0

Fonte: Própria

Tabela 19 – Quantitativo de dias por classe para o Sítio 5 do DKASC – Rotulação Visual

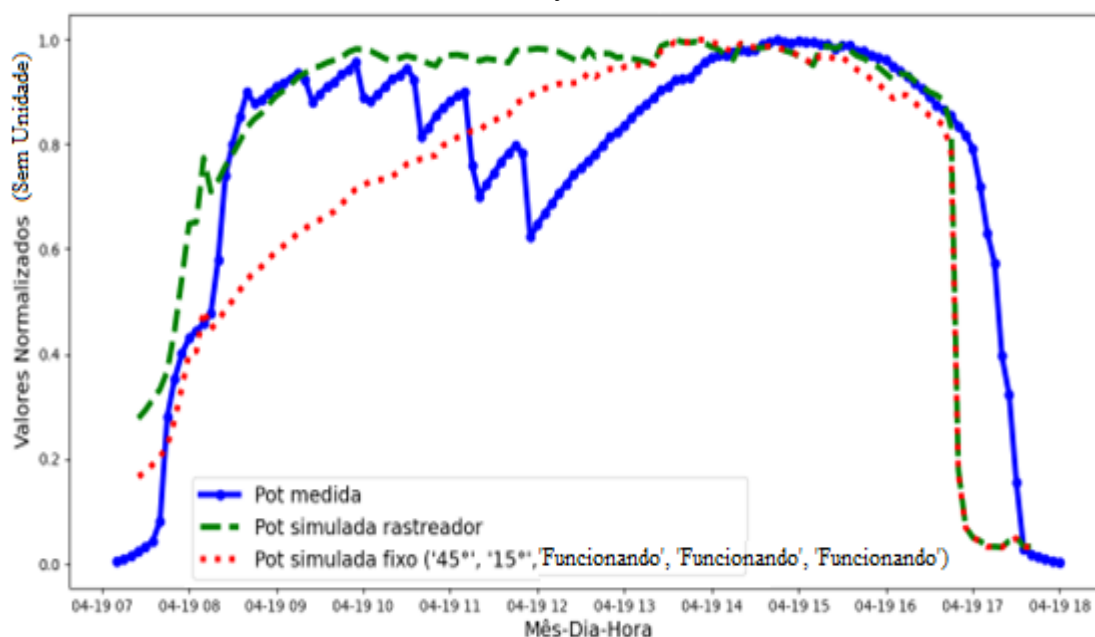
Classe	Quantidade de Dias
Indefinido	489
Funcionamento	578
Falha	373
Dados Ausentes	61
<i>Stow</i>	0

Fonte: Própria

Um aspecto relevante identificado durante a rotulação visual foi a ocorrência de dias nos quais a orientação do rastreador apresentou desalinhamento em múltiplas posições, conforme evidenciado na Figura 56. Na figura, é possível observar que o rastreador permaneceu travado em diferentes posições ao longo do dia. No exemplo ilustrado, o sistema aparentemente

ficou fixo em uma determinada posição durante a manhã e em outra posição durante a tarde. Esses tipos de dias foram classificados como dias de falha durante a inspeção visual. No entanto, é importante destacar que, no caso específico do dia retratado na Figura 55, o algoritmo de classificação erroneamente o classificou como dia de funcionamento.

Figura 55 – Potências medida e simuladas para um inversor com 5 sistemas fotovoltaicos com rastreadores de eixo único, no local 5 no *Desert Knowledge Australia Solar Centre*, em Alice Springs, Austrália Central, em 19 de março de 2011



Fonte: Elaboração Própria.

5.1.3 Método de Classificação

O conjunto de dias analisados pelo método de classificação foram divididos em 2/3 para calibração e 1/3 para validação. Como dito na Seção 5.1.2, a classificação obtida a partir da rotulação visual é utilizada como métrica de avaliação do desempenho do método de classificação.

O método de classificação proposto, denominado de método do limiar móvel, realiza dois procedimentos consecutivos. Primeiro, durante o período de calibração, define os limiares dos atributos, a partir da maximização do f1 m relação aos dias indefinidos e aos dias de falha. Em seguida, seleciona os pares de atributo levando em consideração a consistência de seus desempenhos durante os períodos de calibração e validação.

5.1.3.1 Definição de Limiares dos Atributos

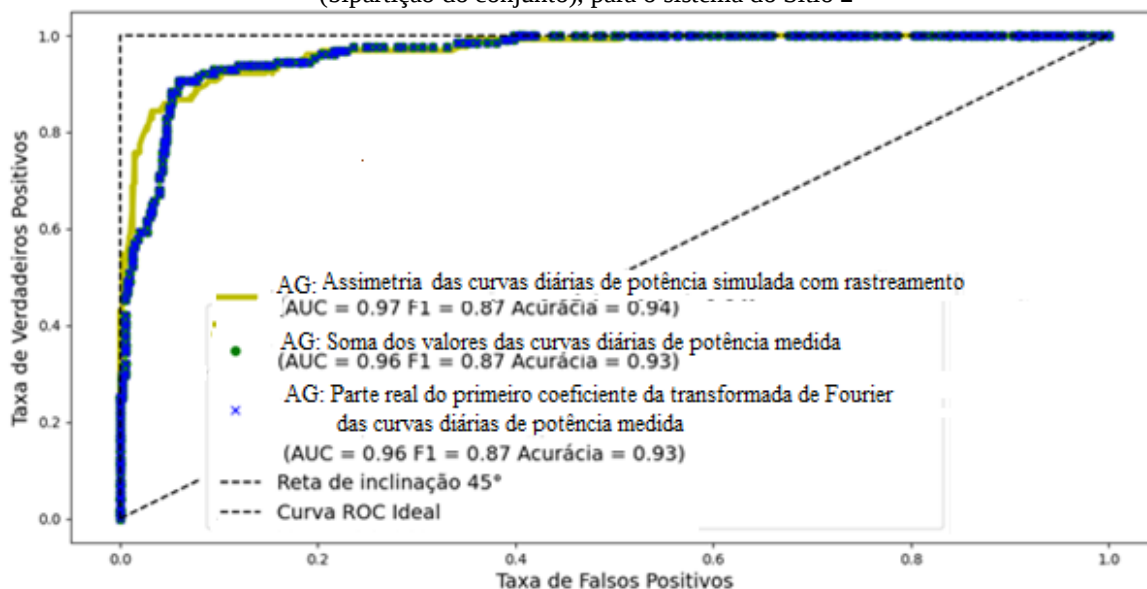
Durante o período de calibração, o primeiro processo do método de classificação é justamente a definição dos limiares dos atributos. Conforme a Seção 4.1.5, os limiares dos atributos são definidos ao se maximizar o $f1$ (Equação 37). Os atributos são avaliados em dois aspectos. Primeiro, quanto a sua capacidade em biparticionar corretamente o conjunto total de dias em indefinidos e definidos. Segundo, em analisar a capacidade desse mesmo atributo em classificar os dias de falha e funcionamento dado o subconjunto de dias sabidamente definidos a partir da rotulação visual. O objetivo disso é justamente identificar quais atributos são melhores na detecção dos dias indefinidos e quais são melhores na detecção dos dias de falha.

Para isso, serão apresentadas, por sítio, duas curvas ROC (Seção 4.1.4) trazendo os três melhores atributos que desempenharam em cada um dos aspectos analisados. Nesse sentido, defina-se AS, AG e AC como as siglas para atributos de similaridade, atributos gerais e atributos de comparação, respectivamente. O agrupamento dos atributos dessa forma foi explicado na Seção 4.1.3. É importante lembrar que os atributos foram extraídos das curvas normalizadas pelos seus valores máximos diários, pois a premissa de detecção de falhas nos rastreadores se baseia no formato da curva e não em sua magnitude.

As Figuras 56 e 57 mostram as curvas ROC dos atributos do sítio 2. Em relação ao aspecto de biparticionar o conjunto de dias em indefinidos e definidos, a Figura 56 exibe a Curva ROC dos três atributos com o maior valor de $f1$. É interessante observar que esses três atributos são atributos gerais. São eles: a assimetria (ou *skewness*) das curvas diárias de potência simulada com rastreamento; a soma dos valores das curvas diárias de potência medida; e a parte real do primeiro coeficiente da transformada de Fourier das curvas diárias de potência medida.

Os três atributos alcançaram resultados semelhantes em termos de valor de $f1$, em torno de 0,87. No entanto, a assimetria apresentou um desempenho ligeiramente melhor, com uma acurácia em torno de 0,94 e uma área sob a curva ROC (AUC) de 0,97. Na Figura 56, é possível observar que os três atributos funcionam bem na classificação dos dias indefinidos, mas apresentam uma ligeira distância em relação à curva ROC ideal.

Figura 56 – Curva ROC dos três atributos com melhor desempenho na classificação dos dias indefinidos (bipartição do conjunto), para o sistema do Sítio 2



Fonte: Elaboração Própria.

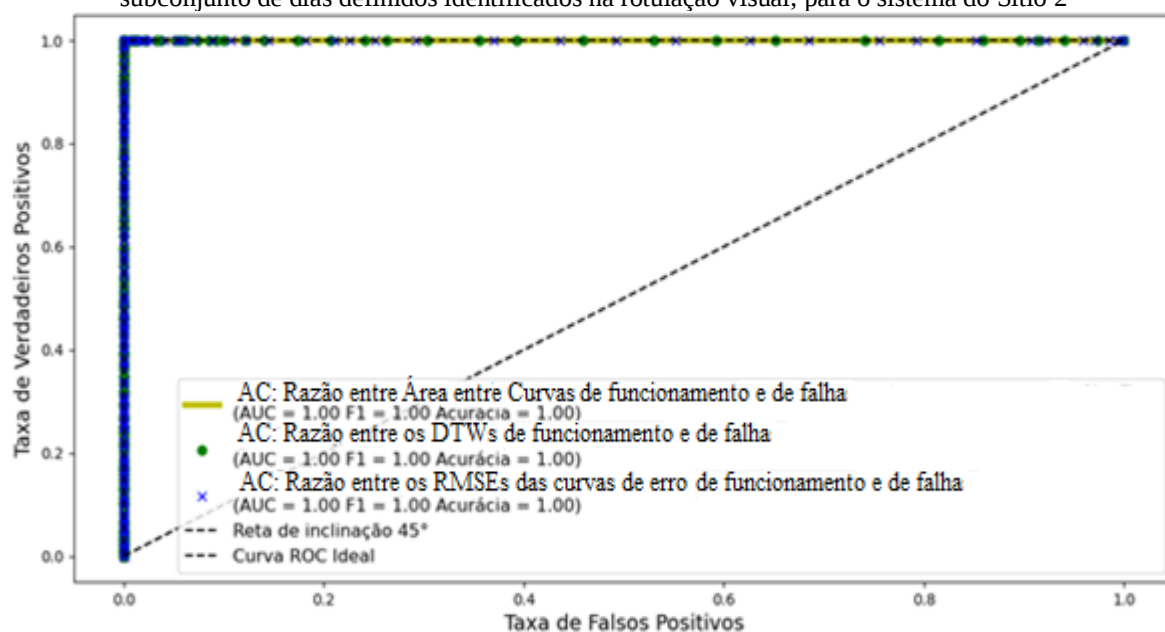
Por sua vez, a Figura 57 apresenta as curvas ROC dos três atributos com os maiores valores de f1 em relação à classificação dos dias de falha no subconjunto de dias indefinidos identificados pela rotulação visual para o sítio 2. É bastante interessante observar como esses atributos conseguem uma curva que se aproxima do ideal, indicando uma clara distinção entre os dias de falha e de funcionamento quando analisados sob a perspectiva desses três atributos, todos eles pertencentes ao grupo de atributos de comparação. Um exemplo dessa distinção pode ser observado na Figura 58, que traz o gráfico com os valores do AC razão entre as Áreas entre Curvas de funcionamento e de falha para o subconjunto de dias indefinidos. No gráfico, é notória a distinção entre os dias de funcionamento e de falha ao se utilizar esse atributo.

No grupo de atributos de comparação, quando, por exemplo, se fala de correlação de funcionamento, refere-se à correlação entre a curva de potência medida e potência simulada para um coletor com rastreamento. Por outro lado, a correlação de falha se refere à correlação entre a potência medida e a potência simulada para um coletor fixo. Nesse sentido, os três atributos com melhor desempenho na classificação de dias de falha para o sítio 2 são: a razão entre as Áreas entre Curvas de funcionamento e de falha, a razão entre os valores de DTW de funcionamento e de falha, e a razão entre os valores de RMSE das curvas de erro de funcionamento e de falha. Em todos eles, a métrica de comparação utilizada foi a razão.

Em relação ao sítio 5, as Figuras 59 e 60 mostram as curvas ROC para definir os limiares dos atributos. Na Figura 59, são apresentados os três atributos com maiores f1 em classificar os dias indefinidos. Diferentemente do sítio 2, os atributos são de similaridade. Inclusive, outras diferenças são justamente que, para o sítio 5, os atributos obtiveram um valor maior de f1,

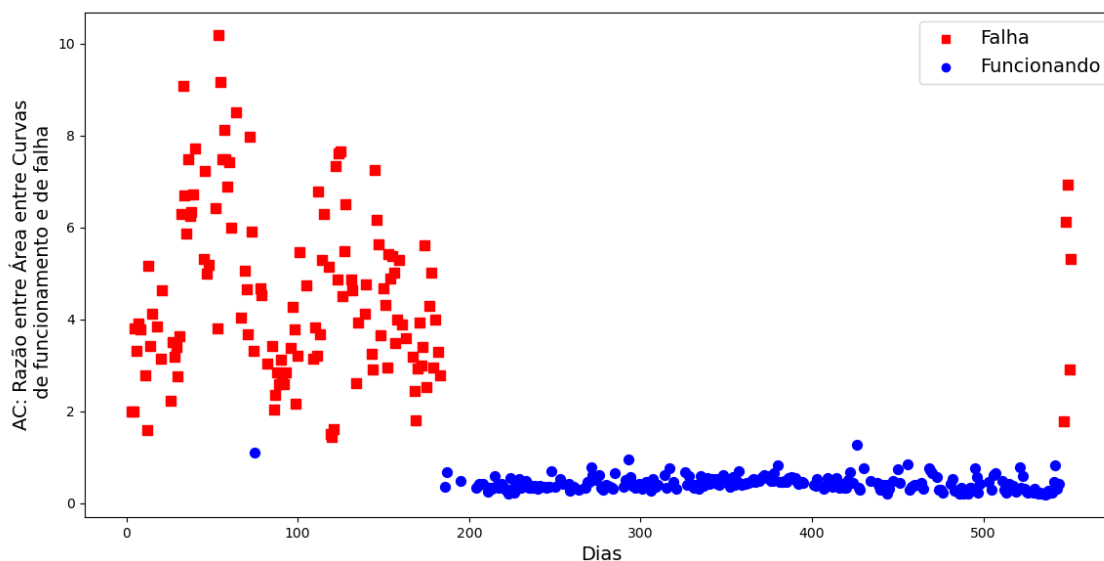
aproximadamente 0,91, e um formato de curva ROC mais próximo do ideal. Os três atributos com maiores f1 em classificar os dias indefinidos foram: o MAE entra a curvas diárias de potências simuladas com rastreamento e com o coletor fixo, a área entre curvas das curvas diárias de potências simuladas com rastreamento e com o coletor fixo e o DTW entre as curvas diárias de potências simuladas com rastreamento e com o coletor fixo.

Figura 57 – Curva ROC dos três atributos com melhor desempenho na classificação dos dias de falha a partir do subconjunto de dias definidos identificados na rotulação visual, para o sistema do Sítio 2



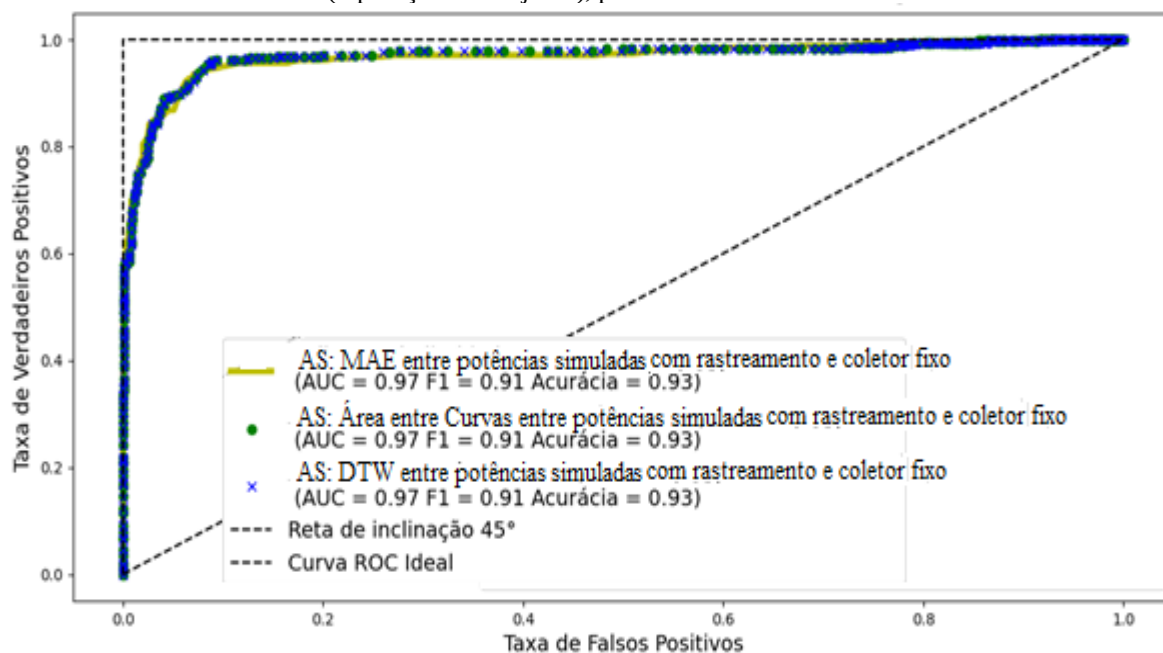
Fonte: Elaboração Própria.

Figura 58 – Gráfico dos valores de AC razão entre as Áreas entre Curvas de funcionamento e de falha para o subconjunto de dias indefinidos, para o sítio 2, durante o período de calibração



Fonte: Elaboração Própria.

Figura 59 – Curva ROC dos três atributos com melhor desempenho na classificação dos dias indefinidos (bipartição do conjunto), para o sistema do Sítio 5

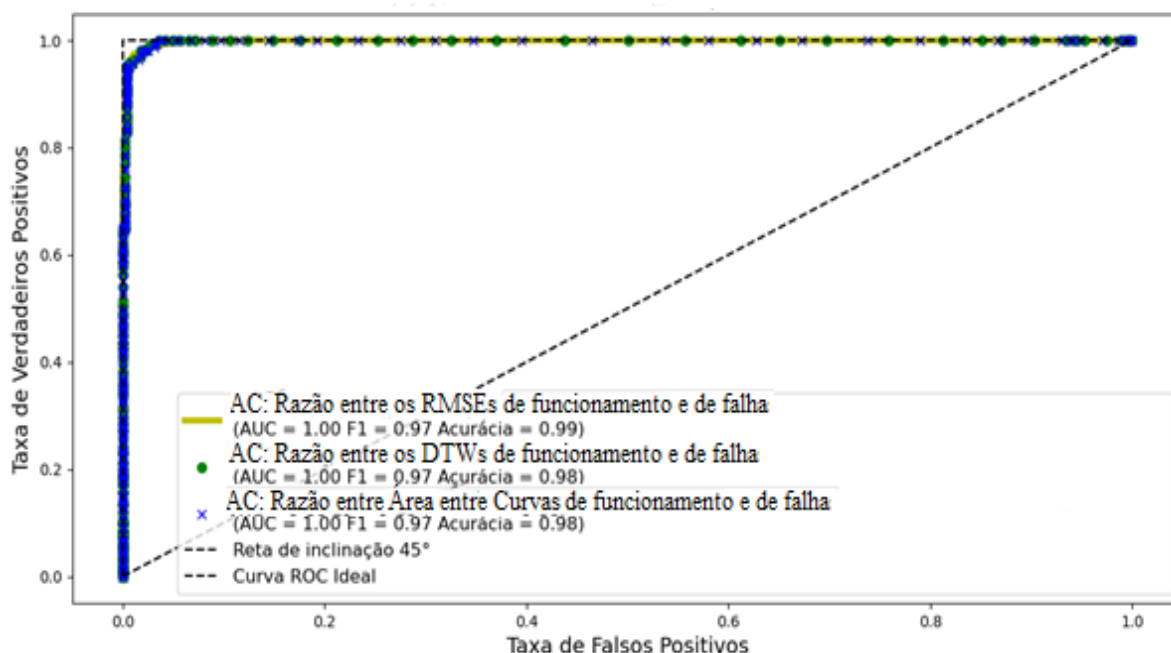


Fonte: Elaboração Própria.

Em relação à classificação dos dias de falha a partir do subconjunto de dias definidos para o sítio 5, os atributos de comparação mais uma vez apresentaram os três maiores valores de f1, seguindo o padrão observado no sítio 2, como ilustrado na Figura 60. Vale ressaltar que, neste caso, as curvas ROC estão muito próximas da curva ideal, embora em menor grau em comparação ao sítio 2. Os três atributos são: a razão entre os valores de RMSE de funcionamento e de falha, a razão entre os valores de DTW de funcionamento e de falha, e a razão entre as áreas das curvas de funcionamento e de falha. Em todos esses atributos, a métrica de comparação utilizada também foi a razão.

Por fim, é importante ressaltar que os atributos com melhor desempenho para cada aspecto do método de detecção, de forma geral, variaram entre os sítios analisados. Essa observação é bastante relevante, uma vez que evidencia a capacidade do método de classificação proposto em se adaptar à realidade específica de cada sistema fotovoltaico, selecionando os atributos mais adequados para a classificação dos dias. Essa flexibilidade é um ponto positivo, uma vez que permite uma abordagem adaptada às características e particularidades de cada sistema, contribuindo para um melhor desempenho do método de classificação proposto.

Figura 60 Curva ROC dos três atributos com melhor desempenho na classificação dos dias de falha a partir do subconjunto de dias definidos identificados na rotulação visual, para o sistema do Sítio 5



Fonte: Elaboração Própria.

5.1.3.2 Seleção de Pares de Atributos

Após estabelecer os limiares para cada atributo que maximizam o valor de f1 tanto na classificação dos dias indefinidos do conjunto total de dias quanto na classificação dos dias de falha do subconjunto de dias definidos, o último procedimento consiste em avaliar o desempenho dos pares de atributos na classificação geral dos dias. O primeiro par do atributo biparticiona o conjunto de dias em indefinidos e definidos, e, a partir do subconjunto de dias definidos classificado pelo primeiro atributo, o segundo classifica dias de falha e funcionamento do sistema de rastreamento.

Cada combinação de par é avaliada durante os períodos de calibração e validação, sendo ordenadas a partir da Métrica de Ranqueamento de Par de Atributo (MRPA) apresentada na Equação 40 e melhor explicada na Seção 4.1.5. O MRPA, dentre outros fatores, leva em conta a consistência do par de atributo em classificar os dias durante o período de calibração e validação, bem como a posição ocupada por eles em relação ao valor de f1 durante a calibração e validação.

Nesse sentido, defina-se as classes 0, 1 e 2 como associadas aos dias indefinidos, de falha e de funcionamento, respectivamente. Mais uma vez, vale destacar que, no grupo de atributos de comparação, quando se fala de um atributo de funcionamento, refere-se ao atributo

extraído entre as curvas de potência medida e de potência simulada para um coletor com rastreamento. Por outro lado, um atributo de comparação de falha se refere ao atributo extraído entre a potência medida e a potência simulada para um coletor fixo. Vale lembrar também que AS, AG e AC significam respectivamente atributos de similaridade, atributos gerais e atributos de comparação.

A Tabela 20 apresenta os três principais pares de atributos com maiores valores de MRPA para o Sítio 2. Os pares foram constituídos por atributos gerais utilizados para a divisão do conjunto de dias em indefinidos e definidos, juntamente com atributos de comparação para a classificação dos dias de falha e funcionamento no subconjunto de dias definidos resultantes. Ao analisar a Tabela 20, pode-se observar que os parâmetros dos pares estão muito próximos entre si. Embora o par 2 tenha ocupado uma posição muito inferior no período de calibração, ele superou o par 3 devido à sua menor diferença absoluta de f1 e da sua posição na validação. No entanto, o par 1 obteve o melhor desempenho devido à combinação favorável de suas posições nos períodos.

Tabela 20 – Os três melhores Pares de Atributos ordenados pelo MRPA para o Sítio 2

Par de Atributos	Posição CAL*1	Posição VAL*1	f1 CAL*1	f1 VAL*1	Diferença absoluta f1	Acurácia CAL*1	Acurácia VAL*1	Diferença absoluta Acurácia	MRPA
Par 1*2	13	4	0,92	0,86	0,06	0,94	0,90	0,03	8,04
Par 2*3	98	1	0,91	0,86	0,05	0,93	0,90	0,03	6,02
Par 3*4	14	5	0,92	0,86	0,06	0,94	0,90	0,03	5,97

Fonte: Elaboração Própria.

*1: CAL significa período de calibração, enquanto VAL período de validação. Significa a posição que um par ocupou durante esses períodos a partir do seu valor de f1 em relação aos demais pares. As primeiras posições representam maiores valores de f1, portanto a posição 1 é ocupada pelo par com maior f1 no período.

*2: O Par 1 é formado respectivamente pelo AG Soma dos Valores das curvas diárias de potência simulada com rastreamento e pelo AC Razão entre os valores de RMSE das curvas de erro de funcionamento e de falha.

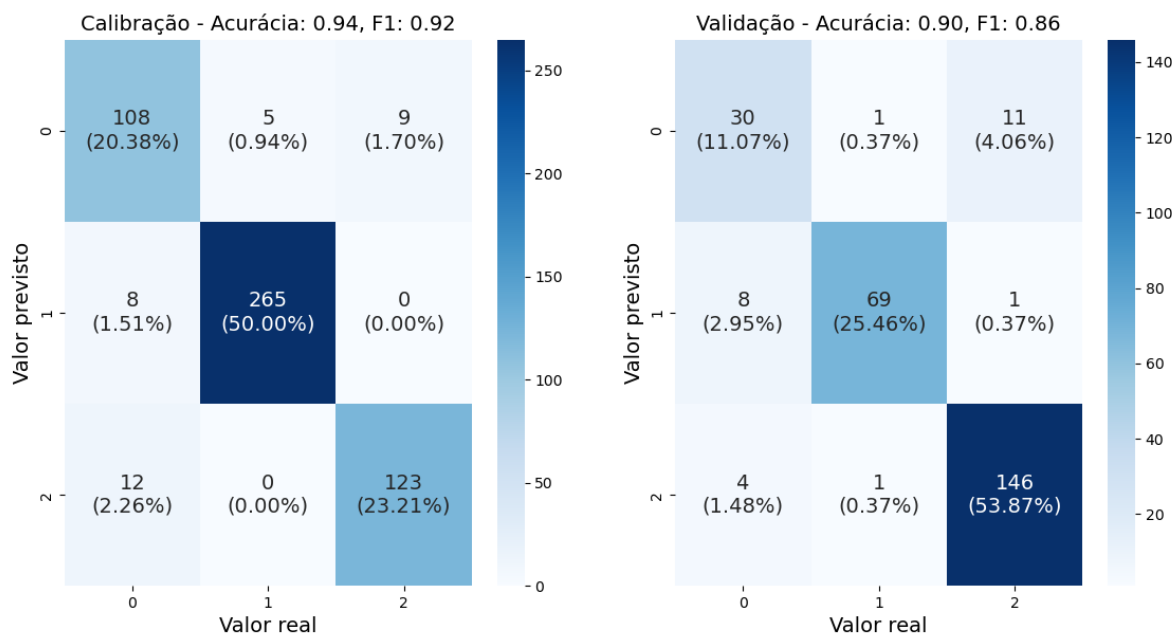
*3: O Par 2 é formado respectivamente pelo AG Soma dos Valores das curvas diárias de potência simulada com rastreamento e pelo AC Razão entre os valores de RMSE de funcionamento e de falha.

*4: O Par 3 é formado respectivamente pelo AG Primeiro Coeficiente da parte real da Transformada de Fourier das curvas diárias de potência simulada com rastreamento e pelo AC Razão entre os valores de RMSE das curvas de erro de funcionamento e de falha.

A Figura 61 apresenta as matrizes de confusão do Par 1 da Tabela 20, correspondentes aos períodos de calibração e validação para o sítio 2. O Par 1 é composto pelos atributos AG Soma dos Valores das curvas diárias de potência simulada com rastreamento e AC razão entre

os valores de RMSE das curvas de erro de funcionamento e de falha. Esse foi o par que obteve o maior valor de MRPA. Ao analisar a classe 2, que representa os dias de falha, pode-se observar os seguintes resultados.

Figura 61 - Matrizes de Confusão para os períodos de calibração e validação do Par 1*1 (maior MRPA) para o sítio 2



Fonte: Elaboração Própria.

O 0 representa os dias indefinidos; o 1, os dias de funcionamento; e o 2, dias de falha.

*1: No sítio 2, o Par 1 possui o maior MRPA e é formado respectivamente pelo AG Soma de Valores das curvas diárias de potência simulada com rastreamento e pelo AC Razão entre os valores de RMSE das curvas de erro de funcionamento e de falha.

Na linha do valor previsto 2 da Tabela 20, o método classificou cerca de 25,47% (135 dias) e 55,72% (151 dias) do total de dias como dias de falha nos períodos de calibração e validação, respectivamente. No entanto, houve uma taxa de erro de 2,26% (12 dias) e 1,85% (5 dias) nessa classificação, o que significa que cerca de 9% e 3% dos dias classificados como dias de falha foram, na verdade, de outras classes. Em relação à coluna do valor real 2, que representa a quantidade total de dias que realmente pertencem à classe de dias de falha, o método classificou aproximadamente 23,21% (123 dias) e 53,87% (146 dias) dos dias como dias de falha nos períodos de calibração e validação, respectivamente. Porém, houve uma taxa de não classificação de cerca de 1,70% (9 dias) e 4,43% (12 dias) do total de dias com falha, respectivamente. Isso significa que aproximadamente 6% e 8% dos dias em que realmente ocorreu uma falha não foram detectados pelo método.

Em resumo, o método de classificação apresentou uma taxa de erro inferior a 10% para os dias classificados como dias de falha e uma taxa de não classificação inferior a 10% para os dias que deveriam ser classificados como dias de falha para o sistema do sítio 2. Em relação à classificação da classe dos dias de funcionamento, vale ressaltar que o comportamento é semelhante ao da classificação dos dias de falha. No entanto, é importante destacar que houve uma taxa de erro mais elevada para a classe de dias indefinidos, especialmente no período de validação, onde cerca de 29% dos valores classificados como dias indefinidos pelo modelo foram incorretos.

A Tabela 21 apresenta os três principais pares de atributos com os maiores valores de MRPA para o Sítio 5. Esses pares foram compostos por atributos de similaridade para a divisão do conjunto de dias em indefinidos e definidos, juntamente com atributos de comparação para a classificação dos dias de falha e funcionamento no subconjunto de dias definidos resultantes. Assim como no Sítio 2, os atributos de comparação demonstraram eficácia na classificação dos dias de falha. Porém, ao contrário do Sítio 2, observa-se uma consistência significativa no desempenho dos pares nos períodos de calibração e validação, ou seja, os parâmetros variaram pouco entre os períodos. Isso se reflete nos altos valores de MRPA, conforme apresentado na Tabela 21. Embora o par 1 tenha ocupado posições inferiores em relação ao par 2, ele apresenta diferenças absolutas no f1 e na acurácia menores, além de ter valores de f1 e acurácia durante a validação que diferem dos valores do par 2 apenas na terceira casa decimal. Portanto, a principal diferenciação entre os pares foi baseada na diferença absoluta entre os parâmetros.

Tabela 21 - Os três melhores Pares de Atributos ordenados pelo MRPA para o Sítio 5

Par de Atributos	Posição CAL*1	Posição VAL*1	f1 CAL*1	f1 VAL*1	Diferença absoluta f1	Acurácia CAL*1	Acurácia VAL*1	Diferença absoluta Acurácia	MRPA
Par 1*2	39	3	0,912	0,911	0,000	0,915	0,918	0,004	5207,737
Par 2*3	34	1	0,912	0,914	0,001	0,916	0,920	0,005	4610,895
Par 3*4	51	5	0,909	0,909	0,000	0,913	0,916	0,004	4307,378

Fonte: Elaboração Própria.

*1: CAL significa período de calibração, enquanto VAL período de validação.

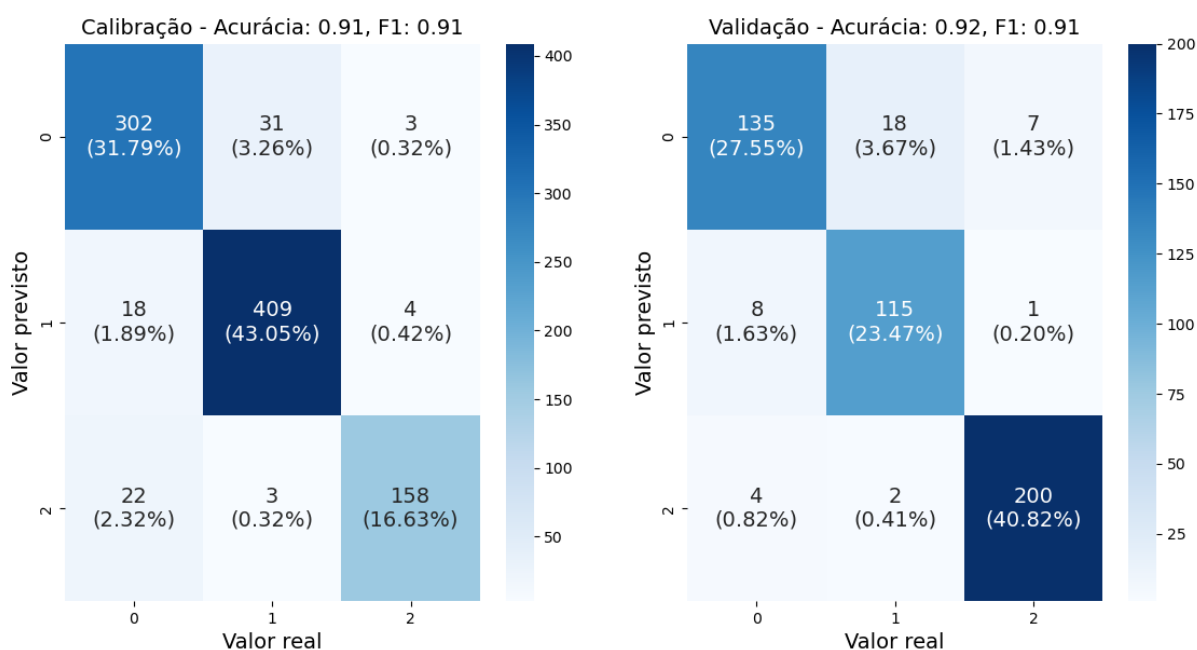
*2: O Par 1 é formado respectivamente pelo AS RMSE entre as curvas diárias de potências simuladas com rastreamento e com o coletor fixo e pelo AC Razão entre os valores de RMSE de funcionamento e de falha.

*3: O Par 2 é formado respectivamente pelo AS MSE entre as curvas diárias de potências simuladas com rastreamento e com o coletor fixo e pelo AC Razão entre os valores de RMSE de funcionamento e de falha.

*4: O Par 3 é formado respectivamente pelo AS MSE entre as curvas diárias de potências simuladas com rastreamento e com o coletor fixo e pelo AC Razão entre as Áreas entre Curvas de funcionamento e de falha.

A Figura 62 apresenta as matrizes de confusão do par de maior MRPA da Tabela 21, o Par 1, correspondentes aos períodos de calibração e validação para o sítio 5. O Par 1 é composto pelos atributos AG Soma dos Valores das curvas diárias de potência simulada com rastreamento e AC razão entre os valores de RMSE das curvas de erro de funcionamento e de falha. Mais uma vez, focou-se a análise para a classificação do método na classe 2.

Figura 62 - Matrizes de Confusão para os períodos de calibração e validação do Par 1*1 (maior MRPA) para o sítio 5



Fonte: Elaboração Própria.

O 0 representa os dias indefinidos; o 1, os dias de funcionamento; e o 2, dias de falha.

*1: No sítio 5, o Par 1 possui o maior MRPA e é formado respectivamente pelo AS RMSE entre as curvas diárias de potências simuladas com rastreamento e com o coletor fixo e pelo AC Razão entre os valores de RMSE de funcionamento e de falha.

Na linha do valor previsto 2, durante a calibração, o método classificou cerca de 19,32% (183 dias) do total de dias como dias de falha, dos quais 2,64% (25 dias) na verdade pertenciam a outras classes. Isso representa que aproximadamente 14% dos dias foram erroneamente classificados pelo método. No período de validação, entretanto, esse percentual foi inferior a 3%, pois o método classificou 42,05% (206 dias) do total como dias de falhas e apresentou uma taxa de erro de 1,23% (6 dias). Esses valores mostram diferenças em relação ao sítio 2, onde os percentuais de erro durante os períodos de calibração e validação foram inferiores a 10%.

Agora, em relação à coluna do valor real 2, que representa a quantidade total de dias de falha, apenas 4% dos dias não foram detectados pelo método nos dois períodos. Isso significa

que, durante a calibração e a validação, a porcentagem real de dias de falha foi de 17,37% (165 dias) e 42,45% (208 dias), respectivamente, com o método deixando de detectar 0,74% (7 dias) e 1,63% (8 dias) desses dias.

Em outras palavras, para o sítio 5, o método de classificação demonstrou uma alta eficiência em detectar os dias de falha, deixando apenas uma porcentagem inferior a 2% de dias em que realmente ocorreu uma falha e que não foram detectados nos períodos de calibração e validação, o que é significativamente menor do que o observado no sítio 2. No entanto, embora tenha apresentado uma taxa de erro baixa em relação aos dias que o método identificou como falha durante a validação, essa taxa atingiu 15% durante a calibração, um valor maior do que o encontrado no sítio 2. Ainda assim, o método mostrou-se particularmente eficiente para a identificação dos dias de falha.

Percebe-se, portanto, que durante a avaliação foi dada mais importância à eficiência do par de atributos de maior MRPA na detecção dos dias de falha. Isso ocorre porque o objetivo do método de detecção é inferir os índices probabilísticos de confiabilidade, e, nesse contexto, os dias de falha são a classe de interesse primordial. Portanto, é evidente que o método demonstrou eficiência nessa tarefa.

5.2 Ajuste dos Dados de Ocorrência de Falha para Inferência dos Índices Probabilísticos de Confiabilidade

A distribuição temporal das ocorrência de falhas é necessária para a inferência dos índices probabilísticos de confiabilidade, isto é, saber em quais dias ocorreram falhas ou não no sistema de rastreamento durante o período de tempo analisado. Assim, a partir da classificação dos dias, o tempo entre falhas, tempo de reparo e tempo de operação serão empregados, conforme os pressupostos demonstrados na Figura 45 da Seção 4.2.

A partir disso, o primeiro ponto da inferência dos índices probabilísticos de confiabilidade é justamente determinar se o sistema será modelado a partir de um Processo Ordinário de Renovação (ORP) ou por um Processo de Renovação Alternada (ARP), conforme explicado na Seção 4.2. Uma boa métrica para isso é a média aritmética das proporções entre tempo de reparo $t_{R,i}$ e tempo entre falhas x_i , justamente para identificar a porcentagem temporal do reparo por ciclo de renovação.

Nesse sentido, utilizando os dados de ocorrência de falhas oriundos da rotulação visual, essas porcentagens foram de 40,92% e 33,05% para os sítios 2 e 5 do DKASC, respectivamente. É possível observar, portanto, que o tempo de reparo durante um ciclo de renovação não é desprezível, assim o sistema será modelado a partir de um Processo de Renovação Alternada.

O principal parâmetro de um ARP é justamente a disponibilidade do sistema. Para isso, faz-se necessário estimar as distribuições relacionadas aos tempos de operação e reparo. Nesse contexto, primeiramente serão estimados os coeficientes da Weibull de dois parâmetros para os processos relacionados aos tempos de operação e de reparo a partir da metodologia explicada na Seção 4.2. A estimativa dos coeficiente se dará utilizando tanto o conjunto de ocorrências de falha oriundos da rotulação visual, como do método do limiar móvel justamente para verificar se a utilização dos dados oriundos do método proposto neste trabalho altera os índices probabilísticos se comparado com a referência, que é a rotulação visual. Após o ajuste dos coeficientes e a análise comparativa entre conjuntos de dados, os índices probabilísticos serão inferidos.

5.2.1 Ajuste dos Coeficientes de Weibull e Comparação dos Conjuntos de Dados

A primeira análise é dos tempos de reparo. Assim, o primeiro passo é ordenar os tempos de reparo oriundos das ocorrências de falhas da rotulação visual e do método do limiar móvel, para assim calcular os valores estimados da CDF pela aproximação de Bernard do ranqueamento da mediana, $\hat{F}_{R,i}$. Para valores de tempo de reparo repetidos, será tomada a mediana do subconjunto. Os valores $\hat{F}_{R,i}$ serão utilizados como referência para os ajustes dos coeficientes da CDF de Weibull de dois parâmetros. Os parâmetros estatísticos de avaliação mostrados nos resultados desta seção sempre serão a comparação de $\hat{F}_{R,i}$ em relação a CDF de Weibull com os coeficientes ajustados (Equação 42), e não de seus fatores linearizados.

A Tabela 22 apresenta os tempos de reparo ordenados para os sítios 2 e 5, separados de acordo com a origem dos dados, seja da Rotulação Visual ou do Método Limiar Móvel. Nessa tabela, também é fornecida a quantidade de vezes que cada valor de tempo de reparo se repetiu no conjunto analisado. É importante destacar que aos valores nulos de tempo de reparo foi atribuído o valor de 0,5, tendo em vista que, durante o processo de ajuste linear, é necessário calcular o logaritmo natural, que não está definido para o valor zero. Além disso, é interessante observar que os tempos de reparo dentro de um mesmo sítio apresentam valores muito próximos entre si. De maneira geral, a maioria dos valores de tempo de reparo se concentra em intervalos

pequenos, indicando que a maioria das ocorrências de reparo ocorreu em períodos de tempo curtos.

Tabela 22 - Ordenamento dos Tempos de Reparo e inferência dos valores da CDF estimada pela aproximação de Bernard do ranqueamento da mediada, por sítio e por conjunto de origem dos dados de falha (Rotulação Visual e Método Limiar Móvel)

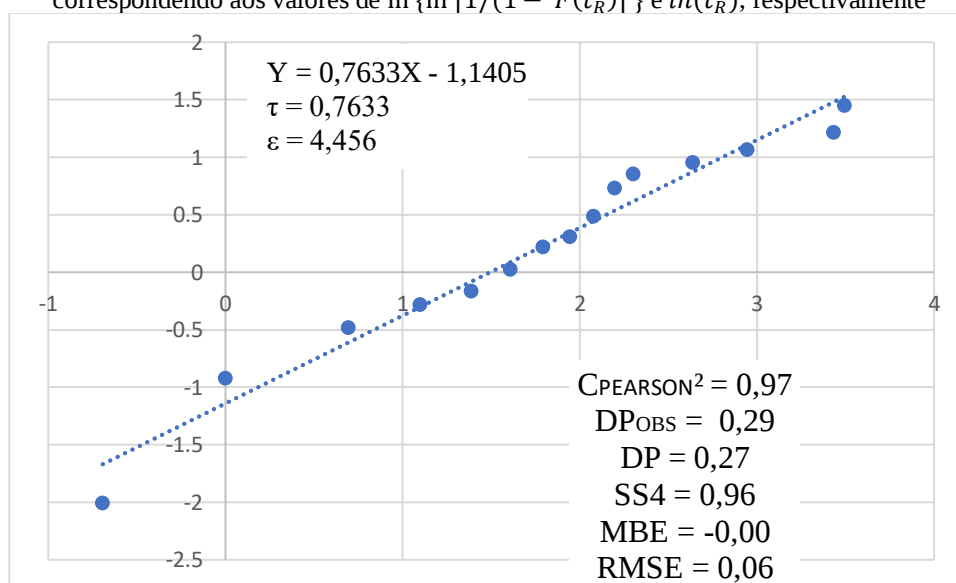
Sítio 2						Sítio 5					
Rotulação Visual			Método Limiar Móvel			Rotulação Visual			Método Limiar Móvel		
$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$
0,5	12	0,13	0,5	11	0,12	0,5	23	0,18	0,5	31	0,21
1	8	0,33	1	8	0,31	1	10	0,43	1	13	0,51
2	5	0,46	2	3	0,43	2	10	0,58	2	6	0,63
3	2	0,53	3	4	0,50	3	3	0,68	3	4	0,70
4	2	0,57	4	2	0,56	4	4	0,74	4	6	0,77
5	5	0,64	5	3	0,61	6	3	0,79	5	1	0,82
6	2	0,71	6	4	0,69	7	1	0,82	6	1	0,83
7	1	0,74	7	2	0,75	8	2	0,84	7	1	0,84
8	5	0,80	8	2	0,79	12	2	0,87	8	1	0,86
9	2	0,87	9	2	0,83	13	1	0,90	9	1	0,87
10	1	0,90	10	1	0,86	14	1	0,91	12	1	0,88
14	1	0,93	11	2	0,89	18	1	0,93	13	2	0,90
19	1	0,95	15	1	0,92	20	1	0,94	14	1	0,92
31	1	0,97	17	1	0,94	21	1	0,96	18	1	0,94
33	1	0,99	22	1	0,96	35	1	0,97	21	1	0,95
			31	1	0,99	67	1	0,99	24	1	0,96
									27	1	0,98
									67	1	0,99

Fonte: Elaboração Própria.

*¹ Qtd representa a quantidade de vezes que determinado valor de tempo de reparo se repetiu no conjunto.

Para o sítio 2, as Figuras 63 e 64 apresentam o sistema cartesiano cujo eixo das ordenadas é composto pelos valores de $\ln \left[\ln \left(\frac{1}{1 - F(t_R)} \right) \right]$ e cujo eixo das abscissa é composto pelos valores de $\ln(t_R)$, justamente os fatores de linearização da CDF de Weibull de dois parâmetros, apresentados na Seção 4.2. Nesse sentido, se os pontos que representam o par $(t_{R,i}; \hat{F}_{R,i})$ apresentarem um comportamento linear o ajuste dos coeficientes será bom.

Figura 63 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas da Rotulação Visual para o Sítio 2, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente



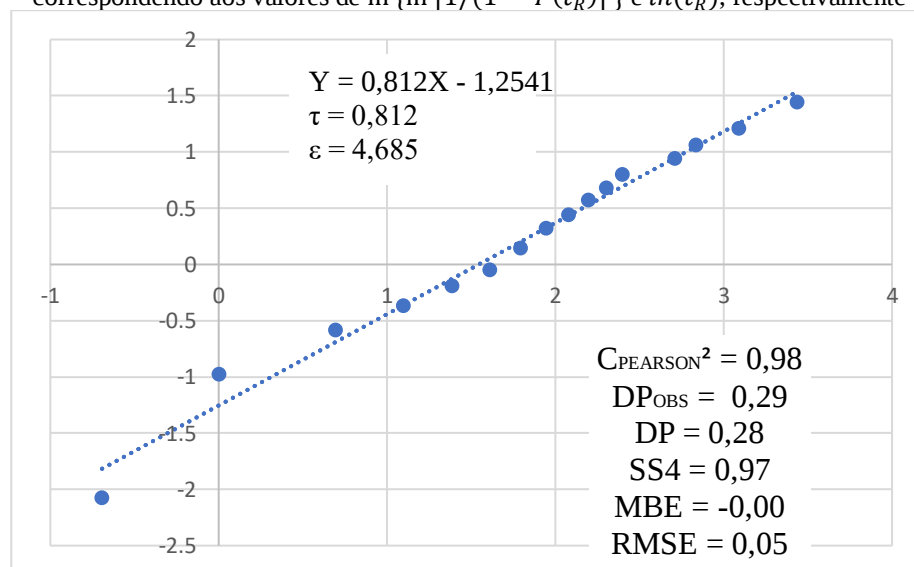
Fonte: Elaboração Própria.

É possível ver nas figuras que o comportamento dos pares ordenados apresenta um comportamento bastante próximo do linear. Inclusive, isso é corroborado pelos parâmetros estatísticos de avaliação do ajuste, mostrados nas figuras. No sítio 2, para os tempos de reparo oriundos da Rotulação Visual (Figura 63) o SS4 foi de 0,96, enquanto que para os oriundos da Método Limiar Móvel (Figura 64) esse valor foi de 0,97. Os valores do SS4 estão muito próximos de 1, indicando a qualidade do ajuste. Também é interessante notar que o RMSE e o MBE também apresentaram valores pequenos. Outro ponto a se destacar é que os fatores de escala τ e de forma ε para ambos conjuntos de dados deram valores semelhantes.

Por sua vez, as Figuras 65 e 66 mostram os ajustes para os tempos de reparo obtidos pela Rotulação Visual e pelo Método do Limiar Móvel, respectivamente, no sítio 5. Assim como no sítio 2, o valor de SS4 é muito próximo de 1, aproximadamente 0,95 em ambos os casos. O RMSE e o MBE também apresentaram valores baixos, reforçando a qualidade do ajuste. É importante destacar que, mais uma vez, os coeficientes estimados tiveram valores

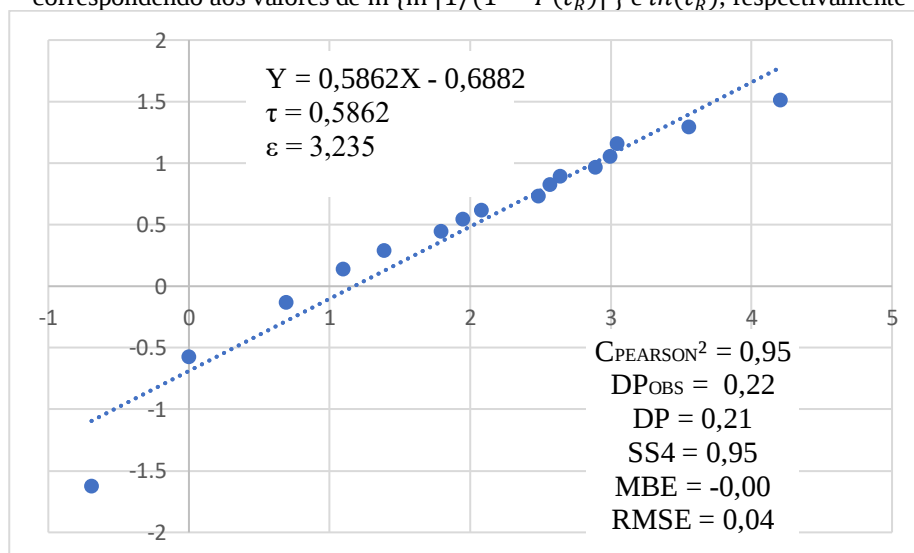
próximos entre si para os diferentes conjuntos de tempo de reparo, sugerindo que o comportamento do tempo de reparo está bem representado pelo método desenvolvido.

Figura 64 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas da Método do Limiar Móvel para o Sítio 2, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente



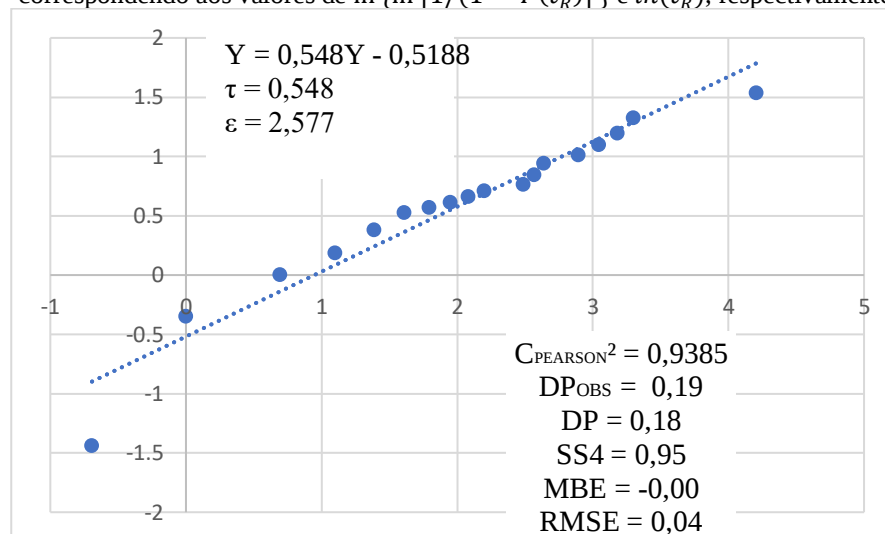
Fonte: Elaboração Própria.

Figura 65 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas da Rotulação Visual para o Sítio 5, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente



Fonte: Elaboração Própria.

Figura 66 – Representação do ajuste linear dos coeficientes τ e ε da distribuição de Weibull para os Tempos de Reparo estimados a partir do conjunto de falhas do Método do Limiar Móvel para o Sítio 5, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_R))] \}$ e $\ln(t_R)$, respectivamente



Fonte: Elaboração Própria.

Por fim, a Tabela 23 apresenta os parâmetros estatísticos da comparação entre as CDFs ajustadas para os tempos de reparo oriundos da Rotulação Visual e do Método Limiar Móvel. Nesse caso, o SS4 é aproximadamente 1, indicando que os tempos de reparo estimados pelo método desenvolvido conseguem representar bem a dinâmica do sistema.

Tabela 23 – Comparação das CDFs ajustadas aos Tempos de Reparo oriundos dos dados da Rotulação Visual e do Método Limiar Móvel

Estatísticos	CDF Rotulação Visual vs CDF Método Limiar Móvel	
	Sítio 2	Sítio 5
$C_{PEARSON}$	1,00	1,00
$DP_{Rotulação}$	0,23	0,16
$DP_{Método}$	0,23	0,15
SS4	1,00	0,99
RMSE	0,01	0,02
MBE	0,00	0,01

Fonte: Elaboração Própria.

Em relação ao tempo de operação, a Tabela 24 apresenta o resultado do processo de construção dos valores $\hat{F}_{R,i}$ utilizados como referência para os ajustes dos coeficientes da CDF de Weibull. É interessante notar que o conjunto de dados para o sítio 5 possui mais elementos em comparação ao sítio 2, o que é explicado pelo fato de que a quantidade de dias

analisados para o sítio 5 é quase o dobro da do sítio 2. Outro ponto relevante é que, principalmente para o sítio 5, o conjunto de tempos de operação oriundos do Método Limiar Móvel é maior do que aquele oriundo da Rotulação Visual. Por fim, assim como nos tempos de reparo, é possível notar que a maioria dos tempos de operação se concentram em valores pequenos, o que sugere que a frequência de operação em períodos curtos é mais alta.

Tabela 24 - Ordenamento dos Tempos de Operação e inferência dos valores da CDF estimada pela aproximação de Bernard do ranqueamento da mediada, por sítio e por conjunto de origem dos dados de falha (Rotulação Visual e Método Limiar Móvel)

Sítio 2						Sítio 5						Continua
Rotulação Visual			Método Limiar Móvel			Rotulação Visual			Método Limiar Móvel			
$t_{R,i}$	Qtd*1	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd*1	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd*1	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd*1	$\hat{F}_{R,i}$	
2	22	0,226721	2	19	0,200413	2	24	0,19	2	26	0,177419	
3	12	0,57085	3	13	0,530992	3	11	0,45	3	13	0,439516	
4	7	0,763158	4	7	0,737603	4	7	0,59	4	8	0,580645	
5	4	0,874494	5	1	0,820248	5	6	0,69	5	5	0,668011	
6	1	0,925101	6	2	0,85124	6	2	0,75	6	2	0,715054	
8	1	0,945344	7	2	0,892562	8	4	0,80	7	1	0,735215	
82	1	0,965587	8	1	0,923554	12	1	0,84	8	2	0,755376	
364	1	0,98583	36	1	0,944215	23	1	0,85	12	1	0,775538	
			47	1	0,964876	29	1	0,87	13	1	0,788978	
			364	1	0,985537	30	1	0,88	16	1	0,802419	
						36	1	0,90	21	1	0,81586	
						40	1	0,91	23	1	0,829301	
						63	1	0,93	25	1	0,842742	

Fonte: Elaboração Própria.

*1 Qtd representa a quantidade de vezes que determinado valor de tempo de reparo se repetiu no conjunto

Tabela 25 - Ordenamento dos Tempos de Operação e inferência dos valores da CDF estimada pela aproximação de Bernard do ranqueamento da mediada, por sítio e por conjunto de origem dos dados de falha (Rotulação Visual e Método Limiar Móvel)

Sítio 2						Sítio 5						Conclusão
Rotulação Visual			Método Limiar Móvel			Rotulação Visual			Método Limiar Móvel			
$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$	$t_{R,i}$	Qtd* ¹	$\hat{F}_{R,i}$	
						143	1	0,94	27	1	0,856183	
						183	1	0,96	31	1	0,869624	
						218	1	0,97	37	1	0,883065	
						226	1	0,99	43	1	0,896505	
									44	1	0,909946	
									52	1	0,923387	
									62	1	0,936828	
									63	1	0,950269	
									104	1	0,96371	
									193	1	0,977151	
									240	1	0,990591	

Fonte: Elaboração Própria.

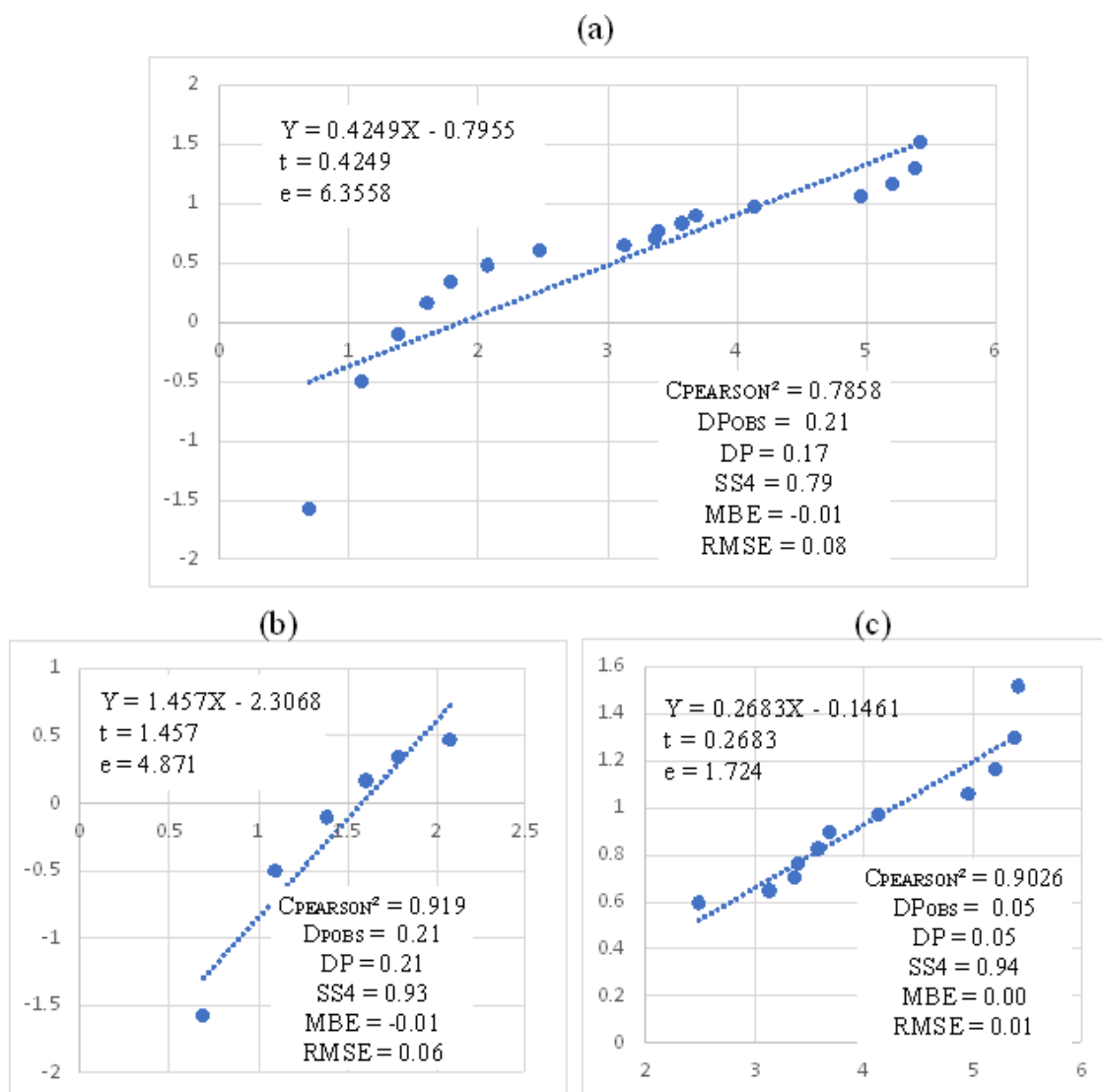
*¹ Qtd representa a quantidade de vezes que determinado valor de tempo de reparo se repetiu no conjunto

Partindo para o ajuste dos coeficientes para os tempos de operação, foi notado que o ajuste linear não apresentou bons resultados, conforme pode ser visto na Figura 67a, na qual os valores de SS4 e $C_{PEARSON}^2$ apresentaram aproximadamente de 0,79 para os dados oriundos da Rotulação Visual do Sítio 5. Esse fenômeno foi possível verificar para os dois conjuntos de dados de ambos os sítios.

A princípio, isso é um indicativo de que a distribuição do conjunto não segue um comportamento monotônico. Assim, já que se está utilizando uma distribuição de Weibull de dois parâmetros, seria necessário particionar o conjunto e ajustar localmente os coeficientes para assim representar a não monotonicidade dos dados. Nesse sentido, ao particionar os conjuntos, o resultado dos ajustes locais gerou distribuições com fator de escala τ maiores que

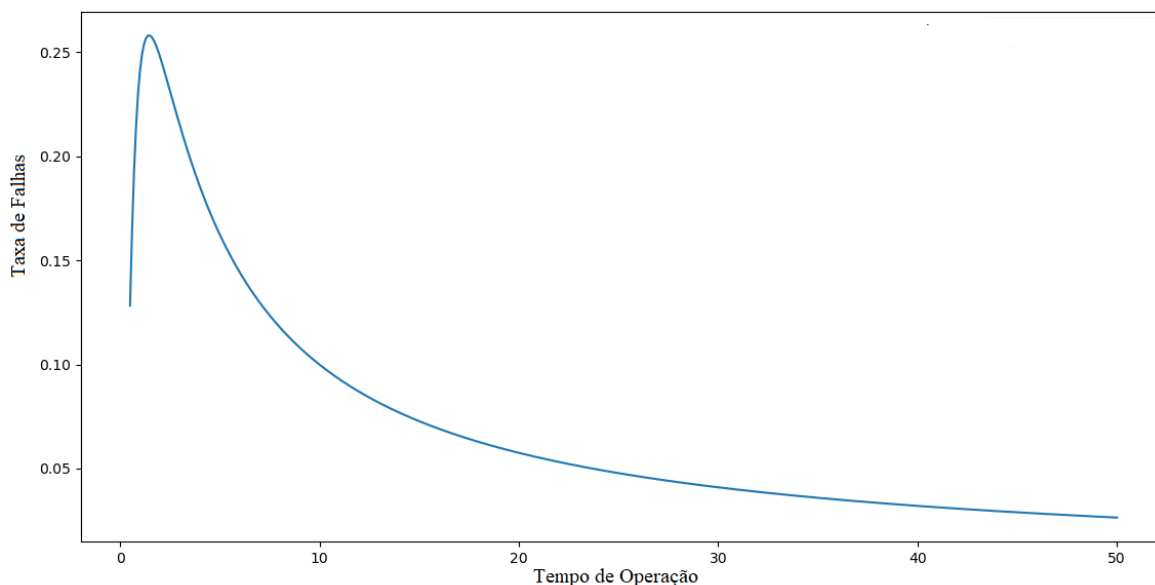
1 para a primeira parte do conjunto e menores que 1 para a segunda parte, conforme pode ser visto na Figura 67b e 67c. Isso, porém, representa uma inconsistência física, pois indica que a função taxa de falhas no início apresenta um comportamento ascendente para depois descender, em contraposição ao que afirma a teoria por trás da curva banheira. (Fatores de escala menores que 1 indicam um comportamento decrescente, maiores que 1 ascendente e iguais a 1 constante). Então, nesse sentido, a partir da informação dos dados, a Função Taxa de Falha teria um formato semelhante ao da Figura 68, obtida a partir de um ajuste não linear dos dados a uma CDF de Weibull de 4 parâmetros modificada exponencialmente (SARHAN; APALOO, 2013).

Figura 67 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 5 oriundos dos dados da Rotulação Visual, e representação dos ajustes locais a partir da partição das amostras, com os eixos y e x correspondendo aos valores de $\ln \{1/(1 - F(t_o))\}$ e $\ln(t_o)$, respectivamente



Fonte: Elaboração Própria.

Figura 68 – Função Taxa de Falhas para o Tempo de Operação para o Sítio 5 obtida do ajuste da distribuição de Weibull de 4 parâmetros modificada com os dados originados da Rotulação Visual sem tratamento



Fonte: Elaboração Própria.

Nesse sentido, a Figura 68 apresenta pistas do entendimento físico da problemática. É possível perceber que durante a primeira partição a função taxa de falhas está tendendo a um valor máximo a partir do qual ela irá começar a cair. Contudo, esse valor máximo deveria estar representado no conjunto de dados desde o início, o que não está e por isso na primeira partição o método ascende até ele. É possível inferir, então, que o que ocorre, na verdade, é que o método de ajuste não consegue representar bem que existe esse valor máximo desde o início.

Assim sendo, para abordar essa questão, adotou-se o seguinte procedimento: primeiramente, utilizaram-se os coeficientes ajustados na primeira partição para calcular os valores da função taxa de falhas na região de ascendência, permitindo estimar o valor máximo. Com base nos coeficientes, o cálculo da função taxa de falhas foi feito por meio da Equação 45. O valor máximo foi então estimado utilizando o terceiro quartil associado aos valores de $\lambda(t_{O,i})$. Em seguida, utilizou-se o tempo de operação associado ao terceiro quartil de $\lambda(t_{O,i})$ para determinar o valor de $F(t_{O,i})$ associado, empregando os coeficientes ajustados na primeira partição e inserindo-os na Equação 42. Por fim, o conjunto dos valores de tempo de operação da primeira partição será substituído pelo da CDF associado ao terceiro quartil.

A Tabela 26 traz um exemplo desse procedimento para o conjunto de dados de tempos de operação oriundos da Rotulação Visual do Sítio 5. Os coeficientes de escala τ e de forma ε são 1,457 e 4,871, conforme pode ser visto na Figura 67b. Os valores de $\lambda(t_{O,i})$ e de $F(t_{O,i})$ foram calculados utilizando as Equações 45 e 42, respectivamente. E por fim, os pares

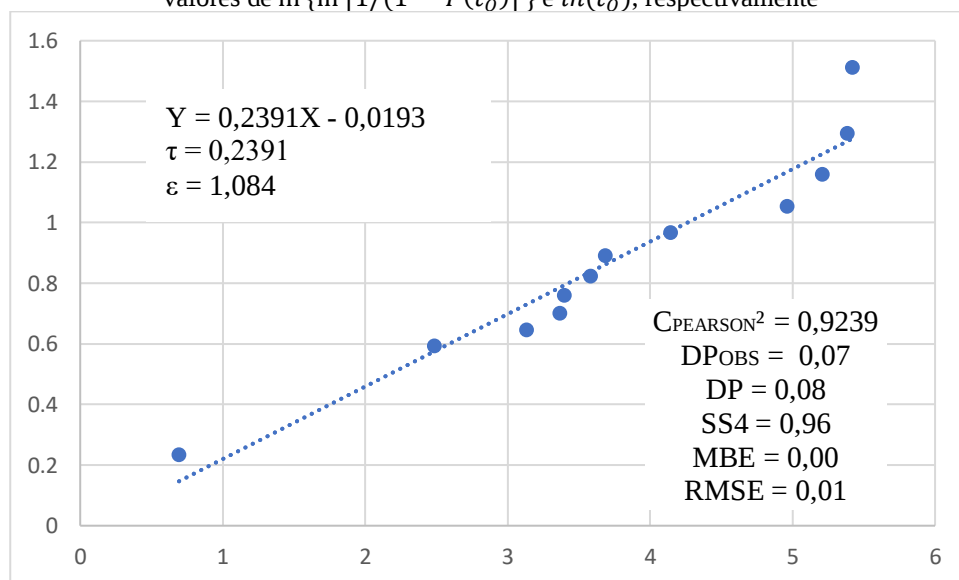
ordenados da primeira partição foram substituídos por (2;0,72). Assim, o ajuste dos coeficientes já parte de uma estimativa do valor máximo da função taxa de falhas. A Figura 69 apresenta o resultado do ajuste após o tratamento dos dados para o conjunto discutido neste parágrafo, evidenciando que o SS4 e a $C_{PEARSON}^2$ tiveram uma significativa melhora.

Tabela 26 – Valores do ajuste ascendente da primeira partição do conjunto de tempos de operação oriundos da Rotulação Visual para o Sítio 5

Tempo de Operação, $t_{o,i}$	Valor da função taxa de falhas $\lambda(t_{o,i})$	Valor associado da CDF, $F(t_{o,i})$	3° Quartil	
			$\lambda(t_{o,i})$	$F(t_{o,i})$
2	0,20	0,24		
3	0,24	0,39		
4	0,27	0,53	0,32	0,72
5	0,30	0,65		
6	0,33	0,74		

Fonte: Elaboração Própria.

Figura 69 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 5 oriundos dos dados da Rotulação Visual, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_o))] \}$ e $\ln(t_o)$, respectivamente



Fonte: Elaboração Própria.

É importante destacar, porém, que essa inconsistência física apenas apareceu porque a partição do conjunto de amostras para ajuste local acarretou um comportamento de ascendência seguido de descendência da função taxa de falhas. Se, contudo, fosse o oposto, isso indicaria

que a função estaria seguindo o comportamento da curva banheira. Assim, os ajustes locais permitiriam representar as duas regiões da função taxa de falhas.

Para todos os conjuntos de tempos de operação oriundos da Rotulação Visual e do Método do Limiar Móvel nos sítios 2 e 5, fez-se necessário realizar a partição das amostras devido ao indicativo de não monotonicidade, evidenciado pelos parâmetros estatísticos do ajuste. A Tabela 27 apresenta os valores do SS4 do ajuste global. Especialmente para o sítio 2, o ajuste linear mostrou-se inefetivo.

Durante a partição, foi observada a inconsistência física relacionada à função taxa de falha, com fatores de escala assumindo valores maiores que 1 na primeira partição e menores que 1 na segunda partição, representando a ascendência e descendência previamente mencionadas. Então, o tratamento dos dados da primeira partição foi aplicado, e as amostras da primeira partição foram substituídos pelo par associado ao valor da CDF que representa a estimativa do valor máximo da função taxa de falhas, como pode ser visto na última coluna da Tabela 27.

Tabela 27 - Análise das partições dos conjuntos de Tempos de Operação e aplicação do procedimento de tratamento dos dados

		SS4	Fator de	Fator de	Intervalo	$\lambda(t_o)$	Par ordenado
		Conjunto	escala τ	escala τ	de t_o da	máximo	$(t_o; F(t_o))$
		Total	Partição 1	Partição 2	Partição 1	estimado	da Partição 1
Sítio 2	Rotulação Visual	0,41	2,28	0,11	[2;6]	0,79	(2; 0,90)
	Método Limiar Móvel	0,51	1,97	0,15	[2;6]	0,66	(2; 0,81)
	Rotulação Visual	0,79	1,46	0,27	[2;8]	0,32	(2; 0,72)
Sítio 5	Método Limiar Móvel	0,85	1,68	0,36	[2;6]	0,37	(2; 0,67)

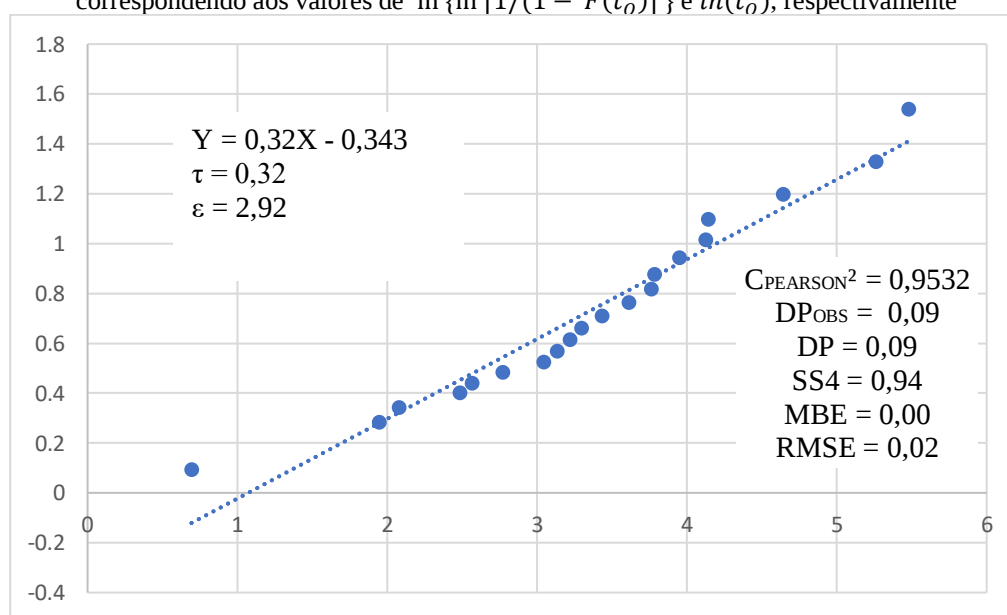
Fonte: Elaboração Própria.

É interessante notar que os períodos de operação relacionados à primeira partição estão consistentemente entre 2 dias e 6 ou 8 dias de operação, conforme evidenciado pelos intervalos

de t_o apresentados na tabela. Os dias contidos nesses intervalos representam períodos curtos de operação e contêm uma grande quantidade de amostras nos conjuntos, conforme pode ser visto nas primeiras linhas da Tabela 24. Esse fato reforça a ideia apresentada anteriormente para o tratamento dos dados da partição 1, tendo em vista que, devido aos períodos de operação serem curtos e as frequências de ocorrência de amostras serem elevadas, a função taxa de falha apresenta, nesses intervalos, valores elevados desde o início que não estão sendo adequadamente representados pela metodologia de ajuste. Portanto, a hipótese de substituir os valores da primeira partição por um valor associado à estimativa do valor máximo da função taxa de falhas se mostra válida.

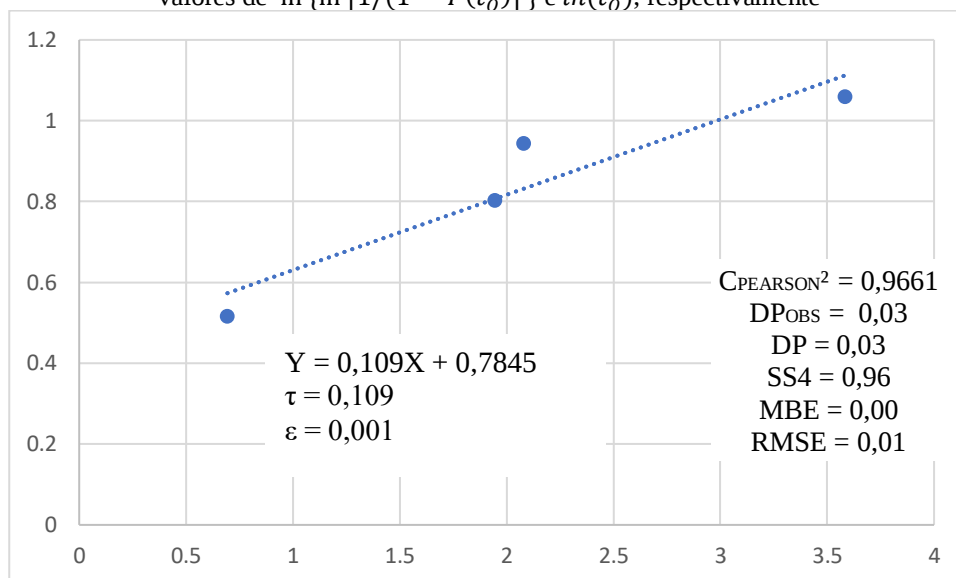
Os resultados dos ajustes de coeficientes da CDF de Weibull para os demais conjuntos de dados de tempos de operação dos sítios 2 e 5, após o tratamento dos dados relacionados à partição 1, são apresentados nas Figuras 70, 71 e 72. Observa-se que o tratamento dos dados permitiu que os ajustes se mostrassem eficazes, o que é evidenciado pelos valores dos parâmetros estatísticos de avaliação. No conjunto de tempos de operação oriundos da Rotulação Visual para o sítio 2, é possível perceber uma quantidade menor de dados em comparação aos demais conjuntos, pois o tratamento de dados da partição 1 resultou em uma redução total do conjunto de 8 para 4 amostras. Apesar dessa redução, nota-se que a tendência da função taxa de falhas se manteve, mostrando uma descendência a partir de um valor alto de taxa de falha.

Figura 70 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 5 oriundos dos dados do Método do Limiar Móvel, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_o))] \}$ e $\ln(t_o)$, respectivamente



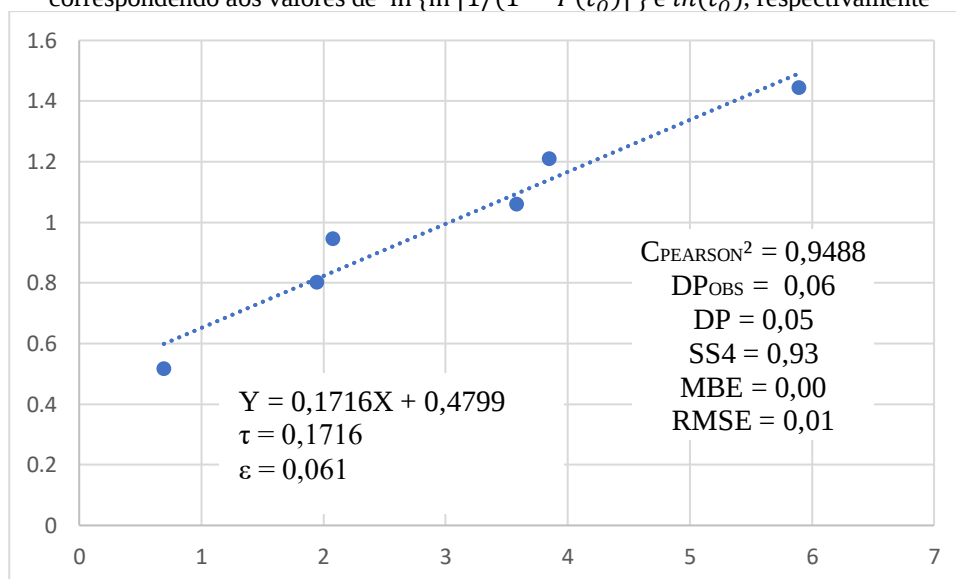
Fonte: Elaboração Própria.

Figura 71 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 2 oriundos dos dados da Rotulação Visual, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_o))] \}$ e $\ln(t_o)$, respectivamente



Fonte: Elaboração Própria.

Figura 72 – Representação do ajuste dos coeficientes de Weibull para os Tempos de Operação do Sítio 2 oriundos dos dados do Método do Limiar Móvel, após o tratamento dos dados, com os eixos y e x correspondendo aos valores de $\ln \{ \ln [1/(1 - F(t_o))] \}$ e $\ln(t_o)$, respectivamente



Fonte: Elaboração Própria.

A Tabela 28 apresenta o resultado da comparação entre os ajustes dos coeficientes da CDF de Weibull para os tempos de operação dos sistemas de rastreamento dos sítios 2 e 5, utilizando as ocorrências de falha da Rotulação Visual e do Método do Limiar Móvel. Nesse caso, os parâmetros de avaliação indicam que os tempos de operação estimados pelo método desenvolvido conseguem representar bem a dinâmica do sistema.

Tabela 28 – Comparação das CDFs ajustadas aos Tempos de Operação oriundos dos dados da Rotulação Visual e do Método Limiar Móvel

Estatísticos	CDF Rotulação Visual vs CDF Método Limiar Móvel	
	Sítio 2	Sítio 5
$C_{PEARSON}$	0,99	1,00
$DP_{Rotulação}$	0,05	0,16
$DP_{Método}$	0,06	0,15
SS4	0,97	0,99
RMSE	0,01	0,02
MBE	0,00	0,01

Fonte: Elaboração Própria.

Utilizando as ocorrências de falha oriundas da Rotulação Visual e do Método do Limiar Móvel, foram ajustados os coeficientes da distribuição de Weibull de 2 parâmetros para os sítios 2 e 5. Ao verificar a eficácia desses ajustes, procedeu-se com a comparação das descrições das CDF de Weibull obtidas a fim de verificar a eficácia do método desenvolvido neste trabalho em representar a dinâmica do sistema. Os resultados obtidos apontaram para ajustes similares entre o método desenvolvido e a referência, que é a Rotulação Visual, evidenciando a capacidade do método em modelar os dados analisados. Assim, com base nas análises e comparações realizadas, foi possível concluir que o método é eficiente na representação dos tempos de operação e reparo dos sistemas, proporcionando uma boa descrição das suas distribuições.

5.2.2 Índices Probabilísticos de Confiabilidade dos Sistemas

Os coeficientes empregados na distribuição de Weibull para cada sistema de rastreamento dos sítios 2 e 5 foram estimados seguindo o processo descrito na seção anterior. As análises revelaram que a utilização do Método do Limiar Móvel desenvolvido neste trabalho é eficiente na representação das distribuições dos tempos de reparo e operação dos sistemas. Assim, as Curvas de Manutenção, Funções Taxa de Reparo, Curvas de Confiabilidade e Funções Taxa de Falhas apresentadas nesta seção utilizaram os coeficientes obtidos a partir dos dados do método de classificação. Em relação ao sítio 2, os fatores de escala τ e de forma ε da distribuição de Weibull correspondem aos valores apresentados nas Figuras 64 e 72. Já para o sítio 5, como o conjunto de dados disponíveis abrangeu um intervalo maior que os 1501 dias

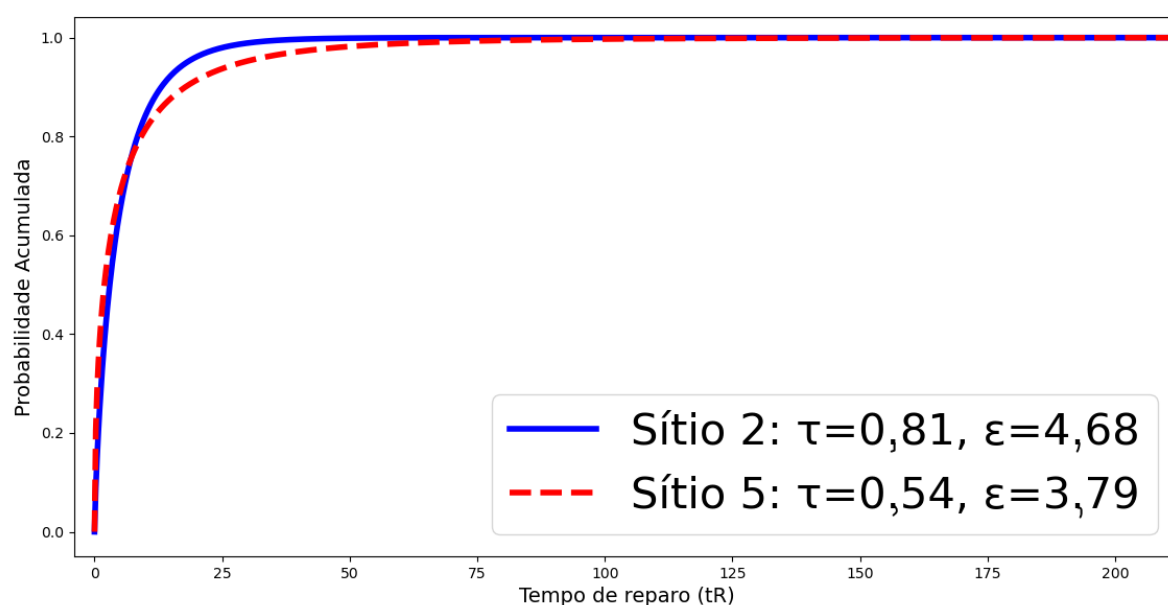
da Rotulação Visual, indo de setembro de 2008 a julho de 2022, a partir da classificação obtida pelo Método do Limiar Móvel, os demais dias foram incluídos e utilizados para um novo ajuste dos parâmetros de Weibull, o qual foi empregado nesta seção para a inferência dos índices probabilísticos do sistema do Sítio 5.

A Figura 73 apresenta as Curvas de Manutenção para os rastreadores dos sítios 2 e 5. A Curva de Manutenção apresenta a probabilidade de que um reparo tenha ocorrido até o tempo t_R . É possível notar que, para o sítio 2, a probabilidade de que um reparo ocorra em menos tempo é levemente maior que para o sítio 5.

Por sua vez, a Figura 74 apresenta as Funções Taxa de Reparo para os sítios 2 e 5. O declínio associado à Função Taxa de Reparo do sítio 5 é maior do que a do sítio 2, indicando que o rastreador do sítio 2 necessita de menos reparos. Em ambos os sítios, as taxas de reparo apresentam um comportamento decrescente, indicando que a quantidade de reparos diminui ao longo do tempo e também que a maioria dos reparos estão concentrados em períodos de tempo pequenos.

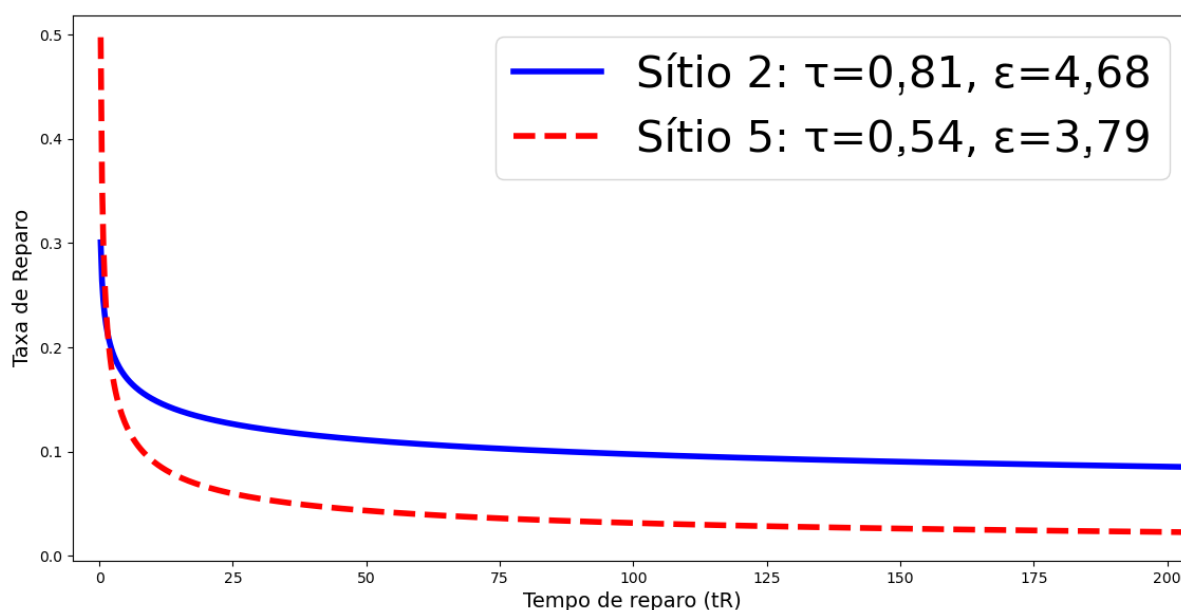
Por fim, as médias dos tempos de reparo, μ_R , calculadas pela Equação 50, foram de 5,26 e 6,66 dias para os sítios 2 e 5, respectivamente. Esses valores demonstram que, no que concerne ao reparo, os dois sistemas apresentam comportamentos parecidos.

Figura 73 – Curva de Manutenção do sistema de rastreamento dos Sítios 2 e 5



Fonte: Elaboração Própria.

Figura 74 - Função Taxa de Reparo do sistema de rastreamento dos Sítios 2 e 5



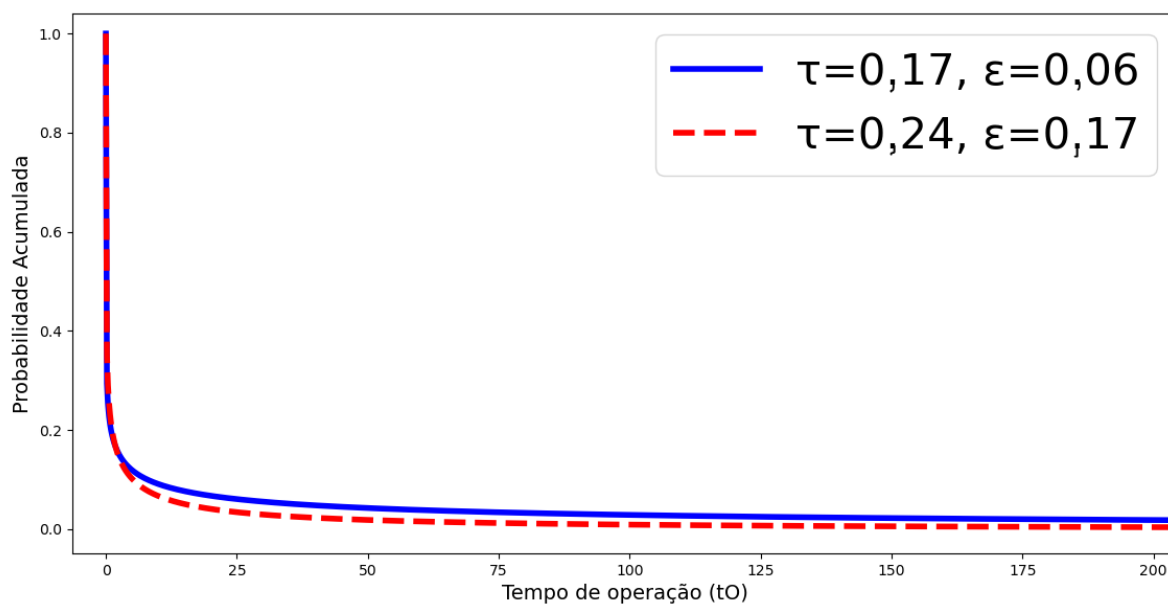
Fonte: Elaboração Própria.

Abordando agora os tempos de operação, a Figura 75 apresenta as Curvas de Confiabilidade para os sistemas de rastreamento dos sítios 2 e 5. Ambas as Curvas de Confiabilidade revelam que os sistemas apresentam uma alta concentração de períodos curtos de operação. Nesse sentido, a expectativa de que o sistema esteja funcionando apropriadamente após intervalos de operação na escala das dezenas é muito baixa. Por exemplo, para ambos os sítios, a probabilidade de que o sistema esteja funcionando até um tempo de operação de 10 dias é menor que 20%.

Esse cenário de curtos intervalos de operação é reforçado pelo fato de que as Funções Taxa de Falha, apresentadas na Figura 76, são predominantemente influenciadas pelas falhas de nascimento, ou seja, as falhas que ocorrem logo no início da operação. Assim, é possível sugerir que o declive acentuado no início das curvas é resultado de problemas associados às fases de projeto, instalação do equipamento ou também problemas na fabricação.

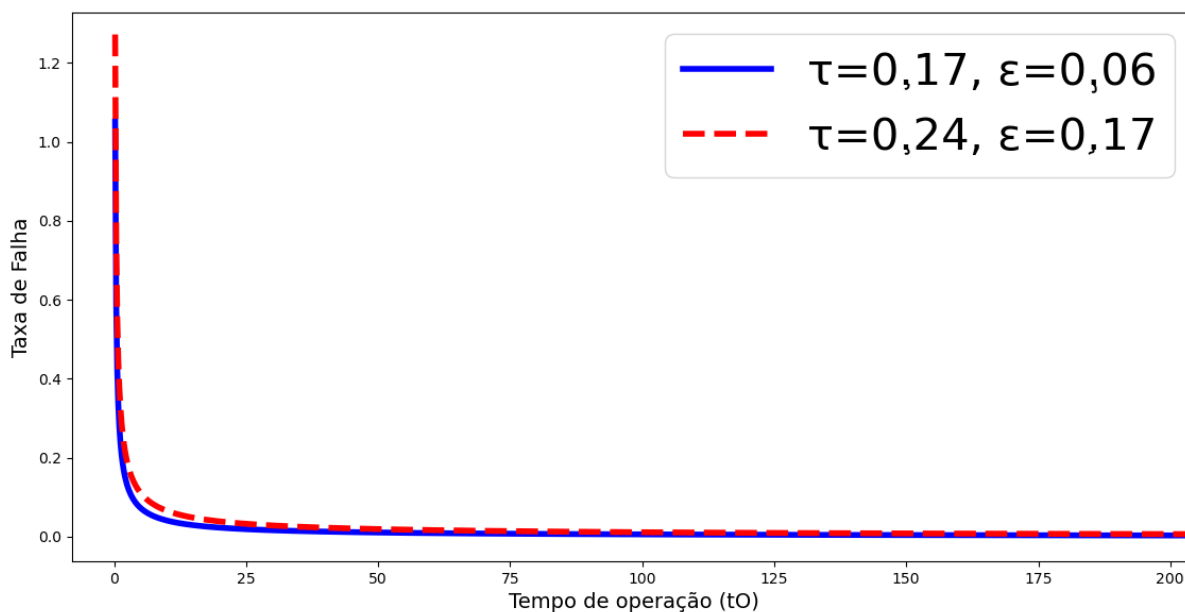
Um indicativo da baixa operacionalidade do rastreador de eixo único do sítio 5 é relatado em 20 de abril de 2016 pela equipe de manutenção (conforme descrito na Seção 4.3), que indica a ocorrência de diversas falhas no equipamento, as quais não puderam ser solucionadas. Esse relato reforça a observação de que o sistema continuou a funcionar de forma inadequada, o que pode ter contribuído para o maior número de falhas e a menor confiabilidade observada nas Curvas de Confiabilidade e Funções Taxa de Falha.

Figura 75 - Curva de Confiabilidade do sistema de rastreamento do Sítio 2



Fonte: Elaboração Própria.

Figura 76 - Função Taxa de Falhas do sistema de rastreamento do Sítio 2



Fonte: Elaboração Própria.

Uma situação parecida acontece com o rastreador de eixo duplo do sítio 2. O sistema funcionou por aproximadamente dois anos e meio, sendo posto de forma fixa a partir de então. Segundo visto na Seção 4.3, os relatos da equipe de manutenção do DKASC indicam que,

durante aproximadamente toda a metade do ano de 2012, o sistema de rastreamento vinha apresentando problemas que culminaram no fim de sua vida útil.

Assim, os relatos da equipe de manutenção corroboram com a informação obtida pela análise da Curva de Confiabilidade e Função Taxa de Falhas para os sítios 2 e 5. Ambos os sistemas, durante o período analisado, operaram com frequência significativa de falhas, contribuindo assim para o fato de que os períodos de operacionalidade dos equipamentos fossem mais curtos e estivessem fortemente relacionados a falhas de nascimento, conforme mencionado no parágrafo anterior. Essas informações evidenciam que as falhas iniciais, ocorridas logo no início da operação dos sistemas tiveram um papel preponderante na confiabilidade geral dos rastreadores.

Por fim, as médias dos tempos de operação, μ_O , foram calculadas pela Equação 49, resultando em 13,83 e 4,63 dias para os sistemas de rastreamento dos sítios 2 e 5, respectivamente. Nota-se que o tempo médio de operação do rastreador do sítio 2 é quase três vezes maior do que o do sítio 5. Essa discrepância pode ser explicada pelo fato de que os dados analisados para o sítio 5 referem-se a aproximadamente 14 anos de operação, enquanto para o sítio 2 foram analisados apenas 2 anos. Esse maior período de operação para o sítio 5 pode ter contribuído para um maior desgaste do equipamento, impactando seu tempo médio de operação.

Nesse contexto, a partir das médias dos tempos de reparo μ_R e dos tempos de operação μ_O , calculou-se a disponibilidade estática dos sistemas pela Equação 48, resultando nos valores aproximados de 71% e 41 %, para os sítios 2 e 5, respectivamente. Esse resultado é bastante interessante, pois indica que a disponibilidade do rastreador de eixo duplo do sítio 2 é maior do que a do conjunto de cinco rastreadores de eixo único do sítio 5. Essa diferença pode ser atribuída ao fato de que o sítio 5 possui uma maior quantidade de rastreadores, o que implica que o sistema será considerado como falho mesmo que apenas um dos equipamentos apresente falha. Nesse sentido, o sítio 5 tem uma probabilidade maior de ser classificado como indisponível, o que impacta diretamente em sua disponibilidade geral. É importante considerar essa peculiaridade na interpretação dos resultados e ao fazer comparações entre os sistemas de rastreamento.

6. CONCLUSÕES

O principal objetivo deste trabalho foi desenvolver uma metodologia de detecção de falhas de rastreadores fotovoltaicos com vistas à inferência dos índices probabilísticos de confiabilidade. O método desenvolvido, denominado Método do Limiar Móvel, se mostrou capaz de representar bem a distribuição dos períodos de interesse, no caso, os tempos de operação e reparo. As análises realizadas a partir desses índices foram corroboradas com as informações dos relatos da equipe de manutenção do DKASC.

As principais contribuições inovadoras deste trabalho estão relacionadas ao método de detecção de falhas e à aplicação da teoria de confiabilidade a rastreadores fotovoltaicos.

Em relação ao método de detecção proposto, o Método do Limiar Móvel apresentou-se como uma evolução em relação às metodologias apresentadas nos trabalhos de Ruth e Muller (2018) e Anderson *et al.* (2022). Essa evolução se deve ao fato de o Método do Limiar Móvel utilizar uma variedade de atributos para descrever as características da relação entre potência medida e potência simulada, bem como empregar uma abordagem não subjetiva na determinação dos limites de classificação.

Esses pontos permitiram que o método proposto se adaptasse às características do conjunto de análise, obtendo métricas de classificação mais satisfatórias. Por exemplo, ao analisar os sistemas de rastreamento do sítio 2, observou-se que o par de atributos com maior ranqueamento segundo a MRPA consistia em um atributo de caráter geral para determinação dos dias indefinidos associado a um atributo de comparação para identificação dos dias de falha. Já para o sítio 5, o par de atributos de maior ranqueamento era composto por um atributo de similaridade associado a um atributo de comparação.

A adaptação dos atributos de acordo com as particularidades de cada sistema demonstra a capacidade do Método do Limiar Móvel em se ajustar de forma satisfatória às características únicas de cada rastreador. Essa flexibilidade na escolha dos atributos são aspectos essenciais para garantir uma detecção de falhas mais eficaz. Além disso, o Método do Limiar Móvel mostrou-se vantajoso ao não depender de uma abordagem subjetiva na determinação dos limites de classificação, o que evita possíveis vieses e incertezas na identificação de falhas e torna o método mais robusto e confiável. Nesse sentido, com base nas vantagens apresentadas pelo Método do Limiar Móvel, fica evidente que a metodologia demonstra um potencial significativo de ser aplicada em outros fenômenos que estudem a comparação de similaridade entre formatos de curvas.

Outro ponto inovador do método de detecção de falhas em relação aos abordados anteriormente na literatura é a conceituação da classe de dias indefinidos, ou seja, dias para os quais não é possível realizar uma classificação precisa entre funcionamento e falha. A identificação desses dias indefinidos é uma parte fundamental da metodologia, uma vez que eles representam o escopo do método, ou seja, os dias cujas características nas curvas de potência medida e simulada impossibilitam uma diferenciação clara entre falha e funcionamento.

No entanto, é importante ressaltar que, até o momento, foi realizada apenas uma divisão simples do período de análise em 2/3 para calibração e 1/3 para validação. Essa é, portanto, uma perspectiva para trabalhos futuros, uma vez que a utilização de diferentes dinâmicas para a divisão dos períodos de calibração e validação pode testar a robustez do modelo e a eficácia da classificação dos dias. Apesar disso, os resultados obtidos na inferência dos índices probabilísticos de confiabilidade demonstram que a divisão simples dos conjuntos de calibração e validação se mostrou suficiente para obter resultados confiáveis.

Outra consideração relevante é que o Método do Limiar Móvel para detecção de falhas tem como objetivo principal representar adequadamente a dinâmica dos tempos de operação e reparo, visando a inferência dos índices probabilísticos de confiabilidade. Nesse aspecto, o método se mostrou significativa eficiência. No entanto, caso o objetivo primário fosse exclusivamente a classificação dos dias em funcionamento ou falha, o método poderia ser incrementado com o uso de técnicas de aprendizado de máquina, como o KNN e árvores de decisão. Essas técnicas podem contribuir para uma classificação mais precisa e refinada dos dias. Nesse contexto, uma perspectiva interessante para trabalhos futuros é justamente buscar desenvolver um método não supervisionado que minimize ao máximo a necessidade de rotulação visual para calibração dos modelos. Isso seria especialmente relevante para o acompanhamento prático de usinas fotovoltaicas, permitindo uma detecção automática de falhas e funcionamento, com aplicação em tempo real e maior autonomia do sistema de rastreamento.

Também, é válido notar que o método proposto neste trabalho demonstra sua adaptabilidade aos diferentes contextos das usinas fotovoltaicas, desde que a simulação de potência leve em conta as características específicas do sistema, como a utilização da técnica do *backtracking* e a presença de sombreamentos sobre o sistema, por exemplo. A técnica do *backtracking*, em particular, afeta diretamente a cinemática dos rastreadores apresentada neste trabalho, e, portanto, precisa ser bem modelada para garantir a eficácia do modelo. Essas

observações são importantes para garantir que o método seja aplicável em diversas situações e ofereça resultados confiáveis relacionados a detecção de falhas nos sistemas de rastreamento.

O segundo ponto inovador deste trabalho foi a aplicação da teoria de confiabilidade para a inferência dos índices probabilísticos de confiabilidade dos sistemas de rastreamento. Especificamente, a análise da proporção entre o tempo de reparo e o tempo total de operação por ciclo de renovação, que ficou entre 30% e 40%, levou à necessidade de utilizar um modelo de Processo de Renovação Alternada. Esse modelo é capaz de inferir a disponibilidade do sistema a partir das distribuições dos tempos de reparo e operação. No entanto, é fundamental destacar que, em usinas fotovoltaicas convencionais, essa proporção de tempo de reparo por ciclo deve ser significativamente menor. Pois, quanto mais tempo o rastreador estiver inoperante, menor será o fator de capacidade da usina, resultando em perdas significativas na produção de energia, que podem estar entre 15% a 20%. Em usinas que possuam equipes de manutenção dedicadas, essa proporção de tempo de reparo por ciclo tende a diminuir drasticamente, já que as intervenções de reparo podem ser mais ágeis e eficientes.

Em relação ao ajuste dos dados de ocorrência de falha para a inferência dos índices probabilísticos de confiabilidade, a metodologia empregada foi a do ajuste dos coeficientes de uma CDF de Weibull aos valores estimados da CDF, obtidos pela aproximação de Bernard do ranqueamento da mediana, através da minimização do erro médio quadrático por meio de um ajuste linear. Essa abordagem levou em conta a não monotonicidade do conjunto de dados, permitindo a realização de ajustes locais a partir de partições. Entretanto, durante a análise, foi observada uma limitação nos procedimentos adotados, resultando em uma inconsistência física no conjunto de tempos de operação. Essa incongruência foi tratada mediante um ajuste dos dados. Embora o tratamento tenha sido eficaz em resolver o problema, é importante destacar que ele foi realizado de forma não automatizada, o que representa um ponto de melhoria na metodologia. Nesse sentido, são necessários estudos mais aprofundados para entender as causas da inconsistência, buscar formas de automatizar e aprimorar o processo de tratamento dos dados.

Por sua vez, a utilização dos dados de ocorrência de falha a partir do Método do Limiar Móvel demonstrou capacidade em representar a distribuição dos tempos de operação e reparo dos sistemas estudados, evidenciando a capacidade do método de detecção em ser aplicado para a inferência dos índices probabilísticos de confiabilidade. As informações obtidas através das Curvas de Manutenção, Funções Taxa de Reparo, Curvas de Confiabilidade e Funções Taxa de Falhas inferidas apresentaram uma relação qualitativa coerente com os relatos fornecidos pelas equipes de manutenção dos sistemas dos sítios 2 e 5.

Com o preenchimento de uma lacuna na literatura, a metodologia proposta para a inferência dos índices probabilísticos de confiabilidade se destaca por possibilitar o planejamento de manutenções centradas na confiabilidade. Isso pode proporcionar melhorias significativas no desempenho geral do sistema, otimizando os recursos e alocando os esforços de manutenção de forma mais estratégica.

Para concluir, os resultados evidenciam que a dinâmica relacionada aos tempos de reparo de ambos os sistemas foi muito parecida, inclusive com médias com valores muito próximos, e que as curvas de confiabilidade e função de taxa de falhas apresentaram uma preponderância de falhas de nascimento. Essas observações abrem espaço para a possibilidade de modelar os índices probabilísticos dos rastreadores de forma mais geral e adaptável, sem depender de dados específicos de ocorrência de falhas ou do uso de um método de detecção de falhas como proposto neste trabalho.

Um exemplo promissor seria a determinação de que os rastreadores solares seguem uma determinada distribuição de probabilidade, cujos parâmetros adquirem valores distintos com base em características locais, como o tipo de sistema de rastreamento e especificidades das usinas, como potência nominal e outros fatores relevantes. Tal modelo generalizado seria extremamente útil na fase de projeto das usinas fotovoltaicas, permitindo uma estimativa confiável dos índices de confiabilidade dos rastreadores antes mesmo de sua implementação.

Nesse contexto, o método proposto neste trabalho desempenharia um papel importante na validação desses modelos. Pois, através do Método do Limiar Móvel, seria possível inferir os índices probabilísticos de confiabilidade dos rastreadores com base em dados facilmente disponíveis nas usinas, que são a potência e a radiação solar, auxiliando na verificação da eficácia dos modelos generalizados.

REFERÊNCIAS

ABD EL-AAL, Abou El Maaty M. *et al.* Modeling and optimizing the size of the power conditioning unit for photovoltaic systems. **Journal of Solar Energy Engineering, Transactions of the ASME**, [s. l.], v. 128, n. 1, p. 40–44, 2006.

ACUÑA, Luceny Guzmán *et al.* Modelling autonomous hybrid photovoltaic-wind energy systems under a new reliability approach. **Energy Conversion and Management**, [s. l.], v. 172, p. 357–369, 2018.

ADES. **Datasheet: Seguidor solar 2 ejes de 265 m2 Modelo V10-24 M_EXP**. [S. l.], 2023. Disponível em: https://archive.org/details/projectcatalogueades_v_web. Acesso em: 10 jul. 2023.

ANDERSON, Kevin *et al.* A Method for Estimating Time-Series PV Production Loss from Solar Tracking Failures. **IEEE Journal of Photovoltaics**, [s. l.], v. 12, n. 1, p. 119–126, 2022.

AZEVEDO, Rafael Valença. **Um problema de decisão simultânea da política de substituição e da quantidade de sobressalentes para sistemas sujeitos a reparos imperfeitos utilizando algoritmos genéticos multiobjetivos e simulação discreta de eventos**. 2013. Dissertação de Mestrado - UFPE, Recife, 2013.

BARBOSE, Galen *et al.* Tracking US photovoltaic system prices 1998-2012: A rapidly changing market. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 23, n. 6, p. 692–704, 2015.

BAZOVSKY, I. **Reliability Theory and Practice**. 1. ed. Mineola, New York, EUA: Dover Publications, Inc., 2004.

BENAMMAR, Samir; TEE, Kong Fah. Structural reliability analysis of a heliostat under wind load for concentrating solar power. **Solar Energy**, [s. l.], v. 181, p. 43–52, 2019.

BIGLARBEGLIAN, Mehrdad *et al.* Scalable Reliability Monitoring of GaN Power Converter through Recurrent Neural Networks. *Em: , 2018. IEEE Energy Conversion Congress and Exposition (ECCE)*. [S. l.: s. n.], 2018. p. 7271–7277.

BILLINTON, R.; WENYUAN, Li. Composite system reliability assessment using a Monte Carlo approach. *Em: , 1991. Third International Conference on Probabilistic Methods Applied to Electric Power Systems*. [S. l.: s. n.], 1991. p. 53–57.

BOLINGER, Mark; SEEL, Joachim; ROBSON, Dana. **Empirical Trends in Project Technology, Cost, Desempenho, and PPA Pricing in the United States-2019 Edition**. [S. l.: s. n.], 2019.

BOUSSHOUA, B; ELMAOUIHAB, A. Smart Grid Reliability Using Reliable Block Diagram Case Study: Adrar's Isolated Network of Algeria. *Em: 2019. 5th International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET-2019)*. [S. l.: s. n.], 2019. p. 1–6.

BRANCH, Mary Ann; COLEMAN, Thomas F; LI, Yuying. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. **SIAM Journal on Scientific Computing**, [s.l.], v. 21, n. 1, p. 1–23, 1999. Disponível em: <http://www.siam.org/journals/sisc/21-1/28910.html>.

BRENNAND, Leonardo José de Petribú; PAULA, José Guilherme Souza de. **Avaliação da qualidade de dados observacionais de vento e radiação solar**. 2018. - UFPE, Recife, 2018. BSQ SOLAR. **Datasheet BSQ-D280/53 CPV System Abridged Data Sheet**. [S. l.], 2023. Disponível em: <https://archive.org/details/datasheet-bsq-d-280-53-v-2.2-2-t>. Acesso em: 3 jul. 2023.

BUCHIN, Maike. **On the Computability of the Fréchet Distance Between Triangulated Surfaces Dissertation zur Erlangung des Doktorgrades**. 2007. [s. l.], 2007. Disponível em: <http://dx.doi.org/10.17169/refubium-6111>. Acesso em: 10 jun. 2023.

CAMPBELL, M *et al.* Minimizing utility-scale PV power plant LCOE through the use of high capacity factor configurations. *Em:* , 2009, Philadelphia, PA, USA. **34th IEEE Photovoltaic Specialists Conference (PVSC)**. Philadelphia, PA, USA: [s. n.], 2009. p. 421–426.

CANALES, Fausto A. *et al.* Cost-reliability analysis of hybrid pumped-battery storage for solar and wind energy integration in an island community. **Sustainable Energy Technologies and Assessments**, [s. l.], v. 44, 2021.

CAO, J.; LIN, J. A study on formulation of objective functions for determining material models. **International Journal of Mechanical Sciences**, [s. l.], v. 50, n. 2, p. 193–204, 2008.

CARRASCO, Luis Miguel *et al.* Reliability of a 13 000-SHS photovoltaic rural electrification programme. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 21, n. 5, p. 1136–1145, 2013.

CASTET, Jean Francois; SALEH, Joseph H. Satellite and satellite subsystems reliability: Statistical data analysis and modeling. **Reliability Engineering and System Safety**, [s. l.], v. 94, n. 11, p. 1718–1728, 2009.

CHA, Ji Hwan; FINKELSTEIN, Maxim. **Point Processes for Reliability Analysis: Shocks and Repairable Systems**. 1. ed. [S. l.]: Springer, 2018-. ISSN 978-3-319-73540-5. Disponível em: <http://www.springer.com/series/6917>.

CHICCO, Davide; WARRENS, Matthijs J.; JURMAN, Giuseppe. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. **PeerJ Computer Science**, [s. l.], v. 7, p. 1–24, 2021.

CHONG, K. K.; WONG, C. W. General formula for on-axis sun-tracking system and its application in improving tracking accuracy of solar collector. **Solar Energy**, [s. l.], v. 83, n. 3, p. 298–305, 2009.

CHOPRA, Prem S; WOLOSEWICZ, Ronald M. APPLICATION OF RELIABILITY, MAINTAINABILITY, AND AVAILABILITY ENGINEERING TO SOLAR HEATING

AND COOLING SYSTEMS. *Em:* , 1980. **Annual Reliability and Maintainability Symposium**. [S. l.: s. n.], 1980.

CHRIST, Maximilian *et al.* Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). **Neurocomputing**, [s. l.], v. 307, p. 72–77, 2018.

CHRIST, Maximilian *et al.* **TSfresh**. [S. l.], 2023. Disponível em: <https://tsfresh.readthedocs.io/en/latest/>. Acesso em: 10 jun. 2023.

COLLINS, Elmer *et al.* Reliability and availability analysis of a fielded photovoltaic system. *Em:* , 2009. **34th IEEE Photovoltaic Specialists Conference (PVSC)**. [S. l.: s. n.], 2009. p. 2316–2321.

COOK, Richard J.; LAWLESS, Jerald F. **The Statistical Analysis of Recurrent Events**. 1. ed. [S. l.]: Springer, 2007.

COSSU, Silvestro; BACCOLI, Roberto; GHIANI, Emilio. Utility scale ground mounted photovoltaic plants with gable structure and inverter oversizing for land-use optimization. **Energies**, [s. l.], v. 14, n. 3048, 2021.

CRISTALDI, Loredana; KHALIL, Mohamed; FAIFER, Marco. Markov process reliability model for photovoltaic module failures. *Em:* , 2015. **International Conference on Renewable Energy Research and Applications (ICRERA)**. [S. l.: s. n.], 2015. p. 203–208. Disponível em: www.imeko.org.

CUSTODIO, Rogério; ANDRADE, João Carlos de; AUGUSTO, Fábio. O AJUSTE DE FUNÇÕES MATEMÁTICAS A DADOS EXPERIMENTAIS. **Química Nova**, [s. l.], v. 20, n. 2, p. 219–225, 1997.

DGS. **Planning and Installing Photovoltaic Systems: A guide for installers, architects and engineers**. 2. ed. [S. l.]: Earthscan , 2008.

DKA. **Desert Knowledge Australia Centre. Download Data**. [S. l.], 2023. Disponível em: <http://dkasolarcentre.com.au/download?location=alice-springs>. Acesso em: 16 nov. 2021.

DOBOS, Aron P. **PVWatts Version 5 Manual**. [S. l.: s. n.], 2014. Disponível em: <http://developer.nrel.gov>.

ELERATH, Jon G. HCPV tracker accelerated reliability tests. *Em: , 2010. 35th IEEE Photovoltaic Specialists Conference*. [S. l.: s. n.], 2010. p. 481–486.

ELERATH, Jon G. Reliability assessment of high-concentration photovoltaic trackers. *Em: , 2011. Proceedings - Annual Reliability and Maintainability Symposium*. [S. l.: s. n.], 2011. p. 1–6.

ELERATH, Jon G. Solar Tracker Architecture: Effects on Availability and Power Production. *Em: , 2018. Annual Reliability and Maintainability Symposium (RAMS)*. [S. l.: s. n.], 2018. p. 1–6.

ELERATH, Jon G. Solar Tracker Effectiveness: It's All About Availability. *Em: , 2017. IEEE International Telecommunications Energy Conference (INTELEC)*. [S. l.: IEEE, 2017. p. 156–162.

ELERATH, Jon G; SPENCER, Mark; HORNE, Steve. Demonstrating reliability in HCPV systems. *Em: , 2011. 37th IEEE Photovoltaic Specialists Conference*. [S. l.: IEEE, 2011. p. 3547–3552.

EL-SAYED, M A H. RELIABILITY ANALYSIS OF SOLAR ENERGY SYSTEMS. *Em: , 1986, Bahrain. Solar Energy Prospect in the Arab World, Second Arab International Solar Energy Conference*. Bahrain: Pergamon Press, Oxford, Engl, 1986. p. 26–30. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-0022954459&partnerID=40&md5=c9d38f1013dfbaa1db6da703cae25eba>.

EL-SAYED, Mohamed A H. RELIABILITY EVALUATION OF EGYPTIAN AND JORDANIAN INTERCONNECTED POWER SYSTEMS. *Em: , 1996. Proceedings of IEEE. AFRICON '96*. [S. l.: s. n.], 1996. p. 151–156.

EPE. **Leilão de Energia de Reserva de 2014: Participação dos Empreendimentos Solares Fotovoltaicos**. [S. l.: s. n.], 2014. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes->

dados-abertos/publicacoes/PublicacoesArquivos/publicacao-94/NT_EPE-DEE-NT-150_2014.pdf. Acesso em: 15 nov. 2021.

EPE. **PROJETOS FOTOVOLTAICOS NOS LEILÕES DE ENERGIA: Características dos empreendimentos participantes nos leilões de 2013 a 2018.** [S. l.: s. n.], 2018. Disponível em: https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-110/topico-418/EPE-DEE-NT-091_2018-r0.pdf. Acesso em: 15 nov. 2021.

FARA, Laurentiu *et al.* Reliability analysis of photovoltaic systems with leds integrated in lighting applications. *Em:* , 2017. **ISES Solar World Congress 2017 - IEA SHC International Conference on Solar Heating and Cooling for Buildings and Industry 2017, Proceedings.** [S. l.]: International Solar Energy Society, 2017. p. 486–497.

FAWCETT, Tom. ROC Graphs: Notes and Practical Considerations for Data Mining Researchers. **Machine Learning**, [s. l.], v. 31, n. 1, p. 1–38, 2004. Disponível em: http://www.purl.org/NET/tfawcett/software/ROC_algs.tar.gz.

FERNANDEZ SAVARI, George *et al.* A Comprehensive Review on the Modern Power System Reliability Assessment. **International Journal of Renewable Energy Research**, [s. l.], v. 11, n. 4, 2021. Disponível em: <https://www.researchgate.net/publication/357366009>.

FIZANTI, Causia Agusti; MUSYAFI, Ali; MUHAMMAD, Khamim Asy'ari. Analysis of RAM (Reliability, Availability, Maintainability) Production of Electric Voltage from 48 v PV (Photovoltaic) at Pantai Baru Pandansimo, Indonesia. *Em:* 2020. **E3S Web of Conferences**. [S. l.]: EDP Sciences, 2020.

FORTUNE BUSINESS INSIGHTS. **Solar Tracker Market Size, Share & COVID-19 Impact Analysis, By Type (Photovoltaic (PV), Concentrated Solar Power (CSP)), By Movement (Single Axis, Dual Axis), By Application (Utility, Non-Utility) and Regional Forecasts, 2020-2027.** [S.l.], 2020. Disponível em: <https://web.archive.org/web/20210513195239/https://www.fortunebusinessinsights.com/industry-reports/infographics/solar-tracker-market-100448>. Acesso em: 17 nov. 2021.

FORTUNE BUSINESS INSIGHTS. **Solar Tracker Market Size, Share & COVID-19 Impact Analysis, By Type (Photovoltaic (PV), Concentrated Solar Power (CSP)), By Movement (Single-Axis, Dual Axis), By Application (Utility, Non-Utility), and Regional Forecast, 2021-2028.** [S.l.], 2021. Disponível em: <https://web.archive.org/web/20210820155639/https://www.fortunebusinessinsights.com/industry-reports/solar-tracker-market-100448>. Acesso em: 17 nov. 2021.

FRAAS, Lewis; PARTAIN, Larry. **Solar Cells and Their Applications**. 2. ed. [S. l.]: Wiley, 2010.

FUENTES, M. *et al.* Application and validation of algebraic methods to predict the behaviour of crystalline silicon PV modules in Mediterranean climates. **Solar Energy**, [s. l.], v. 81, n. 11, p. 1396–1408, 2007.

GAMBOGI, William *et al.* Sequential Stress Testing to Predict Photovoltaic Module Durability. *Em: , 2018. 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC)*. [S. l.: s. n.], 2018. p. 1593–1596.

GARCÍA, Miguel *et al.* Solar-tracking PV plants in Navarra: A 10 MW assessment. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 17, n. 5, p. 337–346, 2009.

GARRO, A.; BARRARA, F. Reliability analysis of residential photovoltaic systems. **Renewable Energy and Power Quality Journal**, [s. l.], v. 1, n. 9, p. 1197–1202, 2011.

GAUTAM, Nalin K; KAUSHIKA, N D. RELIABILITY EVALUATION OF SOLAR PHOTOVOLTAIC ARRAYS. **Solar Energy**, [s. l.], v. 72, n. 2, p. 129–141, 2002. Disponível em: www.elsevier.com/locate/solener.

GEORGILAKIS, Pavlos S.; KATSIGIANNIS, Yiannis A. Reliability and economic evaluation of small autonomous power systems containing only renewable energy sources. **Renewable Energy**, [s. l.], v. 34, n. 1, p. 65–70, 2009.

GRANATA, Jennifer E. *et al.* Design for reliability: A low concentration PV case study. *Em: ,* 2012. **38th IEEE Photovoltaic Specialists Conference**. [S. l.: s. n.], 2012. p. 1739–1743.

GRAND VIEW RESEARCH. **Solar Tracker Market Size, Share & Trends Analysis Report By Technology (Solar Photovoltaic, Concentrated Solar Power, Concentrated Photovoltaic), By Type, By Application, And Segment Forecasts, 2022 - 2030**. [S. l.], 2023. Disponível em: <https://web.archive.org/web/20230323145917/https://www.grandviewresearch.com/industry-analysis/solar-tracker-industry>. Acesso em: 27 jun. 2023.

GRAND VIEW RESEARCH. **Solar Tracker Market Size, Share & Trends Analysis Report By Technology (Solar PV, Concentrator PV), By Product (Single Axis, Dual Axis), By Application, By Region, And Segment, Forecasts 2019 - 2025**. [S. l.], 2019. Disponível em: <https://web.archive.org/web/20190701222920/https://www.grandviewresearch.com/industry-analysis/solar-tracker-industry>. Acesso em: 27 jun. 2023.

GUO, Yang; MAGALHÃES, Pedro H. **Increase PV Energy Yield with Superior Stow Parameters Enabled by Rigid Trackers**. [S. l.: s. n.], 2022.

HADIDIAN MOGHADDAM, Mohammad Jafar *et al.* Optimal sizing and energy management of stand-alone hybrid photovoltaic/wind system based on hydrogen storage considering LOEE and LOLE reliability indices using flower pollination algorithm. **Renewable Energy**, [s. l.], v. 135, p. 1412–1434, 2019.

HAFEZ, A. Z.; YOUSEF, A. M.; HARAG, N. M. Solar tracking systems: Technologies and trackers drive types – A review. **Renewable and Sustainable Energy Reviews**, [s. l.], v. 91, p. 754–782, 2018.

HAMDY, M A; BESHIR, M E; ELMASRY, S E. Reliability Analysis of Photovoltaic Systems. **Applied Energy**, [s. l.], v. 33, p. 253–263, 1989.

HONG, Taehoon *et al.* A preliminary study on the 2-axis hybrid solar tracking method for the smart photovoltaic blind. *Em: ,* 2016. **Energy Procedia**. [S. l.]: Elsevier Ltd, 2016. p. 484–490.

HORNE, S. Concentrating photovoltaic (CPV) systems and applications. *Em: CONCENTRATING SOLAR POWER TECHNOLOGY*. [S. l.]: Elsevier, 2012. p. 323–361.

IEA. **Renewable electricity - Forecast Summary**. [S. l.], 2022.

JACQUELIN, Jean. A Reliable Algorithm for the Exact Median Rank Function. **IEEE Transactions on Electrical**, [s. l.], v. 28, n. 2, p. 168–171, 1993.

JÄGER-WALDAU, A. **PV Status Report 2018**. [S. l.: s. n.], 2018. Disponível em: <https://ec.europa.eu/jrc>.

JAHANNOOSH, Mariye *et al.* New hybrid meta-heuristic algorithm for reliable and cost-effective designing of photovoltaic/wind/fuel cell energy system considering load interruption probability. **Journal of Cleaner Production**, [s. l.], v. 278, 2021.

JANAMALA, Varaprasad; SREENIVASULU REDDY, D. Coyote optimization algorithm for optimal allocation of interline –Photovoltaic battery storage system in islanded electrical distribution network considering EV load penetration. **Journal of Energy Storage**, [s. l.], v. 41, 2021.

JORDAN, Dirk C. *et al.* PV field reliability status—Analysis of 100 000 solar systems. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 28, n. 8, p. 739–754, 2020.

KALBFLEISCH, John D; PRENTICE, Ross L. **The Statistical Analysis of Failure Time Data Second Edition WILEY-INTERSCIENCE**. 2. ed. [S. l.]: John Wiley & Sons, Inc, 2002.

KHAN, M Shuaib; PASHA, G R. The Plotting of Observations for the Inverse Weibull Distribution on Probability Paper Transmuted family of distributions-development of View project The Plotting of Observations for the Inverse Weibull Distribution on Probability Paper. **Journal of Advanced Research in Statistics and Probability**, [s. l.], v. 1, n. 1, p. 11–22, 2009. Disponível em: <https://www.researchgate.net/publication/270393940>.

KIM, Jaeun *et al.* A review of the degradation of photovoltaic modules for life expectancy. **Energies**, [s. l.], v. 14, n. 14, 2021.

KLISE, Geoffrey T. PV Reliability Operations Maintenance (PVROM) Database: Data Collection & Analysis Insights. *Em: , 2015. Conference: Proposed for presentation at the Intersolar North America 2015.* [S. l.: s. n.], 2015. Disponível em: <https://www.osti.gov/biblio/1337988>. Acesso em: 16 nov. 2021.

KUITCHE, Joseph M.; TAMIZH-MANI, Govindasamy; PAN, Rong. Failure modes effects and criticality analysis (FMECA) approach to the crystalline silicon photovoltaic module reliability assessment. *Em: , 2011. Reliability of Photovoltaic Cells, Modules, Components, and Systems IV.* [S. l.]: SPIE, 2011. p. 81120L.

LEWIS, E E *et al.* **Introduction to Reliability Engineering**. 2. ed. [S. l.]: John Wiley & Sons, INC, 1994.

LILLO-BRAVO, Isidoro *et al.* Impact of energy losses due to failures on photovoltaic plant energy balance. **Energies**, [s. l.], v. 11, n. 2, 2018.

LONG, Charles N.; ACKERMAN, Thomas P. Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. **Journal of Geophysical Research Atmospheres**, [s. l.], v. 105, n. D12, p. 15609–15626, 2000.

LONG, C N; DUTTON, E G. **BSRN recommended QC tests**. [S. l.: s. n.], 2010. Disponível em: https://epic.awi.de/id/eprint/30083/1/BSRN_recommended_QC_tests_V2.pdf. Acesso em: 14 jul. 2023.

MAIHULLA, Anas Sani; YUSUF, Ibrahim. Desempenho analysis of photovoltaic systems using (RAMD) analysis. **Journal of the Nigerian Society of Physical Sciences**, [s. l.], v. 3, n. 3, p. 172–180, 2021.

MIRANDA, Deigo Rodrigues de. **Procedimento de Classificação e Regressão Aplicado ao Site Adaptation Da Radiação Solar**. 2023. - UFPE, Recife, 2023.

MORADI, Hadis; ABTAHI, Amir; MESSENGER, Roger. Annual Desempenho Comparison Between Tracking and Fixed Photovoltaic Arrays. *Em: , 2016. 2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC).* [S. l.]: IEEE, 2016. p. 3179–3183.

MORDOR INTELLIGENCE. **SOLAR TRACKER MARKET SIZE & SHARE ANALYSIS - GROWTH TRENDS & FORECASTS (2023 - 2028)**. [S. l.], 2022. Disponível em:

<https://web.archive.org/web/20220809080239/https://www.mordorintelligence.com/industry-reports/solar-tracker-market>. Acesso em: 17 jun. 2023.

MORE, Jorge J. The Levenberg-Marquardt algorithm: Implementation and theory. *Em: , 1977. Numerical analysis: proceedings of the biennial Conference held at Dundee*. [S. l.: s. n.], 1977. p. 105–116.

MORRIS, Russell W.; FIFE, John M. Using probabilistic methods to define reliability requirements for high power inverters. *Em: , 2009. Reliability of Photovoltaic Cells, Modules, Components, and Systems II*. [S. l.]: SPIE, 2009. p. 74120G.

MOUSAZADEH, Hossein *et al.* A review of principle and sun-tracking methods for maximizing solar systems output. **Renewable and Sustainable Energy Reviews**, [s. l.], v. 13, n. 8, p. 1800–1818, 2009.

NARVARTE, L.; LORENZO, E. Tracking and ground cover ratio. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 16, n. 8, p. 703–714, 2008.

NELSON, Wayne. **APPLIED LIFE DATA ANALYSIS**. 1. ed. [S. l.]: John Wiley & Sons, 1982.

NELSON, Wayne B. **Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications**. 1. ed. [S. l.]: ASA SIAM, 2003.

NEXTRACKER. **NX Horizon: Smart Solar Tracking System**. [S. l.], 2021. Disponível em: https://web.archive.org/web/20210302162738/https://info.nextracker.com/hubfs/Datasheets%202020/nxt_nx_horizon_datasheet%20Aug%202020.pdf. Acesso em: 15 nov. 2021.

NREL. **Annual Technology Baseline**. [S. l.], 2023. Disponível em: <https://atb-archive.nrel.gov/electricity/2018/index.html>. Acesso em: 27 jun. 2023.

O'CONNOR, Patrick D. T.; KLEYNER, Andre. **Practical Reliability Engineering**. 5. ed. [S. l.]: John Wiley & Sons, 2012.

ONS. **Boletim Mensal de Geração Solar Fotovoltaica Agosto/2021**. [S. l.: s. n.], 2021. Disponível em: <http://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/Boletim%20Mensal%20de%20Gera%C3%A7%C3%A3o%20Solar%202021-08.pdf>. Acesso em: 17 nov. 2021.

OPREA, Simona Vasilica *et al.* Photovoltaic Power Plants (PV-PP) Reliability Indicators for Improving Operation and Maintenance Activities. A Case Study of PV-PP Agigea Located in Romania. **IEEE Access**, [s. l.], v. 7, p. 39142–39157, 2019.

OWEN-BELLINI, Michael *et al.* Advancing reliability assessments of photovoltaic modules and materials using combined-accelerated stress testing. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 29, n. 1, p. 64–82, 2021.

PAPATHANASIOU, Olga. **Market Survey on Solar Trackers 2021**. [S. l.: s. n.], 2021. Disponível em: <https://taiyangnews.info/technology/the-basics-of-hsat-systems/>. Acesso em: 2 jul. 2023.

PETITJEAN, François; KETTERLIN, Alain; GANÇARSKI, Pierre. A global averaging method for dynamic time warping, with applications to clustering. **Pattern Recognition**, [s. l.], v. 44, n. 3, p. 678–693, 2011.

PETTRIBÚ, L *et al.* PROCEDIMENTO OBJETIVO PARA A GARANTIA DE QUALIDADE DE DADOS DE RADIAÇÃO SOLAR. [s. l.], v. 21, p. 11–67, 2017.

PEYGHAMI, Saeed; PALENSKY, Peter; BLAABJERG, Frede. An overview on the reliability of modern power electronic based power systems. **IEEE Open Journal of Power Electronics**, [s. l.], v. 1, p. 34–50, 2020.

POLITEC. **Datasheet: Seguidor Solar de 1 eixo**. [S. l.], 2021. Disponível em: <https://web.archive.org/web/20210820180041/http://www.politec.eng.br/wp-content/uploads/2014/10/Seguidor-Solar-Politec.pdf>. Acesso em: 15 nov. 2021.

RABL, Ari. **Active Solar Collectors and Their Applications**. 1. ed. New York: Oxford University Press, Inc, 1985.

RACHARLA, Suneetha; RAJAN, K. Solar tracking system—a review. **International Journal of Sustainable Engineering**, [s. l.], v. 10, n. 2, p. 72–81, 2017.

RAGHUWANSHI, Santosh Singh; ARYA, Rajesh. Reliability evaluation of stand-alone hybrid photovoltaic energy system for rural healthcare centre. **Sustainable Energy Technologies and Assessments**, [s. l.], v. 37, 2020.

RAMASAMY, Vignesh *et al.* **U.S. Solar Photovoltaic System and Energy Storage Cost Benchmarks: Q1 2021**. [S. l.: s. n.], 2022. Disponível em: www.nrel.gov/publications.

RECA-CARDEÑA, Juan; LÓPEZ-LUQUE, Rafael. Design Principles of Photovoltaic Irrigation Systems. *Em: ADVANCES IN RENEWABLE ENERGIES AND POWER TECHNOLOGIES*. [S. l.]: Elsevier, 2018. v. 1, p. 295–333.

RELIASOFT. **Introduction to Repairable Systems**. [S. l.], 2015a. Disponível em: https://reliawiki.com/index.php/Introduction_to_Repairable_Systems. Acesso em: 18 jul. 2023.

RELIASOFT. **Recurrent Event Data Analysis**. [S. l.], 2015b. Disponível em: https://reliawiki.com/index.php/Recurrent_Event_Data_Analysis. Acesso em: 11 jun. 2023.

REMUND, Jan *et al.* Worldwide Linke turbidity information. *Em: 2003, Goteborg, Sweden. Proceedings of ISES Solar World Congress*. Goteborg, Sweden: International Solar Energy Society, 2003. p. 16–19. Disponível em: <https://hal.archives-ouvertes.fr/hal-00465791>.

REN21. **Renewables 2020: Global Status Report**. [S. l.: s. n.], 2020.

REN21. **Renewables 2023 Global Status Report - Renewables in Energy Supply**. [S. l.: s. n.], 2023.

RENEWABLE ENERGY WORLD. **Solar Trackers: Facing the Sun**. [S. l.], 2009. Disponível em:

https://web.archive.org/web/20201001000000*/https://www.renewableenergyworld.com/2009/06/01/solar-trackers-facing-the-sun/. Acesso em: 15 nov. 2021.

RHIF, Ahmed. A position control review for a photovoltaic system: Dual axis Sun Tracker. **IETE Technical Review**, [s. l.], v. 28, n. 6, p. 479–485, 2011.

ROTAR, Raul *et al.* Global energy production computation of a solar-powered smart home automation system using reliability-oriented metrics. **Energies**, [s. l.], v. 14, n. 9, 2021.

RUTH, Dan; MULLER, Matthew. A methodology to analyze photovoltaic tracker uptime. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 26, n. 7, p. 491–501, 2018.

SÁNCHEZ-BARROSO, Gonzalo *et al.* Markov chains estimation of the optimal periodicity for cleaning photovoltaic panels installed in the dehesa. **Renewable Energy**, [s. l.], v. 179, p. 537–549, 2021.

SARHAN, Ammar M.; APALOO, Joseph. Exponentiated modified Weibull extension distribution. **Reliability Engineering and System Safety**, [s. l.], v. 112, p. 137–144, 2013.

SAYED, A. *et al.* Reliability, availability and maintainability analysis for grid-connected solar photovoltaic systems. **Energies**, [s. l.], v. 12, n. 7, 2019.

SCHENKELBERG, F. Reliability modeling and accelerated life testing for solar power generation systems. *Em: RELIABILITY CHARACTERISATION OF ELECTRICAL AND ELECTRONIC SYSTEMS*. [S. l.]: Elsevier Inc., 2015. p. 215–250.

SEI. **TRACKING THE SUN**. [S.l.], 2016. Disponível em: <https://web.archive.org/web/20210304125736/https://www.saurenergy.com/solar-energy-articles/tracking-the-sun>. Acesso em: 16 nov. 2021.

SEME, Sebastijan *et al.* Solar photovoltaic tracking systems for electricity generation: A review. **Energies**, [s. l.], v. 13, n. 6, 2020.

SHIVARAMA KRISHNA, K.; SATHISH KUMAR, K. Generation system reliability assessment incorporating photovoltaic energy. *Em: LECTURE NOTES IN ELECTRICAL ENGINEERING*. [S. l.]: Springer Verlag, 2017. v. 436, p. 647–654.

SINGH, Chanan; JIRUTITIJAROEN, Panida; MITRA, Joydeep. **Electric Power Grid Reliability Evaluation: Models and Methods**. Hoboken, New Jersey: John Wiley & Sons, Inc., 2019.

SINGH, Smriti; SAKET, R. K.; KHAN, Baseem. A comprehensive review of reliability assessment methodologies for grid-connected photovoltaic systems. **IET Renewable Power Generation**, [s. l.], 2023.

SONAWANE, Pramod R. *et al.* Reliability and Criticality Analysis of a Large-Scale Solar Photovoltaic System Using Fault Tree Analysis Approach. **Sustainability**, [s. l.], v. 15, n. 5, p. 4609, 2023.

SPERTINO, Filippo; CORONA, Fabio. **Monitoring and checking of desempenho in photovoltaic plants: A tool for design, installation and maintenance of grid-connected systems**. [S. l.: s. n.], 2013.

STEMBER, L H. A Methodology for Photovoltaic System Reliability & Economic Analysis. **IEEE TRANSACTIONS ON RELIABILITY**, [s. l.], n. 3, 1982.

STEMBER, L H. Reliability considerations in the design of solar photovoltaic power systems. **Solar Cells**, [s. l.], v. 3, n. 3, p. 269–285, 1981.

STI NORLAND. **STI-H1250: Centralized Single-Axis Tracker, Technical Data**. [S. l.], 2021. Disponível em: https://web.archive.org/web/20210820175942/https://www.stinorland.com/sites/default/files/sti_norland_-_data_sheet_sti_h1250.pdf. Acesso em: 15 nov. 2021.

SYAMSUDDIN, A. The reliability study of Gili Trawangan photovoltaic system using RBD (Reliability Block Diagram) method. **IOP Conference Series: Materials Science and Engineering**, [s. l.], v. 1098, n. 4, p. 042051, 2021.

TAYLOR, Karl E. Summarizing multiple aspects of model desempenho in a single diagram. **Journal of Geophysical Research Atmospheres**, [s. l.], v. 106, n. D7, p. 7183–7192, 2001.

THERISTIS, Marios; PAPAZOGLOU, Ioannis A. Markovian reliability analysis of standalone photovoltaic systems incorporating repairs. **IEEE Journal of Photovoltaics**, [s. l.], v. 4, n. 1, p. 414–422, 2014.

VANTAGE MARKET RESEARCH. **Solar Tracker Market - Global Industry Assessment & Forecast**. [S.l.], 2023. Disponível em: <https://web.archive.org/web/20230131161858/https://www.vantagemarketresearch.com/industry-report/solar-tracker-market-1800>. Acesso em: 17 jun. 2023.

VÁZQUEZ, Manuel; REY-STOLLE, Ignacio. Photovoltaic module reliability model based on field degradation studies. **Progress in Photovoltaics: Research and Applications**, [s. l.], v. 16, n. 5, p. 419–433, 2008.

WAITE, E; MAVEC, J; WOLOSEWICZ, R. R. Wolosewicz Evaluation of Reliability of Operational Solar-Energy Systems. **Journal of Solar Energy Engineering**, [s. l.], v. 104, n. 41, 1982. Disponível em: <http://solarenergyengineering.asmedigitalcollection.asme.org/>.

WANG, Zhi Ming; YANG, Jian Guo. Numerical method for Weibull generalized renewal process and its applications in reliability analysis of NC machine tools. **Computers and Industrial Engineering**, [s. l.], v. 63, n. 4, p. 1128–1134, 2012.

WITOWSKI, Katharina; STANDER, Nielen. Parameter identification of hysteretic models using Partial Curve Mapping. *Em: , 2012. 12th AIAA Aviation Technology, Integration and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*. [S. l.: s. n.], 2012.

WOHLGEMUTH, John H *et al.* LONG TERM RELIABILITY OF PHOTOVOLTAIC MODULES. *Em: , 2006. IEEE 4th World Conference on Photovoltaic Energy Conference*. [S. l.: s. n.], 2006. p. 2050–2053.

YAN, Weian; LIU, Weidong; KONG, Wenqi. Reliability evaluation of PV modules based on exponential dispersion process. **Energy Reports**, [s. l.], v. 7, p. 3023–3032, 2021.

ZHANG, Peng *et al.* **Reliability assessment of photovoltaic power systems: Review of current status and future perspectives**. [S. l.]: Elsevier Ltd, 2013.