



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m12e/2025/09.01.19.40-TDI

PREVISÃO DE IRRADIÂNCIA E SOBREIRRADIÂNCIA NA AMAZÔNIA COM APRENDIZADO DE MÁQUINA

Vinicius Rogério da Rocha

Tese de Doutorado do Curso de Pós-Graduação em Meteorologia, orientada pelos Drs. Gilberto Fernando Fisch, e Rodrigo Santos Costa, aprovada em 14 de agosto de 2025.

URL do documento original:

<<http://urlib.net/8JMKD2USNRW34T/4E6LK58>>

INPE
São José dos Campos
2025

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

**CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO
INTELECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-
INPE):**

Presidente:

Dr. Thales Sehn Korting - Coordenação-Geral de Ciências da Terra (CGCT)

Membros:

Dr. Antonio Fernando Bertachini de Almeida Prado - Conselho de Pós-Graduação (CPG)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e Ciência Espaciais (CGCE)

Dr. Heyder Hey - Coordenação-Geral de Infraestrutura e Pesquisas Aplicadas (CGIP)

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

EDITORAÇÃO ELETRÔNICA:

Ivone Martins - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m12e/2025/09.01.19.40-TDI

PREVISÃO DE IRRADIÂNCIA E SOBREIRRADIÂNCIA NA AMAZÔNIA COM APRENDIZADO DE MÁQUINA

Vinicius Roggério da Rocha

Tese de Doutorado do Curso de Pós-Graduação em Meteorologia, orientada pelos Drs. Gilberto Fernando Fisch, e Rodrigo Santos Costa, aprovada em 14 de agosto de 2025.

URL do documento original:

<<http://urlib.net/8JMKD2USNRW34T/4E6LK58>>

INPE
São José dos Campos
2025

Dados Internacionais de Catalogação na Publicação (CIP)

Rocha, Vinicius Roggério da.

R582p Previsão de irradiância e sobreirradiância na Amazônia com
aprendizado de máquina / Vinicius Roggério da Rocha. – São
José dos Campos : INPE, 2025.

xxii + 123 p. ; (sid.inpe.br/mtc-m12e/2025/09.01.19.40-TDI)

Tese (Doutorado em Meteorologia) – Instituto Nacional de
Pesquisas Espaciais, São José dos Campos, 2025.

Orientadores : Drs. Gilberto Fernando Fisch, e Rodrigo Santos
Costa.

1. Nuvens. 2. Energia solar. 3. Radiação atmosférica.
4. Inteligência artificial. 5. Sobreirradiância. I.Título.

CDU 551.521.17:681.3.019(811.3)



Esta obra foi licenciada sob uma Licença Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License.



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO



INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

FOLHA DE APROVAÇÃO

Aluno(a): Vinicius Roggério da Rocha

Título: “Previsão de irradiância e sobreirradiância na Amazônia com aprendizado de máquina”

Aprovado(a) pela Banca Examinadora em cumprimento ao requisito exigido para obtenção do Título Doutor (a) em Meteorologia

(assinado eletronicamente)

Dr. Gilberto Fernando Fisch – Orientador/Presidente – INPE

Aprovado Reprovado

(assinado eletronicamente)

Dr. Rodrigo Santos Costa – Orientador (a) – INPE

Aprovado Reprovado

(assinado eletronicamente)

Dra. Simone Marilene Sievert da Costa Coelho – Membro Interno – INPE

Aprovado Reprovado

(assinado eletronicamente)

Dr. António Eduardo de Barros Ruano – Membro Externo – UALG

Aprovado Reprovado

(assinado eletronicamente)

Dra. Marcia Akemi Yamasoe – Membro Externo – USP

Aprovado Reprovado

Este trabalho foi aprovado por:

maioria simples

unanimidade

Declaração de aprovação do membro **Dr. António Eduardo de Barros Ruano** - anexa ao processo.

São José dos Campos 14 de agosto de 2025



Documento assinado eletronicamente por **gilberto fernando fisch (E), Usuário Externo**, em 20/08/2025, às 16:17 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcia Akemi Yamasoe (E), Usuário Externo**, em 20/08/2025, às 16:22 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Simone Marilene Sievert da Costa Coelho, Pesquisadora**, em 21/08/2025, às 07:46 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rodrigo Santos Costa, Tecnologista**, em 25/08/2025, às 09:38 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site <https://sei.mcti.gov.br/verifica.html>, informando o código verificador **13020062** e o código CRC **303A317F**.

*“Por el aire andan plácidas montañas
o cordilleras trágicas de sombra
que oscurecen el día. Se las nombra
nubes. Las formas suelen ser extrañas.
Shakespeare observó una. Parecía
un dragón. Esa nube de una tarde
en su palabra resplandece y arde
y la seguimos viendo todavía.
Qué son las nubes? Una arquitetura
del azar? Quizá Dios las necesita
para la ejecución de Su infinita
obra y son hilos de la trama oscura.
Quizá la nube sea no menos vana
que el hombre que la mira en mañana.”*

Jorge Luis Borges. Nuvens II. Obras completas, volume III.

“Quem é que tem sabedoria para avaliar as nuvens?”

Jó 38:37 (ACF)

AGRADECIMENTOS

A Deus, por me acompanhar em todos os momentos, me dando força, saúde, sabedoria e amor.

À minha mulher, Eliana, por acompanhar e dar apoio em todos os momentos. Aos meus pais, José Reinaldo e Maria Auxiliadora, pelo apoio.

Aos meus orientadores e ao grupo de pesquisa no LABREN, pelas conversas e aprendizado.

Ao corpo docente, técnico e administrativo do programa de Pós-Graduação em Meteorologia do INPE, pela oportunidade da realização deste trabalho.

Ao prof. António Ruano, da UAlg, pela recepção e pelo valioso apoio durante o período do meu doutorado sanduíche, cuja orientação e disponibilidade foram fundamentais para o desenvolvimento deste trabalho.

À prof^a. Marcia Akemi Yamasoe, do IAG/USP, pela oportunidade de trabalho para realização de estágio de docência.

À CAPES, pelo apoio financeiro durante o doutorado regular e sanduíche (processos nº 88887.601400/2021-00 e 88887.885659/2023-00, respectivamente).

RESUMO

Este estudo investigou a caracterização e a previsão de irradiância solar (GHI) em curtíssimo prazo na Amazônia Central, utilizando medidas de irradiância em superfície e estimativas de fração de cobertura de nuvens a partir de uma câmera *all sky*. Foram realizadas análises descritivas das variáveis envolvidas, verificando-se que a variabilidade está positivamente correlacionada com o número de eventos de sobreirradiância e que nuvens baixas são as principais responsáveis por eventos intensos de sobreirradiância e rampas abruptas de GHI, dentro da metodologia proposta. Foi proposta uma nova métrica (o Índice Dinâmico Agregado de Sobreirradiância, DAOI), capaz de capturar padrões temporais em nível de minuto, diário e anual associados à dinâmica das nuvens. Modelos autorregressivos não lineares com entradas exógenas (NARX) foram utilizados, implementados com redes neurais artificiais e otimizados por meio de algoritmos genéticos multiobjetivo (MOGA), com diferentes restrições e combinações de variáveis de entrada. Os modelos superaram a linha de base da persistência, com reduções de RMSE de até 32,5% nos primeiros 10 minutos de previsão, mesmo utilizando significativamente menos informações de entrada do que estudos anteriores - um desafio ampliado pelas características atmosféricas da região tropical. O modelo com melhor desempenho na previsão de GHI em curtíssimo prazo apresentou alta habilidade na detecção de eventos de sobreirradiância, especialmente sob condições de céu variável, mantendo desempenho confiável por até 20–30 minutos à frente (medida-F1 média de 0,65 nesse intervalo). Esses resultados contribuem para o avanço da previsão de variabilidade solar em escalas sub-horárias, com implicações relevantes tanto para a meteorologia aplicada quanto para o planejamento e operação de sistemas fotovoltaicos.

Palavras-chave: Nuvens. Energia solar. Radiação atmosférica. Inteligência artificial. Sobreirradiância.

OVERIRRADIANCE AND IRRADIANCE PREDICTION IN THE AMAZON WITH MACHINE LEARNING

ABSTRACT

This study investigated the characterization and ultra-short-term forecasting of solar irradiance (GHI) in the Central Amazon, using ground-based irradiance measurements and cloud fraction estimates derived from an all sky camera. Descriptive analyses of the involved variables revealed that irradiance variability is positively correlated with the number of overirradiance events and that low-level clouds are the main contributors to both intense overirradiance events and abrupt GHI ramps, according to the proposed methodology. A novel metric was introduced (DAOI, Dynamic Aggregated Overirradiance Index), capable of capturing temporal patterns at minute, daily, and annual scales associated with cloud dynamics. Nonlinear autoregressive models with exogenous inputs (NARX) were implemented using artificial neural networks and optimized through multi-objective genetic algorithms (MOGA), under different constraints and input variable combinations. The models outperformed the persistence baseline, achieving RMSE reductions of up to 32.5% in the first 10 minutes of forecasting, despite relying on significantly fewer input variables than previous studies - a challenge further compounded by the complexity of the tropical atmosphere. The best-performing model for ultra-short-term GHI forecasting demonstrated high skill in detecting overirradiance events, particularly under variable sky conditions, maintaining reliable performance up to 20–30 minutes ahead (average F1-score of 0.65 in this interval). These findings contribute to the advancement of sub-hourly solar variability forecasting, with relevant implications for both applied meteorology and the planning and operation of photovoltaic systems.

Keywords: Clouds. Solar energy. Atmospheric radiation. Artificial intelligence. Overirradiance.

LISTA DE FIGURAS

	<u>Pág.</u>
1.1 Representação esquemática dos quatro mecanismos propostos que modulam a variabilidade da irradiância solar na superfície. Da subfigura (a) à (d), os mecanismos são: escape frontal (<i>forward escape</i>), escape para baixo (<i>downward escape</i>), escape lateral (<i>side escape</i>) e realce por albedo (<i>albedo enhancement</i>). A radiação solar incidente é representada pela seta grossa voltada para baixo; as demais setas indicam, de forma aproximada, a direção média e a intensidade relativa da radiação espalhada. O albedo da superfície é representado pela tonalidade mais clara. Apenas as setas relevantes para cada mecanismo estão representadas.	8
2.1 Localização da estação de pesquisa T3 instalada durante experimento GoAmazon.	34
2.2 Parte do sítio experimental T3 e instrumentos (da esquerda para a direita): Total Sky Imager (TSI), fotômetro solar e plataforma Sky Radiation, composta por piranômetros sombreados para medição da irradiância difusa e não sombreados para medição da irradiância global (GHI).	35
2.3 Piranômetro Eppley PSP.	36
2.4 Total Sky Imager (TSI).	37
2.5 Fotômetro solar (modelo CIMEL).	39
2.6 Pluviômetro do tipo Tipping Bucket (modelo 260-2500), utilizado em experimentos de campo para medição da precipitação.	40
2.7 Fluxograma da metodologia desenvolvida neste estudo.	54
3.1 Integral diária da irradiância solar global (H diária) e média mensal (H mensal).	60
3.2 Distribuições dos valores diários e mensais de GHI (acima) e CF (abaixo), representadas por diagramas de caixa, com indicação da mediana, dos quartis e dos valores atípicos (<i>outliers</i>).	62
3.3 Valores médios diários de AOD (acima) e precipitação acumulada (abaixo) para todo o período analisado.	64
3.4 Ciclo diurno médio de GHI, CSI (eixo esquerdo, W/m ²) e CF (eixo direito, %), com base em dados de 2014–2015.	65
3.5 Ciclo diurno da média das rampas de irradiância (em W/m ²).	67
3.6 DARR diários (pontos) e médias mensais (linhas), com diferenças temporais de 1, 2 e 5 minutos (azul, verde e vermelho, respectivamente).	68

3.7	Número de eventos de sobreirradiância <i>versus</i> DARR, com linha de regressão linear (em vermelho) e respectivas equação ajustada e coeficiente de correlação de Pearson.	70
3.8	DAOI calculado para dias e meses considerando o CSI como referência, em função do tempo, utilizando médias de 1, 2 e 5 minutos.	72
3.9	Ciclo diurno com a soma de eventos de sobreirradiância por minuto, utilizando o CSI como referência, e período de dados suspeitos em vermelho.	74
3.10	Ciclo diurno da magnitude da sobreirradiância, representado pelas médias por minuto das diferenças positivas GHI–CSI (DAOI).	75
3.11	Histograma das magnitudes de sobreirradiância sem restrição (azul, com valores indicados sobre cada barra) e com restrição de $GHI > 1000 \text{ W/m}^2$ (vermelho), com CSI como referência.	77
3.12	Histograma com a distribuição do número de eventos de sobreirradiância conforme sua duração, em minutos, utilizando o CSI como referência, sem (azul) e com restrição para eventos com $GHI > 1000 \text{ W/m}^2$ (vermelho).	78
3.13	Imagens do céu total correspondentes aos três maiores eventos de sobreirradiância em 2 de março de 2015, nos horários 16:27, 16:28 e 16:29 UTC (da esquerda para a direita).	81
3.14	Distribuição percentual mensal dos principais tipos de nuvem observados através do produto ARSCL.	83
3.15	Distribuição das rampas de irradiância solar global (GHI) com módulo maior ou igual a 100 W/m^2 para cada tipo de nuvem, com os valores das medianas escritos nas respectivas caixas.	86
3.16	Evolução dos erros (MAE, linha contínua, e RMSE, linha tracejada) das previsões com frequência de 2 minutos para cada um dos modelos de persistência: baseado no instante anterior (PersistPrev), no mesmo horário do dia anterior (PersistDay) e combinação sazonal (PersistCombined).	88
3.17	Comparação do erro absoluto médio (MAE, a) e do erro absoluto médio normalizado (nMAE, b) ao longo dos passos de previsão considerados neste estudo. São apresentados os casos definidos na Tabela 2.4, incluindo o modelo de persistência (PersistPrev), bem como resultados de outros estudos (identificados pelo sobrenome do primeiro autor e ano de publicação).	90

3.18	Comparação do erro quadrático médio (RMSE, a) e do erro quadrático médio normalizado (nRMSE, b) ao longo dos passos de previsão considerados neste estudo. São apresentados os casos definidos na Tabela 2.4, incluindo o modelo de persistência (PersistPrev) e resultados de outros estudos (identificados pelo sobrenome do primeiro autor e ano de publicação). Fonte: elaboração própria.	91
3.19	Métricas de desempenho da detecção de eventos de sobreirradiância pelo modelo ghicf30cr para diferentes passos de previsão.	94
3.20	Distribuição percentual diária dos principais tipos de nuvem observados através do produto ARSCL para o período entre 2 e 15 de abril de 2014.	97

LISTA DE TABELAS

	<u>Pág.</u>
1.1 Classificação das nuvens segundo pressão no topo e espessura óptica, conforme o ISCCP.	6
2.1 Critérios de classificação de tipos de nuvens com base em altura da base, altura do topo e espessura.	42
2.2 Número de registros diurnos sem valores de GHI por dia.	43
2.3 Tamanho dos conjuntos de dados utilizados para treinamento, teste e validação dos modelos em quatro diferentes configurações.	46
2.4 Resumo dos resultados obtidos com os modelos MOGA para diferentes conjuntos de dados. A tabela apresenta os erros quadráticos médios (RMSE) para os conjuntos de treinamento (tr.) e teste (te.), a complexidade média dos modelos (Complex), o número total de soluções não dominadas (ND) e, quando aplicável, o número de modelos preferenciais selecionados (Pref.).	49
3.1 Eventos com maior excesso percentual de GHI em relação ao modelo de céu claro: data (ano, mês, dia, hora UTC), valor de GHI, valor de CF e "aumento de GHI", que representa o quanto o GHI observado excedeu a irradiância estimada para céus claros, em porcentagem.	80
3.2 Distribuição percentual dos tipos de nuvem em três condições distintas: todos os dados com classificação de nuvem, apenas casos com sobreirradiância (OI), e sobreirradiância com GHI superior a 1000 W/m^2 ($\text{OI} / \text{GHI} > 1000$), em porcentagens.	85
3.3 Desempenho médio do modelo ao longo dos horizontes de previsão considerados, com erros absolutos (MAE) e quadráticos médios (RMSE), tanto em valores absolutos (W/m^2) quanto normalizados (nMAE, nRMSE).	92
3.4 Desempenho médio do modelo na detecção de eventos de sobreirradiância ao longo de diferentes horizontes de previsão.	95
3.5 Métricas de desempenho na detecção de eventos pelo modelo escolhido para cada dia, valores médios de GHI (W/m^2) e CF (%), número de eventos de sobreirradiância (N) e DARR.	97

LISTA DE ABREVIATURAS E SIGLAS

ABNT	–	Associação Brasileira de Normas Técnicas
AERONET	–	<i>Aerosol Robotic Network</i>
AOD	–	<i>Aerosol Optical Depth</i>
ANN	–	<i>Artificial Neural Network</i>
ARSCL	–	<i>Active Remote Sensing of Clouds</i>
ARM	–	<i>Atmospheric Radiation Measurement</i>
ASI	–	<i>All Sky Image</i>
CAPE	–	<i>Convective Available Potential Energy</i>
CCD	–	<i>Charge-Coupled Device</i>
CEIL	–	<i>Vaisala Laser Ceilometer</i>
CF	–	<i>Cloud Fraction</i>
CH	–	<i>Convex Hull</i>
CIMEL	–	[Marca do fotômetro solar]
CLP	–	Camada Limite Planetária
CNN	–	<i>Convolutional Neural Network</i>
COD	–	<i>Cloud Optical Depth</i>
CSI	–	<i>Clear Sky Index</i>
DAOI	–	<i>Dynamic Aggregate Overirradiance Index</i>
DARR	–	<i>Daily Aggregate Ramp Rate</i>
DNI	–	<i>Direct Normal Irradiance</i>
ENSO	–	<i>El Niño–Southern Oscillation</i>
FN	–	<i>Falso Negativo</i>
FP	–	<i>Falso Positivo</i>
GHI	–	<i>Irradiância Global Horizontal</i>
GoAmazon	–	<i>Green Ocean Amazon</i>
HAEANN	–	<i>Hybrid Attention-based Artificial Neural Network</i>
ISCCP	–	<i>International Satellite Cloud Climatology Project</i>
LS-SVM	–	<i>Least-squares support vector machine</i>
LSTM	–	<i>Long Short-Term Memory</i>
MAE	–	<i>Mean Absolute Error</i>
MLP	–	<i>Multi Layer Perceptron</i>
MOGA	–	<i>Multi-Objective Genetic Algorithm</i>
MPL	–	<i>Micropulse Lidar</i>
NAR	–	<i>Nonlinear Autoregressive</i>
NARX	–	<i>Nonlinear Autoregressive with Exogenous Inputs</i>
NASA	–	<i>National Aeronautics and Space Administration</i>
ND	–	<i>Non-Dominated</i>
NOAA	–	<i>National Oceanic and Atmospheric Administration</i>
nMAE	–	<i>Normalized Mean Absolute Error</i>
nRMSE	–	<i>Normalized Root Mean Square Error</i>
OMM	–	Organização Meteorológica Mundial
ONI	–	<i>Oceanic Niño Index</i>
PCHIP	–	<i>Piecewise Cubic Hermite Interpolating Polynomial</i>

PSA	–	Plataforma Solar de Almería
PSP	–	<i>Precision Spectral Pyranometer</i>
RBF	–	<i>Radial Basis Function</i>
RF	–	<i>Random Forest</i>
RMSE	–	<i>Root Mean Square Error</i>
RR	–	<i>Ramp Rate</i>
RWP	–	<i>Radar Wind Profiler</i>
SKYRAD	–	<i>Sky Radiation</i>
SVM	–	<i>Support Vector Machines</i>
TN	–	<i>True Negative</i>
TOA	–	<i>Top of Atmosphere</i>
TP	–	<i>True Positive</i>
TSI	–	<i>Total Sky Imager</i>
UTC	–	<i>Universal Time Coordinated</i>
VN	–	<i>Verdadeiro Negativo</i>
VP	–	<i>Verdadeiro Positivo</i>
WACR	–	<i>W-band Cloud Radar</i>
ZCIT	–	Zona de Convergência Intertropical

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
1.1 Contexto e relevância da energia solar	1
1.2 Variabilidade da irradiância	3
1.3 Contextualização meteorológica	5
1.3.1 Ciclo de vida das nuvens na região amazônica	9
1.4 Clima da Amazônia Central	14
1.5 Fundamentação da estrutura de modelagem	17
1.5.1 ApproxHull	17
1.5.2 MOGA	18
1.5.3 NAR(X)	21
1.6 Trabalhos relacionados de previsão de irradiância	22
1.7 Sobreirradiância: definição e importância	26
1.8 Motivação	31
1.9 Objetivo geral e específicos	31
1.10 Estrutura da tese	32
2 METODOLOGIA	33
2.1 Dados	33
2.1.1 Irradiância Global Horizontal	35
2.1.2 Fração de cobertura de nuvens	36
2.1.3 Profundidade Óptica do Aerossol e precipitação	38
2.1.4 Classificação de nuvens	40
2.2 Pré-processamento	42
2.2.1 Atributos para modelagem	43
2.2.2 Métrica para representação de sobreirradiância	44
2.3 Estrutura de modelagem	45
2.3.1 Métricas de desempenho	50
2.3.2 Modelos de persistência	52
2.3.3 Fluxograma do processo metodológico	53
2.3.4 Análise estatística dos resultados	53
3 RESULTADOS	57
3.1 Variações sazonais e interanuais	58

3.2	Caracterização da variabilidade	66
3.3	Caracterização de eventos de sobreirradiância	70
3.4	Caracterização dos tipos de nuvens	81
3.5	Avaliação de desempenho dos modelos de previsão	87
3.6	Detecção de eventos de sobreirradiância com base nas previsões	93
4	CONCLUSÕES	99
	REFERÊNCIAS BIBLIOGRÁFICAS	105
	APÊNDICE A - ALGORITMO DE CONSTRUÇÃO DO	
	APPROXHULL	123

1 INTRODUÇÃO

1.1 Contexto e relevância da energia solar

O aumento da demanda e as limitações no acesso à energia continuam sendo desafios importantes em diversas regiões do mundo, especialmente em países localizados em áreas tropicais. Nesses contextos, a energia solar se apresenta como uma alternativa promissora para suprir o déficit energético e promover o desenvolvimento sustentável. A expansão da energia solar como fonte de geração elétrica nas últimas décadas é um fenômeno global (REN21, 2025). No Brasil, a capacidade instalada de geração fotovoltaica saltou de 1.160 MW em 2017 para 53.588 MW no final de 2024, correspondendo a aproximadamente 23% da matriz elétrica brasileira (ASSOCIAÇÃO BRASILEIRA DE ENERGIA SOLAR FOTOVOLTAICA (ABSOLAR), 2025). Esse crescimento exponencial reflete tanto os avanços tecnológicos quanto os incentivos regulatórios e econômicos voltados para a diversificação da matriz energética.

Com relação à região amazônica, o Atlas Brasileiro de Energia Solar (PEREIRA et al., 2017) evidencia que a elevada nebulosidade limita o potencial de geração solar. Ainda assim, mesmo as áreas menos favoráveis do Brasil apresentam recurso solar superior ao de países mais distantes dos trópicos como a Alemanha (REN21, 2025), que já se destaca mundialmente em investimentos e capacidade instalada, o que demonstra a viabilidade da tecnologia também na Amazônia. Trata-se de uma região já fortemente pressionada pelo desmatamento, o que torna inadequada a abertura de grandes extensões de floresta para implantação de usinas centralizadas. Como permanece a necessidade de suprimento de energia elétrica, isso leva os sistemas isolados a dependerem de duas alternativas problemáticas: (i) ligação através de longas linhas de transmissão, com elevado custo e impacto ambiental, ou (ii) o abastecimento por combustíveis fósseis, oneroso e poluente. Nesse contexto, soluções descentralizadas, como sistemas fotovoltaicos em pequenas áreas próximas aos centros consumidores, utilizando-se de espaços já degradados ou mesmo de usinas flutuantes, configuram alternativas a serem avaliadas, por reduzirem a necessidade de novos desmatamentos. É fundamental, contudo, que tais estratégias sejam analisadas à luz das especificidades locais, dado que a solução adequada para um ponto da região não necessariamente se aplica a outro.

Um ponto importante a ser considerado é que a geração solar fotovoltaica é altamente dependente da disponibilidade de irradiância solar, a qual sofre variações em diversas escalas espaciais e temporais, influenciadas tanto por fatores astronômicos (como a posição aparente do Sol devido aos movimentos da Terra) quanto por

condições meteorológicas e climáticas. Essa variabilidade representa um desafio significativo para a integração segura e eficiente da energia solar nos sistemas elétricos. À medida que a participação da energia solar aumenta, intensificam-se as discussões sobre os impactos dessa variabilidade na qualidade e confiabilidade do fornecimento elétrico, assim como sobre as estratégias para sua mitigação e o aprimoramento do conhecimento sobre o potencial solar disponível, fundamental para o aproveitamento eficiente da tecnologia e o planejamento energético de longo prazo.

Para que a operação e o planejamento do recurso solar sejam efetivos, é essencial contar com mecanismos confiáveis de previsão da irradiância solar global à superfície. A previsão é uma ferramenta estratégica, pois permite antecipar flutuações na geração fotovoltaica e viabilizar ações preventivas e corretivas por parte dos operadores do sistema. Entre as principais aplicações da previsão de irradiância solar no setor elétrico, destacam-se (KUHN et al., 2018; MARTINS et al., 2022b):

- Planejamento do despacho dos geradores convencionais;
- Cálculo de reservas operacionais;
- Planejamento do intercâmbio de energia entre áreas de controle;
- Especificação de limites operacionais para usinas fotovoltaicas;
- Otimização da operação de sistemas de armazenamento de energia;
- Redução da capacidade nominal necessária de armazenamento;
- Controle das funcionalidades dos inversores;
- Gerenciamento ativo da demanda;
- Planejamento de manutenções programadas;
- Apoio às estratégias de comercialização de energia;
- Evitar episódios de *curtailment* (corte de geração de energia elétrica devido a limitações na capacidade de escoamento da energia gerada para outras regiões).

As previsões de irradiância podem ocorrer em diferentes escalas temporais: de curtíssimo prazo (alguns minutos a uma hora), curto prazo (de algumas horas até uma

semana), médio prazo (de semanas até um ano) e longo prazo (vários anos) (LAOUAFI et al., 2022). As previsões de curtíssimo prazo apresentam aplicação direta na operação de sistemas fotovoltaicos e na gestão da rede elétrica. Nesse intervalo, previsões de 2 a 10 minutos permitem que o operador do empreendimento adote medidas imediatas de proteção dos equipamentos, como o ajuste ou desligamento de inversores durante episódios de sobreirradiância. Já horizontes intermediários, de 20 a 30 minutos, oferecem suporte à operação do sistema, possibilitando maior preparo para oscilações relevantes na geração. Por fim, previsões de até 60 minutos são especialmente úteis no planejamento do despacho energético, auxiliando o operador do sistema elétrico a acionar fontes alternativas e evitar instabilidades na rede.

A variabilidade da irradiância é determinada principalmente pela geometria solar e pela cobertura de nuvens. A geometria solar, por sua vez, é influenciada por fatores astronômicos como a distância entre a Terra e o Sol, a declinação solar (ambos associados ao movimento de translação) e o ângulo horário, relacionado à rotação da Terra. Elementos atmosféricos adicionais, incluindo carga de aerossóis, conteúdo de água precipitável e albedo da superfície (por meio do retroespalhamento da radiação refletida), afetam a irradiância solar da superfície como um fator de segunda ordem. Os valores de irradiância e as participações de cada elemento no balanço de energia apresentam uma variabilidade em diferentes escalas de tempo. A variabilidade de escala anual apresenta uma sazonalidade associada à inclinação do eixo de rotação terrestre em relação ao plano de translação, além de fatores climáticos. Na escala diária, está associada ao movimento de rotação da Terra em torno do seu próprio eixo. Já a variabilidade de escala horária ou sub-horária está associada aos fenômenos atmosféricos, como a passagem de nuvens, ocorrência de chuvas e elevadas concentrações de aerossóis (naturais ou antropogênicos).

Apesar dos avanços recentes, a previsão da irradiância solar ainda enfrenta desafios importantes, especialmente em regiões tropicais como a Amazônia. Dentre os principais obstáculos, destacam-se a alta variabilidade do recurso solar e a ocorrência de episódios de sobreirradiância - ambos diretamente relacionados a processos atmosféricos complexos, como a formação, evolução e estrutura das nuvens.

1.2 Variabilidade da irradiância

Uma das formas mais comuns de se quantificar a variabilidade de curto prazo da irradiância solar incidente na superfície descritas na literatura é através de índices baseados no conceito de rampa (LAVE et al., 2013). Um evento de rampa, quantificado pela métrica conhecida como taxa de rampa (RR, *Ramp Rate*), consiste na

diferença da irradiância entre dois instantes consecutivos, podendo ser caracterizado por atributos como magnitude (em unidades de irradiância), duração (em unidades de tempo), direção (aumento ou diminuição da irradiância) e impulso (razão entre irradiância e tempo) (HANSEN et al., 2010).

Com o objetivo de avaliar a disponibilidade e a estabilidade do recurso solar sob diferentes condições atmosféricas, Rocha et al. (2022) apresentam uma revisão dos principais índices já propostos na literatura para quantificação da variabilidade da irradiância solar global horizontal na superfície (GHI, *Global Horizontal Irradiance*). Esse estudo faz uso da Taxa de Rampa Agregada Diária (DARR, *Daily Aggregate Ramp Rate*), índice inicialmente introduzido por Haaren et al. (2012) e que representa a soma dos valores absolutos das variações de irradiância entre instantes consecutivos ao longo de um dia, normalizada por uma constante $C = 1000 \text{ W/m}^2$ - as diferenças entre esses instantes foram consideradas como 1, 2 e 5 minutos no atual trabalho. O DARR, cuja formulação é apresentada na Equação 1.1, permite quantificar a variabilidade da irradiância sem a necessidade de um modelo de céu claro como referência, o que representa uma vantagem prática importante em termos de aplicabilidade.

$$DARR = \sum_{t=1}^{N-1} \frac{|GHI_t - GHI_{t-1}|}{C} \quad (1.1)$$

Ainda no trabalho de Rocha et al. (2022), observa-se que a relação entre a irradiância solar e sua variabilidade forma uma estrutura em 'ponta de flecha' quando se representa o índice de claridade K_t - razão entre GHI e a irradiância no topo da atmosfera (TOA, *Top of Atmosphere*) - em função do DARR, indicando que dias com o mesmo nível médio de irradiância podem apresentar variabilidade muito distinta. Esse resultado reforça que métricas de magnitude, como GHI ou K_t , não são suficientes para inferir a variabilidade de curto prazo, destacando a importância da inclusão de índices específicos como o DARR para uma caracterização mais robusta e independente do recurso solar.

Um modelo pode ter bom desempenho em estatísticas médias (como o erro quadrático médio, por exemplo), mas ainda assim apresentar outro comportamento em representar a variabilidade temporal, especialmente as rampas rápidas e frequentes. Nesse contexto, a definição de um limiar para a taxa de rampa permite converter o problema contínuo da previsão de rampas em uma tarefa de classificação binária, o que viabiliza a aplicação de métricas de classificação (acurácia, precisão, etc.) na

avaliação específica da ocorrência desses eventos.

O estudo de [Logothetis et al. \(2022\)](#) avalia o uso de câmeras *all sky* para previsão de rampas de irradiância solar, mostrando que esses sistemas podem identificar de 26% a 92% dos eventos reais, com desempenho variável entre diferentes câmeras, e sugere que a combinação de métodos físicos e de aprendizado profundo pode aprimorar a previsão desses eventos críticos. O limiar aqui é definido com base nos valores de taxa de rampa que excedem o percentil 99 das variações observadas em dias de céu claro, caracterizando como rampas apenas as flutuações que superam a variabilidade esperada sob condições atmosféricas estáveis. O estudo de [Zuo et al. \(2023\)](#) propõe um modelo de rede neural convolucional acoplada a uma outra rede com memória de curto e longo prazo (CNN-LSTM, *Convolutional Neural Network - Long Short-Term Memory*) que integra imagens *all sky* (ASI, *All Sky Image*) e séries históricas de GHI para prever a irradiância nos próximos 10 minutos. Nesse caso, os autores calcularam as taxas de rampa normalizadas pela irradiância de céu claro e fixaram o limiar em 0,05, dividindo em rampas fortes e fracas. Por outro lado, [Nouri et al. \(2024\)](#) propõem um limiar mais dinâmico, ajustado de acordo com a componente de irradiância analisada e com o impacto operacional desejado, podendo ser baseado em valores típicos, percentuais da irradiância máxima ou quantis da distribuição histórica de taxas de rampa.

1.3 Contextualização meteorológica

A GHI resulta da atenuação da irradiância extraterrestre ao atravessar a atmosfera. Esta última corresponde ao fluxo de energia solar no TOA sobre uma superfície perpendicular aos raios solares, sendo determinada por fatores puramente astronômicos, como a posição aparente do Sol e a distância Terra-Sol ao longo do ano. Esse fluxo, conhecido como constante solar, tem valor médio de aproximadamente 1361 W/m^2 ([NATIONAL AERONAUTICS AND SPACE ADMINISTRATION \(NASA\), 2025](#)), representando a irradiância incidente fora da atmosfera terrestre sob condições médias e em superfície normal ao Sol. ([YAMASOE; CORRÊA, 2016](#)). No entanto, a quantidade efetiva de radiação que atinge a superfície depende de processos de atenuação atmosférica, incluindo absorção e espalhamento por moléculas gasosas, aerossóis (partículas sólidas ou líquidas em suspensão na atmosfera) e, principalmente, nuvens. Nesse contexto, a radiação global que chega ao solo pode ser decomposta em duas componentes: a direta, que atinge a superfície sem ser desviada em seu trajeto a partir do sol, e a difusa, que resulta do espalhamento da radiação em diferentes direções antes de atingir o solo.

Segundo a classificação da Organização Meteorológica Mundial (OMM), as nuvens são categorizadas conforme a altura da base em baixas (*stratus*, *stratocumulus*, *cumulus*, *cumulonimbus*), médias (*nimbostratus*, *altostratus*, *altocumulus*) e altas (*cirrus*, *cirrostratus*, *cirrocumulus*) (WORLD METEOROLOGICAL ORGANIZATION (WMO), 2017). A nuvem *cumulus congestus* (gênero *cumulus* e espécie *congestus*) apresenta destaque no contexto desse estudo por também apresentar grande desenvolvimento vertical. Segundo (ROSSOW; SCHIFFER, 1999), no âmbito do projeto ISCCP (*International Satellite Cloud Climatology Project*), as nuvens podem ser classificadas com base em duas propriedades físicas principais observáveis por sensores remotos: a pressão no topo da nuvem e a espessura óptica. A pressão no topo da nuvem está relacionada à altitude da formação, enquanto a espessura óptica expressa a capacidade da nuvem em atenuar a radiação solar. A combinação dessas duas variáveis permite distinguir diferentes tipos de nuvem em termos de altura e densidade óptica, fornecendo uma base objetiva para sua categorização em dados globais de satélite. A Tabela 1.1 apresenta a classificação das principais classes de nuvens utilizadas no ISCCP.

Tabela 1.1 - Classificação das nuvens segundo pressão no topo e espessura óptica, conforme o ISCCP.

Tipo de nuvem	Pressão no topo (hPa)	Espessura óptica
<i>Stratus</i>	1000–680	0–3,6
<i>Stratocumulus</i>	1000–680	3,6–23
<i>Cumulus</i>	1000–680	23–379
<i>Altostratus</i>	680–440	0–3,6
<i>Altostratus</i>	680–440	3,6–23
<i>Nimbostratus</i>	680–440	23–379
<i>Cirrus</i>	440–50	0–3,6
<i>Cirrostratus</i>	440–50	3,6–23
<i>Nuvens profundas</i>	440–50	23–379

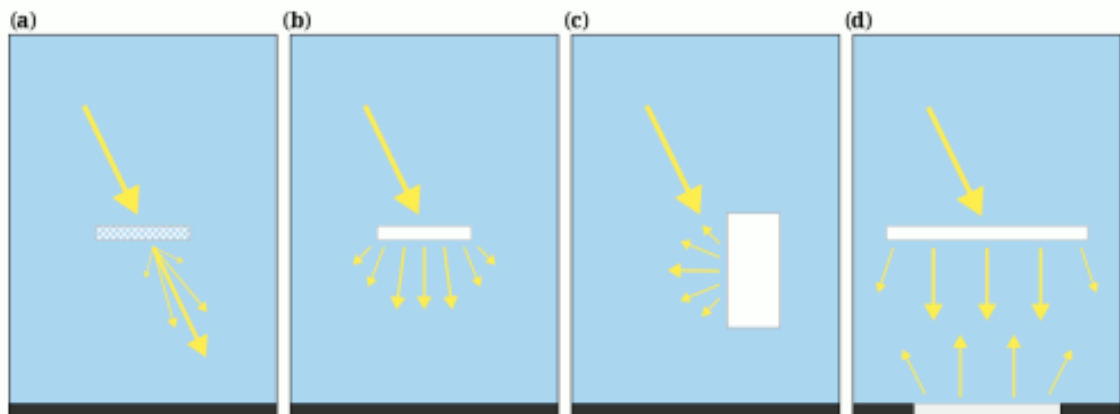
Fonte: Adaptado de Rossow e Schiffer (1999).

A profundidade óptica das nuvens (COD, *Cloud Optical Depth*) representa uma medida da atenuação da radiação solar ao longo do percurso óptico, sendo determinada por um conjunto de propriedades microfísicas, geométricas e ópticas. Entre os fatores que a influenciam, destacam-se a distribuição de tamanho das gotículas, a espessura vertical da nuvem e o índice de refração do meio. De modo análogo, a profundidade óptica dos aerossóis (AOD, *Aerosol Optical Depth*) quantifica o grau

de extinção da radiação causado por partículas suspensas na atmosfera. No caso das nuvens, a profundidade óptica, associada à sua geometria e ao conteúdo de água ou gelo, é um dos principais determinantes da forma como essas formações interagem com a radiação solar.

O estudo de Mol e Heerwaarden (2025) realizou experimentos numéricos com o objetivo de identificar os principais mecanismos responsáveis pela variabilidade da irradiância solar na superfície sob diferentes condições de nebulosidade, fundamentadas em observações de uma estação nos Países Baixos. Os resultados indicaram que quatro mecanismos principais explicam a maior parte das variações observadas: (a) em nuvens opticamente finas ($COD < 6$), predomina o espalhamento para frente, capaz de gerar aumentos superiores a 50% em relação à irradiância de céu limpo, como observado em campos de *altocumulus*; (b) em nuvens planas e opticamente espessas ($COD > 6$), o espalhamento difuso descendente gera máximos de irradiância sob a própria base da nuvem; (c) em nuvens com estrutura vertical, o espalhamento lateral domina até que sombras formadas por nuvens tipo bigorna limitem esse efeito; e (d) sob cobertura espessa, o albedo superficial intensifica os fluxos radiativos por meio de múltiplos espalhamentos, contribuindo com 10% a 60% da irradiância total adicional, dependendo da refletância da superfície. A Figura 1.1 ilustra esquematicamente os quatro mecanismos principais responsáveis pela variabilidade da irradiância solar em superfície sob diferentes tipos de cobertura de nuvens.

Figura 1.1 - Representação esquemática dos quatro mecanismos propostos que modulam a variabilidade da irradiância solar na superfície. Da subfigura (a) à (d), os mecanismos são: escape frontal (*forward escape*), escape para baixo (*downward escape*), escape lateral (*side escape*) e realce por albedo (*albedo enhancement*). A radiação solar incidente é representada pela seta grossa voltada para baixo; as demais setas indicam, de forma aproximada, a direção média e a intensidade relativa da radiação espalhada. O albedo da superfície é representado pela tonalidade mais clara. Apenas as setas relevantes para cada mecanismo estão representadas.



Fonte: Adaptado de Mol e Heerwaarden (2025).

As propriedades ópticas e morfológicas das nuvens exercem papel fundamental na modulação da irradiância solar incidente na superfície, ao influenciar tanto a componente direta quanto a difusa da radiação. Essa influência varia significativamente com o tipo de nuvem, sua espessura óptica, distribuição espacial e evolução temporal. A seguir, são descritas (conforme Mol e Heerwaarden (2025)) as principais características radiativas associadas a diferentes tipos de nuvens, com ênfase nos mecanismos que favorecem a variabilidade da irradiância superficial e a ocorrência de eventos de sobreirradiância:

- *Cumulus* rasos: caracterizam-se por rápidas e frequentes transições entre sombra e iluminação solar direta, resultando em alta variabilidade de irradiância em curtas escalas espaciais e temporais. As sombras projetadas por esses sistemas são intensas (com reduções de 50% a 80% em relação à CSI), enquanto as aberturas entre nuvens permitem incrementos abruptos da irradiância direta.
- *Alto cumululus*: apesar de também produzirem transições rápidas, o aumento da fração de cobertura tende a reduzir a frequência dos eventos de in-

cremento de irradiância. A irradiância difusa sob essas formações é, em geral, mais elevada do que sob outras configurações nebulosas e mantém relativa estabilidade durante os picos de aumento de irradiância. Como são nuvens opticamente finas, a componente direta da irradiância não é totalmente bloqueada, o que possibilita, em momentos de breves aberturas entre nuvens, a superposição de radiação direta e difusa intensificada - cenário propício à ocorrência de eventos extremos de sobreirradiância.

- *Cumulonimbus*: devido à sua espessura óptica elevada e desenvolvimento vertical acentuado, esses sistemas atuam como barreiras radiativas, bloqueando quase totalmente a irradiância direta e atenuando substancialmente a difusa. No entanto, projetam sombras extensas e podem aumentar a componente difusa nos flancos iluminados pelo sol, especialmente por espalhamento lateral em situações de geometria solar favorável.
- *Stratus e stratocumulus*: em geral, são nuvens espessas e extensas que produzem irradiância difusa homogênea e baixa, com pouca variabilidade ao longo do tempo, exceto em situações de dissipação, formação recente, ou advecção de suas bordas. Nessas circunstâncias, especialmente na presença de lacunas ou transições abruptas entre regiões nubladas e céu claro, podem ocorrer picos de irradiância de curta duração e intensidade variável. Entretanto, essas situações são menos frequentes, dado o tamanho típico e a persistência dessas formações.

Em conjunto, esses diferentes regimes de nebulosidade condicionam a distribuição espaço-temporal da irradiância em superfície, influenciando diretamente não apenas o balanço de energia local, mas também a dinâmica de processos atmosféricos em múltiplas escalas.

1.3.1 Ciclo de vida das nuvens na região amazônica

O ciclo típico de vida das nuvens em atmosferas tropicais é fortemente influenciado pela abundância de calor e umidade disponíveis na superfície, além dos processos convectivos intensos que caracterizam essas regiões. As chamadas nuvens quentes formam-se quando uma parcela de ar úmido sobe na atmosfera e se resfria adiabaticamente até atingir o nível de condensação por levantamento — ponto em que sua temperatura se iguala ao ponto de orvalho e o vapor d'água se condensa sobre núcleos higroscópicos, originando gotículas líquidas. Essas nuvens são compostas apenas por água líquida e tendem a se desenvolver em níveis mais baixos da troposfera, onde as

temperaturas são relativamente altas. Em contraste, nuvens que alcançam altitudes maiores ou que se formam em ambientes mais frios podem conter também cristais de gelo, sendo classificadas como nuvens mistas ou frias, a depender da proporção entre os dois estados da água. A evolução de qualquer tipo de nuvem depende de múltiplos fatores, incluindo a umidade ambiente, a intensidade das correntes ascendentes, o entranhamento de ar seco e as propriedades físico-químicas dos aerossóis atmosféricos, que podem atuar como núcleos de condensação ou de congelamento (HOBBS, 1993).

A interação entre aerossóis e nuvens constitui um componente crítico do sistema climático, dada sua capacidade de modificar propriedades microfísicas e macroscópicas das nuvens, com implicações diretas e indiretas no balanço radiativo da Terra (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2023). Mais aerossóis atuando como núcleos de condensação promovem o aumento do número de gotículas de nuvem e a consequente redução de seu raio efetivo (considerando um mesmo conteúdo de água líquida), o que intensifica o albedo da nuvem, conforme proposto por Twomey (1977). Essa alteração pode desencadear ajustes secundários na nuvem, como o aumento do tempo de vida da nuvem e a supressão ou o atraso da precipitação, aumentando o conteúdo integrado de água líquida da nuvem e/ou a fração de cobertura (ALBRECHT, 1989). No entanto, observações indicam que, em determinadas condições, sobretudo em nuvens não precipitantes, o acréscimo de aerossóis pode levar à redução de seu conteúdo de água líquida e da fração de cobertura (LEBSOCK et al., 2008), devido à intensificação da evaporação das pequenas gotículas e à maior mistura com o ar ambiente, promovendo a dissipação das nuvens. Além disso, aerossóis absorvedores podem aquecer camadas atmosféricas, aumentando a estabilidade vertical e reduzindo a nebulosidade e a refletividade das nuvens (ACKERMAN et al., 2000; KAUFMAN; NAKAJIMA, 1993). Assim, a resposta das nuvens à presença de aerossóis é complexa e depende de fatores como fase da nuvem, regime de precipitação e características do aerossol. Uma análise detalhada da influência dos aerossóis nas propriedades de nuvens não está dentre os objetivos deste estudo, podendo ser consultada em Rocha (2011).

A seguir, descreve-se em mais detalhe o ciclo de vida das nuvens tropicais com base em conceitos consolidados da literatura atmosférica (Holton (2004), por exemplo) e em evidências observacionais específicas para a Amazônia Central, utilizando dados da mesma campanha experimental do atual estudo (como Carneiro e Fisch (2020), Carneiro et al. (2023), Giangrande et al. (2023), Gonçalves et al. (2022), Henkes et al. (2021), Öktem et al. (2023), Machado et al. (2018), Tian et al. (2021)). Basicamente,

o ciclo de vida das nuvens convectivas compreende três etapas: (i) formação, (ii) maturação e (iii) dissipação.

(i) A formação de nuvens está intimamente ligada ao aquecimento radiativo da superfície e aos processos de mistura dentro da camada limite planetária (CLP). A radiação líquida incidente na superfície é convertida principalmente em fluxos de calor latente e sensível, sendo o calor latente o componente predominante do balanço de energia ao longo do ano na Amazônia Central, responsável por cerca de 70% do fluxo energético total (CARNEIRO et al., 2023). Durante a madrugada e as primeiras horas da manhã, a CLP apresenta condições predominantemente estáveis, caracterizadas por ar mais frio junto à superfície devido ao resfriamento radiativo noturno. Essa estratificação térmica é rompida à medida que o Sol se eleva e aquece o solo, iniciando o processo de erosão da camada limite noturna, no qual a CLP evolui rapidamente para uma camada bem misturada.

Em dias com nuvens *cumulus* rasas, observam-se maiores fluxos de calor sensível, que promovem crescimento mais intenso da CLP, favorecendo o entranhamento de ar seco da troposfera livre e, conseqüentemente, reduzindo a umidade relativa e elevando a base das nuvens (TIAN et al., 2021). Ainda assim, a presença de alta umidade residual desde a noite anterior pode suavizar o resfriamento noturno e sustentar maiores fluxos de calor sensível nas primeiras horas do dia, contribuindo para uma transição da CLP de um regime estável para um regime convectivo e facilitando a formação precoce de nuvens rasas (HENKES et al., 2021). A instabilidade condicional da atmosfera, diagnosticada pelo gradiente negativo de temperatura, cria um ambiente propício à elevação de parcelas de ar úmidas. A flutuabilidade positiva dessas parcelas permite atingir o nível de convecção livre, especialmente quando velocidades verticais iniciais e o conteúdo de umidade são suficientes para vencer o entranhamento (TIAN et al., 2021). Uma vez iniciada a condensação, a liberação de calor latente reforça os movimentos verticais, sustentando a formação de nuvens convectivas.

As condições atmosféricas durante a estação seca (como menor umidade relativa nos baixos níveis, nebulosidade reduzida durante a noite e nas primeiras horas do dia e maior aquecimento radiativo da superfície) favorecem menor cobertura de nuvens rasas pela manhã e maiores valores de energia potencial disponível para convecção (CAPE, *Convective Available Potential Energy*) nos baixos níveis, em comparação à estação chuvosa. Como resultado, as células convectivas tendem a apresentar correntes ascendentes mais intensas logo nos estágios iniciais de vida

(GIANGRANDE et al., 2023). Além disso, anomalias em variáveis de grande escala, como temperatura, umidade atmosférica e movimentos verticais médios, têm papel importante na formação e dissipação das nuvens, sendo influenciadas por condições do dia anterior. Essa "memória" da dinâmica atmosférica de larga escala pode ser útil para estimar a fração de cobertura de nuvens e o conteúdo de água associado (GONÇALVES et al., 2022).

(ii) A maturação ocorre conforme o aquecimento da superfície se intensifica entre o meio e o final da manhã (até umas 12:00 HL), a convecção na CLP se fortalece, favorecendo o crescimento vertical das nuvens e sua transição de *cumulus* rasos para *cumulus congestus* e, eventualmente, no período da tarde, em nuvens do tipo *cumulonimbus*. Esse estágio de maturação é marcado por forte desenvolvimento vertical, presença de correntes ascendentes organizadas e liberação contínua de calor latente ao longo da coluna atmosférica, que realimenta a convecção. A relação entre a radiação líquida e a taxa de evapotranspiração apresenta resposta linear, refletindo o papel da radiação como principal força motriz para os fluxos superficiais na Amazônia (CARNEIRO et al., 2023).

O crescimento das nuvens convectivas é influenciado por fatores dinâmicos como o cisalhamento vertical do vento, que atua ao inclinar as correntes ascendentes dentro da nuvem. Essa inclinação desloca a região de precipitação para fora da coluna principal, impedindo que o resfriamento evaporativo das gotas de chuva sufoque o movimento ascendente. Como resultado, reduz-se o entranhamento de ar seco e favorece-se o desenvolvimento de convecção mais profunda e persistente (TIAN et al., 2021). Além disso, *cumulus congestus* e nuvens profundas (resultantes da convecção profunda) apresentam diferentes distribuições de altura do topo, associadas a diferenças no cisalhamento vertical de ventos na média troposfera. Após o atingimento do nível de convecção livre, a umidade da troposfera livre e o entranhamento de ar ambiental tornam-se determinantes para a profundidade final da nuvem (TIAN et al., 2021).

Durante a estação seca, o desenvolvimento convectivo tende a ocorrer de forma mais precoce e intensa, com as células atingindo o pico de intensidade e tamanho em estágios mais iniciais do ciclo de vida, em comparação com as da estação chuvosa (GIANGRANDE et al., 2023). Nessas situações, observam-se correntes ascendentes intensas já nos primeiros momentos da formação da célula, especialmente em regiões com refletividade acima de 35 dBZ, muitas vezes superando o nível de congelamento (GIANGRANDE et al., 2023). Em contraste, as tempestades da estação chuvosa

costumam ser mais duradouras e associadas a uma maior incidência de correntes ascendentes moderadas (2–5 m/s) nos níveis superiores da troposfera (GIANGRANDE et al., 2023).

A fricção gerada pelos movimentos verticais convectivos (*cumulus friction*) pode alterar o campo de ventos na baixa troposfera ao redistribuir o momento atmosférico, especialmente nas camadas próximas da superfície. Essa redistribuição influencia a organização espacial das células convectivas ao modificar o cisalhamento do vento ou ao induzir novas circulações locais que reforçam ou inibem o crescimento convectivo. Essas nuvens afetam significativamente os fluxos de radiação de onda curta incidente na superfície, modulando o balanço de energia e o acoplamento entre superfície e atmosfera.

(iii) A dissipação das nuvens ocorre geralmente no final da tarde e início da noite, quando o fluxo de radiação solar diminui e o fornecimento de energia térmica à CLP é interrompido. A redução do fluxo de calor sensível e do aporte de ar úmido pelas correntes ascendentes enfraquece gradualmente a convecção, promovendo a estabilização da atmosfera. Nuvens convectivas profundas, como os *cumulonimbus*, perdem sustentação dinâmica e começam a se desorganizar, resultando em precipitação residual, formação de núcleos frios e subsidência associada à evaporação da chuva - especialmente pronunciada durante a estação seca, quando há maior entranhamento de ar seco, intensificação da evaporação e consequente reforço de correntes descendentes com maior instabilidade negativa (GIANGRANDE et al., 2023).

Essas correntes descendentes persistentes nas porções superiores das nuvens, localizadas ao redor dos núcleos de precipitação, ocorrem com mais frequência durante a estação seca, enquanto na estação chuvosa observa-se menor incidência de descensos intensos em altos níveis (GIANGRANDE et al., 2023). Essa assimetria sazonal na dissipação das células convectivas é reflexo das diferenças nas condições de umidade atmosférica, estabilidade térmica e taxas de entranhamento entre os períodos úmido e seco (TIAN et al., 2021).

Nas camadas mais altas, mesmo após o colapso das correntes principais de convecção, a presença de gelo e cristais pode manter estruturas de nuvens cirriformes por períodos prolongados. Em alguns casos, remanescentes da convecção diurna persistem na forma de nuvens altas e difusas ao longo da noite, sendo influenciadas por anomalias em variáveis de grande escala (como umidade relativa, estabilidade e vento) em relação às condições do dia anterior, as quais condicionam a manutenção ou dispersão das formações de nuvens (GONÇALVES et al., 2022).

Na superfície, a redução da cobertura de nuvens associada ao enfraquecimento convectivo favorece o resfriamento radiativo noturno. Esse processo conduz à formação de camadas estáveis na CLP e, em condições de vento fraco e alta umidade, pode resultar na ocorrência de nevoeiros rasos ao amanhecer (HENKES, 2021). Durante a estação seca, o ciclo diurno da evapotranspiração não acompanha de forma imediata as variações na temperatura do ar e no déficit de pressão de vapor, refletindo uma resposta com atraso e não linear da vegetação ao forçamento radiativo ao longo do dia. Esse descompasso entre os picos de radiação e de evapotranspiração evidencia a complexidade das interações entre a superfície e a atmosfera, a qual também se manifesta nas maiores incertezas nas estimativas de fluxo de calor latente observadas nesse período, mesmo com um fechamento energético mais robusto (até 82%) (CARNEIRO et al., 2023).

Por fim, o enfraquecimento das correntes ascendentes, o aumento da estabilidade e a reorganização da estrutura atmosférica ao longo da noite preparam o ambiente para o reinício do ciclo convectivo no dia seguinte.

Em regiões tropicais, a variabilidade temporal da cobertura de nuvens é alta, com predominância de processos convectivos de difícil previsão (LAING; EVANS, 2011). Essa dinâmica gera flutuações rápidas e intensas da irradiância solar à superfície, em escalas que variam de minutos a horas. A variação espacial e temporal da fração de cobertura de nuvens (CF), associada a propriedades como profundidade óptica, conteúdo de água líquida e distribuição do tamanho das gotículas, modula a quantidade de radiação transmitida à superfície. Adicionalmente, a elevada umidade atmosférica e a presença de aerossóis influenciam a transmissividade da coluna atmosférica, introduzindo variabilidade nos sinais de GHI e CF e aumentando a complexidade da modelagem preditiva desses parâmetros.

1.4 Clima da Amazônia Central

A região central da Amazônia, onde se localiza o sítio experimental deste estudo, apresenta clima tropical úmido do tipo Af, segundo a classificação de Köppen-Geiger (BECK et al., 2018), caracterizado por alta umidade relativa e chuvas abundantes ao longo do ano. A estação chuvosa, entre novembro e março, distingue-se por intensa nebulosidade e precipitação, enquanto a estação seca, de maio a setembro, apresenta menor cobertura de nuvens e redução significativa da umidade atmosférica. Os meses de abril e outubro funcionam como períodos de transição entre esses dois regimes (FISCH et al., 1998). Alterações recentes, atribuídas ao desmatamento e ao aquecimento global, têm resultado na tendência de prolongamento da estação

seca (BOTTINO et al., 2024), o que tem implicações diretas na disponibilidade de irradiância e no balanço energético da região.

A precipitação e a nebulosidade na Amazônia Central são controladas por mecanismos multiescalares que incluem a convecção diurna induzida pelo aquecimento solar, linhas de instabilidade oriundas do litoral atlântico e sistemas convectivos de meso e larga escala, frequentemente associados à penetração de frentes frias oriundas do sul do Brasil (MOLION, 1987). Essas variáveis também são moduladas na região por fenômenos de grande escala como o ciclo El Niño-Oscilação Sul (ENSO, *El Niño–Southern Oscillation*), a oscilação de Madden-Julian e ondas de Kelvin acopladas convectivamente (JONES; HOREL, 1990; MAYTA et al., 2018; MAYTA et al., 2021).

Episódios de El Niño, em particular, estão associados à redução significativa da precipitação, especialmente nas porções central e norte da bacia, como documentado nos eventos de 2015–2016 (MARENGO; ESPINOZA, 2015; JIMÉNEZ-MUÑOZ et al., 2016; MARENGO et al., 2017; PANISSET et al., 2018). A explicação está no aquecimento anômalo das águas do Pacífico tropical leste, que altera a circulação de Walker e favorece a subsidência (movimento descendente do ar) sobre a Amazônia (ROPELEWSKI; HALPERT, 1987). Isso inibe o desenvolvimento de convecção profunda ao reduzir a umidade atmosférica na troposfera média e alta, e ao aumentar a estabilidade atmosférica geral (TOLEDO et al., 2024). Embora a elevação das temperaturas superficiais possa aumentar a evapotranspiração e a instabilidade local nas camadas inferiores, a falta de reposição de umidade atmosférica por convecção e transporte horizontal impede que o ar ascenda com intensidade suficiente para formar nuvens de grande desenvolvimento vertical.

Além disso, a alteração na posição da Zona de Convergência Intertropical (ZCIT), frequentemente deslocada para o norte durante eventos de El Niño, reduz a penetração de umidade atmosférica sobre a Amazônia. Com isso, o ambiente se torna menos favorável à formação de nuvens convectivas profundas, o que tende a aumentar a frequência de céu limpo ou de nuvens rasas. Esse padrão pode estar associado a episódios de sobreirradiância, pois favorece a incidência direta da radiação solar ao longo do dia, com menor atenuação por cobertura de nuvens densas.

Regionalmente, grandes rios e florestas são capazes de modular a convecção rasa e o transporte de calor e massa quando há maior incidência de radiação solar e gradientes térmicos mais acentuados entre áreas alagadas e em terra firme, fenômeno esse conhecido como brisa fluvial e que ocorre especialmente durante a estação seca

ou anos com maior área alagada (SANTOS et al., 2019; DIAS et al., 2004).

A Amazônia Central, inserida em uma região tropical úmida, é caracterizada por forte instabilidade convectiva, elevados teores de umidade atmosférica e intensa formação de nuvens ao longo do dia. Essa dinâmica atmosférica favorece a ocorrência de um ciclo diurno bem definido de nebulosidade e precipitação, com manhãs tipicamente mais limpas, desenvolvimento de nuvens cumuliformes a partir do final da manhã e frequente ocorrência de chuvas convectivas durante a tarde. A formação de nuvens convectivas na região segue uma progressão vertical que inclui fases de nuvens *rasas* (com topos de até 3 km), *cumulus congestus* (entre 3 e 8 km) e nuvens convectivas profundas (com topos acima de 8 km) (WORLD METEOROLOGICAL ORGANIZATION (WMO), 2017). A duração e a intensidade de cada fase dependem das condições termodinâmicas locais, o que implica em significativa variabilidade espacial e temporal na atenuação da radiação solar (HOLTON, 2004).

Com base em séries temporais de fração de cobertura de nuvens obtidas por satélite, Sena et al. (2018) identificaram um declínio sistemático na nebulosidade nas últimas décadas, particularmente pronunciado às 08:00 Hora Local (HL), com tendência de até -6% por década na Amazônia central e oriental. Essa redução está associada ao aumento da temperatura da superfície, intensificação da estabilidade atmosférica e diminuição do transporte de umidade para a região. Esse fato foi observado também por dados de reanálise e observações *in situ* (ARIAS et al., 2010).

As características das nuvens e da precipitação também variam sazonalmente. Durante a estação seca, as nuvens convectivas tendem a ser menores, mais isoladas e com maior intensidade de precipitação por unidade de área, enquanto na estação chuvosa a cobertura de nuvens é mais extensa, porém com taxas de precipitação localmente menos intensas (MACHADO et al., 2018). O total acumulado, no entanto, é maior durante a estação chuvosa. Ainda segundo os autores, a heterogeneidade do uso e cobertura do solo influencia a microfísica das nuvens: nuvens sobre áreas florestais tendem a conter gotículas maiores do que sobre áreas desmatadas ou urbanas, com maiores taxas de precipitação observadas em altitudes mais elevadas, especialmente na estação seca. Adicionalmente, a presença da pluma de poluição de Manaus afeta as propriedades microfísicas das nuvens, resultando em maior número de gotículas de menor tamanho e crescimento mais lento (CECCHINI et al., 2016).

O ciclo diurno da irradiância global e da convecção também apresenta diferenças sazonais marcantes. No período chuvoso, observa-se maior variabilidade e menores médias de irradiância, refletindo a intensa formação de nuvens entre 11:30 e 13:30

HL, com dissipação da convecção profunda a partir das 16:30 HL (MACHADO et al., 2002). Esse comportamento é associado à advecção de umidade atmosférica pelos ventos de nordeste, comuns na estação chuvosa, que favorecem o disparo da convecção ao longo do dia (SATYAMURTY et al., 2012; MARTIN et al., 2016). Por sua vez, durante a estação seca, as condições de céu apresentam menor variabilidade, favorecendo maior previsibilidade da irradiância.

Adicionalmente, Macedo e Fisch (2018) analisaram a variabilidade temporal das frações direta e difusa da radiação solar, correlacionando-as com a cobertura de nuvens espessas e finas, precipitação, aerossóis (AOD) e focos de incêndio. Este estudo, que utilizou o mesmo conjunto de dados da presente Tese, revela que a composição óptica e microfísica das nuvens é sensível à presença de partículas em suspensão, afetando o balanço entre radiação direta e difusa. Esse equilíbrio é especialmente importante para a previsão da irradiância em superfície, uma vez que a fração difusa tende a aumentar sob cobertura de nuvens finas ou na presença de aerossóis, enquanto a radiação direta predomina em dias com céu limpo — condição típica de episódios de sobreirradiância.

1.5 Fundamentação da estrutura de modelagem

Para desenvolver modelos preditivos robustos com base em técnicas de aprendizado de máquina, três etapas principais devem ser seguidas: seleção cuidadosa dos dados para formar conjuntos consistentes de treinamento, teste e validação; identificação dos preditores de entrada mais adequados, juntamente com a determinação da topologia da rede; e, em seguida, estimativa dos parâmetros da rede. As metodologias adotadas para cada uma dessas etapas são apresentadas a seguir.

1.5.1 ApproxHull

A presença de amostras que definem o domínio de operação dos modelos no conjunto de treinamento é essencial para garantir um bom desempenho preditivo. Como o fecho convexo, ou envoltória convexa (CH, *Convex Hull*), representa a menor região convexa capaz de envolver completamente um conjunto de pontos em um plano ou espaço, é desejável que os pontos que compõem esse envoltório sejam incorporados ao conjunto de treinamento. Algoritmos de envoltório convexo permitem determinar o número mínimo de pontos que definem essa região. Em geral, esses algoritmos apresentam complexidade não polinomial em tempo e espaço, o que implica que o tempo de execução cresce rapidamente com o aumento do número de pontos. Visando melhorar essa questão, Khosravani et al. (2016) propuseram o algoritmo

denominado ApproxHull, que permite uma operação eficiente mesmo em espaços de alta dimensionalidade. Os benefícios do uso do ApproxHull já foram demonstrados em diversas aplicações (ver, por exemplo, (LAOUALI et al., 2022)).

O ApproxHull é um algoritmo projetado para construir uma aproximação do envoltório convexo de um conjunto de pontos em um espaço Euclidiano. Seu objetivo é gerar uma aproximação que seja eficiente em termos computacionais (tempo e espaço), mantendo a precisão necessária. Seja $P = p_1, p_2, \dots, p_n$ um conjunto de n pontos em \mathbb{R}^d . O envoltório convexo $\text{Conv}(P)$ de P é definido como o menor conjunto convexo que contém todos os pontos de P .

O algoritmo busca ser simultaneamente eficiente e preciso por meio de refinamentos iterativos do envoltório convexo. O processo do ApproxHull está detalhado no algoritmo (em pseudocódigo) no Apêndice A. Antes de realizar as operações descritas, o ApproxHull realiza um pré-processamento no conjunto de dados, removendo colunas duplicadas, linhas redundantes e combinações lineares detectadas.

O processamento com o ApproxHull é fundamental para selecionar as amostras mais representativas do conjunto de dados. Ao permitir que o modelo se concentre nas características mais relevantes, contribui para uma previsão mais eficiente e precisa. Sem essa etapa, a escolha aleatória de pontos poderia aumentar a variabilidade nas entradas, comprometendo a estabilidade do modelo e elevando as taxas de erro, além de reduzir sua capacidade preditiva. Além disso, o custo computacional seria mais elevado, já que o modelo teria de processar um número maior de amostras, muitas das quais sem contribuição efetiva para as previsões. Embora o modelo possa operar sem ApproxHull, espera-se que seu desempenho geral (especialmente no que diz respeito à generalização e à precisão) seja comprometido.

1.5.2 MOGA

Algoritmos Genéticos Multiobjetivos (MOGA, *Multi-Objective Genetic Algorithm*) são técnicas de otimização projetadas para lidar com problemas que envolvem múltiplos objetivos conflitantes. Esses algoritmos buscam encontrar um conjunto de soluções ótimas (ou não-dominadas), conhecido como fronteira de Pareto, em vez de uma única solução ótima. Em termos práticos, isso significa que tais soluções representam compromissos entre critérios conflitantes. Um exemplo clássico é a escolha entre custo e desempenho de um equipamento: reduzir o custo tende a diminuir o desempenho, enquanto aumentar o desempenho geralmente implica em maior custo.

Assim, uma solução é considerada não-dominada se não existir outra solução que seja melhor em todos os objetivos e, pelo menos, igual em um deles. Essas soluções não-dominadas são geralmente consideradas as mais "equilibradas" ou "ótimas" em relação aos diferentes objetivos. As operações do algoritmo MOGA são explicadas em detalhe em (FERREIRA; RUANO, 2011), mas serão aqui resumidas.

O MOGA projeta modelos de Redes Neurais Artificiais (ANNs, *Artificial Neural Network*) estáticos ou dinâmicos para fins de classificação, aproximação ou previsão. Essas redes consistem em camadas de unidades chamadas neurônios, que recebem entradas, aplicam uma transformação matemática (geralmente uma soma ponderada seguida de uma função de ativação) e propagam o resultado para as camadas seguintes, de modo a aprender relações complexas entre variáveis. Dentre os modelos de ANN estão redes do tipo RBF (*Radial Basis Function*), Perceptrons Multicamadas e Máquinas de Vetores de Suporte por Mínimos Quadrados (LS-SVM, *Least-squares support vector machine*), que compartilham a característica de terem seus parâmetros separáveis em termos lineares e não lineares. Diversos métodos construtivos estão disponíveis na literatura para problemas de regressão com redes neurais (ver, por exemplo, (KWOK; YEUNG, 1997), (SÁ et al., 2022)). Até onde se tem conhecimento, o MOGA, conforme introduzido por (FERREIRA; RUANO, 2011), é o único aplicável a modelos com parâmetros separáveis. A vantagem de trabalhar com tais modelos reside, por um lado, em sua aplicabilidade a uma ampla gama de problemas e, por outro, em sua elevada precisão quando comparada a alternativas.

Entre os modelos compatíveis com o MOGA, as redes RBF se destacam pela simplicidade estrutural e eficiência de aprendizado, sendo amplamente adotadas desde sua proposta original por (BROOMHEAD; LOWE, 1988). Desde então, devido ao aprendizado mais rápido associado a essas redes, seu uso tem-se ampliado em diversas aplicações. Para compreender seu funcionamento, considere a formulação matemática de uma rede neural artificial do tipo propagação direta (*feedforward*) com uma camada oculta de neurônios pode ser representada pela seguinte expressão:

$$\hat{y}(\mathbf{x}_k, \mathbf{w}) = u_0 + \sum_{i=1}^n u_i \varphi_i(\mathbf{x}_k, \mathbf{v}_i) \quad (1.2)$$

em que $\mathbf{w} = [u, v]^T$ é o vetor de parâmetros do modelo, $u = [u_0, u_1, \dots, u_n]$ é o vetor de parâmetros lineares escalares, $v = [v_1, \dots, v_n]$ é composto pelos vetores de parâmetros não lineares v_i , cada um associado a um neurônio, e \mathbf{x}_k representa os vetores de entrada.

Na ANN RBF, normalmente se utiliza a função Gaussiana para a $i^{\text{ésima}}$ função de base ou neurônio, $\varphi_i(\mathbf{x}_k, \mathbf{v}_i)$:

$$\varphi_i(\mathbf{x}_k, \mathbf{v}_i) = \exp\left(-\frac{1}{2\sigma_i^2}|\mathbf{x}_k - \mathbf{c}_i|^2\right) \quad (1.3)$$

Nesse caso, $v_i = [c_i, \sigma_i]$ é o vetor de parâmetros não lineares, onde c_i é um ponto no espaço de entrada (o centro da função Gaussiana), σ_i é o espalhamento correspondente e $|\cdot|$ é a norma euclidiana. As saídas de todos os neurônios são, então, combinadas linearmente (como na Equação 1.2) para produzir a saída da rede.

O treinamento consiste na determinação dos parâmetros w por meio da minimização da função erro:

$$\Omega(\mathbf{X}, \mathbf{w}) = |y - \hat{y}(\mathbf{X}, \mathbf{w})|^2 \quad (1.4)$$

em que \mathbf{X} representa o conjunto completo de dados de entrada utilizados no treinamento do modelo, podendo ser representado como uma matriz $n \times d$, onde cada linha corresponde a uma amostra de entrada x_k e cada coluna a uma das d características (features) de entrada.

Para modelos com parâmetros separáveis em lineares e não lineares, (1.4) pode ser reescrita como:

$$\Omega(\mathbf{X}, \mathbf{w}) = |y - \Gamma(\mathbf{X}, \mathbf{v})u|^2 \quad (1.5)$$

em que $\Gamma(\mathbf{X}, \mathbf{v}) = [\varphi(x_1, \mathbf{v}), \varphi(x_2, \mathbf{v}), \dots, \varphi(x_m, \mathbf{v})]^T$, sendo sua solução ótima global com respeito a u dada por (o símbolo $+$ indica a operação de pseudoinversa):

$$u^* = \Gamma^\dagger(\mathbf{X}, \mathbf{v})y \quad (1.6)$$

Substituindo (1.6) em (1.5), obtém-se um novo critério dependente apenas de \mathbf{v} (parâmetros não lineares):

$$\Psi(\mathbf{X}, \mathbf{w}) = \left|y - \Gamma(\mathbf{X}, \mathbf{v})\Gamma^\dagger(\mathbf{X}, \mathbf{v})y\right|^2 \quad (1.7)$$

A etapa de estimação de parâmetros no MOGA consiste em otimizar os parâmetros de um modelo específico (com um conjunto de entrada e número de neurônios fixado). Essa etapa explora inicialmente a propriedade de separação de parâmetros (equações 1.5 a 1.7) e, em seguida, utiliza o algoritmo de Levenberg-Marquardt (LEVENBERG, 1944; MARQUARDT, 1963) para minimizar a Equação 1.7.

A parte genética envolve a seleção do número de neurônios e das variáveis de entrada a serem utilizadas, a partir de um conjunto completo de características. Essa seleção é feita com base nos objetivos definidos pelo usuário no escopo do MOGA.

1.5.3 NAR(X)

MOGAs também podem ser utilizados para projetar modelos do tipo NAR e NARX. Modelos NAR (*Nonlinear Autoregressive*) são redes neurais recorrentes empregadas na previsão de séries temporais com base exclusivamente nos valores passados da variável-alvo, ou seja, na sua componente endógena. Nesse contexto, “endógeno” refere-se a algo que é explicado a partir de si próprio, sem recorrer a informações externas. Quando o modelo é estendido para incluir também variáveis exógenas — isto é, fatores externos que exercem influência sobre a variável de interesse — ele passa a ser denominado modelo NARX (*Nonlinear Autoregressive with Exogenous Inputs*). O termo autorregressivo, por sua vez, indica que o modelo utiliza a relação entre valores passados e futuros da série para construir a previsão. A estrutura NARX é particularmente adequada para sistemas dinâmicos, nos quais o comportamento futuro depende não apenas do histórico da variável principal, mas também de fatores externos que modulam sua evolução ao longo do tempo.

Em problemas de previsão de séries temporais, assume-se geralmente que as amostras são equidistantes no tempo. No modelo NARX, a saída futura é estimada com base nos valores passados da variável endógena y e da variável exógena u (assumindo uma única variável de entrada). Considerando dois atrasos para as variáveis endógena e exógena, a previsão de um passo à frente no instante $k + 1$, baseada nas informações disponíveis até o instante k , é dada por:

$$\hat{y}(k + 1)|k = f(y(k), y(k - 1), u(k), u(k - 1)) \quad (1.8)$$

Essa previsão pode ser estendida iterativamente para horizontes de previsão maiores. Por exemplo, a previsão de dois passos à frente é dada por:

$$\hat{y}(k+2)|k = f(\hat{y}(k+1)|k, y(k), \hat{u}(k+1)|k, u(k)) \quad (1.9)$$

Esse processo pode ser repetido até atingir um horizonte de previsão PH , mesmo quando todas as entradas do modelo forem valores previstos. Conhecida como previsão iterativa (ou recursiva), essa abordagem consiste na reaplicação de um único modelo (originalmente treinado para prever um passo à frente) utilizando suas próprias saídas como entradas para estimar múltiplos passos futuros. Alternativamente, podem ser adotadas estratégias como previsão direta (em que um modelo independente é treinado para cada instante $k+h$) ou previsão multi-saída (que estima simultaneamente todos os PH valores futuros), com o objetivo de reduzir a acumulação de erros e a propagação da incerteza, embora à custa de maior complexidade ou menor capacidade de generalização.

Para organizar essas previsões, constrói-se uma matriz de previsão PM_y , em que cada linha representa um instante de tempo e cada coluna contém os atrasos e previsões da variável y . De modo semelhante, pode-se definir uma matriz PM_u para a variável exógena u , que pode ser modelada por meio de uma abordagem NAR simples (assumindo apenas um atraso):

$$\hat{u}(k+1) = f(u(k)) \quad (1.10)$$

A previsão pode incluir também variáveis estáticas z , que não apresentam dinâmica temporal e devem ser fornecidas ao modelo no instante de previsão. Nesse caso, o modelo torna-se:

$$\hat{y}(k+1) = f(y(k), y(k-1), u(k), u(k-1), \hat{z}(k+1)) \quad (1.11)$$

Esse arcabouço permite a avaliação do desempenho dos modelos ao longo de uma janela deslizante, utilizando as previsões armazenadas nas matrizes PM_y e PM_u , e comparando-as com os valores reais medidos da série temporal.

1.6 Trabalhos relacionados de previsão de irradiância

O aprendizado de máquina (*machine learning*) é o estudo científico de algoritmos e modelos estatísticos que os sistemas de computadores usam para executar uma tarefa específica sem serem explicitamente programados (MAHESH, 2020). Eles são

treinados a partir de um conjunto de dados, com o objetivo de extrair padrões relevantes que permitam ao modelo generalizar para novos dados e realizar previsões ou classificações com maior acurácia.

A previsão de irradiância solar tem sido amplamente investigada nos últimos anos, refletindo o crescente interesse do tema. Revisões bibliográficas recentes, como as de Antonanzas et al. (2016), Barhmi et al. (2024), Chu et al. (2021), Martins et al. (2022a), Khair et al. (2025), Kumari e Toshniwal (2021), Lin et al. (2023), já oferecem um panorama abrangente sobre os métodos, modelos e desafios relacionados à previsão de irradiância e de geração solar, abrangendo desde modelos de aprendizado de máquina até técnicas baseadas em imagens do céu com fluxo óptico para diferentes passos e horizontes de previsão. Os estudos referenciados a seguir são destacados não apenas por sua relevância metodológica, mas também porque seus dados e condições de previsão são comparáveis aos utilizados no presente trabalho, servindo, assim, como base para comparações contextuais e de desempenho.

Dentre os métodos baseados em séries temporais e inteligência artificial, destaca-se o trabalho de Ferreira et al. (2012), com dados de Faro (Portugal). Os autores empregaram redes neurais artificiais com funções de base radial, otimizadas por algoritmos genéticos multiobjetivo, para realizar previsões multietapas de cobertura de nuvens, radiação solar e temperatura do ar com até quatro horas de antecedência. Hoyos-GÓMEZ et al. (2022) também avaliaram previsões com um dia de antecedência utilizando apenas dados históricos de irradiância, demonstrando que modelos LSTM superam abordagens mais simples, especialmente sob condições de cobertura de nuvens variáveis. Os autores utilizaram dados de 12 estações na Colômbia (1 na Amazônia, 4 próximas ao Pacífico e as demais nos Andes). De forma semelhante, Shirazi et al. (2024) compararam o desempenho de sete algoritmos de regressão estatística e aprendizado de máquina, incluindo *Ridge Regression* e descida de gradiente estocástica, para previsão de irradiância de curto prazo, destacando as vantagens de cada método em diferentes cenários, utilizando dados da Bélgica.

Modelos baseados em aprendizado profundo também têm sido amplamente explorados. Richardson et al. (2019) investigaram o uso de diversos algoritmos como *Random Forests*, *Gradient Boosted Trees* e redes neurais profundas para previsão de irradiância a partir de imagens capturadas por câmeras olho-de-peixe, aplicando inicialmente técnicas de fluxo óptico e, posteriormente, adotando o estimador *H2ODeepLearning*. Os dados utilizados foram provenientes de três localidades: *National Renewable Energy Laboratory* (NREL), Colorado; San Antonio, Texas; e Ilhas Canárias, Es-

panha. De modo semelhante, [Hendriks et al. \(2024\)](#) utilizaram redes LSTM para prever irradiância com até 20 minutos de antecedência, obtendo desempenho robusto sob diversas condições de cobertura de nuvens, com dados da Plataforma Solar de Almería, no sul da Espanha.

Uma parte significativa dos avanços recentes tem-se concentrado em métodos que integram informações visuais obtidas por câmeras *all sky*, capazes de capturar em tempo real a dinâmica da cobertura de nuvens. [Cheng e Yu \(2014\)](#) aplicaram modelos de regressão baseados em vetores de atributos extraídos de ASI obtidas em Taiwan, produzindo previsões tipicamente com 5 minutos de antecedência. [Nouri et al. \(2021\)](#) propuseram modelos híbridos que combinam persistência com previsões espaciais baseadas em ASI, alcançando desempenho superior ao da persistência pura em condições de cobertura de nuvens variáveis. Posteriormente, [Nouri et al. \(2023\)](#) desenvolveram uma estrutura de *nowcasting* probabilístico não paramétrico, na qual a incerteza da previsão é modelada explicitamente com base na variabilidade recente da irradiância normal direta (DNI, *Direct Normal Irradiance*). O método utiliza erros históricos de previsão, estratificados por tempo de antecipação e classes de variabilidade da irradiância, para construir tabelas empíricas de quantis. Estas são então utilizadas para gerar intervalos de previsão ao redor dos *nowcasts* determinísticos, possibilitando uma representação probabilística dos valores esperados de irradiância em horizontes de tempo curtos. Ambos os estudos utilizaram dados da Plataforma Solar de Almería.

Outras abordagens recentes têm avançado no uso de redes neurais convolucionais (CNN, *Convolutional Neural Network*) e técnicas de pré-processamento de imagens para extrair padrões relevantes à previsão. Por exemplo, [Zuo et al. \(2023\)](#) empregaram CNNs aplicadas a imagens redimensionadas e medições convertidas em índices de céu claro, utilizando o método de persistência como referência e dados do Colorado, EUA. [Logothetis et al. \(2022\)](#) classificaram as condições de nuvens em seis categorias e compararam o desempenho de múltiplos modelos baseados em ASI (como LSTM, MLP e CNN) sob esses diferentes cenários, utilizando dados da Plataforma Solar de Almería. Complementarmente, [Straub et al. \(2024\)](#) propuseram um método que transforma imagens do céu não distorcidas em mapas de índice de céu limpo, utilizando regressão por *gradient boosting* para associar cada pixel da imagem a valores de GHI, com base em variáveis locais e altura da base das nuvens obtidas em múltiplos locais em Freiburg, Alemanha.

A fusão de múltiplas fontes de dados também tem sido explorada como estratégia

para melhorar a acurácia das previsões. [López-Cuesta et al. \(2023\)](#) combinaram quatro modelos baseados em ASI com dois modelos baseados em satélite e um modelo utilizando dados históricos de superfície, realizando combinação de modelos para prever GHI e DNI com resolução temporal de um minuto e até 90 minutos à frente, em Sevilha (sul da Espanha). [Samu et al. \(2022\)](#) propuseram um sistema que gera mapas de irradiância com resolução espacial de 8 km x 8 km e resolução temporal de um minuto, combinando previsões de GHI e DNI com classificações simplificadas de variabilidade atmosférica. Os autores utilizaram dados da Austrália (Perth), Alemanha (Oldenburg e Julich), Portugal (Évora) e Espanha (La Africana e PSA). [Marzouq et al. \(2020\)](#) propuseram uma estrutura de previsão baseada em Rede Neural Artificial Híbrida com Atenção (HAEANN, *Hybrid Attention-based Artificial Neural Network*), na qual algoritmos genéticos são empregados para otimizar automaticamente tanto o histórico de entrada quanto a arquitetura da rede, gerando múltiplos modelos para horizontes de previsão de até 6 horas à frente. Utilizando dados de irradiância solar de 28 cidades do Marrocos e avaliando três cenários distintos (cidade individual, agrupamento regional e conjunto completo), os modelos propostos demonstraram bom desempenho e alta capacidade de generalização.

O estudo de [Sousa et al. \(2020\)](#) avaliou a acurácia de previsões pontuais e intervalares de séries temporais univariadas de irradiância solar utilizando ANN. O modelo foi testado com dados coletados a cada 10 minutos em Belém, na região leste da Amazônia, entre dezembro de 2015 e novembro de 2016, medidos por um piranômetro CM11 da Kipp & Zonen. As redes foram configuradas para realizar a previsão de apenas um passo à frente - ou seja, 10 minutos. Dos modelos estudados de previsões pontuais, o que apresentou melhor desempenho foi a rede de Elman, obtendo o erro quadrático médio normalizado (nRMSE) igual a 8,2% e o erro absoluto médio normalizado (nMAE) igual a 5,9%.

[Vats \(2024\)](#) avaliaram a previsão da irradiância solar em quatro cidades do estado do Amazonas (Manaus, São Gabriel da Cachoeira, Tabatinga e Humaitá), utilizando dados satelitais da NASA (*POWER Project*) no período de janeiro de 2013 a novembro de 2022. Os modelos foram implementados com diferentes agregações temporais: diária, semanal e mensal. Considerando a maior frequência (diária) e a modelagem para Manaus, a melhor performance foi com o algoritmo *Adapt Boost*, com erro absoluto médio (MAE) igual a 0,327.

Apesar dos avanços observados na literatura quanto à previsão de irradiância solar, somente esses últimos dois estudos foram encontrados contemplando dados proveni-

entes da região amazônica, e nenhum referente à caracterização de eventos extremos. Ainda assim, [Sousa et al. \(2020\)](#) não usa dados da Amazônia Central e faz somente uma previsão para 10 minutos, enquanto que [Vats \(2024\)](#) não faz previsão de curtíssimo prazo, sendo o passo de previsão de 1 dia. Essa lacuna é especialmente relevante diante das peculiaridades atmosféricas e da alta variabilidade associada ao regime de nuvens da região. Assim, o presente trabalho busca contribuir para o preenchimento dessa lacuna, ao propor e avaliar abordagens específicas baseadas em dados observacionais obtidos na Amazônia Central, ampliando o entendimento e a aplicabilidade de métodos de previsão em contextos tropicais úmidos.

1.7 Sobreirradiância: definição e importância

O aumento da implantação de sistemas fotovoltaicos em larga escala em regiões tropicais tem evidenciado os efeitos de condições ambientais extremas (como alta temperatura, umidade atmosférica e irradiância solar intensa) sobre o desempenho e a confiabilidade desses sistemas. Dentre os fenômenos atmosféricos que mais impactam a operação de geradores solares está a sobreirradiância, definida como picos de GHI medidos na superfície que excedem os valores esperados para condições de céu limpo, com base em referências astronômicas ou modelos de radiação ([PFISTER et al., 2003](#)). Esses surtos de curta duração estão geralmente associados a campos de nuvens isoladas (*broken clouds*) ou céu parcialmente nublado e constituem um desafio particular para a previsão da irradiância solar.

Do ponto de vista operacional, episódios de sobreirradiância podem impactar os sistemas fotovoltaicos de diferentes formas ([CHEN et al., 2010](#); [HORNER et al., 2023](#)):

- Sobrecarga de inversores e perdas por *clipping*: quando a potência gerada excede a capacidade do inversor, provocando descarte de energia ou sobrecorrentes que acionam mecanismos de proteção, resultando em perdas imediatas na conversão de corrente contínua (DC, *Direct Current*) para corrente alternada (AC, *Alternating Current*) — DC/AC;
- Sobreaquecimento e degradação de componentes elétricos: fusíveis, cabos, conexões e soldas podem sofrer desgaste térmico acelerado, levando a falhas prematuras e redução da confiabilidade do sistema;
- Alterações elétricas e de desempenho dos módulos: mudanças nas curvas I-V (relação entre corrente I e tensão V de um módulo fotovoltaico, utilizadas para caracterizar seu comportamento elétrico em diferentes condições

de irradiância e temperatura) e a variabilidade espectral podem afetar de forma diferenciada distintas tecnologias, gerando diferenças de rendimento entre módulos conectados em série e exigindo ajustes de projeto;

- Aumento de custos de operação e manutenção: decorrente da necessidade de revisões mais frequentes para mitigar falhas e prolongar a vida útil dos componentes.

As primeiras pesquisas dedicaram-se à avaliação das características de reflexão da irradiância solar ao redor das paredes e topos de *cumulus* usando simulações computacionais (DAVIES, 1978; MCKEE; COX, 1974), medições de aeronaves (REYNOLDS et al., 1978) e testes de laboratório (DAVIS et al., 1983). No entanto, o estudo de Segal e Davis (1992) foi um dos primeiros a avaliar a magnitude e as características temporais da então chamada "*cumulus solar-irradiance reflection*" (CSIR) com GHI medida no nível do solo e suas possíveis implicações.

A ocorrência de sobreirradiância está relacionada a diferentes mecanismos ópticos que atuam em escalas espaciais e temporais reduzidas. Um deles é o chamado *cloud enhancement*, no qual a radiação solar incidente é intensificada por múltiplas reflexões entre nuvens próximas e o solo, mesmo sem bloqueio direto do disco solar (GUEYMARD, 2017). Nuvens rasas do tipo *cumulus*, por exemplo, geram bordas altamente brilhantes devido ao espalhamento da luz solar, amplificando temporariamente a irradiância sob determinadas condições. Nuvens do tipo *altocumulus* também podem produzir efeito semelhante, embora de menor intensidade devido à maior altitude (PFISTER et al., 2003). Além disso, cristais de gelo presentes em nuvens cirriformes podem contribuir para a intensificação da radiação por meio de efeitos de lente atmosférica (THUILLIER et al., 2013).

Outro mecanismo associado é o *cloud lensing*, que ocorre quando a borda de uma nuvem atua como uma lente, concentrando os raios solares em uma região estreita por alguns segundos. Esse efeito pode resultar em incrementos de irradiância de até 30% em relação aos valores de céu limpo (BALFOUR, 2011), sendo frequentemente observado em nuvens do tipo *cumulus*, cuja geometria e propriedades ópticas favorecem a focalização da radiação (ZEHNER et al., 2010). A distribuição tridimensional do campo de radiação, em conjunto com a estrutura e a morfologia das nuvens, torna-se, portanto, um fator determinante para a ocorrência e intensidade desses eventos.

A referência utilizada para caracterizar eventos de sobreirradiância varia entre os

estudos. Alguns autores definem o fenômeno com base na irradiância estimada por modelos de céu claro (CSI, *Clear Sky Irradiance*) (PFISTER et al., 2003), valores calculados para o TOA como referência (SOUZA et al., 2024) ou ambos (NATSIS et al., 2024). Embora determinados eventos ocorram com intensidades inferiores a 1000 W/m² - valor padrão adotado para certificação de módulos fotovoltaicos (INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC), 2021) - e possam não representar riscos operacionais, episódios mais intensos e duradouros podem causar falhas em sistemas elétricos. Em geral, eventos de sobreirradiância têm duração inferior a 30 segundos, mas análises com diferentes janelas temporais revelam tanto ocorrências breves quanto persistentes. Eventos de sobreirradiância na escala de minutos podem criar condições suficientes para a queima de fusíveis do tipo temporizado (*slow-blow*), mesmo projetados para suportar sobrecorrentes por curtos intervalos de tempo (RÜTHER et al., 2017). Episódios com duração superior a cinco minutos, sobretudo em condições de elevada temperatura de operação, podem levar à falha desses dispositivos mesmo quando corretamente dimensionados conforme as especificações do fabricante (NASCIMENTO et al., 2019). Além disso, ainda que episódios de curta duração não sejam necessariamente críticos de forma isolada, sua repetição ao longo do tempo pode induzir efeitos cumulativos, com implicações na confiabilidade e durabilidade dos sistemas fotovoltaicos.

Alguns estudos relataram eventos de sobreirradiância em diferentes lugares do mundo. Emck e Richter (2008) investigaram eventos de sobreirradiância nas montanhas andinas do sul do Equador (4°S), entre 1500 e 3400 m de altitude, em 1998–2001, registrando valores extraordinários de irradiância global na horizontal de até 1832 W/m², sendo atribuídos exclusivamente a efeitos radiativos de nuvens. Tapakis e Charalambides (2014) analisaram dados de GHI medidos em Chipre, durante a primavera e o verão de 2010, observando valores superiores a 1500 W/m² em sete ocasiões, com mais de 20 dias apresentando irradiância acima de 1200 W/m², indicando que condições parciais de nebulosidade podem sustentar períodos prolongados de sobreirradiância. Yordanov et al. (2015) relataram o evento mais extremo de sobreirradiância medido até então na Noruega meridional, com pico de 1600 W/m² e potencial para atingir até 1700 W/m², causado por forte espalhamento frontal da luz solar por nuvens *altocumulus* finas e fragmentadas, sugerindo que valores ainda mais altos podem ocorrer em latitudes tropicais.

O estudo de Natsis et al. (2024) analisou as características dos eventos de sobreirradiância em Tessalônica (Grécia), com base em uma série temporal de GHI com 30 anos de duração e resolução de 1 Hz. Os eventos foram identificados a partir da

comparação com uma referência de irradiância sob céu limpo gerada por modelo de transferência radiativa, apoiado por dados locais de aerossóis. A análise revelou um aumento consistente, ao longo do tempo, tanto na ocorrência quanto na intensidade dos episódios, tendência que ainda não havia sido documentada na literatura. Esses eventos foram mais frequentes no final da primavera e início do verão do hemisfério norte. Eventos com duração superior a 10 minutos foram muito raros (menor que 8%). Além disso, em alguns casos, a irradiância registrada superou os valores esperados no TOA para a mesma posição solar, com picos de até 204 W/m^2 (em 75% dos casos, os excessos foram inferiores a 40 W/m^2).

No Brasil, [Piacentini et al. \(2011\)](#) investigaram a ocorrência de eventos de sobreirradiância solar em Recife, PE, entre fevereiro de 2008 e janeiro de 2009, cujo maior valor extremo registrado foi de 1477 W/m^2 . Os autores propuseram que a ocorrência de sobreirradiância estaria associada a condições específicas de cobertura de nuvens, incluindo: (a) a presença de campo de nuvens *cumulus* alternado com regiões de céu claro, permitindo simultaneamente a incidência direta da radiação solar e sua reflexão nas bordas das nuvens; (b) uma cobertura de nuvens entre aproximadamente 50% e 90%, favorecendo a existência de lacunas por onde a radiação direta pode atravessar; e (c) nuvens densas o suficiente para intensificar a componente difusa da radiação por meio de múltiplos espalhamentos. [Almeida et al. \(2014\)](#) documentaram o evento mais extremo de sobreirradiância já registrado no Brasil até então em altitudes inferiores a 1000 m, com pico de 1590 W/m^2 durante 32 segundos em São Paulo, SP, destacando que valores acima de 1000 W/m^2 podem representar mais de 25% da irradiância diária, mesmo sob condições nubladas.

Uma lista com os maiores eventos de sobreirradiância global horizontal já registrados foi compilada por [NASCIMENTO et al. \(2019\)](#), que observaram que sete dos dez maiores ocorreram no Brasil, todos em regiões de baixa altitude. Esse fato torna os registros ainda mais notáveis, considerando que altitudes mais baixas estão associadas a uma coluna atmosférica mais espessa, o que, em geral, promove maior absorção e espalhamento da radiação solar, reduzindo a probabilidade de eventos intensos de sobreirradiância. Essas regiões coincidem com áreas onde estão sendo implantadas usinas fotovoltaicas de grande porte, evidenciando a vulnerabilidade do território nacional. O estudo também destacou que dados com resolução de 1 minuto são adequados para modelagem e dimensionamento de sistemas fotovoltaicos, embora medições em alta frequência (subminuto) possam revelar eventos adicionais que, de outro modo, passariam despercebidos. Outro ponto importante a se destacar desse estudo é que os dados analisados não incluíram a região amazônica, permanecendo,

portanto, lacunas importantes sobre a ocorrência dos eventos de sobreirradiância nessa área.

Nesse contexto, Braga et al. (2020), utilizando dados de irradiância com resolução de 1 Hz, observaram que a sobreirradiância é muito mais comum em baixas altitudes do que se supunha anteriormente. Além disso, verificaram que uma fração significativa dos eventos possui duração superior a um minuto e que mais da metade dos episódios extremos ocorre em sucessão rápida, sugerindo que esse fenômeno tem sido historicamente subestimado tanto em frequência quanto em severidade. Dados de cinco locais de teste - Florianópolis e Brotas de Macaúbas, no Brasil; Bernburg, na Alemanha; Albuquerque, nos Estados Unidos; e Loughborough, no Reino Unido - foram apresentados e analisados.

O trabalho de Fernandes (2023) teve como foco a previsão de eventos de sobreirradiância na cidade de Natal (Rio Grande do Norte, Brasil). Foram avaliados quatro algoritmos de aprendizado de máquina: *Random Forest* (RF), *Support Vector Machines* (SVM), Perceptron Multicamadas (MLP, *Multi Layer Perceptron*) e redes neurais do tipo LSTM. O desempenho dos modelos foi avaliado por meio de métricas padrão de classificação (acurácia, precisão, revocação e medida-F1), com acurácia variando entre 70% e 80%. O modelo *Random Forest* apresentou o melhor desempenho geral. O objetivo do modelo não foi prever o instante exato de início da sobreirradiância, mas sim indicar, com até cinco minutos de antecedência, a probabilidade estatística de ocorrência de um evento.

Além de questões ligadas à meteorologia e à geração de energia solar, o fenômeno de sobreirradiância também apresenta implicações em outras áreas do conhecimento. Entre elas, destacam-se o monitoramento do índice ultravioleta (FEISTER et al., 2015; TIBA; LEAL, 2017), o conforto térmico humano e a biometeorologia. Nesta última, a sobreirradiância pode estar associada ao aumento da fração difusa da radiação atmosférica (RODERICK et al., 2001), o que favorece a penetração da luz em ambientes com dossel vegetal denso, beneficiando a vegetação de sub-bosque. Vale lembrar que muitas espécies apresentam maior eficiência fotossintética em condições de irradiância moderada, tipicamente na faixa de 100–200 W/m² (MONTEITH, 1976). No entanto, essa possível interação deve ser interpretada com cautela, uma vez que sua relevância ecofisiológica depende de fatores como duração do evento, composição espectral e arquitetura do dossel. Apesar dessa diversidade de aplicações potenciais, o fenômeno permanece pouco explorado na literatura científica, em especial no contexto de regiões tropicais úmidas como a Amazônia.

1.8 Motivação

Como evidenciado na revisão apresentada, variabilidade sub-horária da irradiância solar em regiões tropicais úmidas é fortemente influenciada por processos convectivos, dinâmica de nuvens de pequeno porte e alta umidade atmosférica, tornando sua previsão um desafio para modelos tradicionais. Além disso, os mecanismos que favorecem episódios de sobreirradiância ainda são pouco compreendidos na região amazônica, apesar da importância desse comportamento para a geração fotovoltaica e para o balanço de energia na superfície, especialmente em escalas de minutos.

Nesse contexto, esta tese tem como objetivo responder às seguintes questões científicas:

- É possível desenvolver modelos de previsão em curtíssimo prazo (até 60 minutos) de irradiância solar global em superfície (GHI), com alta resolução temporal, baseando-se apenas em variáveis locais como GHI e fração de cobertura de nuvens estimada por câmeras *all sky* (CF), com resultados comparáveis ou superiores aos da literatura?
- Até que ponto modelos de previsão de irradiância solar são capazes de antecipar com precisão eventos de sobreirradiância em uma região tropical úmida, considerando os padrões temporais característicos (intra-horários, diários e sazonais) e as limitações físicas e operacionais dos diferentes horizontes de previsão?

1.9 Objetivo geral e específicos

O objetivo geral é o de aprofundar o entendimento sobre os padrões temporais e a previsibilidade de eventos de sobreirradiância solar em ambiente tropical úmido, por meio da caracterização estatística e do desenvolvimento de modelos de previsão com uso de aprendizado de máquina em curtíssimo prazo baseados em irradiância global e fração de cobertura de nuvens.

Com este objetivo principal, pretende-se especificamente:

- Analisar o comportamento de GHI, de CF, da variabilidade (por meio de rampas de GHI e do índice DARR) e dos tipos de nuvens, com o objetivo de avaliar suas inter-relações, sua contribuição para a ocorrência de eventos de sobreirradiância e seu impacto no desempenho do modelo preditivo desenvolvido;

- Caracterizar o comportamento temporal de eventos de sobreirradiância solar, incluindo o desenvolvimento de um novo índice temporalmente agregado (DAOI);
- Desenvolver e avaliar modelos de previsão de irradiância em curtíssimo prazo com base em medições de GHI e na CF, assim como analisar a capacidade desses modelos em prever eventos de sobreirradiância, como uma consequência emergente de sua habilidade preditiva.

Para atingir tais objetivos, serão utilizados dados observacionais com alta resolução temporal (1 minuto), obtidos a partir de medições de irradiância global em superfície (GHI) e estimativas de cobertura de nuvens (CF) derivadas de imagens de uma câmera *all sky*.

1.10 Estrutura da tese

A tese está estruturada em quatro capítulos principais, além da introdução (que aborda a motivação e uma revisão de literatura). O segundo capítulo descreve os dados utilizados e os procedimentos de preparação e transformação necessários para viabilizar as etapas de modelagem, incluindo o cálculo de variáveis auxiliares e a construção dos vetores de entrada, assim como o desenvolvimento dos modelos preditivos, detalhando as estratégias de otimização multiobjetivo e a estruturação das redes neurais aplicadas. No terceiro capítulo, são analisadas as variações temporais das variáveis estudadas e suas inter-relações, discutidos os resultados da caracterização temporal dos eventos de sobreirradiância, a proposta do índice DAOI e a avaliação da capacidade dos modelos em antecipar tais eventos, considerando diferentes horizontes e condições atmosféricas. Por fim, o último capítulo reúne as conclusões da pesquisa e aponta perspectivas para estudos futuros.

2 METODOLOGIA

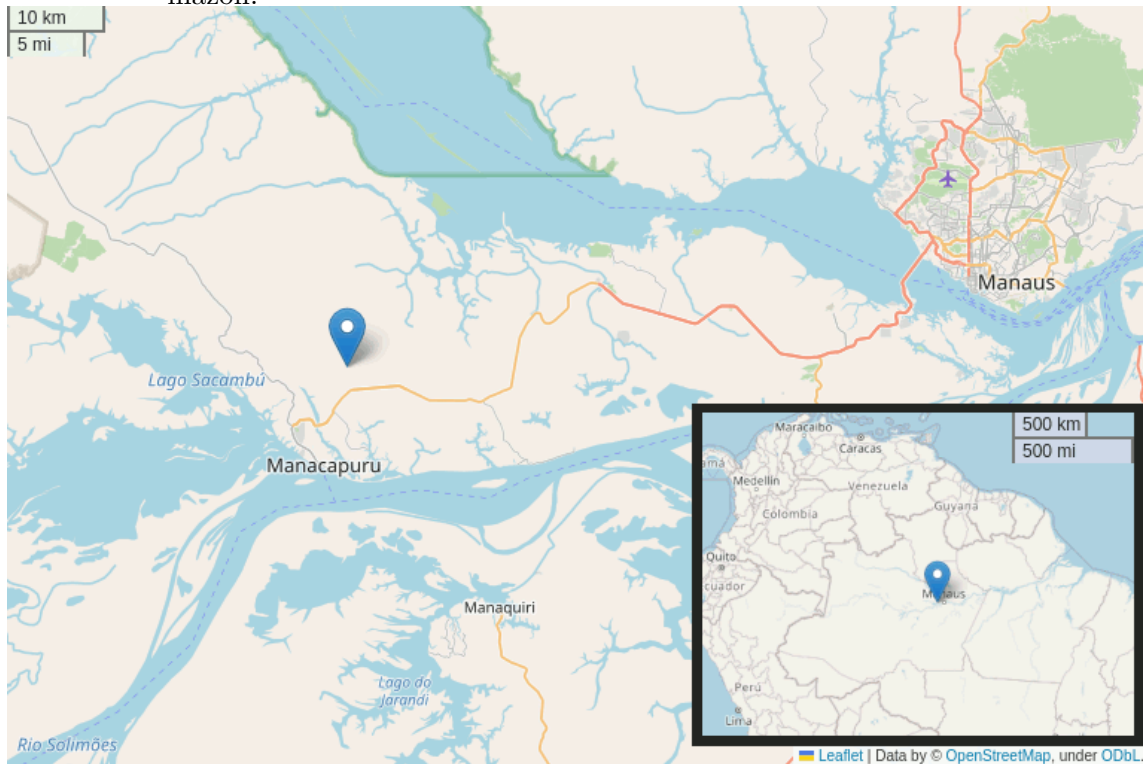
Os dados utilizados neste estudo foram coletados na região central da Bacia Amazônica, conforme descrito na subseção Dados. Séries temporais de GHI e CF foram processadas usando o algoritmo ApproxHull para definir os conjuntos de treinamento, teste e validação, conforme detalhado na subseção Processamento de Dados. Um Algoritmo Genético Multiobjetivo (MOGA, *Multi-Objective Genetic Algorithm*) foi aplicado para desenvolver e selecionar os modelos. Posteriormente, um conjunto foi formado a partir dos melhores modelos não dominados, resultando em previsões mais robustas e estáveis. Os valores previstos foram comparados aos observados para calcular métricas de desempenho tanto para a tarefa de previsão quanto para a detecção de eventos de sobreirradiância.

2.1 Dados

Os dados utilizados neste estudo foram obtidos durante a campanha experimental *Green Ocean Amazon* (GoAmazon 2014/5), realizada nos anos de 2014 e 2015 na região central da Amazônia (MARTIN et al., 2016). A iniciativa científica teve como principal objetivo compreender os ciclos de vida dos aerossóis e das nuvens na Bacia Amazônica, com ênfase especial em suas interações com a precipitação. Além disso, a campanha investigou como as emissões de poluentes oriundos de uma megacidade tropical (no caso, Manaus, AM) influenciam esses processos atmosféricos, possibilitando a análise de cenários que variam desde condições atmosféricas praticamente limpas até situações de elevada carga poluente. Essa abordagem permitiu a coleta de informações fundamentais para o aprimoramento de modelos climáticos e a previsão dos impactos das atividades humanas sobre o clima amazônico, tanto em escala regional quanto global (MARTIN et al., 2016).

A estação de coleta de dados utilizada neste trabalho, identificada como T3, esteve operacional exclusivamente durante a campanha e foi instalada em uma área de pastagem cercada por floresta nativa com dossel de aproximadamente 35 metros de altura. Sua localização geográfica situa-se ao norte da zona urbana do município de Manacapuru (AM), ao sul da margem esquerda do rio Solimões, nas coordenadas 3,21°S e 60,6°W (fuso horário UTC-4), conforme ilustrado na Figura 2.1.

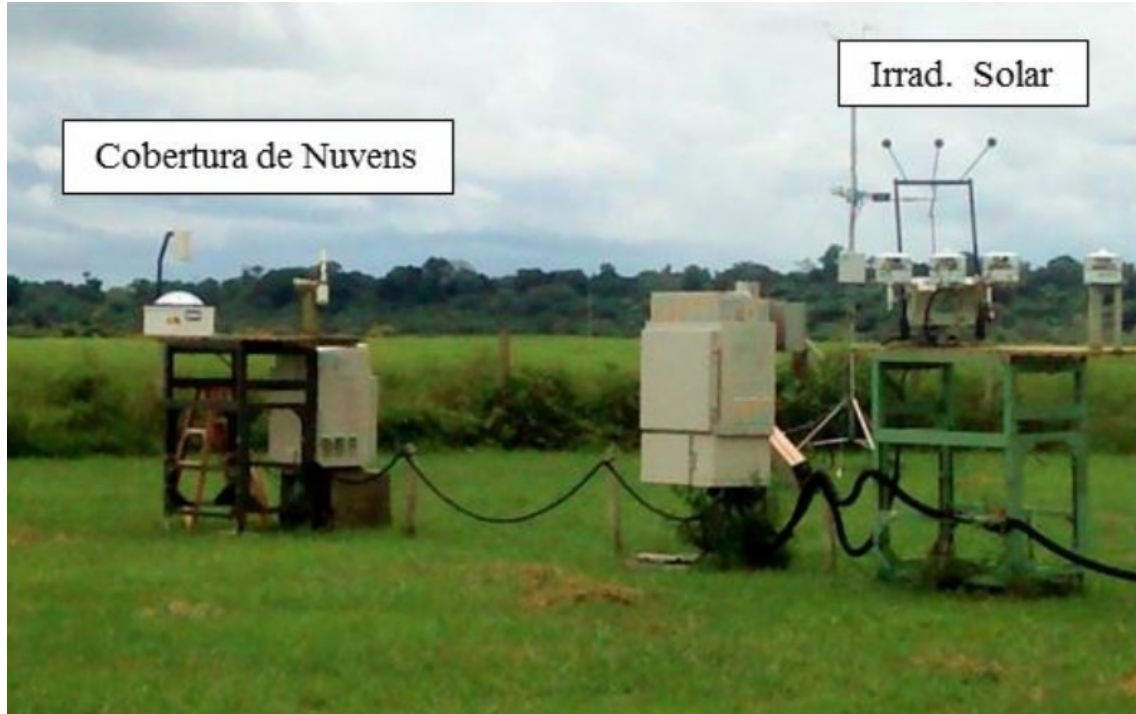
Figura 2.1 - Localização da estação de pesquisa T3 instalada durante experimento GoAmazon.



Fonte: Rocha et al. (2025).

Os instrumentos instalados na estação T3 foram projetados para medir variáveis relacionadas ao ciclo hidrológico, aos fluxos de energia na camada limite atmosférica e a outros parâmetros meteorológicos e radiativos. As observações coletadas estão disponíveis publicamente no repositório da *Atmospheric Radiation Measurement (ARM) Climate Research Facility* (FLYNN, 2014; SENGUPTA et al., 2014; KYROUAC et al., 2014).

Figura 2.2 - Parte do sítio experimental T3 e instrumentos (da esquerda para a direita): Total Sky Imager (TSI), fotômetro solar e plataforma Sky Radiation, composta por piranômetros sombreados para medição da irradiância difusa e não sombreados para medição da irradiância global (GHI).



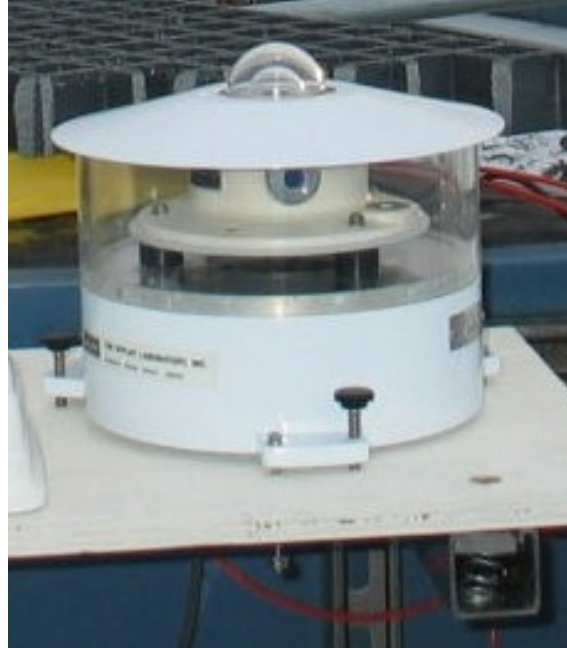
Fonte: Macedo (2016).

2.1.1 Irradiância Global Horizontal

Os dados de irradiância global horizontal (GHI, Global Horizontal Irradiance) usados neste estudo foram coletados por meio da plataforma SKYRAD (*Sky Radiation*), que integra vários sensores radiométricos (ANDREAS et al., 2018). Estes incluem um piranômetro ventilado e sem sombra para medir a irradiância hemisférica total (global), piranômetros sombreados para irradiância difusa e um pireliômetro para irradiância normal direta. O primeiro, que mede o GHI, foi um PSP (*Precision Spectral Pyranometer*) com ventilador, fabricado pelo *The Eppley Laboratory, Inc.* As medições são feitas a cada 2 segundos e, em seguida, a média é calculada em intervalos de 1 minuto. Os sinais brutos, inicialmente registrados em milivolts (mV), são convertidos para W/m^2 usando um fator de calibração específico para cada sensor. Os instrumentos operam em uma faixa espectral de 0,3 a 3,0 micrômetros, com uma incerteza estimada de $\pm 4\%$ até $-(4\% + 20 W/m^2)$ para ângulo zenital menor que 80° (ANDREAS et al., 2018). Isso significa que, em situações ideais, a

incerteza está em torno de $\pm 4\%$. Na prática, o erro não é simétrico e há tendência à subestimativa, ou seja, os valores reportados podem estar até 4% abaixo do valor real, com mais 20 W/m² de viés negativo adicional.

Figura 2.3 - Piranômetro Eppley PSP.



Fonte: Adaptado de NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA) (2024).

2.1.2 Fração de cobertura de nuvens

A fração de cobertura de nuvens (CF, *Cloud Fraction*) utilizada neste trabalho foi obtida por meio do instrumento TSI (*Total Sky Imager*), fabricado pela *Yankee Environmental Systems Inc* (Figura 2.4). (LONG et al., 2001) O TSI é uma câmera do tipo *all sky* que registra imagens hemisféricas do céu utilizando uma câmera CCD (*Charge-Coupled Device*) de estado sólido orientada para baixo sobre um espelho hemisférico aquecido. Esse espelho reflete a abóbada celeste para a lente do sistema, permitindo a captura da imagem completa do céu. Para evitar a saturação da imagem causada pela radiação solar direta, o equipamento conta com uma faixa de sombra (*shadow band*), cuja posição é continuamente ajustada com base em efemérides solares para bloquear o disco solar durante todo o dia. Conforme observado em trabalhos anteriores, imagens de câmeras *all sky* são representativas da atmosfera local em um raio aproximado de 1 km e em escalas temporais de até

30 minutos (ANTONANZAS et al., 2016), sendo amplamente utilizadas em estudos sobre variabilidade da nebulosidade e previsão de irradiância solar (CHENG; YU, 2014; LOGOTHETIS et al., 2022; LÓPEZ-CUESTA et al., 2023; NOURI et al., 2021).

Figura 2.4 - Total Sky Imager (TSI).



Fonte: Adaptado de ARM (2014).

O processamento das imagens é realizado por um software do próprio fabricante que executa, inicialmente, uma máscara para eliminar obstruções fixas na imagem, como o corpo da câmera, seu suporte e a própria faixa de sombra. Além disso, é recortada uma região circular central com o intuito de selecionar somente a região de interesse, removendo a região que não pertence ao reflexo do céu. A estimativa de CF baseia-se na classificação dos pixels restantes da imagem processada. Para isso, um algoritmo analisa as relações de cor de cada pixel a fim de determinar se representa céu limpo ou nebulosidade. Além dessa abordagem cromática, o software avalia o gradiente de brilho ao longo da faixa de sombra para inferir a presença de nuvens finas ou espessas encobrindo o sol. A combinação dessas duas informações resulta na fração percentual de cobertura total, expressa como um valor inteiro entre 0 e 100 (MORRIS, 2000). Essa metodologia automatizada permite substituir abordagens tradicionais de observação visual por parte de observadores meteorológicos,

conferindo maior frequência temporal e objetividade às medições.

As imagens utilizadas neste estudo foram coletadas com frequência de 30 segundos, com resolução de 640×480 pixels, mas foram utilizadas no trabalho atual somente as observações ocorridas nos instantes dos registros de GHI, que apresentam frequência de 1 minuto. O período de coleta ocorreu entre março de 2014 e novembro de 2015. No entanto, a maioria das imagens de março de 2014 foi afetada por um problema de alinhamento da faixa de sombra e, portanto, foi excluída. O problema foi resolvido apenas nos últimos dias do mês e, portanto, apenas os dados de 1^o de abril de 2014 a 30 de novembro de 2015 foram considerados para a modelagem.

2.1.3 Profundidade Óptica do Aerossol e precipitação

A coleta de dados de profundidade óptica de aerossóis (AOD, *Aerosol Optical Depth*) foi realizada por meio do fotômetro solar CIMEL (CSPHOT, *Cimel Sun Photometer*, Figura 2.5), um radiômetro automático multicanal amplamente utilizado em estudos atmosféricos. Os dados estão disponíveis junto da base de dados do ARM (GIBLER et al., 2014) mas o instrumento também integra a rede global AERONET (*Aerosol Robotic Network*), coordenada pela NASA (HOLBEN et al., 1998), e é projetado para realizar medições da irradiância solar direta e da radiância do céu em ângulos específicos de espalhamento. As observações são feitas em comprimentos de onda discretos, distribuídos entre as faixas do visível e do infravermelho próximo, utilizando filtros espectrais controlados por microprocessador para garantir precisão e estabilidade nas medições. O intervalo de amostragem é de aproximadamente 20 segundos, permitindo alta resolução temporal ao longo do dia.

Figura 2.5 - Fotômetro solar (modelo CIMEL).



Fonte: Adaptado de Franco et al. (2025).

A profundidade óptica dos aerossóis é determinada a partir das medidas espectrais da radiação solar direta, aplicando-se a lei de Beer-Lambert para quantificar a atenuação da luz solar provocada pelas partículas em suspensão ao longo da coluna atmosférica (HOLBEN et al., 1998). Para a presente análise, foram utilizados os dados de AOD no comprimento de onda de 500 nm, por se tratar de uma faixa dentro da região do visível onde a atmosfera livre de aerossóis apresenta alta transparência. Esse comprimento de onda é amplamente adotado em estudos climatológicos e de aerossóis por ser sensível à presença de partículas e representar uma referência para análises de atenuação radiativa.

A medição da AOD permite quantificar a carga de aerossóis presente na atmosfera e monitorar suas variações temporais, sendo essencial para a compreensão de processos como transporte de partículas, qualidade do ar e impactos no balanço radiativo.

Como parte da metodologia de obtenção dos valores de AOD, foram considerados apenas os dados coletados em condições atmosféricas livres de interferência por nuvens, assegurando que a atenuação observada fosse atribuída exclusivamente aos aerossóis. Uma visão mais abrangente da variabilidade da AOD e de CF na Amazônia Central e em toda a bacia amazônica pode ser consultada em [Martins et al. \(2018\)](#). Além disso, dados de precipitação foram obtidos a partir de pluviômetros automáticos instalados na mesma estação (modelo *Novalynx 260-2500E-12 Tipping Bucket Rain Gauge*, Figura 2.6), complementando a análise atmosférica ao fornecer informações sobre o ciclo hidrológico local.

Figura 2.6 - Pluviômetro do tipo Tipping Bucket (modelo 260-2500), utilizado em experimentos de campo para medição da precipitação.



Fonte: NOVALYNX CORPORTION (2024).

2.1.4 Classificação de nuvens

A classificação de nuvens tem sido, historicamente, realizada por meio de observação visual feita por profissionais treinados, baseando-se na aparência morfológica das formações nebulosas. No entanto, essa abordagem está sujeita a limitações inerentes à subjetividade humana, como a variabilidade na experiência do observador e a ambiguidade na identificação de propriedades macrofísicas das nuvens. Com o

avanço das técnicas de sensoriamento remoto, tornou-se possível automatizar a classificação de nuvens com base em características físicas objetivamente mensuráveis, como através de imagens de satélite (SETZER; SANO, 1998) e outros instrumentos, como explicado a seguir.

Neste estudo, a classificação de tipos de nuvem foi realizada no produto ARSCL (*Active Remote Sensing of Clouds*), desenvolvido pela *Atmospheric Radiation Measurement (ARM) Climate Research Facility* (JOHNSON; JENSEN, 1996). Esse produto integra dados de diferentes instrumentos de sensoriamento remoto ativos. O primeiro é o radar de nuvem de banda W (WACR, *W-band Cloud Radar*), que fornece perfis de refletividade capazes de detectar a estrutura vertical das nuvens com alta resolução. O segundo é o radar perfilador de vento (RWP, *Radar Wind Profiler*), que mede perfis verticais da velocidade do vento e também contribui com informações de refletividade úteis para a detecção de nuvens. O terceiro instrumento é o ceilômetro (CEIL, *Vaisala Laser Ceilometer*), que emite pulsos de luz no infravermelho próximo e registra o sinal de retroespalhamento, possibilitando a detecção de até três camadas de base de nuvens com alcance vertical de até 7.700 metros e resolução temporal de 16 segundos. Além disso, um Lidar (MPL, *Micropulse Lidar*) complementa a identificação das bases de nuvem com maior sensibilidade para partículas menores.

A classificação de nuvens com o produto ARSCL é realizada a partir da combinação dos dados desses sensores, considerando a refletividade dos radares e as alturas da base das nuvens fornecidas pelos instrumentos ópticos. A metodologia segue os critérios estabelecidos por McFarlane et al. (2013) e Burleyson et al. (2015), classificando as nuvens com base na altura da base, altura do topo e espessura vertical. Sete classes de nuvem são utilizadas, apresentadas na Tabela 2.1 juntamente com os critérios de classificação adotados.

Tabela 2.1 - Critérios de classificação de tipos de nuvens com base em altura da base, altura do topo e espessura.

Tipo de nuvem	Altura da base (km)	Altura do topo (km)	Espessura (km)
Nuvens rasas	< 3	< 3	sem restrição
<i>Congestus</i>	< 3	3–8	$\geq 1,5$
Convecção profunda	< 3	> 8	≥ 5
<i>Alto cumululus</i>	3–8	3–8	< 1,5
<i>Altostratus</i>	3–8	3–8	$\geq 1,5$
<i>Cirrostratus/anvil</i>	3–8	> 8	$\geq 1,5$
<i>Cirrus</i>	> 8	> 8	sem restrição

Fonte: Adaptado de McFarlane et al. (2013), Burleyson et al. (2015).

A classificação fornecida pelo ARSCL é disponibilizada em resolução temporal de 30 segundos, com informações verticais distribuídas em 600 níveis de altura.

2.2 Pré-processamento

O processo de preparação dos dados para modelagem envolveu a análise conjunta de duas variáveis principais: a GHI, considerada a variável endógena, e a CF, tratada como variável exógena. Variável endógena é aquela que se deseja modelar, prever ou explicar em um sistema - seu comportamento depende das outras variáveis do modelo. Variável exógena é aquela que influencia a variável endógena, que ajuda a explicar ou prever o comportamento da variável endógena.

Apenas nove dias não consecutivos ao longo do período de estudo apresentaram falhas pontuais nas medições diurnas de GHI com frequência de 1 minuto, conforme detalhado na Tabela 2.2, o que indica uma perda de dados mínima. Apenas períodos com dados simultâneos de GHI e CF foram utilizados para o treinamento e avaliação dos modelos. Lacunas com menos de 10 registros consecutivos ausentes foram preenchidas por interpolação: aplicou-se o método de Akima (AKIMA, 1970) para a GHI e interpolação PCHIP (FRITSCH; BUTLAND, 1984) para a CF, ambos considerados adequados para séries temporais por preservarem a forma e evitarem oscilações artificiais.

Tabela 2.2 - Número de registros diurnos sem valores de GHI por dia.

Data	Número de falhas
2014-01-31	52
2014-11-01	65
2014-11-11	233
2015-02-05	56
2015-09-04	60
2015-10-16	9
2015-10-17	69
2015-10-18	6
2015-11-01	17

Fonte: Produção do autor.

As séries temporais de GHI e CF foram utilizadas com resolução de 1 minuto para as análises descritivas e para a caracterização dos eventos de sobreirradiância, com o intuito de preservar a maior riqueza temporal possível dos dados observados. No entanto, para as etapas de modelagem e previsão, as séries de GHI e CF foram posteriormente reamostradas para intervalos de 2 minutos, por meio do cálculo da média aritmética de cada par de valores consecutivos. A adoção de um passo temporal de 2 minutos, em vez de 1 minuto, foi inicialmente motivada por contribuir para a suavização de ruídos de alta frequência e variações abruptas que poderiam comprometer o desempenho dos modelos. Com isso, definiu-se um horizonte de previsão de 1 hora (30 passos), o qual se mostrou suficiente para representar a dinâmica atmosférica com menor ruído e maior viabilidade computacional. Além de ampliar o intervalo de previsão inicialmente desejado de 30 minutos - permitindo uma análise mais completa da degradação do desempenho ao longo do tempo - essa configuração favorece a comparação com estudos que adotam horizontes semelhantes ou superiores.

2.2.1 Atributos para modelagem

A modelagem de séries temporais requer que a estrutura dos dados de entrada represente adequadamente a dinâmica temporal do fenômeno analisado. Nesse contexto, a criação de vetores de atraso (*lags*) é fundamental, pois permite que o modelo aprenda tanto dependências de curto prazo quanto padrões recorrentes (BROWNLEE, 2020; MORETTIN; TOLOI, 2006). Cada vetor de entrada é composto por observações passadas da variável de interesse, organizadas em diferentes janelas temporais, que são utilizadas como conjunto de atributos (*features*) para prever o valor futuro de-

sejado. Cada vetor de entrada inclui:

- Os 30 valores mais recentes anteriores ao instante atual t , isto é, do instante $t - 30\Delta t$ até $t - \Delta t$;
- O valor observado exatamente 24 horas antes, ou seja, em $t - 24h$;
- Uma janela de 10 valores anteriores e 10 valores posteriores ao instante $t - 24h$, isto é, de $t - 24h - 10\Delta t$ até $t - 24h - \Delta t$, e de $t - 24h + \Delta t$ até $t - 24h + 10\Delta t$.

Essa estrutura de atrasos permite que o modelo capture tanto as dependências temporais de curto prazo quanto os padrões diurnos intrínsecos aos dados, sendo Δt o intervalo de amostragem dos dados (no caso deste estudo, 2 minutos). Os vetores de atraso de GHI e CF compuseram o conjunto de atributos a ser utilizado nos modelos, juntamente com os respectivos valores-alvo (variáveis que se desejam prever no modelo), resultando em um total de 51 atributos por variável.

2.2.2 Métrica para representação de sobreirradiância

Além do preparo dos dados realizado para o treinamento dos modelos, procedimentos específicos foram aplicados com o objetivo de apoiar a análise da variabilidade da irradiância em escalas de tempo muito curtas. Tais procedimentos buscaram preservar a alta resolução temporal e minimizar os efeitos de suavização que poderiam dificultar a interpretação de flutuações rápidas. Enquanto os modelos de previsão utilizaram dados reamostrados em intervalos de 2 minutos, as análises exploratórias desta seção se basearam em médias de 1, 2 e 5 minutos, a fim de avaliar como a resolução temporal afeta a representação de eventos extremos e transitórios. Para essa análise, também foram consideradas duas referências: TOA e CSI. Os valores de CSI foram calculados conforme modelo proposto por [Ineichen e Perez \(2002\)](#), que estima a irradiância a partir da extraterrestre no topo da atmosfera, corrigida pela atenuação conforme o ângulo zenital solar e um parâmetro de turbidez atmosférica. Esse parâmetro é derivado de climatologias médias de vapor d'água, aerossóis, albedo de superfície e outros constituintes, funcionando como uma medida efetiva da espessura óptica da atmosfera e integrando os principais processos de absorção e espalhamento da radiação. Posteriormente, e em alguns casos, a análise também foi restrita aos casos em que os valores de GHI ultrapassaram 1000 W/m^2 .

Totais de eventos de sobreirradiância em escala de minuto, diária e mensal fornecem uma perspectiva de frequência absoluta com interpretação direta, sendo importante

para avaliar a distribuição temporal e o risco operacional. Além dessa grandeza, os estudos revisados para este trabalho tratavam outras dimensões relevantes, como a intensidade dos eventos (em W/m^2) e a sua duração (em minutos), mas de forma isolada, o que dificulta a caracterização integrada em períodos mais longos, como dias ou meses. Para avaliar de forma mais precisa o impacto energético médio da sobreirradiância, foi desenvolvido o Índice Dinâmico Agregado de Sobreirradiância (DAOI, *Dynamic Aggregate Overirradiance Index*), que representa o excesso médio absoluto da irradiância medida em relação a uma referência, considerando apenas os períodos em que ocorre sobreirradiância. O DAOI contorna a limitação da simples contagem de eventos, que pode mascarar a relevância de poucos episódios intensos entre muito fracos, ao ponderar cada evento de acordo com sua magnitude. Isso permite identificar minutos, dias ou meses com maior risco operacional ou potencial energético, mesmo quando ocorrem poucos eventos. Matematicamente, o índice é calculado como a soma das diferenças $GHI_t - GHI_{Ref,t}$ para todos os instantes t pertencentes ao conjunto $S = t \mid GHI_t > GHI_{Ref,t}$ (ou seja, os instantes em que houve sobreirradiância), dividida pelo número de eventos $N_s = |S|$ (conforme Equação 2.1). A irradiância de referência Ref pode ser estimada a partir de um modelo de céu claro (CSI) ou calculada para o topo da atmosfera (TOA).

$$DAOI_{Ref} = \frac{1}{N_s} \sum_{t \in S} (GHI_t - GHI_{Ref,t}) \quad (2.1)$$

2.3 Estrutura de modelagem

Conforme apresentado na Seção 1.5, o processo teve início com a aplicação do algoritmo ApproxHull para selecionar as amostras mais representativas, formando os conjuntos de treinamento, teste e validação. Em seguida, o MOGA foi utilizado para identificar os preditores de entrada, definir a topologia da rede e estimar seus parâmetros. Como modelo preditivo, adotou-se uma rede NAR(X), adequada à abordagem proposta.

Devido ao grande tamanho do conjunto de dados, o algoritmo ApproxHull foi primeiramente aplicado separadamente aos dados de cada mês para identificar os vértices dos respectivos cascos convexos. Em seguida, o ApproxHull foi aplicado novamente ao conjunto combinado de todos os pontos CH mensais. Uma vez determinados os pontos CH, o ApproxHull garante que eles sejam explicitamente incluídos no conjunto de treinamento durante a geração dos conjuntos de treinamento, teste e validação a serem usados no MOGA.

Como já mencionado na Seção 1.5.2, a saída do MOGA não é um único modelo, mas sim um conjunto de modelos não-dominados. Um modelo específico pode ser posteriormente selecionado pelo usuário; alternativamente, como adotado neste trabalho, é possível selecionar um número pré-definido de modelos e empregar um *ensemble*. A seleção é realizada conforme o procedimento descrito em (RUANO; RUANO, 2024). A saída do *ensemble* corresponde à mediana das saídas dos modelos selecionados.

Foram desenvolvidos dois grupos de modelos para previsão de irradiância: (i) modelos autorregressivos (NAR), que utilizam apenas valores passados de GHI - referidos como ghi1 (1 passo) e ghi30 (30 passos); e (ii) modelos com variáveis exógenas, que combinam valores passados de GHI e de CF - referidos como ghicf1 e ghicf30. Esses modelos foram utilizados para prever a irradiância com horizontes que variam de 1 passo à frente (2 minutos) até 30 passos à frente (60 minutos). Os pontos selecionados foram utilizados para compor os conjuntos de treinamento, teste e validação (60%, 20% e 20%, respectivamente), com o número de amostras em cada conjunto resumido na Tabela 2.3 - como são inicializações aleatórias para cada conjunto, os números são diferentes.

Tabela 2.3 - Tamanho dos conjuntos de dados utilizados para treinamento, teste e validação dos modelos em quatro diferentes configurações.

Conjunto	ghi1	ghi30	ghicf1	ghicf30
Treinamento	10.566	10.246	11.309	11.094
Teste	3.522	3.415	3.769	3.698
Validação	3.522	3.416	3.771	3.699
Total	17.610	17.077	18.849	19.301

Fonte: Produção do autor.

Dessa forma, cada configuração de entrada (ghi1, ghi30, ghicf1 ou ghicf30) foi construído um ensemble específico, composto pela mediana das previsões de um subconjunto de modelos não-dominados gerados pelo MOGA. Nesse tipo de abordagem, o ensemble não representa variações de cenários físicos (como um modelo atmosférico em que múltiplas simulações são realizadas a partir de diferentes condições iniciais ou parametrizações para capturar a incerteza inerente à evolução atmosférica e fornecer uma faixa de possíveis cenários), mas sim uma agregação estatística de modelos com diferentes características de ajuste e generalização. A escolha pela mediana como saída final visa mitigar o viés individual de cada modelo, aumentar

a robustez da previsão e suavizar variações indesejadas, resultando em estimativas mais estáveis e confiáveis da irradiância.

Dentro desse contexto, o uso de intervalos de previsão (PIs, *Prediction Intervals*) adicionaria uma dimensão probabilística às estimativas determinísticas fornecidas pelo ensemble. Enquanto a mediana assegura estabilidade frente a outliers e vieses individuais dos modelos, os PIs permitiriam quantificar a incerteza residual, fornecendo limites superiores e inferiores plausíveis para cada horizonte de previsão. A largura média dos intervalos serviria como indicador do grau de previsibilidade alcançado pelos modelos selecionados. Já na previsão de sobreirradiância, a combinação entre limites físicos de referência e os limites superiores dos PIs possibilitaria estabelecer limiares adaptativos, reduzindo alarmes falsos e tornando a detecção mais robusta a ruídos instrumentais e às variações de menor relevância. Assim, a associação entre a previsão central (mediana) e os PIs resultaria em um produto mais completo, capaz de ampliar tanto a confiabilidade operacional quanto a interpretação de eventos extremos, o que deverá ser explorado em trabalhos futuros.

As configurações estruturais das redes neurais do tipo RBF, especificamente o número de neurônios na camada oculta e o número de termos de entrada (defasagens), foram tratadas como hiperparâmetros e otimizadas por meio do MOGA. Nesse contexto, os hiperparâmetros correspondem a escolhas feitas antes do treinamento do modelo, responsáveis por definir sua arquitetura e influenciar diretamente seu desempenho. O número de neurônios foi limitado a um intervalo entre 2 e 25, enquanto o número de termos de entrada foi permitido variar de 1 a 30. Esses intervalos foram escolhidos empiricamente, visando equilibrar a capacidade do modelo e a complexidade computacional, ao mesmo tempo em que se evita o sobreajuste (*overfitting*). Modelos com poucos neurônios tendem a ser computacionalmente eficientes, mas podem falhar em capturar padrões complexos dos dados; por outro lado, modelos com muitos neurônios podem ajustar-se excessivamente ao conjunto de treinamento, prejudicando a capacidade de generalização. Uma lógica semelhante aplica-se à escolha do número de defasagens.

A avaliação dos modelos dentro do arcabouço MOGA foi baseada em quatro objetivos: o erro quadrático médio (RMSE) nos conjuntos de treinamento e de teste, a complexidade do modelo (com o intuito de penalizar modelos excessivamente complexos e evitar sobreajuste) e o erro de previsão (baseado nas séries temporais mencionadas anteriormente). Para os algoritmos genéticos, os principais parâmetros incluem: gerações, que correspondem às iterações do processo evolutivo; indivíduos, que re-

presentam as soluções candidatas avaliadas em cada geração; pressão de seleção, que define a intensidade da preferência por soluções de melhor desempenho; taxa de emigrantes aleatórios, a fração de indivíduos substituídos por novos aleatórios a cada geração; e taxa de crossover, que indica a probabilidade de recombinação entre pares de soluções. Foram adotados parâmetros usuais da literatura: 100 gerações, uma população de 100 indivíduos, pressão de seleção igual a 2, taxa de emigrantes aleatórios de 0,10 e taxa de *crossover* de 0,70. Esses parâmetros visam manter a diversidade da população, prevenir convergência prematura e assegurar uma exploração eficaz do espaço de busca.

Cada modelo foi treinado utilizando uma implementação eficiente do algoritmo de Levenberg-Marquardt, com parada antecipada (*early stopping*) para evitar sobreajuste e um número máximo de 50 iterações. Para garantir a robustez dos resultados, cada configuração de modelo foi avaliada em cinco execuções independentes, sendo selecionada a solução com melhor compromisso entre os objetivos, definida como a menor distância até a origem no espaço multiobjetivo. Essa abordagem assegura a confiabilidade dos resultados, mantendo o controle sobre os custos computacionais.

Foi realizada uma segunda execução do MOGA para cada configuração, desta vez incorporando restrições derivadas da análise da primeira execução. Essas restrições correspondem a limites superiores impostos sobre três objetivos: complexidade do modelo, RMSE no conjunto de treinamento e RMSE no conjunto de teste. Os limites normalizados das restrições utilizados em cada caso (cr = com restrição) são apresentados na Tabela 2.4, juntamente com os resultados para os casos sem restrição (sr), também em valores normalizados. A tabela inclui ainda o número de soluções não dominadas (ND, aquelas pertencentes à fronteira de Pareto, ou seja, para as quais não existe outra solução melhor em todos os objetivos simultaneamente) e as soluções preferenciais, que constituem um subconjunto dos modelos ND que satisfazem adicionalmente os três limites de restrição.

Tabela 2.4 - Resumo dos resultados obtidos com os modelos MOGA para diferentes conjuntos de dados. A tabela apresenta os erros quadráticos médios (RMSE) para os conjuntos de treinamento (tr.) e teste (te.), a complexidade média dos modelos (Complex), o número total de soluções não dominadas (ND) e, quando aplicável, o número de modelos preferenciais selecionados (Pref.).

Modelo	RMSE (tr.)	RMSE (te.)	Complex	ND	Pref.
ghi1sr	0,20342	0,14154	546	446	–
ghi1cr	0,21000	0,15000	550	409	254
ghi30sr	0,16708	0,12299	609	468	–
ghi30cr	0,16000	0,12000	600	505	28
ghicf1sr	0,20304	0,18057	519	199	–
ghicf1cr	0,20000	0,18000	500	100	89
ghicf30sr	0,19581	0,16922	573	464	–
ghicf30cr	0,20000	0,17000	550	348	153

Fonte: Produção do autor.

No presente estudo, o MOGA é empregado para definir a estrutura ótima dos modelos NAR(X), selecionando automaticamente os atrasos mais relevantes das variáveis endógena e exógena, bem como a topologia do modelo. Embora a variável exógena possa ser prevista separadamente em outros contextos, aqui ela é utilizada com base nos dados observados. Essa abordagem permite que a otimização realizada pelo MOGA concentre-se exclusivamente na identificação da melhor configuração do modelo NAR(X), sem introduzir incertezas adicionais associadas à previsão da entrada exógena. Como resultado, obtém-se um modelo mais preciso e estável, aplicado às características reais do conjunto de dados analisado.

Na etapa final do estudo, um conjunto de *scripts* foi utilizado para construir redes em cascata de modelos preditivos, explorando diferentes combinações entre modelos NAR e NARX. A estrutura da rede seguiu restrições de projeto, como a proibição de uso de modelos NARX na primeira camada (por dependerem de variáveis exógenas) e de modelos NAR na camada final (por não utilizarem entradas externas). Os modelos foram selecionados a partir das soluções não dominadas ou preferenciais previamente geradas, considerando um nível de significância de 10% para a escolha das entradas relevantes. As redes foram construídas para diferentes horizontes de previsão (de 2 até 60 minutos à frente), permitindo avaliar o desempenho da abordagem em diferentes escalas temporais.

2.3.1 Métricas de desempenho

Os conjuntos de dados gerados com o algoritmo ApproxHull foram utilizados como entrada para os procedimentos do MOGA, os quais envolveram a geração de modelos e a avaliação de métricas de desempenho. Além desses procedimentos, um conjunto adicional, correspondente à série temporal original, foi empregado para avaliação complementar do desempenho dos modelos e apoio aos objetivos do MOGA. Esse conjunto consistiu em uma sequência contínua abrangendo os primeiros 15 dias de abril de 2014, selecionada principalmente pela ausência de lacunas nos dados, condição essencial para o cálculo preciso dos erros de previsão e para a elaboração de gráficos comparativos entre os valores observados e previstos, tanto para análise visual quanto quantitativa. Além disso, esse período ocorre durante a transição entre a estação chuvosa e a seca, oferecendo um cenário representativo, ainda que não extremo, da variabilidade atmosférica local e sem a carga de aerossóis típica da transição da estação seca para a úmida (ROCHA; YAMASOE, 2013). No período de maior carga de aerossol, espera-se uma redução na componente direta e um aumento relativo na componente difusa, o que tende a suavizar a variabilidade instantânea da irradiância. Por outro lado, o aerossol introduz uma fonte adicional de incerteza, já que a atenuação por partículas pode ser confundida com a atenuação por nuvens finas. Assim, a previsibilidade pode ser alterada — em alguns casos favorecida pela suavização, mas em outros comprometida pela ambiguidade entre aerossol e nuvem. Embora não represente toda a sazonalidade da região, fornece uma linha de base consistente para a avaliação inicial dos modelos. Trabalhos futuros podem estender essa análise para períodos típicos das estações seca e chuvosa, a fim de investigar diferenças sazonais no desempenho dos modelos.

Diferentemente dos erros de treino e teste, que avaliam a capacidade do modelo em prever pontos futuros com base em conjuntos de entrada-saída conhecidos, o erro de previsão mensura o desempenho do modelo ao longo de um horizonte de previsão, simulando previsões contínuas em tempo real. Essa métrica captura a acumulação de erros ao longo do tempo quando o modelo realiza previsões iterativas ou encadeadas.

Diversos critérios de desempenho foram utilizados para avaliar a técnica proposta. Para suas definições, considere n o número de observações, y_i o valor observado e \hat{y}_i o valor previsto para a i -ésima observação.

O Erro Quadrático Médio (RMSE, *Root Mean Square Error*) é uma métrica amplamente utilizada para quantificar o erro em modelos de previsão numérica, sendo calculado como a raiz quadrada da média dos erros quadráticos. Já o Erro Médio

Absoluto (MAE, *Mean Absolute Error*) corresponde à média das diferenças absolutas entre os valores observados e previstos, fornecendo uma medida linear que não amplifica a influência de valores discrepantes.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.3)$$

Para permitir comparações entre diferentes conjuntos de dados, tanto o RMSE quanto o MAE podem ser normalizados. A normalização consiste em dividir essas métricas por um valor de referência, como a média da variável (adotada neste estudo). Isso gera o RMSE e o MAE normalizados (nRMSE e nMAE), que expressam o erro relativo à escala dos dados, facilitando a interpretação dos resultados de forma independente das unidades ou magnitudes específicas do problema analisado.

Para avaliar o desempenho do modelo na detecção de eventos de sobreirradiância, utilizou-se a matriz de confusão, que resume os resultados da classificação em quatro categorias: verdadeiros positivos (VP ou TP, *True Positives*), quando anomalias de sobreirradiância são corretamente identificadas; verdadeiros negativos (VN ou TN, *True Negatives*), quando eventos normais são corretamente classificados; falsos positivos (FP, *False Positives*), quando eventos normais são incorretamente classificados como anomalias; e falsos negativos (FN, *False Negatives*), quando anomalias reais não são identificadas.

A acurácia (*accuracy*) mede a proporção de classificações corretas (VP + VN) em relação ao total de casos, indicando a correção geral do modelo. A precisão (*precision*) avalia a proporção de VP entre todos os positivos previstos, refletindo a robustez do modelo na identificação de resultados positivos. A sensibilidade ou revocação (*recall*) quantifica a capacidade do modelo em identificar corretamente os casos positivos entre todos os positivos reais. Por fim, a medida-F1 (*F1-score*) corresponde à média harmônica entre precisão e revocação, sendo especialmente útil em cenários com classes desbalanceadas pois penaliza modelos que só acertam por acaso ou que erram sistematicamente ao não detectar eventos reais. Essas métricas fornecem uma visão abrangente do desempenho do modelo, equilibrando sua capacidade de detecção de anomalias com a exatidão da classificação.

$$\text{Acurácia} = \frac{\text{VP} + \text{VN}}{\text{Total}} \quad (2.4)$$

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}} \quad (2.5)$$

$$\text{Revocação} = \frac{\text{VP}}{\text{VP} + \text{FN}} \quad (2.6)$$

$$\text{Medida-F1} = 2 \cdot \frac{\text{Precisão} \cdot \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (2.7)$$

Essas métricas também são utilizadas para a avaliação da capacidade do modelo em prever rampas, conforme realizado no estudo de [Logothetis et al. \(2022\)](#) mas com a normalização e limiar fixo conforme [Zuo et al. \(2023\)](#). Para isso, foram calculadas as rampas observadas, usando a janela temporal de 2 minutos, e previstas, considerando os valores previstos 2 minutos à frente.

2.3.2 Modelos de persistência

Um modelo de persistência é um modelo de referência simples que assume que o valor futuro de uma variável será igual ao seu valor atual ou a um valor recente (como o valor no mesmo horário do dia anterior). É amplamente utilizado em problemas de previsão de séries temporais, especialmente para variáveis como a irradiância, por fornecer uma linha de base contra a qual se pode avaliar se modelos mais complexos realmente apresentam melhorias significativas ([HYNDMAN; ATHANASOPOULOS, 2021](#)). Existem diferentes tipos de modelos de persistência:

- **Persistência de passo anterior (PersistPrev):** Essa abordagem assume que o valor de GHI no instante $t + k$ será igual ao seu valor no instante t . Trata-se da forma mais simples de persistência e é útil para previsões de curtíssimo prazo. Fórmula: $GHI(t + k) = GHI(t)$, com $k = 2$ minutos no presente estudo.
- **Persistência diária (PersistDay):** Essa abordagem assume que o padrão diário da irradiância se repete. Fórmula: $GHI(t + k) = GHI(t + k - 1440 \text{ min})$, pois 24 horas equivalem a 1440 minutos.
- **Persistência sazonal combinada (PersistCombined):** Esse modelo

combina a persistência de curtíssimo prazo (2 minutos) e de curto prazo (24 horas), utilizando uma média ponderada. Fórmula: $GHI(t) = \alpha \cdot GHI(t - k) + (1 - \alpha) \cdot GHI(t - 24h)$, onde α é um peso ajustável.

O parâmetro α define o peso relativo entre a persistência de curtíssimo prazo e a sazonalidade diária. Para $\alpha = 1$, o modelo equivale ao PersistPrev; para $\alpha = 0$, ao PersistDay. Para determinar o valor ótimo de α no modelo de persistência combinada, foi utilizada uma abordagem de otimização contínua com base na minimização de erro. Especificamente, a escolha de α foi formulada como um problema de otimização univariada, em que se busca o valor que minimiza uma função objetivo definida por uma métrica de erro entre os valores observados de GHI e as previsões do modelo de persistência combinada. A função objetivo pode ser o RMSE, o MAE ou outra métrica de desempenho relevante. A implementação foi feita utilizando a rotina de minimização escalar com restrição (*bounded*) da biblioteca SciPy (VIRTANEN, 2020), amplamente empregada em aplicações científicas em Python. Esse procedimento permite identificar, de forma objetiva e automatizada, o peso que melhor equilibra as contribuições de cada tipo de persistência, resultando em previsões mais acuradas dentro do horizonte considerado.

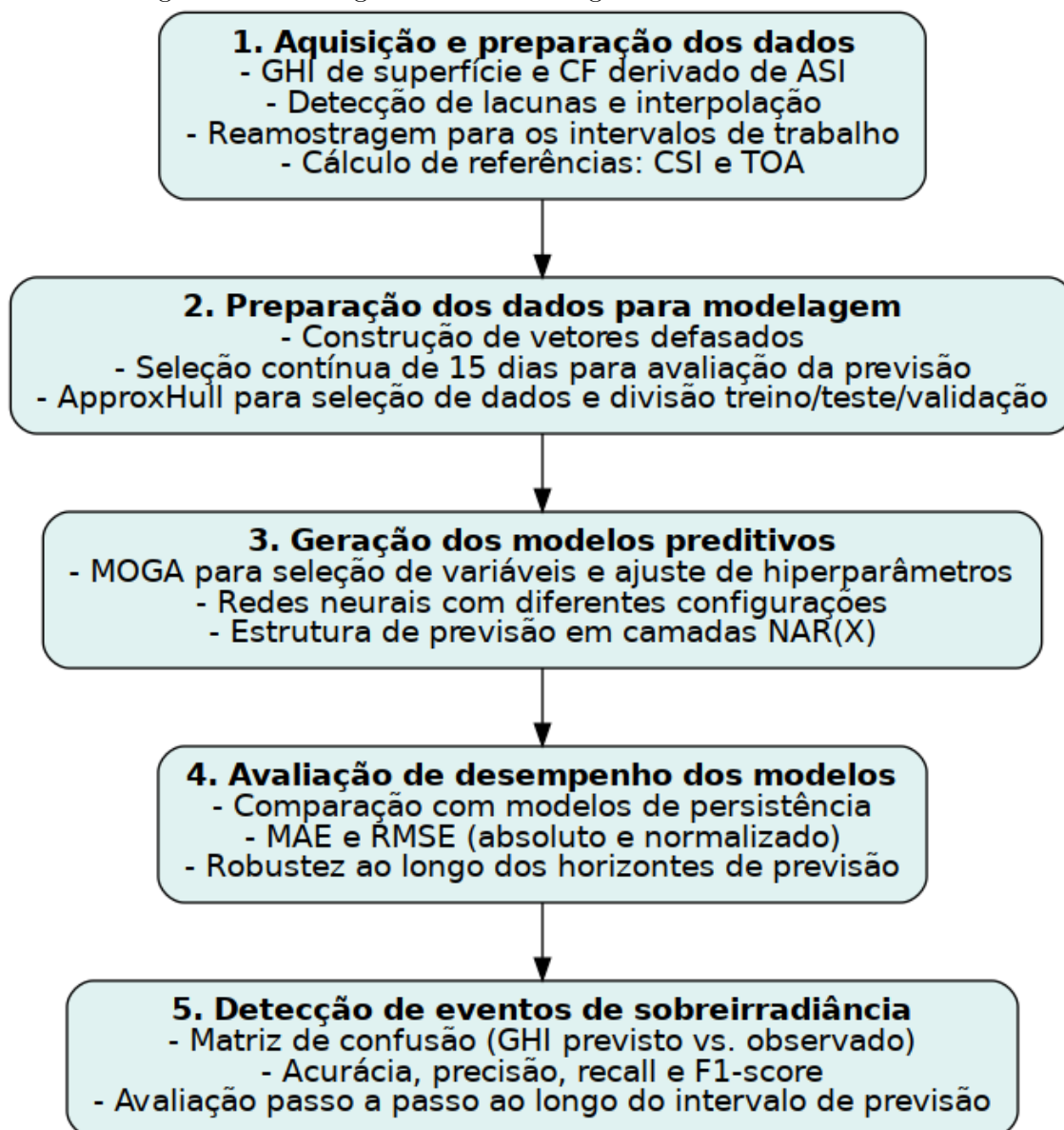
2.3.3 Fluxograma do processo metodológico

A Figura 2.7 apresenta uma visão geral do fluxo metodológico adotado neste trabalho, abrangendo desde a aquisição dos dados até a detecção dos eventos de sobreirradiância.

2.3.4 Análise estatística dos resultados

Além de gráficos de séries temporais (com valores médios em escalas minutar, diária e mensal), também foram utilizadas representações estatísticas complementares, como histogramas e diagramas de caixa (*box-plot*), com o objetivo de sintetizar características importantes das distribuições dos dados. Para avaliar a significância estatística das diferenças entre séries temporais, é importante conhecer o valor-p (*p-value*), que é uma medida estatística que indica a probabilidade de se obter um resultado tão extremo quanto o observado (ou mais), assumindo que a hipótese nula seja verdadeira. Quanto menor o valor-p, menor a evidência a favor da hipótese nula e maior a indicação de que há uma diferença estatisticamente significativa entre os grupos comparados. Valores-p inferiores a um limiar pré-definido (comumente 0,05) são geralmente interpretados como evidência suficiente para rejeitar a hipótese nula em testes estatísticos. Foram aplicados testes estatísticos clássicos, conforme descrito

Figura 2.7 - Fluxograma da metodologia desenvolvida neste estudo.



Fonte: Adaptado de Rocha et al. (2025).

a seguir.

Inicialmente, a normalidade das distribuições foi verificada por meio do teste de Shapiro-Wilk (SHAPIRO; WILK, 1965), amplamente utilizado para avaliar se uma amostra provém de uma distribuição normal (RAZALI; WAH, 2011). Esse teste é particularmente apropriado para amostras de pequeno e médio porte, sendo sensível a desvios de simetria e curtose. Um valor-p inferior a 0,05 indica a rejeição da hipótese nula de normalidade, sugerindo que a distribuição dos dados difere signifi-

cativamente de uma distribuição normal.

Com base no resultado do teste de normalidade, optou-se entre testes paramétricos e não paramétricos para comparar as distribuições de dois grupos independentes (por exemplo, anos distintos). Quando ambas as distribuições apresentaram comportamento aproximadamente normal e variâncias comparáveis, foi utilizado o teste t de Student para duas amostras independentes. Esse teste paramétrico verifica se a diferença entre as médias de dois grupos é estatisticamente significativa, assumindo que os dados seguem distribuições normais e que as amostras são independentes (LEHMANN; ROMANO, 2022).

Por outro lado, quando pelo menos uma das distribuições não atendeu ao critério de normalidade, recorreu-se ao teste de Mann-Whitney U (também conhecido como teste de Wilcoxon para amostras independentes). Trata-se de um teste não paramétrico que compara as medianas de dois grupos independentes, sem pressupor uma distribuição específica dos dados. Esse teste é robusto frente a assimetrias ou *outliers*, e permite avaliar se as distribuições de duas populações diferem de forma estatisticamente significativa (LEHMANN; ROMANO, 2022).

3 RESULTADOS

A Seção 3.1 apresenta uma análise descritiva dos principais padrões sazonais e interanuais observados nas variáveis de interesse, com ênfase em GHI e CF. Os resultados são organizados por meio de gráficos de séries temporais que ilustram as tendências médias nas escalas de minuto, dia e mês, bem como as variações intra e interanuais entre 2014 e 2015, destacando as diferenças entre as estações seca e chuvosa e o impacto do fenômeno El Niño ocorrido em 2015 - de acordo com o que foi apresentado na Seção 1.4.

A Seção 3.2 analisa os padrões estatísticos da variabilidade da irradiância solar em diferentes escalas de tempo, com ênfase na caracterização das rampas de irradiância e do índice DARR, assim como a relação com a sobreirradiância.

A Seção 3.3 aborda os eventos de sobreirradiância. Conforme descrito na metodologia, eles foram identificados com base em duas referências (CSI e TOA), considerando médias em janelas de 1, 2 e 5 minutos e restringindo-se aos casos em que os valores de GHI ultrapassaram 1000 W/m^2 . Os eventos foram agregados em totais diários e mensais; o índice DAOI foi calculado para cada dia e apresentado graficamente para análise. O uso do CSI como referência mostrou-se mais sensível para avaliar o comportamento atmosférico no contexto da irradiância solar incidente à superfície. Além disso, as análises com base no CSI, apoiadas por um conjunto de dados mais robusto, forneceram fundamentos estatísticos mais consistentes e, por isso, foram priorizadas na apresentação dos resultados. Devido ao menor volume de dados disponível nas análises baseadas em TOA, e à alta similaridade de comportamento em relação ao CSI, os gráficos correspondentes não são apresentados de forma detalhada.

A Seção 3.4 contém os gráficos e as análises dos tipos de nuvens classificados no produto ARSCL. Para associar essas classificações às propriedades radiativas superficiais, o tipo de nuvem dominante é atribuído ao instante de observação de acordo com a formação mais baixa presente na coluna atmosférica, conforme o procedimento descrito por (GIANGRANDE et al., 2017). Além disso, também foram consideradas somente as observações ocorridas nos instantes dos registros de GHI que apresentam frequência de 1 minuto - ou seja, somente dados diurnos.

A Seção 3.5 apresenta os resultados dos modelos de previsão de GHI, inicialmente com três abordagens baseadas em persistência que serviram como uma linha de base para avaliação. Em seguida, a performance dos modelos gerados foi analisada considerando a inclusão (ou não) da variável CF como exógena, diferentes horizontes

de previsão e restrições de busca.

Por fim, a Seção 3.6 contém um modelo selecionado para a tarefa de detecção de eventos de sobreirradiância e testado quanto à sua capacidade de antecipar tais ocorrências com base nas previsões geradas.

Conforme explicado na Metodologia (Seção 2.2), as análises estatísticas apresentadas na primeira seção foram realizadas com os dados originais na frequência de 1 minuto, e os modelos de previsão foram treinados com dados reamostrados para uma frequência de 2 minutos. Esta decisão foi particularmente importante para a estimativa de médias diárias e mensais, bem como para a análise de eventos de curta duração, como a sobreirradiância. Essa preocupação também motivou a opção de não aplicar restrições baseadas no ângulo zenital, tanto na geração dos modelos quanto nas análises estatísticas - ainda que seja comum, em estudos de radiação atmosférica, descartar dados quando o ângulo zenital excede 80° , ou mesmo 70° . No entanto, o uso de um limiar de GHI superior a 1000 W/m^2 para análises específicas atua, de forma efetiva, como um filtro que exclui ângulos solares baixos, uma vez que tais valores de GHI tendem a ocorrer apenas quando o Sol se encontra em um ângulo zenital menor.

3.1 Variações sazonais e interanuais

A integral diária da irradiância solar global (H) representa a quantidade total de energia solar recebida por unidade de área ao longo de um dia, sendo tipicamente expressa em megajoules por metro quadrado (MJ/m^2) ou quilowatt-hora por metro quadrado (kWh/m^2). Essa grandeza é obtida por meio da integração temporal de GHI, a qual considera a potência por unidade de área incidente sobre uma superfície horizontal proveniente tanto da radiação direta quanto da difusa. H é uma métrica fundamental em estudos climatológicos e no dimensionamento de sistemas fotovoltaicos, pois reflete a disponibilidade de energia solar em um determinado local. Essa disponibilidade é determinada por fatores como latitude e estação do ano, e modulada por condições atmosféricas variáveis, como a cobertura de nuvens e a presença de aerossóis.

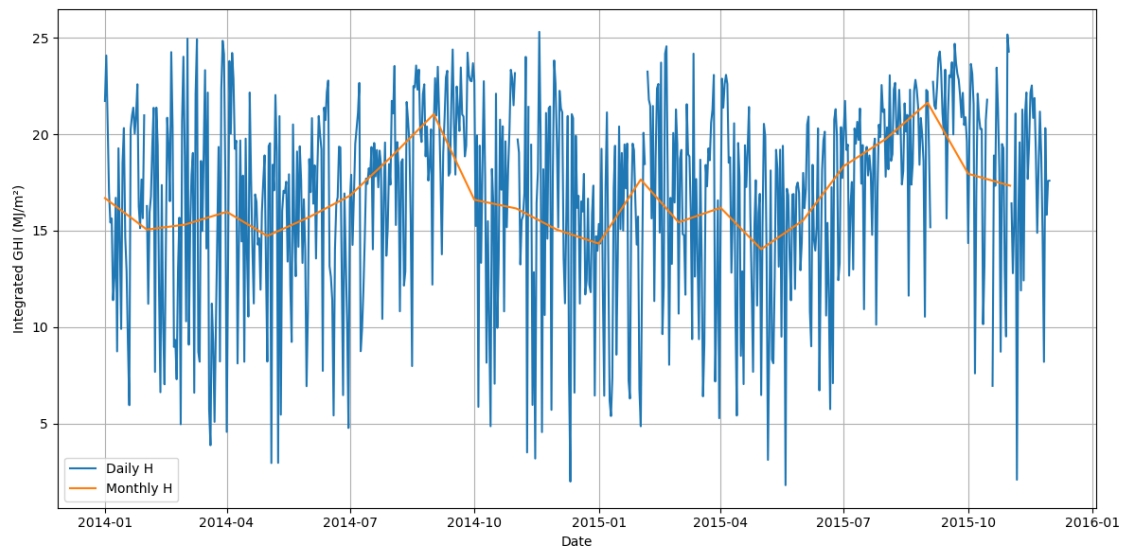
A irradiância solar na parte superior da atmosfera (TOA) varia ao longo do ano em função da inclinação do eixo da Terra e da variação da altura solar ao meio-dia. Os maiores valores de TOA ao meio-dia ocorrem durante o solstício de verão (dezembro no hemisfério sul), quando o Sol atinge sua maior elevação. Os menores valores são registrados nas proximidades do solstício de inverno (junho), quando o ângulo zenital

solar ao meio-dia é maior e a duração do dia é reduzida, diminuindo a irradiância recebida. Embora o Sol não atinja exatamente o zênite durante o solstício de verão, sua elevação ainda é elevada, e a maior duração do dia resulta em valores elevados de irradiância, ligeiramente inferiores aos observados nos equinócios. Importante destacar que, devido à baixa latitude do sítio experimental ($3,21^{\circ}\text{S}$), a redução no TOA ao meio-dia durante o equinócio, em comparação com um ponto localizado exatamente sobre o Equador, é inferior a $0,2\%$. Essa diferença foi estimada com base no cosseno do ângulo zenital solar ao meio-dia.

A variação da irradiância solar TOA ao longo do ano deve-se também à excentricidade da órbita terrestre, que faz com que a Terra esteja mais próxima do Sol no periélio (cerca de 147,1 milhões de km), que ocorreu em 4 de janeiro para os anos de 2014 e 2015 ([UNITED STATES NAVAL OBSERVATORY, 2025](#)), e mais distante no afélio (cerca de 152,1 milhões de km). Com base na lei do inverso do quadrado da distância, isso resulta em uma diferença de aproximadamente $6,9\%$ na irradiância TOA entre os dois extremos ([YAMASOE; CORRÊA, 2016](#)).

A variabilidade anual é ilustrada na Figura 3.1, juntamente com o efeito da nebulosidade, o qual resulta em picos de irradiância em setembro e mínimos em maio, para ambos os anos analisados. Ao se analisar os valores diários dentro de cada mês, observa-se que a variabilidade é menor durante a estação seca, devido à redução da cobertura de nuvens. Os valores de irradiância solar global observados estão em conformidade com a climatologia regional, apresentando máximos da ordem de $20 \text{ MJ/m}^2/\text{dia}$ durante a estação seca e mínimos em torno de $15 \text{ MJ/m}^2/\text{dia}$ na estação chuvosa ([MACEDO; FISCH, 2018](#); [SALATI; MARQUES, 1984](#)).

Figura 3.1 - Integral diária da irradiância solar global (H diária) e média mensal (H mensal).



Fonte: Produção do autor.

(MACEDO; FISCH, 2018) utilizou o mesmo conjunto de dados observacionais do projeto GoAmazon 2014/15, abrangendo irradiância solar (global, difusa e direta), CF, AOD e precipitação. O estudo concentrou-se na caracterização geral da irradiância a partir de valores médios mensais, comparações interanuais, sazonais e variações horárias. No presente trabalho, buscou-se um aprofundamento dessas análises, incorporando medianas de irradiância global e CF nos valores mensais, bem como médias minutais para a avaliação do ciclo diurno. Foram ainda investigados o ciclo diurno e os agregados temporais das rampas de GHI, além da detecção de eventos de sobreirradiância e sua relação com outras variáveis, incluindo tipologia de nuvens. Dessa forma, este estudo se apoia em bases semelhantes, mas amplia a abordagem metodológica, complementando e avançando em relação ao trabalho precursor de (MACEDO; FISCH, 2018) na caracterização da variabilidade da irradiância na Amazônia Central.

A Figura 3.2 apresenta as distribuições dos valores diários e mensais de GHI e de CF para todo o período de dados analisado, representadas por diagramas de caixa (box plots). As médias diárias de GHI foram calculadas de forma padronizada, considerando todas as 1.440 amostras por dia (frequência de 1 minuto), incluindo os períodos noturnos ($GHI = 0$). Essa abordagem tem como objetivo evitar a obtenção de valores médios artificialmente elevados em dias com menor duração de luz solar,

padronizando as médias em relação a diferentes pontos do mundo. Para os dias com dados ausentes, a média foi registrada como NaN (*Not a Number*).

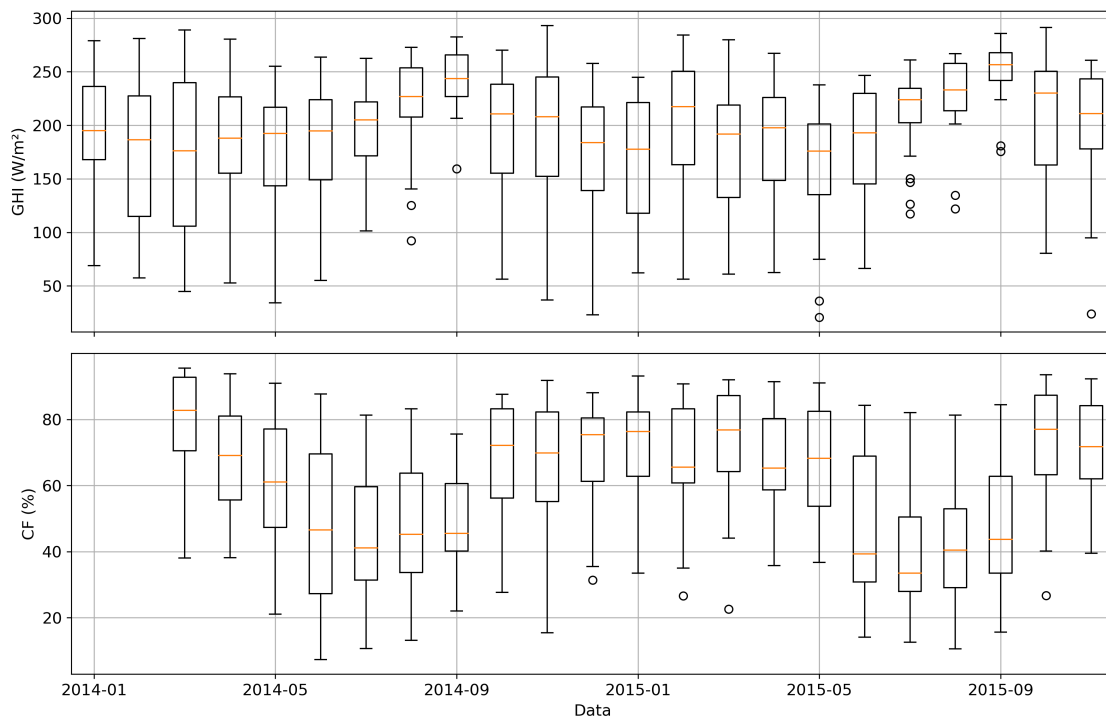
Do ponto de vista sazonal, os valores medianos mensais de GHI aumentam à medida que a estação seca avança, impulsionados pela redução da nebulosidade. Na análise interanual (2014–2015), a maior diferença mensal foi observada em fevereiro (aproximadamente 30 W/m^2), enquanto a menor ocorreu em junho (cerca de 2 W/m^2). No entanto, as diferenças entre os anos de 2014 e 2015 não foram estatisticamente significativas ao nível de 5%. Os resultados apresentados estão em concordância com [Macedo \(2016\)](#), que calculou médias mensais de GHI sem a mencionada padronização, encontrando as maiores diferenças para os meses de janeiro (63 W/m^2) e fevereiro (57 W/m^2), e a menor diferença para o mês de junho (4 W/m^2).

Durante o período analisado nesta pesquisa (2014–2015), houve uma variabilidade interanual significativa associada ao ciclo ENSO. Em 2014, observou-se uma tentativa de formação de El Niño, com o Índice Oceânico Niño (ONI, *Oceanic Niño Index*) atingindo valores próximos ou ligeiramente superiores ao limiar de $+0,5^\circ\text{C}$, mas sem persistência suficiente para caracterizá-lo como um evento oficial segundo os critérios da Administração Nacional Oceânica e Atmosférica dos EUA (NOAA, *National Oceanic and Atmospheric Administration*). Assim, a maior parte de 2014 foi marcada por condições de neutralidade, com incursões ocasionais próximas ao limiar de El Niño ([CLIMATE PREDICTION CENTER, NCEP, NOAA, 2025](#)). O El Niño 2015–2016, por outro lado, foi um dos mais intensos já registrados, com início no primeiro trimestre de 2015 ($\text{ONI} = +0,6^\circ\text{C}$) e pico no último trimestre do ano ($\text{ONI} > +2,0^\circ\text{C}$), persistindo até maio de 2016. Este evento influenciou significativamente a dinâmica atmosférica na Amazônia Central, impactando a formação de nebulosidade e chuvas.

Os menores valores medianos de CF ocorreram durante a estação seca, especialmente entre junho e setembro, como era esperado ([ARIAS et al., 2010](#); [SENA et al., 2018](#)). Para a maioria dos meses, os valores variaram entre 39% e 77%, intervalo compatível com os valores médios mensais de CF reportados por [Machado et al. \(2004\)](#) (42% a 76%) para a região de Manaus com base em dados climatológicos de satélite. Dois meses se destacaram como exceções a esse intervalo: julho de 2015 (33%) e março de 2014 (83%). Esses extremos são coerentes com as condições atmosféricas típicas desses períodos, uma vez que o valor mais baixo foi registrado durante a estação seca em um evento de El Niño, que tende a reduzir a cobertura de nuvens na região amazônica ([MARENGO; ESPINOZA, 2015](#); [SENA et al., 2018](#)). O valor

extremo mais alto ocorreu no auge da estação chuvosa em um ano climático típico, período comumente associado a maiores valores de CF, e também é compatível com o valor de CF médio igual a 80% encontrado em [Alves et al. \(2007\)](#) para o mês de março, calculado para um setor de $2,5^\circ \times 2,5^\circ$, centrado em Manaus, para os anos de 1983 a 1994. Ao comparar os valores medianos de CF da estação seca de um ano típico (2014) com os observados durante um evento de El Niño (2015), nota-se uma redução, particularmente nos meses de junho (de 47% para 39%) e julho (de 41% para 33%). Essa diferença, embora não estatisticamente significativa, está em consonância com a dinâmica atmosférica de grande escala na região, que tende a inibir a formação de nuvens durante eventos de El Niño ([SENA et al., 2018](#)). Dessa forma, os dados sugerem uma possível influência do fenômeno, ainda que não conclusiva do ponto de vista estatístico.

Figura 3.2 - Distribuições dos valores diários e mensais de GHI (acima) e CF (abaixo), representadas por diagramas de caixa, com indicação da mediana, dos quartis e dos valores atípicos (*outliers*).

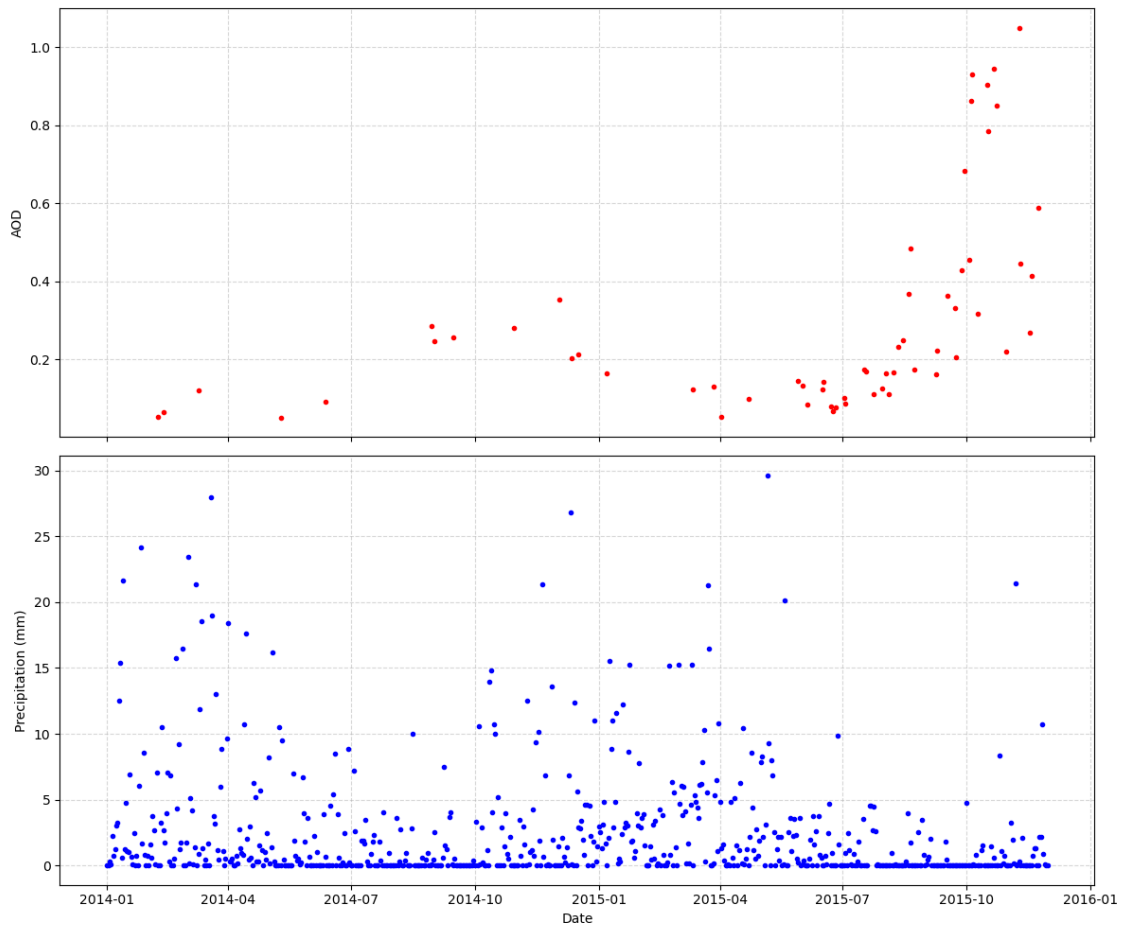


Fonte: Produção do autor.

A Figura 3.3 apresenta os valores médios diários da profundidade óptica de aerossóis (AOD) e da precipitação acumulada diária. Apesar da menor cobertura de nuvens

observada durante a estação seca de 2015 (ano de El Niño), a visibilidade atmosférica foi reduzida, como indicado pelos maiores valores de AOD. Com a diminuição das chuvas, o ambiente tornou-se mais seco e suscetível a queimadas. De fato, o número total de focos de calor em 2015 foi 63% maior que em 2014 no Brasil (MACEDO; FISCH, 2018) e 48% maior no estado do Amazonas (INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE), 2025), o que explica o aumento dos valores de AOD. Esse aumento da carga de aerossóis contribui para explicar a ausência de diferenças estatisticamente significativas na irradiância média entre 2014 e 2015. Embora a redução da cobertura de nuvens durante o El Niño de 2015 pudesse favorecer maior incidência de radiação solar à superfície, os elevados valores de AOD indicam uma atmosfera mais carregada de partículas, principalmente oriundas de queimadas. Esses aerossóis atuam como barreiras à radiação solar direta, intensificando a dispersão e absorção da luz na coluna atmosférica, o que pode neutralizar ou até anular o efeito do céu mais limpo sobre os níveis médios de irradiância observados.

Figura 3.3 - Valores médios diários de AOD (acima) e precipitação acumulada (abaixo) para todo o período analisado.



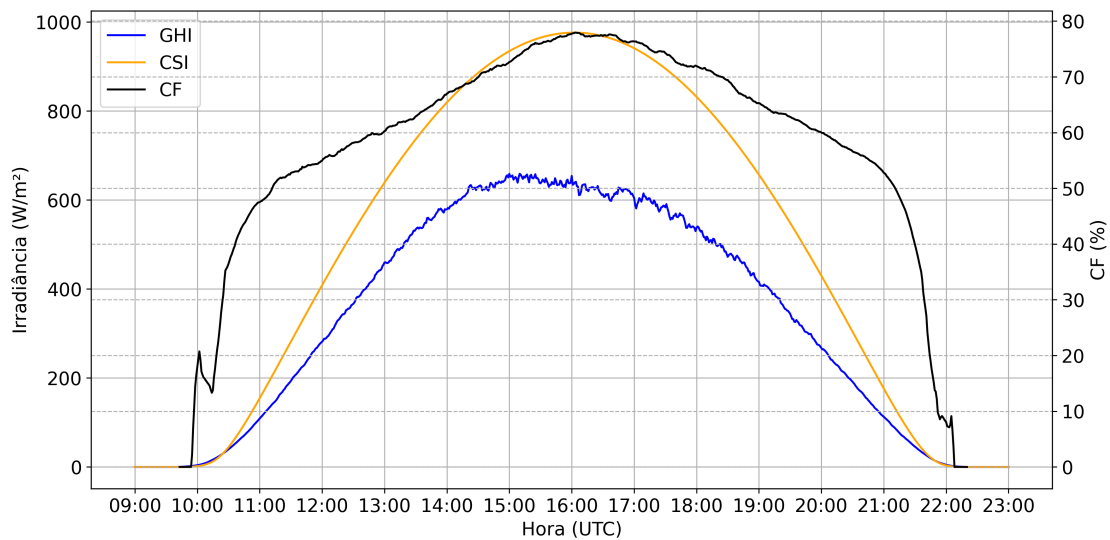
Fonte: Produção do autor.

A Figura 3.4 apresenta a variação diurna média integrada de GHI, CSI e CF, com base nos dados de 2014 e 2015. A curva de GHI segue o padrão típico em forma de sino, determinado pela trajetória solar, com maior variabilidade ao redor do meio-dia solar, provavelmente influenciada pela cobertura de nuvens e por outros fatores atmosféricos que afetam a transmissão da radiação solar. Os valores de CF iniciam o dia em níveis baixos, aumentam rapidamente durante a manhã e atingem seu pico por volta das 16:00 UTC, antes de declinarem gradativamente ao final do dia. Esse comportamento reflete um ciclo diurno bem definido, possivelmente associado a processos convectivos característicos da região de estudo - conforme explicado na Seção 1.3.1. Próximo das 10:00 UTC, existe um aumento de aproximadamente 20 pontos percentuais em menos de 5 minutos e uma queda de 5 pontos nos 10 minutos seguintes para, então, haver um aumento contínuo com o tempo. Esses valores

mais altos de CF próximos ao nascer do sol podem estar relacionados à formação de nevoeiros localizados, comuns sob condições de alta umidade, baixos ventos e atmosfera estável. No entanto, importante considerar que os períodos de nascer e pôr do sol também estão sujeitos a limitações do algoritmo de estimativa de CF, o qual tende a apresentar menor desempenho sob condições de baixa luminosidade (ROCHA et al., 2025). Essa parece ser a principal explicação para o aumento de 2 pontos percentuais no valor médio de CF que ocorre às 22:03 UTC após uma queda contínua e depois retoma a diminuição dos valores de CF até zero.

Durante as horas centrais do dia, particularmente entre 13:00 e 19:00 UTC, observa-se uma divergência pronunciada entre GHI e CSI, coincidindo com os valores elevados de CF. Além disso, a GHI vespertina tende a ser ligeiramente inferior à observada pela manhã, enquanto o oposto ocorre para a CF. Essa divergência evidencia o efeito modulador da cobertura de nuvens sobre a irradiância na superfície: enquanto a CSI segue a geometria solar esperada, a GHI sofre atenuação em razão da maior presença de nuvens.

Figura 3.4 - Ciclo diurno médio de GHI, CSI (eixo esquerdo, W/m^2) e CF (eixo direito, %), com base em dados de 2014–2015.



Fonte: Produção do autor.

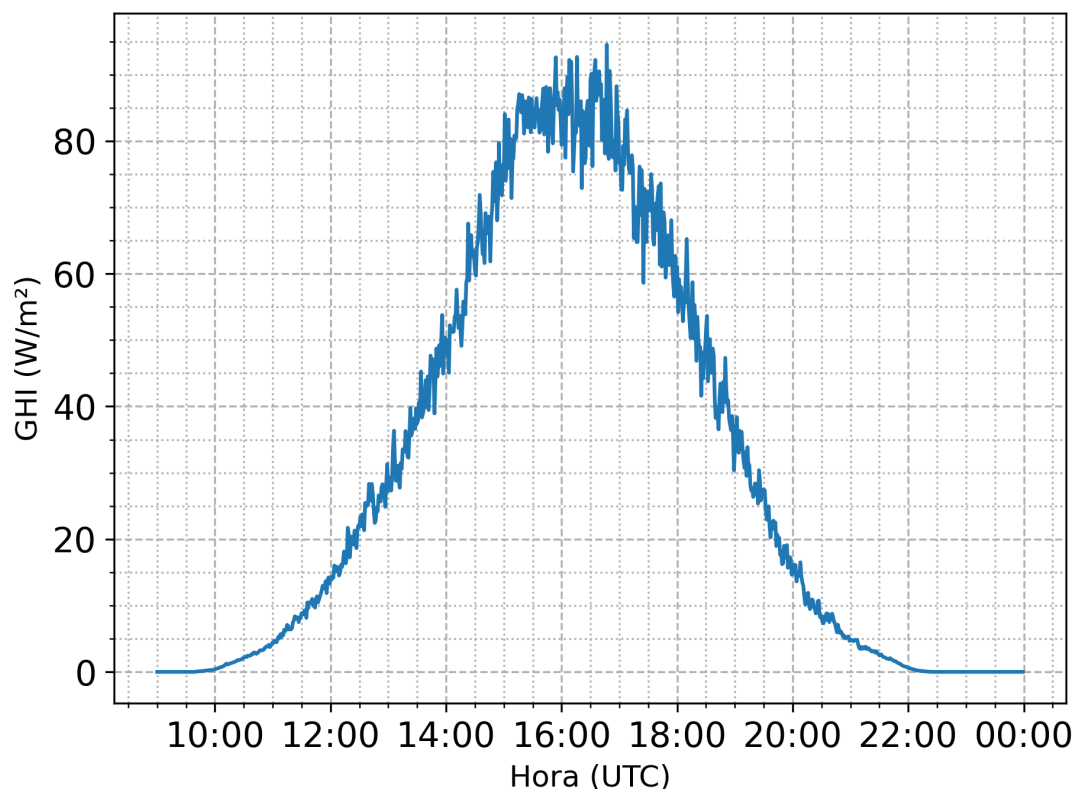
3.2 Caracterização da variabilidade

A distribuição dos valores de taxa de rampa (RR) apresenta um comportamento assimétrico decrescente: maior concentração de eventos de pequena magnitude e frequência decrescente à medida que a variação entre instantes consecutivos aumenta. Esse padrão é esperado e pode ser interpretado como semelhante a uma distribuição do tipo gama ou exponencial, já que grandes rampas de irradiância (resultantes de mudanças atmosféricas abruptas) são naturalmente raras em comparação com pequenas flutuações.

Por outro lado, os valores de DARR, por serem obtidos a partir da soma das diferenças absolutas de irradiância ao longo do dia, tendem a se distribuir de forma mais simétrica em torno de um valor médio. Esse comportamento é compatível com uma distribuição aproximadamente normal (ou gaussiana), resultado da agregação de múltiplos eventos ao longo do tempo - o que, conforme a teoria do limite central, favorece uma concentração de valores próximos da média e redução gradual da frequência nos extremos.

A Figura 3.5 apresenta o tamanho médio das rampas, calculadas nos casos em que existem valores de GHI no minuto seguinte, em seu ciclo diurno, com dados de todo o período estudado. Observa-se uma concentração de rampas de maior magnitude e maior variabilidade à medida que se aproxima o meio-dia solar, com destaque para uma assimetria vespertina, em que as rampas tendem a ser ligeiramente mais intensas. Esse comportamento está associado ao aumento da irradiância direta e ao aumento da nebulosidade - conforme observado na Figura 3.4. A maior granularidade (oscilações rápidas entre valores consecutivos) observada no traçado do gráfico durante o período de pico pode indicar maior variabilidade de nebulosidade, com formação e dissipação frequente de nuvens e passagens intermitentes de sombras. Esses padrões sugerem uma predominância de rampas de curta duração e alta intensidade nesse intervalo do dia.

Figura 3.5 - Ciclo diurno da média das rampas de irradiância (em W/m^2).



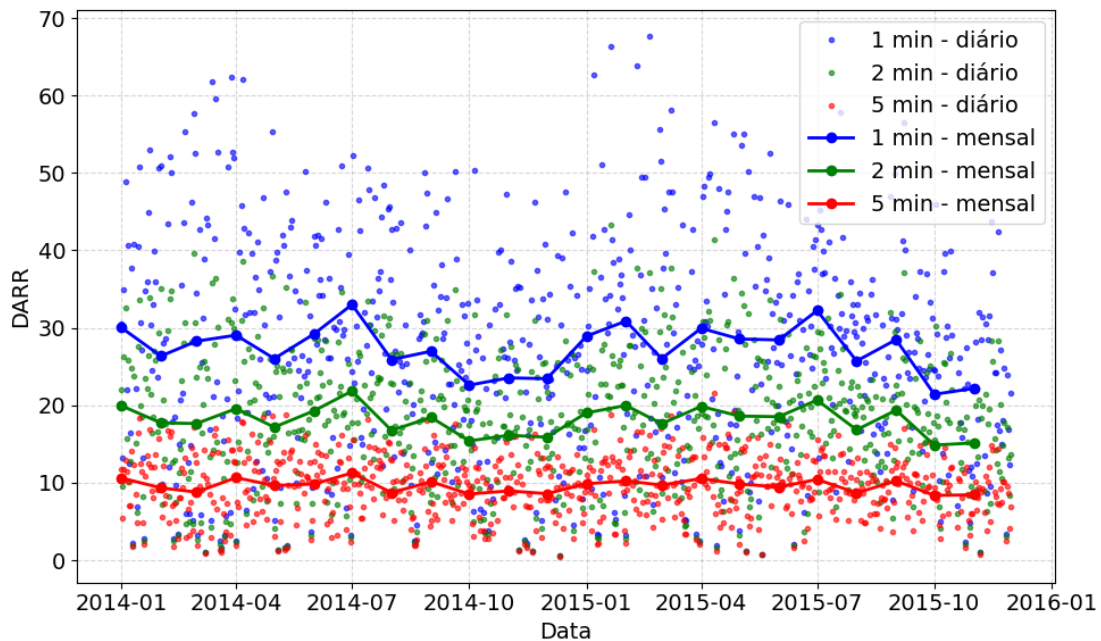
Fonte: Produção do autor.

Valores mais altos de DARR indicam maior variabilidade da irradiância, o que pode estar associado à passagem de nuvens. A Figura 3.6 apresenta as séries diárias e mensais de DARR para três janelas temporais diferentes: 1, 2 e 5 minutos. Observa-se que, quanto maior o intervalo de tempo considerado entre as medições, menores tendem a ser os valores médios (tanto os diários quanto os mensais) de DARR, assim como a dispersão dos valores também diminui. Esse comportamento ocorre porque o uso de intervalos maiores suaviza flutuações rápidas na irradiância, reduzindo a magnitude das variações instantâneas capturadas e, conseqüentemente, os valores médios de DARR. O fato de as médias de 1 minuto apresentarem valores sistematicamente mais altos que as outras médias indica que as variações de irradiância são particularmente intensas em escala sub-minuto.

As curvas das médias mensais mostram um comportamento sazonal, com variações relativamente suaves ao longo do ano. Embora as médias mensais indiquem que julho apresenta os maiores valores de DARR, a análise dos valores diários revela uma

variabilidade considerável ao longo de todo o período. Essa dispersão diária é suficiente para impedir que se estabeleça uma conclusão estatisticamente significativa sobre qual período apresenta os maiores ou menores valores de DARR. Em outras palavras, embora a estação seca tenda a concentrar valores médios mais altos, existem dias em outros meses com valores de variabilidade comparável ou até superior às observadas em julho, tornando a interpretação baseada apenas nas médias limitada. Com isso, observa-se que, embora a estação seca tenha médias de irradiância mais altas, isso não significa que a variabilidade diária seja menor ou maior que na estação úmida. Assim, mesmo durante a estação seca ou a úmida, podem surgir dias com alta variabilidade e dias relativamente estáveis.

Figura 3.6 - DARR diários (pontos) e médias mensais (linhas), com diferenças temporais de 1, 2 e 5 minutos (azul, verde e vermelho, respectivamente).



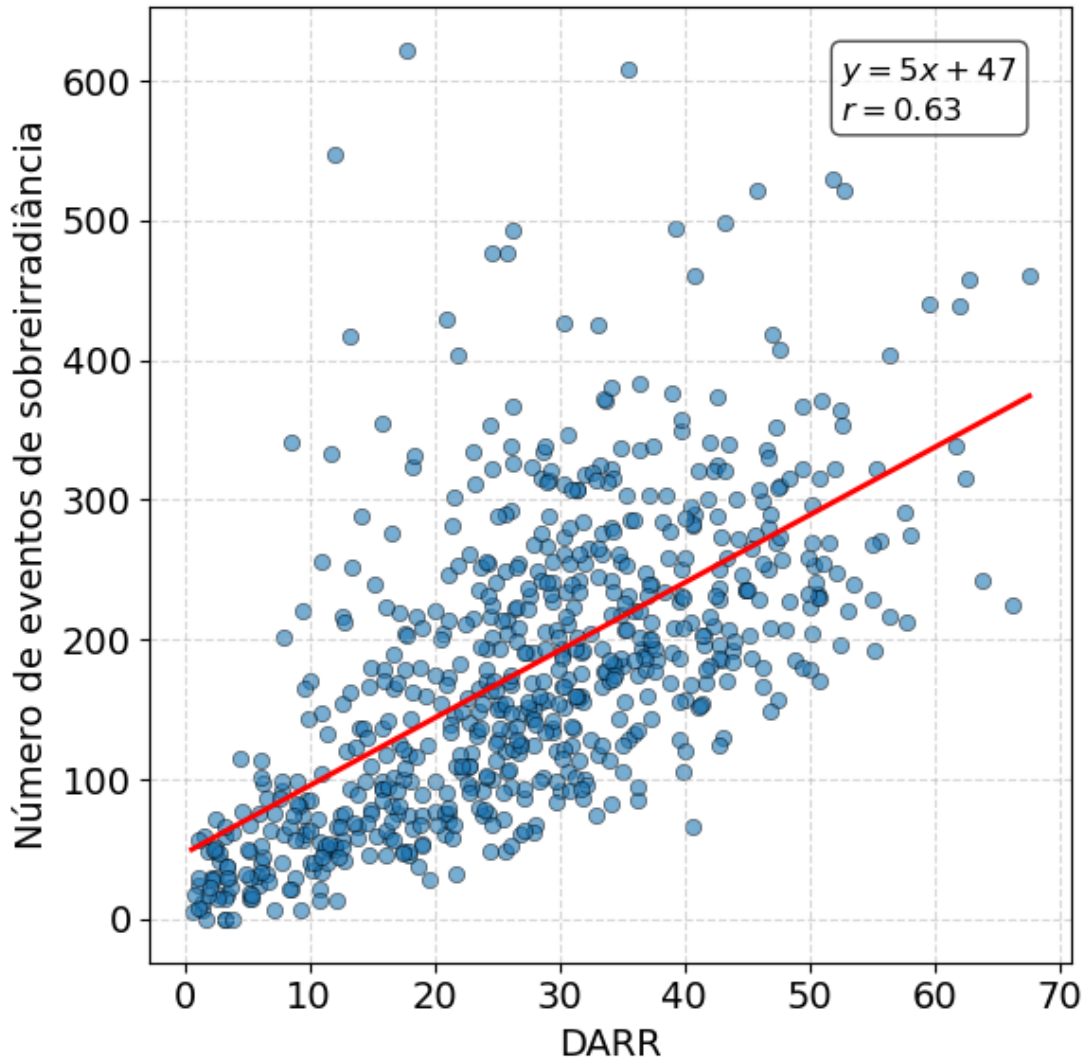
Fonte: Produção do autor.

Considerando todos os valores de 1 minuto (Figura 3.7), observa-se uma correlação positiva de 0,63 entre o índice DARR e o número diário de eventos de sobreirradiação. Esse valor sugere que a variabilidade intradiária da irradiância solar está associada à ocorrência desses eventos, e a relação linear obtida (coeficiente angular aproximadamente 5) indica que aumentos na DARR tendem a intensificar a frequência dos eventos de sobreirradiação. Esse resultado é coerente com o comportamento

físico da atmosfera em regiões tropicais, onde a presença de nuvens convectivas e efeitos ópticos podem induzir flutuações intensas e rápidas na irradiância (MOL; HEERWAARDEN, 2025). Essa observação reforça que a variabilidade de GHI atua como fator contribuinte para a ocorrência de sobreirradiância, particularmente em escalas sub-horárias. Considerando somente valores das estações chuvosa e seca, a correlação altera-se para 0,67 e 0,58, respectivamente. Uma explicação para isso seria que, quando há mais cobertura de nuvens, criam-se mais bordas, mesmo que a irradiância média seja menor, e são criadas muitas oscilações menores, porém mais associadas aos eventos de sobreirradiância.

Apesar de os maiores valores médios de DARR ocorrerem na estação seca, observa-se uma correlação mais forte entre esta variabilidade (DARR) e a ocorrência de eventos de sobreirradiância durante a estação chuvosa. Essa aparente contradição pode ser explicada pelas diferentes naturezas da variabilidade em cada período: na estação seca, predominam rampas intensas associadas a nuvens isoladas em um fundo de céu limpo, que nem sempre resultam em sobreirradiância. Na estação chuvosa, a maior cobertura de nuvens fragmentadas e a presença de condições ópticas favoráveis (como bordas de nuvens e múltiplos espalhamentos) aumentam a probabilidade de ocorrência de sobreirradiância, mesmo em condições de variabilidade menos intensa - o que torna mais significativa a associação entre DARR e os eventos nesse período. Essa interpretação é coerente com a hipótese de que, na região estudada, a estação chuvosa é dominada por sistemas de maior escala (como o movimento de noroeste para sudeste da convecção da América do Sul, segundo Horel et al. (1989)), enquanto a estação seca é mais influenciada por processos convectivos locais e isolados (MACHADO et al., 2018), o que impacta a estrutura das nuvens e os mecanismos ópticos responsáveis pela sobreirradiância.

Figura 3.7 - Número de eventos de sobreirradiância *versus* DARR, com linha de regressão linear (em vermelho) e respectivas equação ajustada e coeficiente de correlação de Pearson.



Fonte: Produção do autor.

3.3 Caracterização de eventos de sobreirradiância

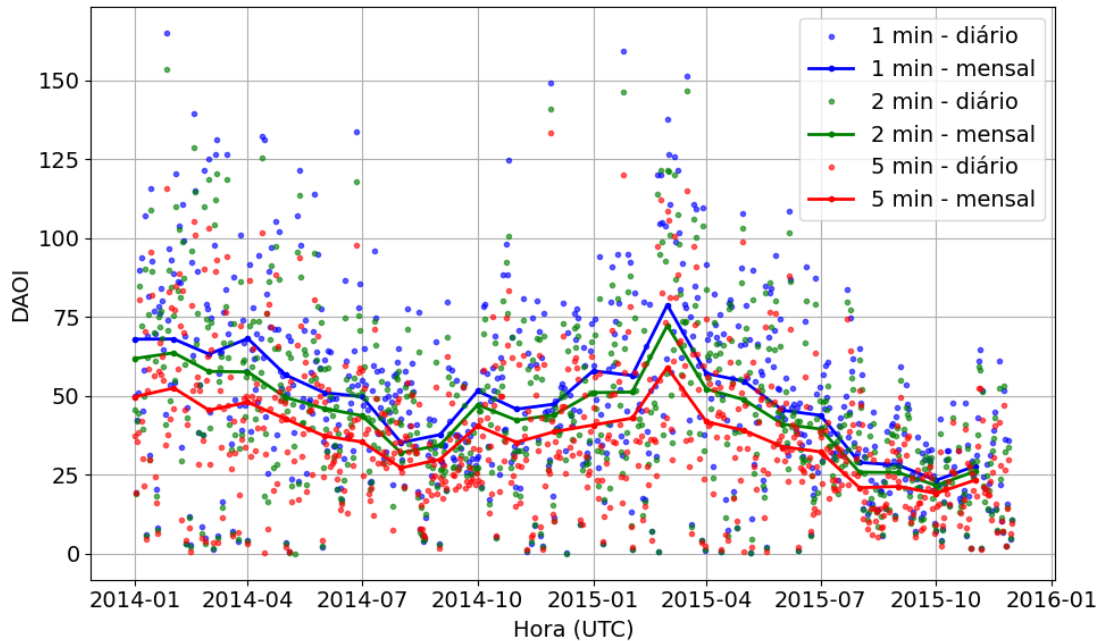
A Figura 3.8 apresenta os valores de DAOI, com pontos isolados representando os valores diários e pontos unidos com linhas indicando as médias mensais, utilizando o CSI como referência. Os valores são representados ao longo do tempo sem restrição de valores de GHI; quando considerados apenas os eventos de sobreirradiância com $GHI > 1000 \text{ W/m}^2$, os padrões gerais se mantêm, mas os valores de DAOI aumentam, pois a filtragem remove os eventos mais fracos - que, apesar de mais

numerosos, contribuem com excedentes menores e reduzem a média. No caso da utilização do TOA como referência, essa restrição chega a suprimir completamente os dias com sobreirradiância - consequência do TOA ser uma referência mais restritiva e já associada a poucos eventos. Essa redução deve-se principalmente à exclusão de diversas ocorrências de sobreirradiância em ângulos zenitais solares mais elevados (uma discussão mais detalhada é apresentada na análise das Figuras 3.10 e 3.11 e na Seção 3.6). A análise foi conduzida com dados médios em intervalos de 1, 2 e 5 minutos, sendo que resoluções temporais mais altas levam à detecção de um maior número de eventos de sobreirradiância.

A Figura 3.8 também permite uma avaliação sazonal da ocorrência de eventos de sobreirradiância em todas as resoluções temporais. Observa-se um declínio acentuado no número de eventos ao longo da estação seca (de maio a setembro), período que geralmente apresenta menor nebulosidade. Em termos de variação interanual, 2014 apresentou mais eventos de sobreirradiância do que 2015, especialmente entre os meses de agosto e novembro. Para esses meses, foram obtidos $\overline{DAOI} = 43 \pm 8$ para 2014 e $\overline{DAOI} = 27 \pm 3$ para 2015 - redução de 36%. Essa diferença tem significância estatística ao nível de 5%, mesmo calculando com DAOI diários ou mensais, com ou sem restrição de $GHI > 1000 \text{ W/m}^2$ ou mesmo selecionando todos os meses. Os meses de 2015 com maior diferença com relação ao ano anterior pertencem ao período em que o El Niño já estava bem estabelecido (CLIMATE PREDICTION CENTER, NCEP, NOAA, 2025). Com isso, existe uma intensificação do movimento de subsidência na parte central da Amazônia, que tende a suprimir a formação de nuvens, reduzindo assim a probabilidade de ocorrência de sobreirradiância.

Apesar dos valores de DAOI durante a estação seca de 2014 e 2015 apresentarem diferenças significativas, os valores medianos de GHI (Figura 3.2) não apresentaram diferença com significância estatística. Isso ocorre porque a sobreirradiância está mais relacionada à variabilidade de curta escala da irradiância e à estrutura das nuvens do que ao valor médio de GHI (PIACENTINI et al., 2011). Em particular, condições atmosféricas mais favoráveis à ocorrência de bordas de nuvens, múltiplos espalhamentos e passagens rápidas de nuvens convectivas podem aumentar a frequência de picos de irradiância momentânea (PIACENTINI et al., 2011), mesmo quando a média diária de GHI se mantém relativamente constante.

Figura 3.8 - DAOI calculado para dias e meses considerando o CSI como referência, em função do tempo, utilizando médias de 1, 2 e 5 minutos.



Fonte: Produção do autor.

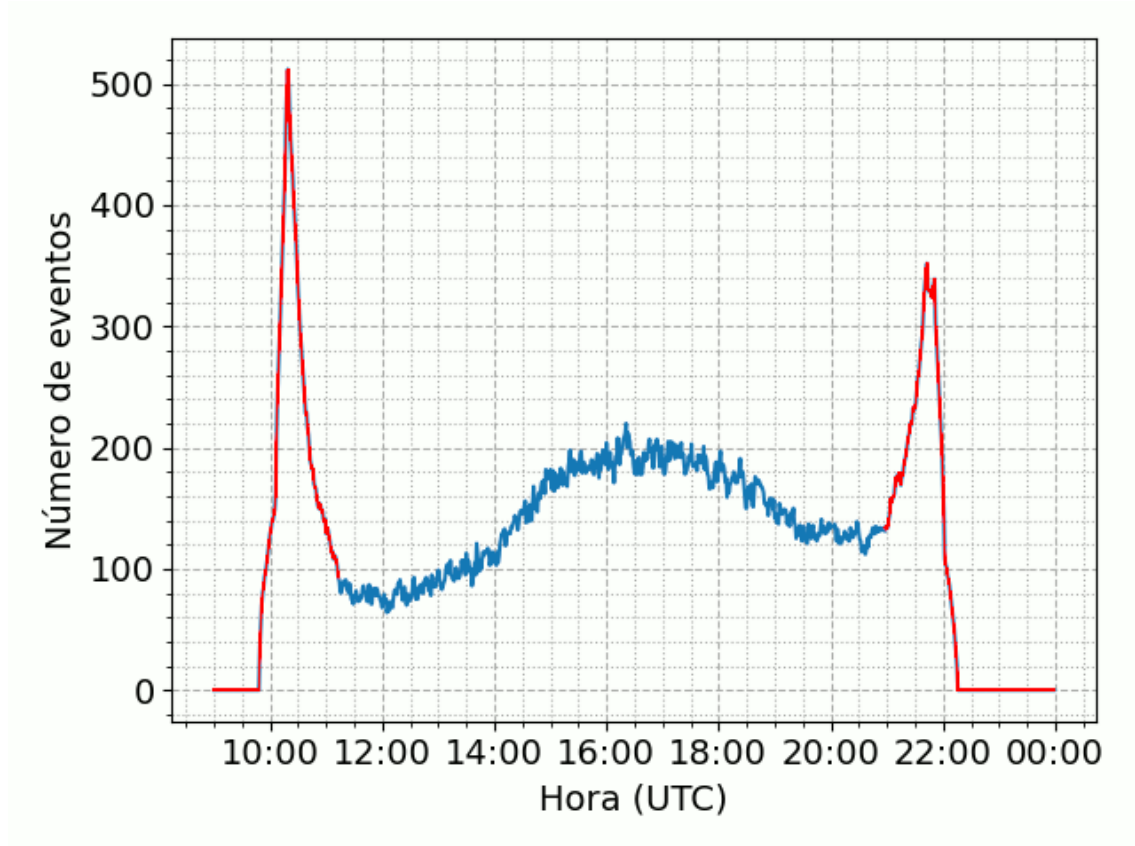
É igualmente importante analisar o comportamento da sobreirradiância ao longo do ciclo diurno - aqui simplificado pela consideração apenas da frequência de 1 minuto. A Figura 3.9 mostra o ciclo diurno do número total de eventos de sobreirradiância por minuto, utilizando o CSI como referência (o TOA apresenta padrão semelhante, mas com menos amostras). Um número elevado de eventos ocorre quando o Sol está mais próximo do horizonte (aproximadamente entre 10:00 e 10:30 UTC e entre 21:30 e 22:00 UTC), pois a irradiância de referência é menor, o que facilita que qualquer valor de GHI a ultrapasse. Nesses casos, mesmo um pequeno aumento de irradiância causado por reflexão em nuvens pode resultar em uma grande diferença relativa, embora não necessariamente significativa em termos de magnitude absoluta ou geração de energia solar. Além disso, observa-se que o pico matutino concentra um número maior de eventos de sobreirradiância (pouco mais de 500) em comparação ao pico vespertino (cerca de 340). Uma possível explicação para essa assimetria é que, nas primeiras horas da manhã, a presença de camadas horizontais de nevoeiro e nuvens do tipo *stratus* proporciona uma área mais extensa para o espalhamento da radiação solar incidente em ângulos rasantes, devido à posição baixa do Sol no horizonte. Essa configuração geométrica favorece uma maior interação entre a radiação solar e essas formações atmosféricas, potencializando a ocorrência de aumentos pontuais de

irradiância na superfície. Como resultado, seria mais provável a detecção de eventos de sobreirradiância no período próximo ao nascer do Sol do que ao entardecer. Contudo, a confirmação dessa hipótese requer uma análise mais detalhada, incluindo a inspeção visual das ASIs, o cruzamento com dados de tipos de nuvens fornecidos pelo produto ARSCL e a incorporação de informações sobre a radiação difusa.

Essas limitações podem afetar a detecção de eventos de sobreirradiância. No entanto, em ângulos zenitais solares maiores que 80° , o erro instrumental é dominado por viés negativo (subestimativa). Assim, a detecção de sobreirradiância nessas condições é conservadora: é improvável que o instrumento gere falsos picos positivos; ao contrário, a tendência é atenuá-los ou perdê-los. Dessa forma, optou-se por manter esses horários na análise para preservar a completude do ciclo diurno e para permitir uma descrição abrangente dos padrões de ocorrência. Essa decisão é especialmente relevante considerando que esses horários representam transições importantes na dinâmica atmosférica e na evolução da convecção e, portanto, podem contribuir com informações valiosas sobre os mecanismos associados à ocorrência de sobreirradiância, mesmo que com cautela na interpretação dos dados. Como existe o risco de que parte dos eventos de sobreirradiância registrados nesses horários seja resultado de artefatos instrumentais, e não de eventos reais, optou-se por destacar essa faixa da curva em vermelho no gráfico. Essa escolha visa segregar visualmente os dados e sinalizar que sua interpretação requer maior cautela.

Excluindo os eventos das primeiras horas da manhã e do final da tarde, observa-se um aumento progressivo no número de eventos ao longo do ciclo diurno, com predominância no período vespertino e mais eventos no final da tarde do que no início da manhã. Esse padrão segue o ciclo diurno de CF (Figura 3.4), o que reforça a influência direta da cobertura de nuvens na ocorrência desses eventos ao longo do dia.

Figura 3.9 - Ciclo diurno com a soma de eventos de sobreirradiância por minuto, utilizando o CSI como referência, e período de dados suspeitos em vermelho.

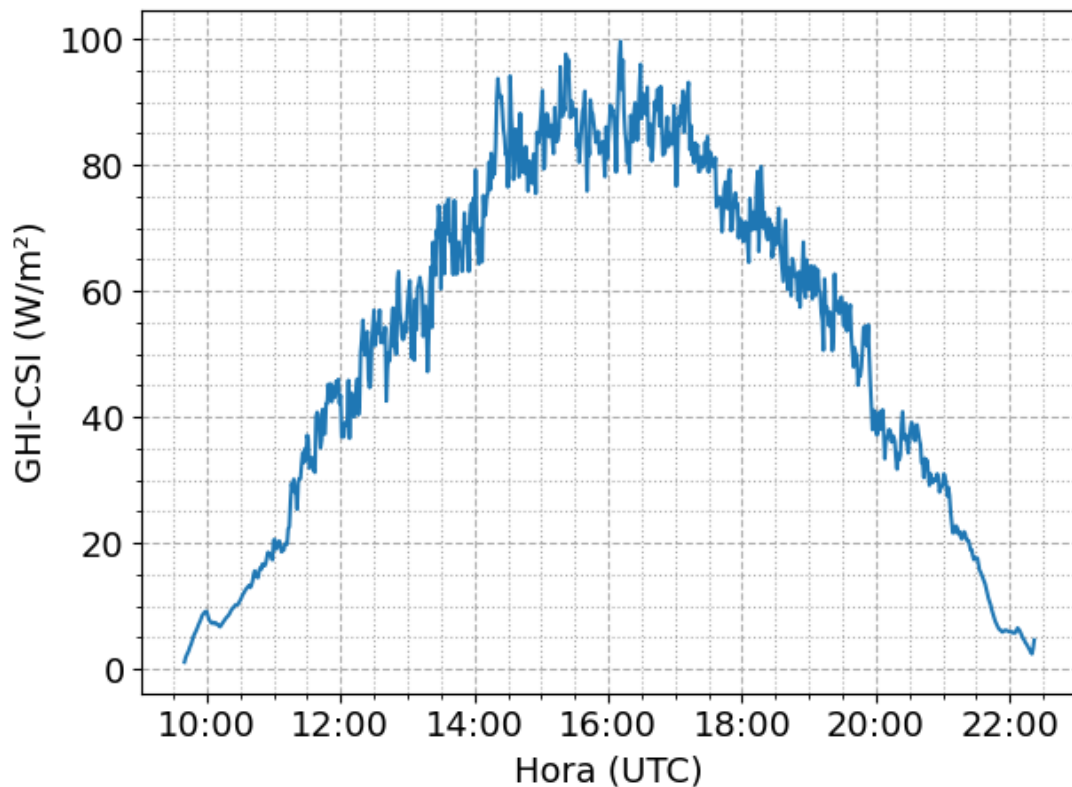


Fonte: Produção do autor.

A Figura 3.10 apresenta o ciclo diurno da magnitude da sobreirradiância, calculada como a diferença positiva entre GHI e CSI, com médias por minuto calculadas ao longo de todo o período de estudo. Em outras palavras, corresponde ao cálculo do DAOI para cada minuto do dia, utilizando todos os dados disponíveis. Observa-se um aumento gradual a partir das primeiras horas da manhã, com valores máximos entre 14:00 e 17:00 UTC (10:00-13:00 HL), seguido por uma redução progressiva em direção ao final da tarde. Uma explicação é que ângulos zenitais solares menores resultam em maior disponibilidade de radiação solar, o que amplifica a magnitude da sobreirradiância. Observa-se também forte similaridade entre esse ciclo e o correspondente de CF, mostrado na Figura 3.4, indicando que a ocorrência de sobreirradiância está fortemente associada à presença de nuvens. Uma explicação é apresentada em Machado et al. (2018) e resumida a seguir, em complemento ao que foi explicado na Seção 1.3.1. Durante a manhã, há um aumento gradual na quantidade de nuvens do tipo *cumulus* raso; à tarde, algumas dessas nuvens se desenvolvem

verticalmente, crescendo à custa dos inúmeros *cumulus* menores presentes anteriormente. Em outras palavras, a convecção profunda tende a estar associada a uma maior instabilidade da atmosfera (CAPE maiores e CINE menores) e a valores mais baixos de CF, apesar da presença de nuvens mais desenvolvidas. Isso ocorre porque a cobertura de nuvens é normalmente observada de forma bidimensional, de modo que nuvens isoladas com grande desenvolvimento vertical podem ocupar uma fração menor do céu visível do que uma camada de nuvens rasas mais espalhadas. Em contraste, a estação chuvosa é caracterizada por uma convecção rasa mais disseminada, resultando em valores mais altos de CF.

Figura 3.10 - Ciclo diurno da magnitude da sobreirradiância, representado pelas médias por minuto das diferenças positivas GHI–CSI (DAOI).



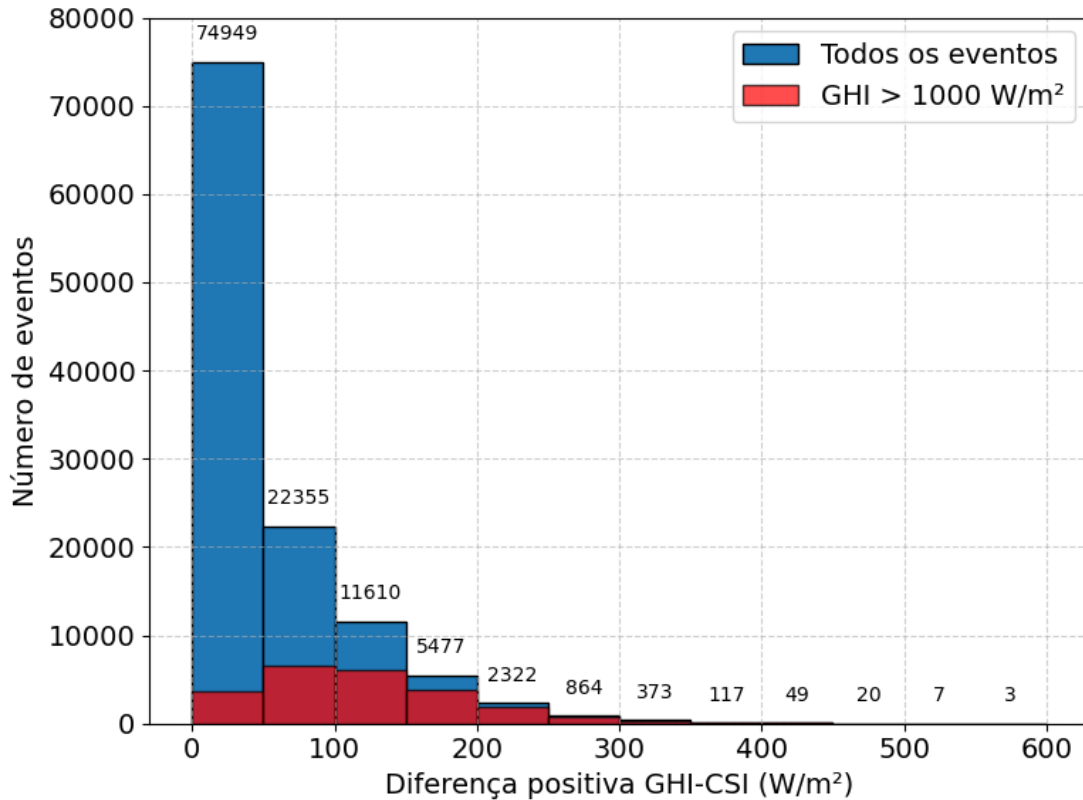
Fonte: Produção do autor.

A distribuição das magnitudes dos eventos de sobreirradiância também deve ser examinada. Para esse fim, a Figura 3.11 apresenta um histograma das diferenças entre GHI e CSI durante todos os eventos de sobreirradiância detectados. O histograma revela a presença de alguns poucos eventos extremos com diferenças superiores a 400

W/m^2 . Esses casos podem estar associados a configurações atmosféricas específicas, como o alinhamento favorável entre o Sol e bordas espessas de nuvens. A grande maioria dos casos está concentrada em diferenças pequenas, com uma frequência decrescente à medida que a magnitude aumenta. Esse comportamento é esperado, uma vez que muitos desses eventos de baixa magnitude ocorrem durante períodos de alto ângulo zenital solar (como visto na Figura 3.9), quando a irradiância de céu limpo é baixa, e até mesmo pequenas flutuações atmosféricas (como nuvens finas ou reflexões) podem gerar excedências modestas em relação à curva de referência.

Ao restringir essa análise apenas aos eventos com GHI superior a $1000 \text{ W}/\text{m}^2$ (barras vermelhas na mesma Figura 3.11), o número total de casos é significativamente menor. A distribuição também se altera de forma perceptível: as ocorrências no primeiro intervalo (diferenças menores) são substancialmente reduzidas, e os valores remanescentes passam a se distribuir de forma mais uniforme entre os intervalos intermediários, formando uma curva semelhante a uma distribuição normal. Essa mudança reflete a exclusão dos períodos de baixa irradiância da análise, permanecendo apenas os eventos que ocorrem próximos ao meio-dia solar, quando valores mais altos de irradiância (e, portanto, maiores potenciais de diferença entre GHI e CSI) são mais prováveis. Isso sugere que, quando o GHI é elevado, os eventos de sobreirradiância tendem a ser relativamente mais intensos e suas magnitudes distribuídas de uma forma aproximadamente simétrica.

Figura 3.11 - Histograma das magnitudes de sobreirradiância sem restrição (azul, com valores indicados sobre cada barra) e com restrição de $GHI > 1000 \text{ W/m}^2$ (vermelho), com CSI como referência.

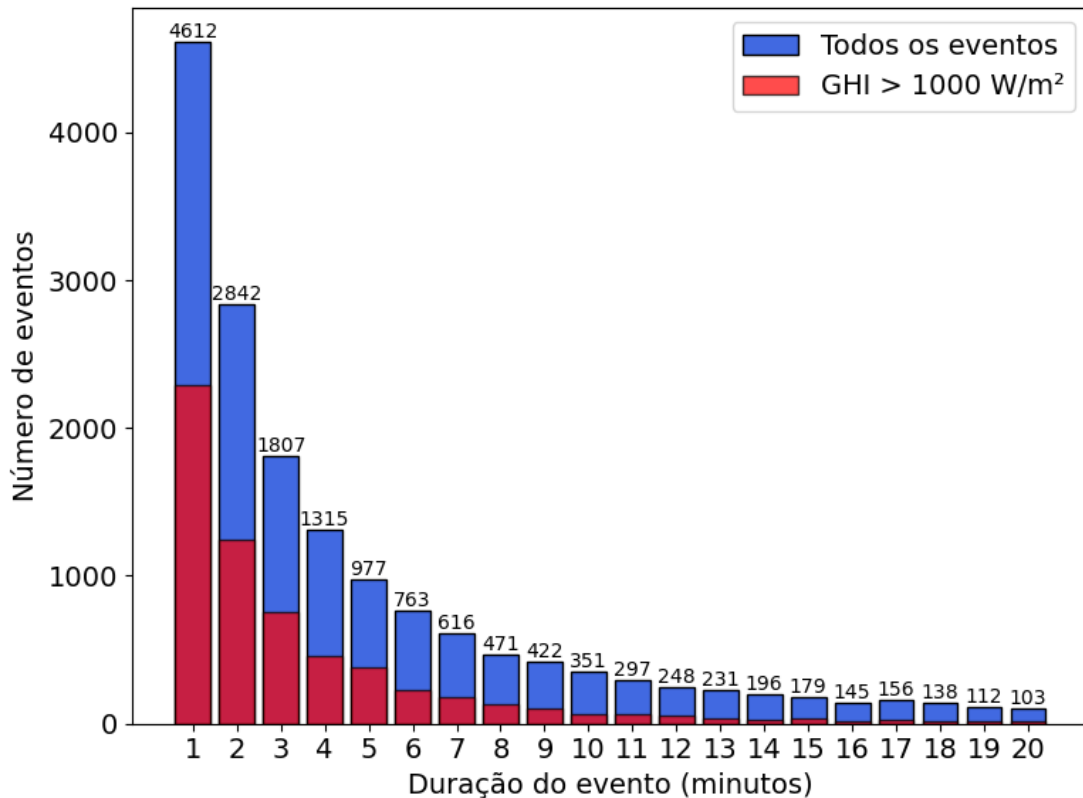


Fonte: Produção do autor.

Eventos de sobreirradiância podem durar vários minutos. A Figura 3.12 apresenta a duração dos episódios de sobreirradiância, definidos como um ou mais minutos consecutivos com irradiância excedendo o valor de referência (CSI, embora os mesmos padrões sejam observados com o TOA). Como esperado em fenômenos transitórios, observa-se que a frequência dos eventos de sobreirradiância decresce com o aumento de sua duração. Como esses eventos dependem de mudanças rápidas na cobertura de nuvens e no ângulo solar, eventos mais longos exigiriam uma configuração persistente e estável de nuvens refletoras, o que se torna progressivamente menos provável com o passar do tempo. Os eventos mais longos (acima de 20 minutos) geralmente se iniciam ao redor dos horários próximos ao nascer e pôr do sol, períodos em que a diferença entre GHI e CSI é relativamente pequena; quando se restringe para os casos em que GHI é maior que 1000 W/m^2 , os eventos basicamente só ocorrem entre 14:00 e 18:00 UTC.

Utilizando a irradiância do TOA como referência, foram identificados 927 eventos de sobreirradiância, dos quais 120 tiveram duração entre 2 e 5 minutos (13%) e 14 duraram 5 minutos ou mais (1,5%). Braga et al. (2020) relataram, para os mesmos intervalos de duração, 12% e 0,17% para a estação de Florianópolis (litoral sul do Brasil) e 6% e 0,19% para Brotas de Macaúbas (região semiárida do nordeste brasileiro). Comparando esses resultados, observa-se que o sítio amazônico deste estudo apresentou uma frequência mais elevada de eventos longos (≥ 5 min) de sobreirradiância, além de uma maior proporção de eventos de duração moderada (2–5 min). Brotas de Macaúbas apresenta, em geral, céus mais limpos, onde a sobreirradiância é rara e tende a ser de curta duração. Já o local de estudo, situado na Amazônia central, apresenta um ambiente com alta variabilidade na cobertura de nuvens convectivas, elevada concentração de vapor d'água e dinâmicas intensas de movimentação vertical, o que favorece tanto a ocorrência quanto a persistência de eventos de reflexão solar em múltiplas direções.

Figura 3.12 - Histograma com a distribuição do número de eventos de sobreirradiância conforme sua duração, em minutos, utilizando o CSI como referência, sem (azul) e com restrição para eventos com $GHI > 1000 \text{ W/m}^2$ (vermelho).



Fonte: Produção do autor.

Os maiores valores de GHI com frequência de 1 minuto foram selecionados (Tabela 3.1). Alguns desses eventos duraram vários minutos consecutivos, mas seus valores médios seriam menores — o que justifica, em parte, a adoção da mais alta resolução temporal disponível. Ainda que os valores ultrapassem a constante solar (1361 W/m^2), essa comparação não é direta, pois GHI inclui radiação difusa e pode ser intensificada por efeitos atmosféricos e reflexivos. Portanto, os valores registrados são compatíveis com a literatura e fisicamente plausíveis. Em [Martins et al. \(2022b\)](#), é apresentada uma tabela com medições extremas de sobreirradiância reportadas na literatura, incluindo informações relevantes sobre as condições de medição, como tempo de resposta dos instrumentos, resolução e localidade. A lista reúne registros obtidos em diversos regimes climáticos ao redor do mundo, incluindo regiões tropicais. Um dos exemplos é Ouro Preto do Oeste (Rondônia), onde uma campanha realizada em janeiro e fevereiro de 1999 registrou aproximadamente 1400 W/m^2 com resolução de 1 minuto, utilizando um piranômetro Eppley. Os valores registrados no atual estudo estariam entre a 15^a e a 25^a posição. No entanto, as frequências de medição nesses estudos são geralmente inferiores a 30 segundos, enquanto os valores aqui relatados correspondem a médias de 1 minuto, sugerindo que a irradiância instantânea pode ter sido ainda mais elevada do que a registrada.

Alguns eventos de sobreirradiância foram observados sob alta fração de cobertura de nuvens ($CF \geq 90\%$), incluindo casos com até 99%. A inspeção visual das imagens do céu total correspondente confirmou a presença de ampla cobertura de nuvens. No entanto, elementos estruturais como a carcaça da câmera, suporte e banda de sombra podem obstruir parcialmente a visualização de pequenas aberturas nas nuvens alinhadas com o disco solar. Essas aberturas, embora não totalmente captadas pelo sistema de imagem, podem permitir a transmissão momentânea do feixe direto até o sensor que mede a GHI. Assim, mesmo sob condições predominantemente encobertas, o movimento dinâmico do campo de nuvens pode favorecer intermitentemente a ocorrência de sobreirradiância.

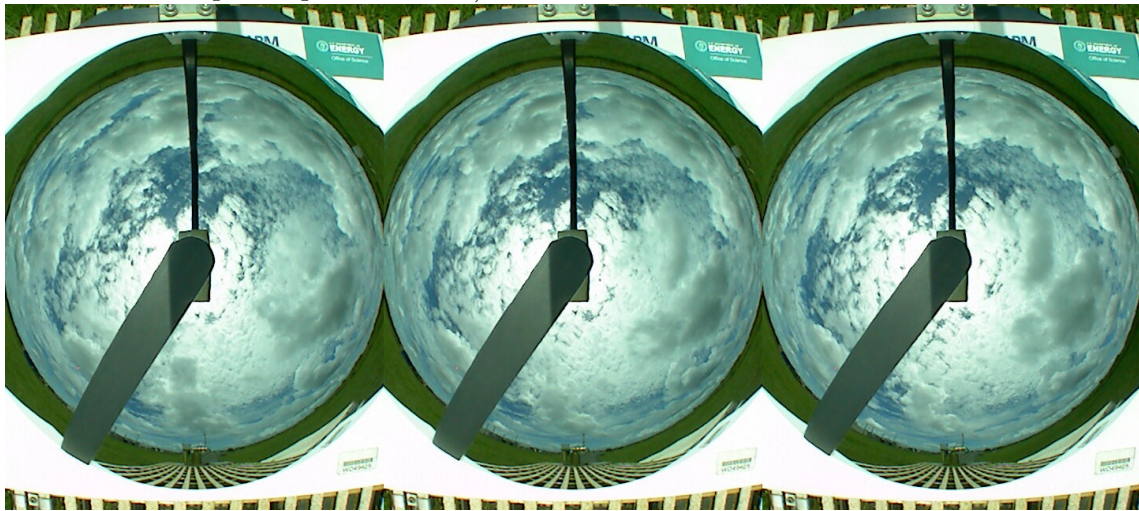
Tabela 3.1 - Eventos com maior excesso percentual de GHI em relação ao modelo de céu claro: data (ano, mês, dia, hora UTC), valor de GHI, valor de CF e "aumento de GHI", que representa o quanto o GHI observado excedeu a irradiância estimada para céus claros, em porcentagem.

#	Data/Horário	GHI (W/m ²)	CF (%)	Aumento de GHI (%)
1	2015-03-02 16:28:00	1590	94	55
2	2015-03-02 16:27:00	1561	94	52
3	2015-03-02 16:29:00	1556	93	52
4	2015-03-24 15:38:00	1497	–	49
5	2015-03-16 15:24:00	1496	92	52
6	2015-03-07 16:13:00	1488	75	46
7	2015-03-02 16:26:00	1477	94	44
8	2015-03-07 16:12:00	1470	74	44
9	2015-03-07 16:10:00	1461	75	43
10	2015-02-23 15:35:00	1458	99	44
11	2014-03-07 15:38:00	1458	–	45
12	2015-03-07 16:04:00	1456	77	43
13	2015-03-07 16:07:00	1454	77	43
14	2015-03-07 16:08:00	1452	76	42
15	2015-03-07 16:05:00	1451	77	42
16	2015-03-07 16:06:00	1450	77	42
17	2015-03-07 16:03:00	1449	78	42
18	2015-03-07 16:11:00	1445	75	42
19	2015-03-07 16:09:00	1444	76	42
20	2015-03-13 16:10:00	1441	99	42

Fonte: Produção do autor.

A maioria dos valores extremos de GHI ocorreu em março de 2015, ano de El Niño, sob condições de alta nebulosidade e próximo ao meio-dia solar. O mês de março marca o final da estação chuvosa, o que contribui para a limpeza da atmosfera em relação aos aerossóis, reduzindo, assim, a extinção da radiação solar causada por esses particulados. Os três principais eventos ocorreram no mesmo dia, em sequência, e suas respectivas imagens do céu total estão apresentadas na Figura 3.13 — essas imagens, obtidas em intervalos de 1 minuto, permitem visualizar a evolução típica do campo de nuvens. O céu está predominantemente coberto por nuvens *altocumulus*, especialmente próximas ao zênite e ao redor do sol, enquanto nuvens cúmulos são observadas em ângulos zenitais mais altos.

Figura 3.13 - Imagens do céu total correspondentes aos três maiores eventos de sobreirradiação em 2 de março de 2015, nos horários 16:27, 16:28 e 16:29 UTC (da esquerda para a direita).



Fonte: Produção do autor.

3.4 Caracterização dos tipos de nuvens

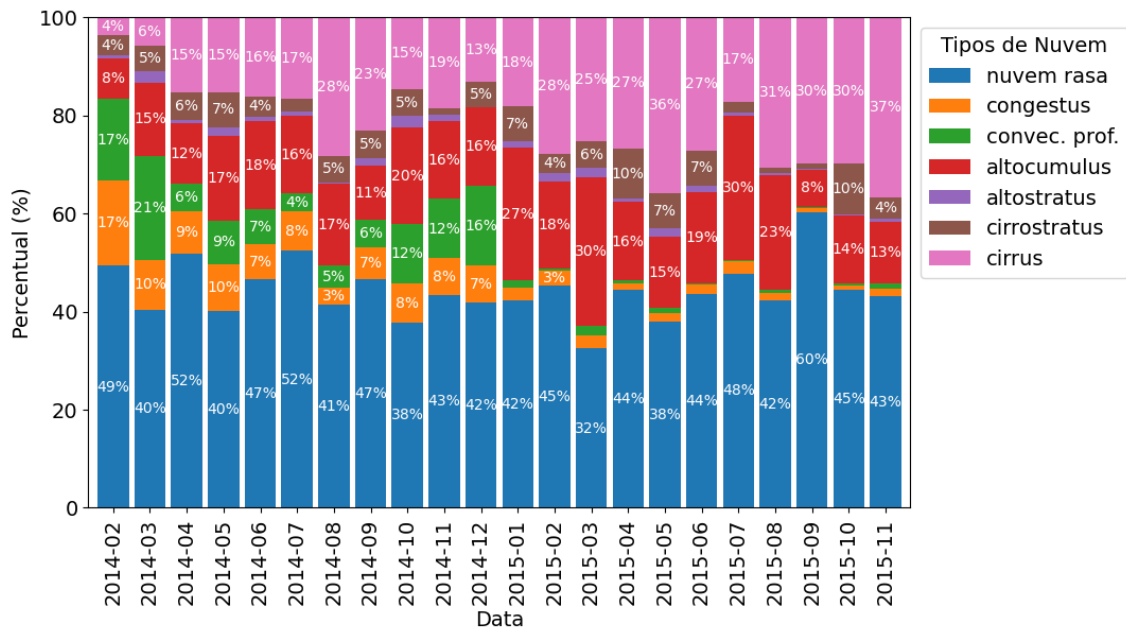
A análise dos tipos de nuvem foi realizada com base nos dados do produto ARSCL, que fornece perfis verticais da cobertura de nuvens. Como os dados são restritos ao ponto zenital, eles representam exclusivamente as nuvens localizadas diretamente acima do sensor, desconsiderando aquelas inclinadas no horizonte. Essa limitação pode introduzir vieses importantes na análise, especialmente em estudos de irradiância solar, já que nuvens periféricas ao disco solar (frequentemente responsáveis por efeitos ópticos associados a eventos de sobreirradiação) não são corretamente representadas pelo perfilamento vertical. A ocorrência de múltiplas camadas de nuvem (*overlap*) não foi considerada neste estudo, mas constitui uma extensão relevante para investigações futuras, dado o seu potencial impacto na modulação radiativa. Tal análise poderia ser conduzida não apenas com o produto ARSCL, que fornece perfis verticais, mas também com classificações derivadas ASIs, as quais permitem observar a cobertura além do ponto zenital e sob diferentes geometrias solares.

A Figura 3.14 apresenta a distribuição percentual dos principais tipos de nuvem observados mensalmente ao longo do período de estudo. Embora não tenha sido identificado um padrão marcante no ciclo diurno do tipo de nuvem predominante em cada dia (não apresentado), a análise mensal revela uma redução significativa na

ocorrência de nuvens verticalmente desenvolvidas (como *cumulus congestus* e nuvens de convecção profunda) em 2015 em relação a 2014. Essa diferença foi particularmente marcada na estação seca, em consonância com os resultados de [Giangrande et al. \(2017\)](#). Essa mudança abrupta pode estar relacionada à configuração do fenômeno El Niño, ainda que, nos primeiros meses de 2015, o evento não estivesse plenamente estabelecido, aliada à forçante de larga escala da convecção na estação chuvosa, o que sugere a necessidade de uma análise mais aprofundada.

Observou-se um aumento na ocorrência de nuvens altas e médias, como *cirrus* e *altocumulus*, sobretudo nos meses mais secos de 2015, acompanhado de uma redução expressiva na frequência de nuvens verticalmente desenvolvidas, como *cumulus congestus* e nuvens de convecção profunda, em comparação com o ano anterior. Essa mudança no padrão de nebulosidade é compatível com um cenário de subsidência de ar seco em altos níveis, frequentemente associado a episódios de El Niño. Como discutido na Seção 1.4, o aquecimento anômalo das águas do Pacífico leste altera a célula de Walker, favorecendo movimento descendente sobre a região amazônica. Isso inibe o desenvolvimento de nuvens convectivas profundas e contribui para a redução da nebulosidade observada neste período. Além disso, outro fator que pode ter contribuído para essa alteração na formação de nuvens foi o aumento significativo da concentração de aerossóis nos meses de outubro e novembro de 2015, quando foram registrados valores de profundidade óptica de aerossóis (AOD) entre 0,4 e 1,0. A presença de aerossóis pode afetar direta e indiretamente os processos de formação e desenvolvimento das nuvens: por um lado, ao espalharem e absorverem a radiação solar, reduzem o aquecimento da superfície e, conseqüentemente, enfraquecem a convecção; por outro, aerossóis absorvedores, como a fuligem resultante da queima de biomassa, podem aquecer camadas atmosféricas superiores, induzindo inversões térmicas que dificultam o crescimento vertical das nuvens. Essa interpretação é compatível com os resultados de [Öktem et al. \(2023\)](#), que, ao reavaliar dados da campanha GoAmazon com critérios objetivos, não encontraram evidências de que aerossóis aumentem a intensidade da convecção em nuvens de fase quente, contrariando a hipótese de fortalecimento e reforçando o papel inibitório da poluição sobre o desenvolvimento vertical das nuvens. Essas alterações nas condições atmosféricas, incluindo a influência do ciclo ENSO e dos aerossóis, ajudam a explicar a menor frequência de nuvens profundas observada ao longo do segundo semestre de 2015.

Figura 3.14 - Distribuição percentual mensal dos principais tipos de nuvem observados através do produto ARSCL.



Fonte: Produção do autor.

Foi também estudado o ciclo diurno da classificação de nuvens, verificando-se o tipo mais frequente de nuvem para cada hora do dia - como o padrão observado foi simples e esperado, optou-se por apresentar apenas a descrição textual. Entre 09:00 e 19:00 UTC, observa-se a predominância de nuvens baixas, inicialmente nevoeiros e nuvens *stratus* com o posterior desenvolvimento típico da convecção diurna na região. Já entre 20:00 e 22:00 UTC, o tipo de nuvem mais frequente passa a ser o *cirrus*. Essas observações concordam com as realizadas por [Giangrande et al. \(2017\)](#) para todos os dados do GoAmazon como a nuvem mais baixa na coluna. Nas primeiras horas do dia, logo após o nascer do sol, predominam nuvens baixas como *stratus* ou nevoeiros, resultantes do resfriamento noturno da camada próxima à superfície, com camada limite noturna tipicamente entre 200 e 250 metros ([CARNEIRO; FISCH, 2020](#)). Posteriormente, ocorre o ciclo convectivo típico da Amazônia Central: durante o dia, o aquecimento solar da superfície úmida intensifica a convecção, formando nuvens baixas do tipo *cumulus*, com o topo da Camada Limite Convectiva entre 1200 e 1400 metros. No fim do dia, com o enfraquecimento da convecção e a dissipação das nuvens mais espessas, restam cristais de gelo nos altos níveis da atmosfera, favorecendo a formação de *cirrus*.

A Tabela 3.2 apresenta a distribuição percentual dos tipos de nuvens sob três condições distintas: todos os registros com classificação disponível, casos de sobreirradiância e, por fim, casos de sobreirradiância com GHI superior a 1000 W/m^2 . A coluna "Total" reflete o comportamento médio observado ao longo do período de estudo, com predominância de nuvens baixas, seguidas por *cirrus* e *altocumulus* - padrão coerente com o ciclo convectivo típico da Amazônia Central.

Ao restringir os dados apenas aos instantes com ocorrência de sobreirradiância, observa-se uma amplificação das desigualdades entre os tipos de nuvem: nuvens baixas, *altocumulus* e *cirrus* tornam-se mais frequentes, enquanto *congestus* e nuvens profundas têm sua participação reduzida. Essa assimetria sugere que certos tipos de nuvem são mais propensos a gerar sobreirradiância. Nuvens baixas do tipo *cumulus*, por exemplo, apresentam bordas bem definidas que favorecem o espalhamento e a focalização da luz solar, amplificando a irradiância ao nível do solo. Já as nuvens altas do tipo *cirrus*, compostas por cristais de gelo, podem atuar como lentes ópticas, concentrando a luz solar em determinados ângulos — embora com menor eficiência (BALFOUR, 2011).

Quando se aplica o critério adicional de $\text{GHI} > 1000 \text{ W/m}^2$, a dominância das nuvens baixas torna-se ainda mais evidente, representando 78% dos casos, ao passo que todos os demais tipos reduzem significativamente sua contribuição. Esse recorte extremo impõe uma condição de irradiância muito elevada, que só pode ser atingida sob condições ópticas muito específicas - geralmente associadas a céu parcialmente nublado com nuvens convectivas baixas próximas ao disco solar. Nessas situações, a geometria solar favorece a ocorrência de efeitos de refração e reflexão direcionada, mais prováveis com nuvens próximas à superfície, devido ao ângulo de incidência mais favorável.

É importante destacar que a participação reduzida de *congestus* e nuvens profundas nos eventos de sobreirradiância pode estar relacionada a uma limitação do método de observação, uma vez que os tipos de nuvem são identificados com base no perfilamento vertical no zênite. Assim, eventos gerados por nuvens inclinadas ou periféricas ao disco solar podem não ser capturados adequadamente, mesmo tendo potencial para provocar aumentos intensos na irradiância medida. Ainda assim, os resultados obtidos reforçam a importância das nuvens baixas no processo de geração de sobreirradiância, sobretudo em condições de alta irradiância direta.

Tabela 3.2 - Distribuição percentual dos tipos de nuvem em três condições distintas: todos os dados com classificação de nuvem, apenas casos com sobreirradiância (OI), e sobreirradiância com GHI superior a 1000 W/m² (OI c /GHI>1000), em porcentagens.

Tipo de Nuvem	Total	OI	OI c/ GHI>1000
1 - Nuvem baixa	44	50	78
2 - <i>Congestus</i>	6	2	1
3 - Nuvem profunda	7	0	1
4 - <i>Alto cumulus</i>	17	21	12
5 - <i>Altostratus</i>	1	1	0
6 - <i>Cirrostratus</i>	5	2	1
7 - <i>Cirrus</i>	20	25	7

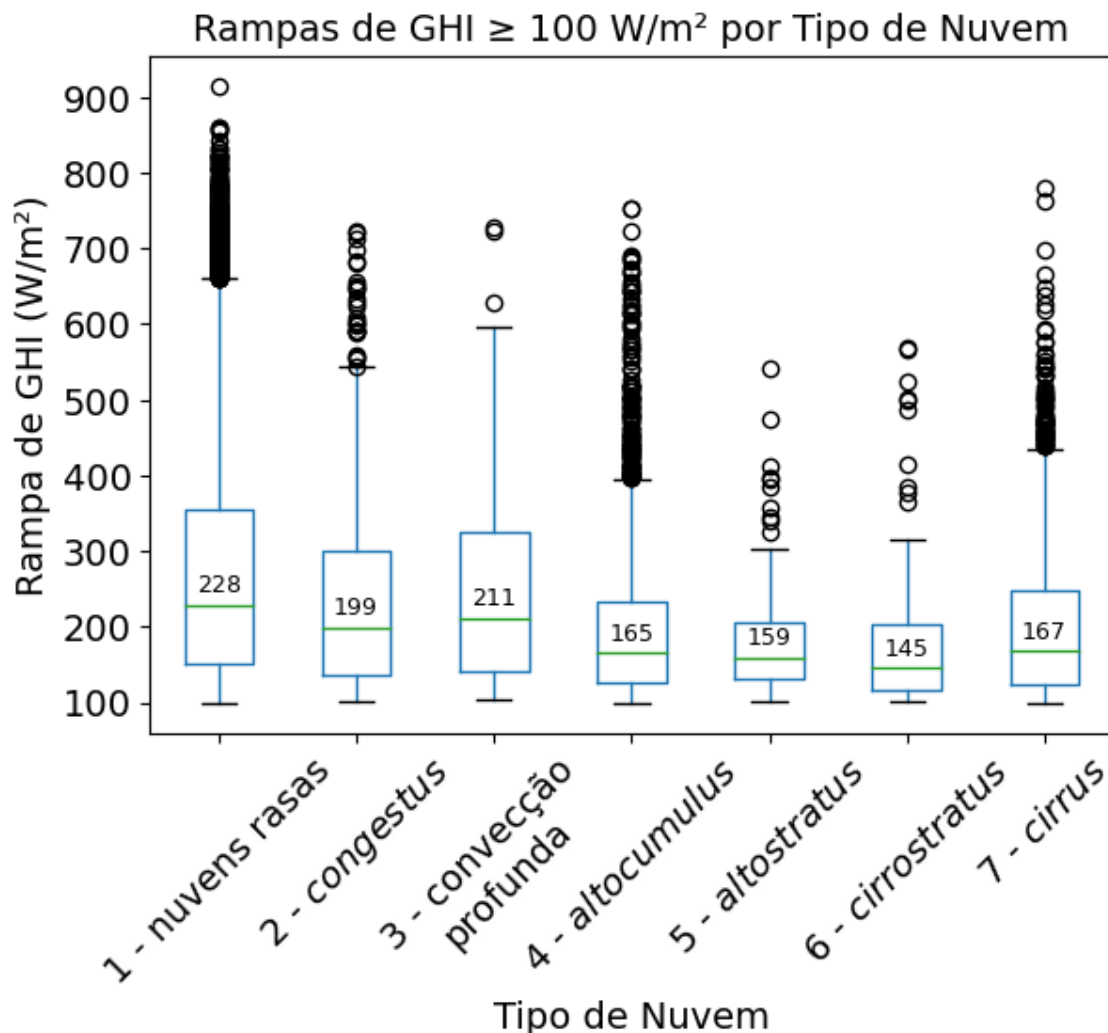
Fonte: Produção do autor.

Para investigar a relação entre os tipos de nuvem e a intensidade das rampas de irradiância, foi realizado um recorte dos dados considerando apenas os casos em que a variação absoluta entre dois instantes consecutivos de GHI foi maior ou igual a 100 W/m². Essa escolha se justifica pelo fato de que aproximadamente 88% das rampas observadas apresentam módulo inferior a esse limiar, de modo que a análise das rampas mais intensas permite concentrar a avaliação nas variações de maior impacto para aplicações energéticas e operacionais, já que oscilações dessa magnitude podem comprometer a estabilidade de sistemas solares em tempo real, exigindo maior capacidade de resposta. A Figura 3.15 apresenta a distribuição dessas rampas para cada tipo de nuvem, com base em classificações obtidas por perfilamento vertical.

A presença marcante de *outliers* nas distribuições de rampas de GHI associadas a tipos como nuvens rasas (número de amostras n = 21.391), *altocumulus* (n = 2.634), *cirrus* (n = 1.430) e *congestus* (n = 573) sugere que essas formações estão relacionadas a uma maior ocorrência de rampas abruptas de irradiância, possivelmente em razão de sua estrutura fragmentada, curta duração ou elevada dinâmica vertical. Nuvens mais baixas (categorias 1 a 3) tendem a apresentar rampas com maior magnitude e maior variabilidade, com medianas situadas pouco acima de 200 W/m². Isso se deve à sua maior espessura óptica e proximidade da superfície, o que provoca sombreamentos mais abruptos à medida que se deslocam sobre o ponto de medição. Nuvens baixas e *congestus* são particularmente associadas a rampas intensas, dada sua opacidade e dinâmica atmosférica.

Por outro lado, tipos como convecção profunda ($n = 70$), *altostratus* ($n = 93$) e *cirrostratus* ($n = 127$) apresentaram menor número de amostras e menos *outliers*, o que pode refletir tanto características físicas mais homogêneas e persistentes (que inibem transições rápidas de irradiância) quanto o efeito estatístico da amostragem reduzida. Essas análises indicam a importância de considerar não apenas o tipo de nuvem, mas também a variabilidade intra-classe e o papel de fatores como a geometria solar e a dinâmica local do sistema nuvem-radiação.

Figura 3.15 - Distribuição das rampas de irradiância solar global (GHI) com módulo maior ou igual a 100 W/m^2 para cada tipo de nuvem, com os valores das medianas escritos nas respectivas caixas.



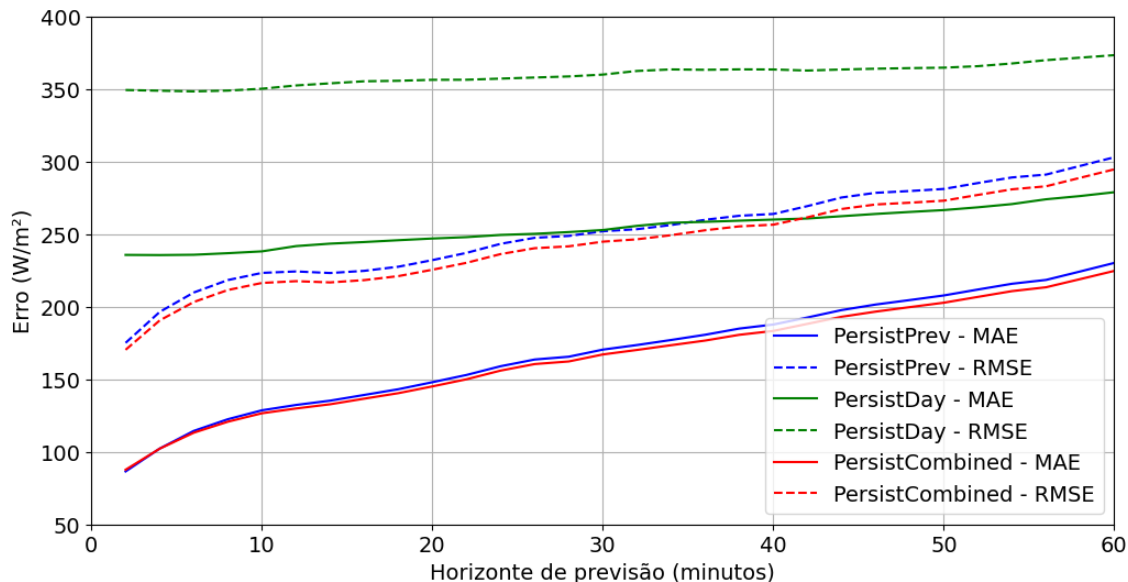
Fonte: Produção do autor.

3.5 Avaliação de desempenho dos modelos de previsão

A Figura 3.16 apresenta a evolução dos erros de previsão (MAE e RMSE) com resolução de 2 minutos para cada um dos modelos de persistência: baseado no instante anterior (PersistPrev), no mesmo horário do dia anterior (PersistDay) e em uma combinação sazonal (PersistCombined). A série temporal contínua dos primeiros 15 dias de abril de 2014 foi utilizada para esse cálculo. O parâmetro α do modelo de persistência combinada foi ajustado com base na minimização do erro entre os valores observados e estimados de GHI, considerando duas métricas: $\alpha_{MAE} = 0,991$ e $\alpha_{RMSE} = 0,925$. Mesmo com uma diferença pequena entre os dois, optou-se por adotar este último como valor final; o α_{MAE} foi testado e produziu resultado semelhante, mas não foi adotado como valor final para manter a coerência da análise, assim optando-se pelo que utiliza o valor que minimiza o RMSE. Ambos os valores de α indicam uma forte predominância da persistência de curtíssimo prazo.

Embora o erro aumente com o horizonte, como é característico em previsões de séries temporais, a taxa de degradação varia entre os modelos avaliados, evidenciando diferenças na capacidade de generalização. O modelo PersistCombined apresentou os menores valores de MAE (entre 88 e 225 W/m²) e de RMSE (entre 170 e 295 W/m²) ao longo de todo o horizonte de previsão, superando ligeiramente o modelo PersistPrev (MAE entre 87 e 230 W/m² e RMSE entre 175 e 303 W/m²). As melhorias médias foram de aproximadamente 2% para o MAE e 3% para o RMSE. Embora discretas, essas vantagens indicam que a inclusão do padrão diurno — por meio da persistência de 24 horas — contribui para mitigar os erros na previsão de curtíssimo prazo, especialmente à medida que o horizonte se alonga. Essa diferença, no entanto, permanece pequena, possivelmente porque, em regiões tropicais altamente convectivas como a Amazônia, as condições atmosféricas tendem a se "reinicializar" a cada dia, o que limita a utilidade da persistência baseada no dia anterior. Ainda assim, em situações em que as características atmosféricas se mantêm de um dia para o outro, a persistência pode oferecer ganhos. Por outro lado, eventos de precipitação intensa podem alterar significativamente a umidade do solo e, conseqüentemente, impactar o desenvolvimento convectivo no dia seguinte (CARNEIRO; FISCH, 2020).

Figura 3.16 - Evolução dos erros (MAE, linha contínua, e RMSE, linha tracejada) das previsões com frequência de 2 minutos para cada um dos modelos de persistência: baseado no instante anterior (PersistPrev), no mesmo horário do dia anterior (PersistDay) e combinação sazonal (PersistCombined).



Fonte: Produção do autor.

As métricas MAE e RMSE (com e sem normalização) foram calculadas para cada passo de previsão até o horizonte total, possibilitando a comparação entre os modelos desenvolvidos neste estudo e aqueles de trabalhos anteriores, brevemente descritos na introdução (Figuras 3.17 e 3.18). Entre os modelos propostos, aqueles com restrições associadas aos objetivos no MOGA (ghi30cr e ghicf30cr) apresentaram melhor desempenho nos primeiros 10 minutos: RMSE de 104 para 163 W/m^2 (ghi30cr, aumento de 59 W/m^2) e de 103 para 162 W/m^2 (ghicf30cr, aumento de 59 W/m^2). No entanto, seus erros aumentaram mais rapidamente com o tempo de previsão quando comparados aos modelos sem restrições: RMSE de 142 para 161 W/m^2 (ghi30sr, aumento de 19 W/m^2) e de 155 para 158 W/m^2 (ghicf30sr, aumento de 3 W/m^2). O aumento do erro é esperado, uma vez que as restrições auxiliam a guiar o processo de busca em direção a soluções que não apenas minimizem o erro, mas também atendam a critérios adicionais relevantes para o problema, produzindo modelos potencialmente mais úteis e confiáveis.

A partir do horizonte de 10 minutos, o melhor desempenho foi observado para o modelo ghicf30sr, gerado sem restrições de busca. Uma possível explicação é que as restrições impostas aos demais modelos podem ter levado o MOGA a priorizar

soluções altamente especializadas no curtíssimo prazo, minimizando os erros nos primeiros passos da previsão. No entanto, essa escolha parece ter ocorrido em detrimento da estabilidade para horizontes de previsão mais longos.

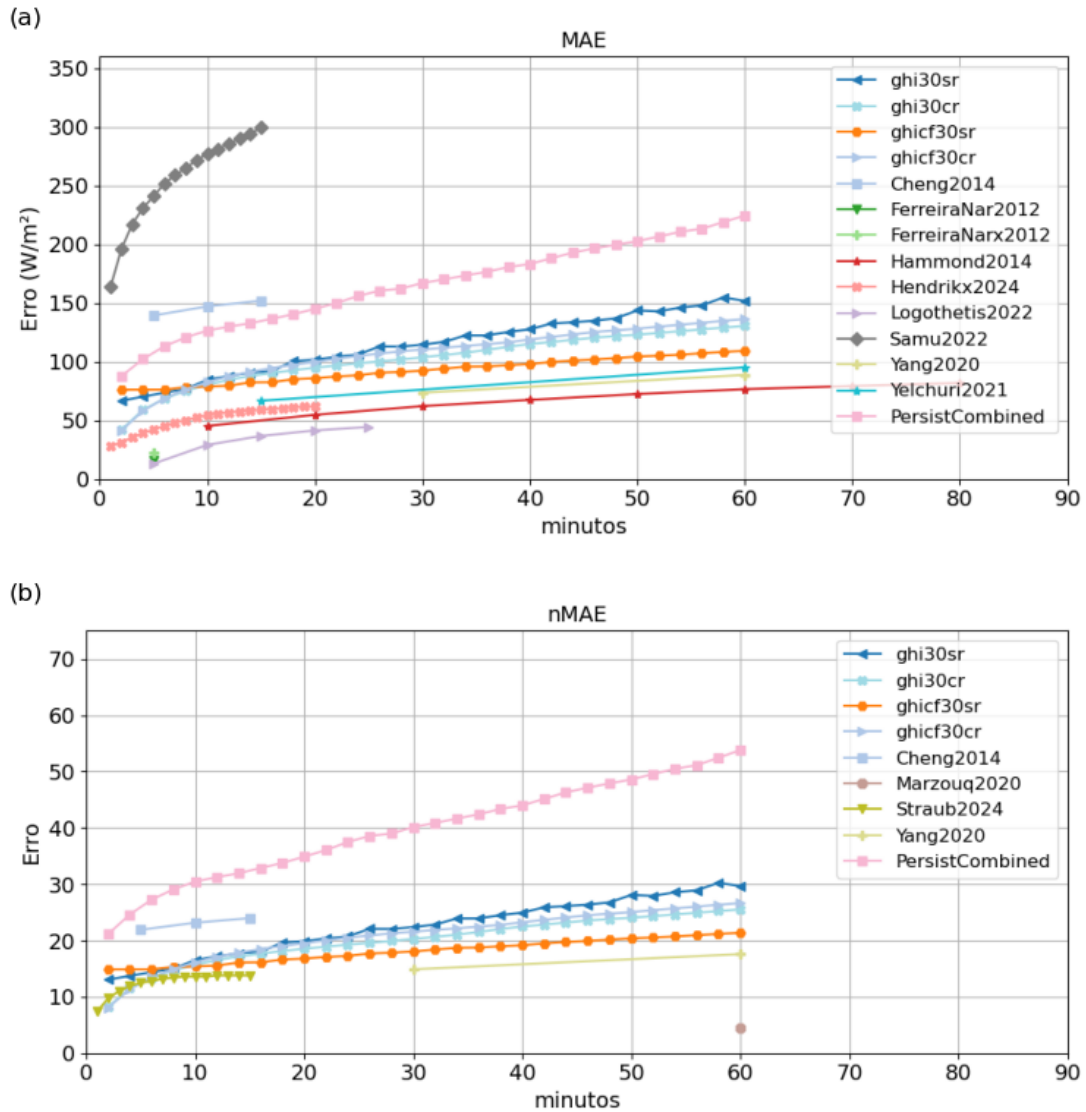
Por outro lado, o modelo sem restrições teve liberdade para explorar uma gama mais ampla de compromissos (*trade-offs*) durante a otimização, o que pode ter resultado em soluções mais equilibradas ao longo do tempo. Em outras palavras, as restrições podem ter conduzido o MOGA a uma abordagem mais conservadora, com menor flexibilidade e maior risco de propagação de erros conforme o tempo avança.

Esse padrão é evidenciado pela análise da evolução do erro ao longo do tempo. Por exemplo, entre os minutos 10 e 60, o RMSE do modelo ghicf30cr (com restrições) aumentou de 162 para 229 W/m^2 — um crescimento de 67 W/m^2 , ou 41%. Já no modelo ghicf30sr, o aumento foi de apenas 36 W/m^2 (de 158 para 194 W/m^2), ou 23%. Isso representa uma taxa média de crescimento do erro quase duas vezes maior no modelo com restrições: 1,34 vs. 0,72 $\text{W}/\text{m}^2/\text{min}$.

Além disso, o modelo com restrições apresentou desempenho inicial significativamente melhor: aos 2 minutos, seu RMSE foi 33% menor do que o do modelo sem restrições. No entanto, essa vantagem se dissipou rapidamente — em apenas 10 minutos, o erro cresceu quase 20 vezes mais no modelo ghicf30cr, indicando uma priorização excessiva do ajuste inicial, com prejuízo à robustez para previsões mais longas. Isso reforça a ideia de que pequenas imprecisões iniciais em modelos muito rígidos podem se propagar rapidamente e comprometer a performance futura.

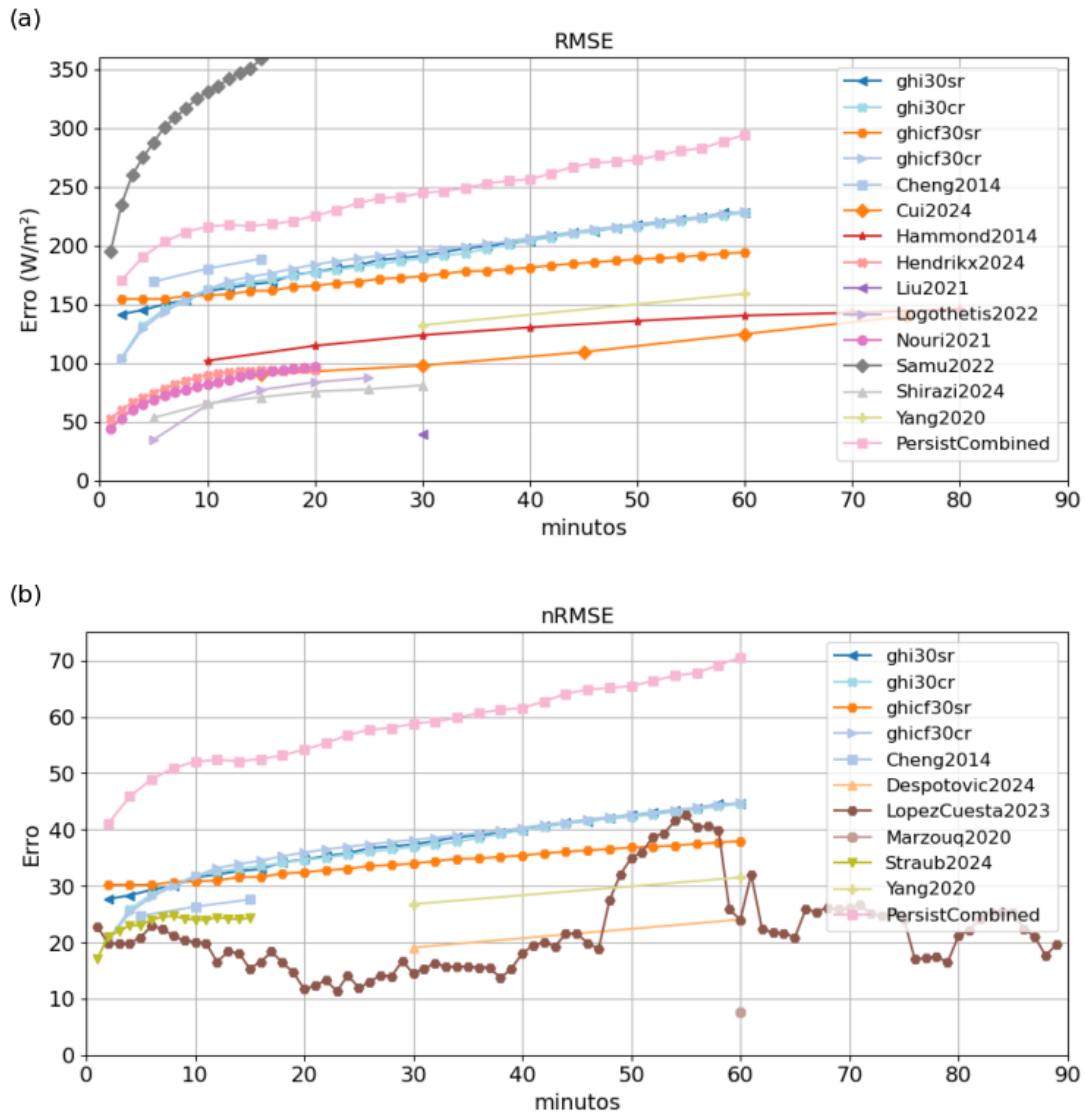
Por fim, o modelo ghicf30sr parece ter alcançado um ponto de equilíbrio vantajoso entre precisão inicial e estabilidade. Embora seu RMSE aos 10 minutos (158 W/m^2) seja apenas 4 W/m^2 menor que o do modelo com restrições, ele manteve uma evolução mais estável do erro, com um crescimento significativamente mais contido ao longo do tempo de previsão.

Figura 3.17 - Comparação do erro absoluto médio (MAE, a) e do erro absoluto médio normalizado (nMAE, b) ao longo dos passos de previsão considerados neste estudo. São apresentados os casos definidos na Tabela 2.4, incluindo o modelo de persistência (PersistPrev), bem como resultados de outros estudos (identificados pelo sobrenome do primeiro autor e ano de publicação).



Fonte: Produção do autor.

Figura 3.18 - Comparação do erro quadrático médio (RMSE, a) e do erro quadrático médio normalizado (nRMSE, b) ao longo dos passos de previsão considerados neste estudo. São apresentados os casos definidos na Tabela 2.4, incluindo o modelo de persistência (PersistPrev) e resultados de outros estudos (identificados pelo sobrenome do primeiro autor e ano de publicação). Fonte: elaboração própria.



Fonte: Produção do autor.

Para resumir o desempenho do modelo ghicf30sr em horizontes de previsão de curtíssimo prazo, a Tabela 3.3 apresenta os valores médios de MAE e RMSE (em termos absolutos e normalizados), calculados cumulativamente para os passos de previsão de até 10, 20 e 30 minutos à frente. No intervalo de até 10 minutos, por exemplo, o modelo alcançou um RMSE médio de $138 W/m^2$, o que representa uma redução

de 32,5% em relação ao valor correspondente da linha de base da persistência (205 W/m²), indicando ganhos relevantes nos primeiros instantes de previsão. Para fins de comparação, também são incluídas as métricas referentes a todo o período de previsão. Os resultados confirmam o aumento dos erros conforme o avanço do horizonte, embora o desempenho permaneça dentro de limites aceitáveis para aplicações que exigem resolução temporal em escala de minutos.

Tabela 3.3 - Desempenho médio do modelo ao longo dos horizontes de previsão considerados, com erros absolutos (MAE) e quadráticos médios (RMSE), tanto em valores absolutos (W/m²) quanto normalizados (nMAE, nRMSE).

Intervalo	MAE	nMAE	RMSE	nRMSE
Até 10 min	65	13	138	27
Até 20 min	80	16	158	31
Até 30 min	89	17	169	33
Todo o período	107	21	191	37

Fonte: Produção do autor.

Uma consideração importante refere-se à consistência física do cenário atmosférico captado pela câmera *all sky* e à sua validade temporal. Assumindo uma velocidade média do vento na base das nuvens (em torno de 1.300-1.400 m) de aproximadamente 2,5-3,0 m/s (valor típico para condições de tempo calmo), a cobertura de nuvens pode deslocar-se cerca de 10.000 m em uma hora. Considerando uma câmera com campo de visão de 140° (ângulo zenital de 70°) e altura da base das nuvens de 1.300 m, o raio projetado da área observada no solo é de aproximadamente 3.570 m em relação ao ponto de observação. Isso implica que, após cerca de 20-25 minutos, o campo de visão da câmera pode ter sido alterado devido à movimentação das nuvens. Devido tanto à advecção das nuvens quanto à rápida evolução local da convecção, a coerência da cena atmosférica observada tende a se degradar após esse intervalo. No entanto, ainda há a possibilidade de que as nuvens que chegam mantenham uma fração de cobertura semelhante. A metodologia atual não prevê CF, mas em aplicações futuras que incluam a previsão da fração de cobertura de nuvens a partir das imagens, esse limite temporal deve ser considerado para evitar previsões fisicamente inconsistentes de CF - uma proposta sugerida na seção de trabalhos futuros.

Embora o uso de valores estimados de CF para cada instante futuro represente uma

vantagem (equivalente a uma previsão perfeita da variável exógena), trata-se de um cenário idealizado, pouco provável de ser replicado em aplicações reais. Portanto, o bom desempenho observado nos primeiros 30 minutos deve ser interpretado como um limite superior otimista, e análises mais detalhadas podem ser razoavelmente restringidas a esse intervalo, no qual o estado atmosférico permanece relativamente coerente com o momento da observação.

Na comparação com outros estudos, embora o método proposto não alcance o melhor desempenho absoluto, destaca-se pelo uso de um conjunto muito mais limitado de informações de entrada — restrito à irradiância de superfície e à fração de cobertura de nuvens estimada a partir de uma única câmera *all sky*. Observa-se que os métodos com os melhores resultados geralmente dependem de conjuntos de dados significativamente mais amplos — incluindo imagens de satélite, saídas de modelos numéricos de previsão do tempo ou múltiplas medições em superfície. A configuração atual simula intencionalmente um cenário de entrada mínima e baixo custo, e, dentro dessa limitação, os resultados permanecem competitivos, especialmente nos primeiros 30 minutos de previsão.

Quanto à previsão de rampas, os testes preliminares indicaram que o modelo tem potencial para identificar eventos intensos, mas ainda apresenta limitações significativas. Como os resultados obtidos foram considerados insatisfatórios e não passaram por análise mais aprofundada, optou-se por não apresentá-los nesta etapa. No entanto, o tema será retomado em trabalhos futuros, com a incorporação de estratégias como a proposta por [Nouri et al. \(2024\)](#), que destaca a importância da seleção de limiares adequados para otimizar a detecção e avaliação de rampas de irradiância.

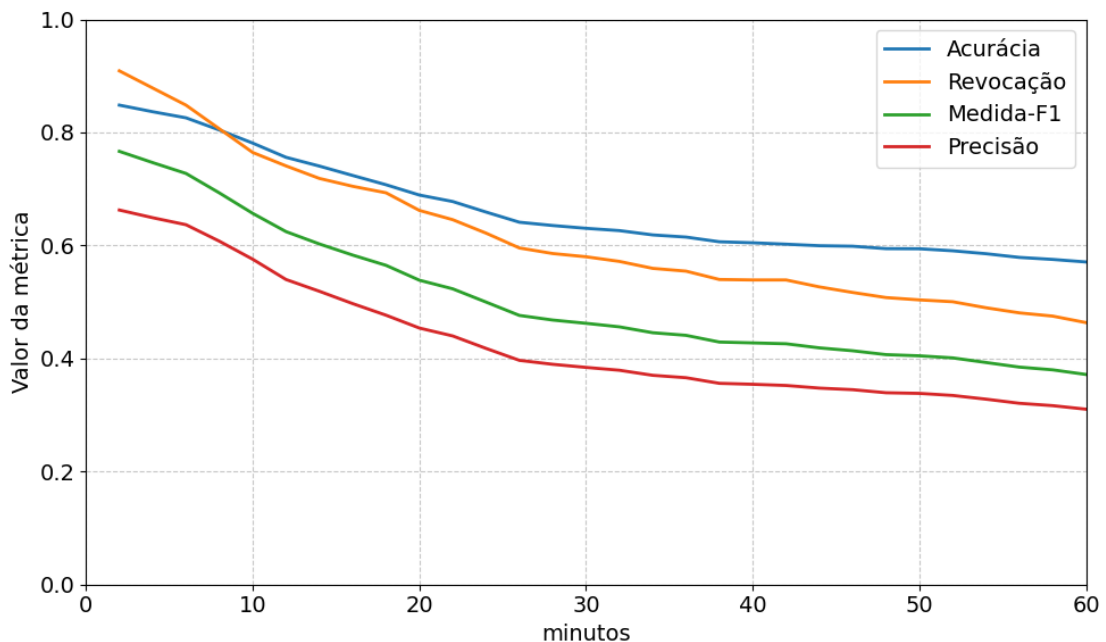
3.6 Detecção de eventos de sobreirradiância com base nas previsões

O modelo ghicf30cr foi selecionado para avaliação de sua capacidade de prever eventos de sobreirradiância. Essa avaliação foi realizada apenas durante os períodos em que o ângulo zenital solar era inferior a 80° . Um dos motivos para essa restrição é que o limiar utilizado (CSI) torna-se muito baixo logo após o nascer do sol e antes do pôr do sol, sendo facilmente ultrapassado. Nesses casos, se a irradiância global horizontal (GHI) prevista for ligeiramente superior ao baixo valor do CSI (mesmo na ausência de um evento físico de sobreirradiância), o algoritmo pode classificá-lo incorretamente como um evento detectado. Contudo, como o GHI real permanece baixo, isso não corresponde a um evento observado, resultando em um falso positivo. Além disso, ruídos e erros instrumentais tornam-se mais significativos em baixos níveis de irradiância, comprometendo ainda mais a confiabilidade da detecção de eventos

nesses horários.

Os valores de VP, VN, FP e FN (definidos na Seção 2.3.1) foram obtidos por meio da comparação entre os valores observados e modelados de GHI em cada instante dentro do período de previsão analisado, e as respectivas métricas de desempenho foram calculadas (Figura 3.19). Os resultados indicam que o desempenho geral do modelo diminui à medida que o horizonte de previsão aumenta. Esse comportamento é esperado devido à perda natural de informação sobre o estado atmosférico futuro, o que leva a maior incerteza, propagação de erros e redução da acurácia na detecção de eventos de sobreirradiância. Como consequência, há aumentos simultâneos nos FN e FP, resultando em quedas consistentes nas métricas de acurácia, precisão, revocação e medida-F1. A Tabela 3.4 apresenta o desempenho médio do modelo na detecção de eventos de sobreirradiância através dessas métricas, considerando diferentes janelas de previsão. Os resultados foram calculados cumulativamente para os horizontes de até 10, 20 e 30 minutos, bem como para todo o período de previsão. Assim, até cerca de 20-30 minutos, ainda há, neste estudo, um compromisso aceitável entre acurácia e sensibilidade.

Figura 3.19 - Métricas de desempenho da detecção de eventos de sobreirradiância pelo modelo ghicf30cr para diferentes passos de previsão.



Fonte: Produção do autor.

Tabela 3.4 - Desempenho médio do modelo na detecção de eventos de sobreirradiância ao longo de diferentes horizontes de previsão.

Intervalo	Acurácia	Precisão	Revocação	Medida-F1
Até 10 min	0,82	0,63	0,84	0,72
Até 20 min	0,77	0,56	0,77	0,65
Até 30 min	0,73	0,51	0,72	0,60
Total	0,66	0,43	0,62	0,50

Fonte: Produção do autor.

Uma análise mais detalhada foi realizada considerando a previsão mais próxima de cada evento, ou seja, aquela feita 2 minutos antes do instante avaliado. O modelo atinge acurácia geral de 0,85, considerando tanto instantes com e sem sobreirradiância. Isso indica um bom desempenho geral; no entanto, como a sobreirradiância é um evento relativamente raro, a acurácia sozinha pode ser enviesada pelo alto número de casos de falsos negativos, ressaltando a importância de considerar também outras métricas de desempenho, como a medida-F1, a revocação e a precisão. A medida-F1 de 0,77, por ser calculada como média harmônica entre precisão e revocação (Seção 2.3.1), indica uma performance robusta do modelo, já que esse tipo de média penaliza discrepâncias acentuadas entre as duas métricas. Isso sugere que, embora a precisão ainda possa ser aprimorada, o modelo consegue manter um bom equilíbrio ao detectar corretamente a maior parte dos eventos sem gerar um número excessivo de falsos alarmes. A revocação de 0,91 reflete uma alta proporção de eventos corretamente identificados em relação ao total de eventos reais, implicando baixa taxa de falsos negativos. Considerando aplicações nas quais a falha em detectar um evento relevante é mais crítica do que o risco de emitir alarmes falsos, como detectar um evento de sobreirradiância, essa é uma característica especialmente desejável. A precisão, por sua vez, indica que 66% dos eventos de sobreirradiância previstos pelo modelo foram corretos, refletindo uma taxa moderada de falsos positivos e um sistema de alerta razoavelmente confiável.

A detecção de eventos também foi avaliada em uma base diária, com o objetivo de investigar a relação entre as condições de nebulosidade e o desempenho geral do modelo (Tabela 3.5). Novamente, foi utilizada a previsão realizada 2 minutos antes do instante avaliado. Em dias com maior número de eventos de sobreirradiância ($N \geq 100$), como 3, 4, 5, 6 e 15 de abril de 2014, o modelo apresenta desempenho médio superior, com precisão de aproximadamente 0,75, revocação de 0,94 e medida-

F1 de 0,83. Esses valores são consideravelmente mais altos do que os obtidos nos demais dias, sugerindo que o modelo tende a funcionar melhor sob condições com ocorrência frequente de eventos. Por outro lado, em dias com poucos eventos de sobreirradiância (como 2, 7, 8, 10 e 12 de abril de 2014, com $N < 60$), o desempenho do modelo é consideravelmente inferior, especialmente em termos de precisão, cuja média fica em torno de 0,46. Nesse contexto, o impacto dos falsos positivos torna-se proporcionalmente mais relevante, resultando também em medida-F1 mais baixa ($\approx 0,59$), apesar de a revocação ainda se manter relativamente alta ($\approx 0,85$). Em dias encobertos (como 9 e 14 de abril de 2014), caracterizados por baixa variabilidade e menor quantidade de nuvens do tipo *cumulus* (vide Figura 3.20), a ocorrência de eventos de sobreirradiância é praticamente nula. Nesses casos, o modelo torna-se conceitualmente pouco útil, pois não há eventos a serem detectados, e sua aplicação pode inclusive resultar em falsos positivos, o que compromete a relevância prática da previsão.

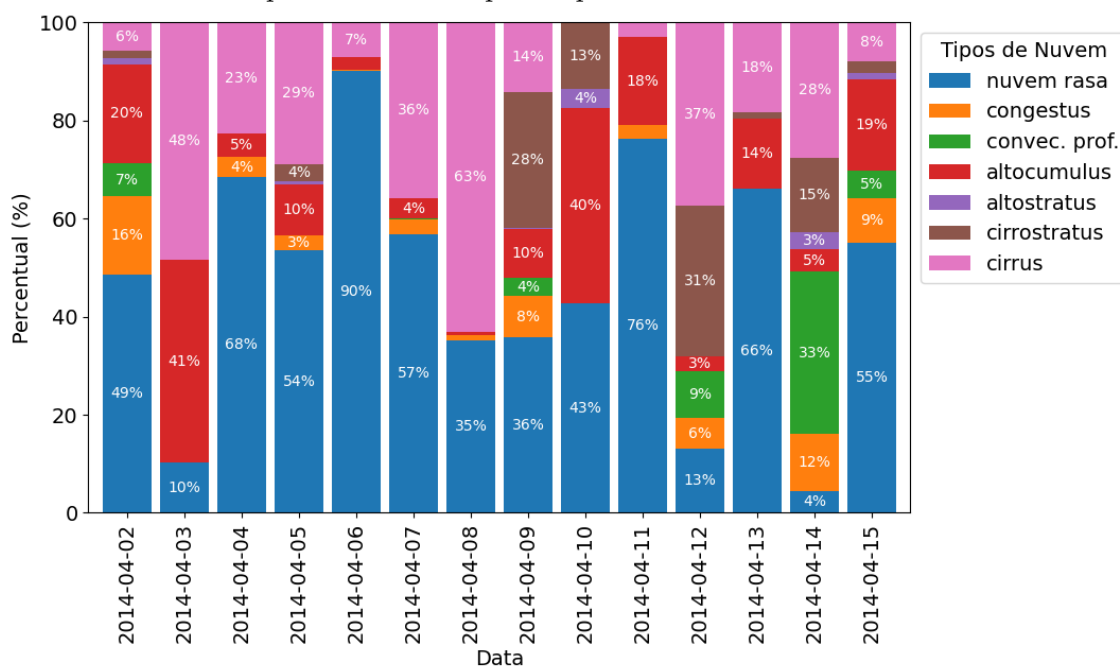
Excluindo esses dias com acurácia artificialmente alta (dias 9 e 14 de abril de 2014), foi calculada a correlação entre os valores de DARR e as respectivas acurácias, para verificar a influência da variabilidade sobre a detecção de eventos de sobreirradiância do modelo. Foi obtido o valor de $r = -0,65$ e com significância estatística, pois $p = 0,0218$ ($p < 0,05$). Isso pode ser interpretado como uma evidência moderada de que uma maior variabilidade diária (DARR) prejudica a acurácia do modelo. No entanto, essa relação não é fortemente determinística. Outros fatores também contribuem para a variabilidade da acurácia, como o número e a distribuição temporal dos eventos, a complexidade da cobertura de nuvens (incluindo casos de multicamadas), a qualidade e disponibilidade dos dados de entrada, além de possíveis vieses do próprio modelo. Assim, a acurácia diária reflete a interação de múltiplas variáveis atmosféricas e operacionais, não se explicando apenas pela variabilidade intradiária da irradiância.

Tabela 3.5 - Métricas de desempenho na detecção de eventos pelo modelo escolhido para cada dia, valores médios de GHI (W/m^2) e CF (%), número de eventos de sobreirradiância (N) e DARR.

Data	Acurácia	Precisão	Revocação	Medida-F1	GHI	CF	N	DARR
2014-04-02	0,915	0,49	0,923	0,64	154	95	26	18
2014-04-03	0,806	0,742	0,977	0,843	275	44	171	22
2014-04-04	0,849	0,778	0,93	0,847	232	62	143	40
2014-04-05	0,84	0,849	0,937	0,891	280	57	222	46
2014-04-06	0,733	0,716	0,891	0,794	265	42	184	62
2014-04-07	0,739	0,374	0,796	0,509	223	75	54	43
2014-04-08	0,805	0,454	0,942	0,613	228	69	52	32
2014-04-09	0,987	0,571	0,8	0,667	94	90	5	7
2014-04-10	0,905	0,435	0,833	0,571	165	81	24	19
2014-04-11	0,792	0,559	0,877	0,683	228	65	81	42
2014-04-12	0,887	0,536	0,75	0,625	167	73	40	32
2014-04-13	0,855	0,694	0,852	0,765	207	62	88	40
2014-04-14	0,994	0,0	0,0	0,0	95	91	1	9
2014-04-15	0,77	0,649	0,938	0,767	229	52	128	32

Fonte: Produção do autor.

Figura 3.20 - Distribuição percentual diária dos principais tipos de nuvem observados através do produto ARSCL para o período entre 2 e 15 de abril de 2014.



Fonte: Produção do autor.

Antes de definir a modelagem final apresentada na Tabela 3.4, foi testado um critério adicional de filtragem dos dados, no qual apenas os instantes com GHI observado acima de 1000 W/m^2 foram considerados - abordagem esta previamente utilizada em outras etapas deste estudo. Essa filtragem teve como efeito positivo a redução expressiva dos falsos positivos. No entanto, também resultou em uma queda substancial no tamanho da amostra e impôs um viés temporal, concentrando a análise em eventos com irradiância muito elevada, próximos ao meio-dia. Dado esse comprometimento na representatividade, optou-se por não adotar esse critério na modelagem final. Ainda assim, o relato desse teste é incluído aqui para evitar dúvidas sobre essa possibilidade metodológica e para reforçar que essa estratégia foi considerada, mas se mostrou inadequada neste contexto.

4 CONCLUSÕES

Inicialmente, foram realizadas análises descritivas das séries temporais de GHI, CF, variabilidade de GHI, ocorrência de eventos de sobreirradiância e tipos de nuvens, com o objetivo de caracterizar o comportamento temporal do sistema atmosférico e fornecer subsídios à etapa de modelagem.

As distribuições diárias e mensais de GHI e CF indicam, respectivamente, um aumento da irradiância e uma redução da nebulosidade durante a estação seca. Diferenças interanuais entre 2014 e 2015 foram identificadas, embora não estatisticamente significativas, sugerindo uma possível influência de fenômenos climáticos como o El Niño sobre a cobertura de nuvens, especialmente nos meses de seca. Durante a estação seca de 2015, o aumento dos focos de queimada elevou os valores de AOD, o que ajudou a explicar a ausência de diferença significativa na irradiância média entre 2014 e 2015, apesar da menor cobertura de nuvens. A análise do ciclo diurno médio revelou que a GHI apresenta maior variabilidade em torno do meio-dia solar, enquanto a CF segue um padrão convectivo típico da região, com aumento progressivo durante a manhã, pico à tarde e posterior declínio. A divergência observada entre GHI e CSI nas horas centrais do dia, associada a altos valores de CF, confirma o papel das nuvens na modulação da irradiância superficial.

A variabilidade de GHI, quantificada pelo índice DARR, está positivamente associada à ocorrência de eventos de sobreirradiância, com correlação geral de 0,63. Essa relação é mais forte durante a estação chuvosa ($r = 0,67$) do que na seca ($r = 0,58$), indicando que a estrutura fragmentada das nuvens e as condições ópticas presentes nesse período (estação chuvosa) favorecem a ocorrência de flutuações rápidas capazes de gerar sobreirradiância. A análise sazonal do DARR indicou maior variabilidade na estação seca, particularmente em julho, o que pode ser atribuído à combinação de alta irradiância direta com nuvens isoladas do tipo *cumulus*. A maior variabilidade intradiária e a concentração de rampas intensas ao redor do meio-dia, especialmente com assimetria vespertina, sugerem que esse é o período mais suscetível à ocorrência de oscilações rápidas e intensas. Além disso, o fato de os valores médios de DARR diminuírem com o aumento da janela temporal confirma que a variabilidade de curtíssimo prazo é relevante para a caracterização da instabilidade atmosférica.

A análise dos eventos de sobreirradiância mostrou maior frequência durante a estação úmida e nos períodos de maior atividade convectiva, especialmente à tarde. Em 2015, observou-se uma redução expressiva desses eventos, atribuída à intensi-

ficção do El Niño, que inibe a formação de nuvens favoráveis à reflexão solar, e aos aerossóis, que dificultaram o desenvolvimento vertical das nuvens. A magnitude da sobreirradiância apresenta um ciclo diurno bem definido, com máximos entre o final da manhã e o início da tarde, coincidindo com o aumento da irradiância solar e o desenvolvimento convectivo das nuvens, refletindo forte dependência do ciclo da cobertura de nuvens. A frequência de ocorrência ao longo do dia também segue esse padrão, embora influências instrumentais nos horários de baixa irradiância — como início da manhã e final da tarde — possam introduzir incertezas; excluindo esses horários, há um aumento progressivo na frequência dos eventos ao longo do dia, com pico vespertino. A distribuição das magnitudes é assimétrica, com predominância de valores modestos, mas eventos sob alta irradiância tendem a ser mais intensos e simetricamente distribuídos. Por fim, a duração dos eventos é majoritariamente curta, com quedas acentuadas na frequência conforme a duração aumenta; episódios mais longos são raros e ocorrem, em geral, sob condições específicas próximas ao nascer ou pôr do sol.

A análise dos tipos de nuvens revelou padrões distintos em escalas sazonal, diurna e sob condições específicas de sobreirradiância. Em 2015, observou-se uma redução expressiva na ocorrência de nuvens verticalmente desenvolvidas (como *cumulus congestus* e convecção profunda), possivelmente associada ao El Niño, e aos aerossóis, que dificultaram o desenvolvimento vertical das nuvens. O ciclo diurno indica predominância de nuvens baixas entre 09:00 e 19:00 UTC, com *cirrus* dominando à noite, refletindo o desenvolvimento típico da convecção na Amazônia Central. Em eventos de sobreirradiância, especialmente sob $GHI > 1000 \text{ W/m}^2$, há predominância de nuvens baixas, que, por sua geometria e espessura óptica, favorecem mecanismos de concentração e espalhamento da luz solar. Já *cirrus* e *altocumulus* também aumentam sua frequência nesses eventos, embora com menor eficiência óptica. A menor ocorrência de nuvens profundas em tais episódios pode decorrer de limitações metodológicas relacionadas ao perfilamento zenital, que não captura nuvens em uma geometria mais propícia para a ocorrência de sobreirradiância. Na análise das rampas intensas (maiores que 100 W/m^2), nuvens baixas e *cumulus congestus* apresentam maiores magnitudes e variabilidade, com sombreamentos abruptos associados à sua opacidade e proximidade da superfície. Por outro lado, nuvens altas como *cirrus* geram rampas menos intensas, mas ainda com presença de *outliers*, indicando que a intensidade das variações de irradiância depende não só do tipo de nuvem, mas também da variabilidade intra-classe e da interação com a geometria solar e a dinâmica atmosférica local.

Como aplicação prática, este estudo propôs e avaliou uma abordagem de modelagem baseada em aprendizado de máquina para previsão de irradiância em horizonte de curtíssimo prazo, com um horizonte de até 60 minutos e resolução temporal de 2 minutos. Os dados de entrada incluíram a irradiância medida por um piranômetro e a fração de cobertura de nuvens estimada a partir de imagens de câmera *all sky*. A metodologia envolveu a geração de modelos individuais por meio de MOGA, posteriormente organizados em redes de modelos com múltiplas camadas e diferentes estruturas de previsão (NAR, NARX). A comparação com modelos clássicos de persistência mostrou que os modelos propostos são competitivos, especialmente nos primeiros 30 minutos do horizonte de previsão, período em que a consistência física da cena atmosférica observada ainda é mantida.

Observou-se que a introdução de restrições no processo de busca do MOGA, embora possa limitar a diversidade das soluções, contribui para a obtenção de modelos mais alinhados aos objetivos definidos, resultando em bom desempenho nas etapas iniciais da previsão. No entanto, modelos sem restrições demonstraram maior robustez em horizontes mais longos, possivelmente por explorarem soluções mais equilibradas entre erro inicial e estabilidade ao longo do tempo. Os resultados indicam que, mesmo com um conjunto relativamente limitado de dados de entrada, é possível obter previsões confiáveis e fisicamente consistentes, desde que sejam respeitados os limites espaciais e temporais do sistema observado.

O modelo com melhor desempenho para a previsão de GHI em curtíssimo prazo foi utilizado para avaliar a detecção de eventos de sobreirradiância. Foram consideradas apenas as janelas temporais com ângulo zenital solar inferior a 80° , a fim de evitar falsos positivos associados ao limiar baixo do CSI próximos ao nascer e ao pôr do sol. Os resultados mostraram que o desempenho do modelo decresce progressivamente com o aumento do horizonte de previsão, sendo mantido um compromisso aceitável entre precisão e sensibilidade até aproximadamente 20 a 30 minutos. Para a previsão mais próxima ao evento (2 minutos antes), o modelo apresentou desempenho robusto, com acurácia de 0,85, medida-F1 de 0,77, revocação de 0,91 e precisão de 0,66, indicando alta capacidade de detecção dos eventos, apesar de uma taxa moderada de falsos positivos. A análise diária revelou que o desempenho melhora em dias com maior ocorrência de sobreirradiância, enquanto em dias nublados ou com baixa variabilidade o modelo tende a produzir mais falsos positivos, evidenciando a influência das condições de nebulosidade sobre sua eficácia.

Dessa forma, a metodologia apresentada representa um avanço na previsão de ir-

radiância com alta resolução temporal (escala de tempo entre 20-30 minutos), ao alcançar desempenho robusto com um número reduzido de preditores. Essa simplicidade estrutural favorece a aplicação operacional dos modelos propostos. Vale destacar que tais resultados foram obtidos em um ambiente atmosférico altamente dinâmico e desafiador (Amazônia Central), onde predominam nuvens convectivas de rápida evolução. Em contraste com muitos estudos realizados em regiões subtropicais ou temperadas, onde a variabilidade atmosférica é mais suave e previsível, este trabalho demonstra a viabilidade de modelos eficazes mesmo sob condições típicas de regiões tropicais, contribuindo para preencher uma lacuna relevante na literatura. Para além dos aspectos operacionais, este estudo contribui para o aprimoramento do conhecimento científico associado à compreensão da variabilidade da irradiância solar em regiões tropicais e oferece perspectivas inéditas sobre a sobreirradiância (fenômeno ainda pouco explorado na literatura), destacando sua relevância para previsões e aplicações em energia solar em ambientes equatoriais.

Em um cenário operacional, a implementação seria baseada em um *ensemble*, com atualização das previsões a cada passo de aquisição de dados (por exemplo, a cada 2 minutos), gerando horizontes móveis de até 60 minutos. Tal produto poderia atuar junto a geradores de energia, no suporte ao planejamento da operação e à mitigação de riscos decorrentes da variabilidade de curto prazo, às distribuidoras, interessadas em reduzir incertezas associadas à intermitência e a eventos de sobreirradiância, e até mesmo ao ONS (Operador Nacional do Sistema) com o sistema implementado localmente em várias usinas, no gerenciamento da operação em tempo real e na definição da reserva de segurança em dias de elevada variabilidade. Uma validação adicional em fluxo contínuo de dados constitui, portanto, a etapa seguinte necessária para a adaptação do método a aplicações reais.

Como sugestão de trabalhos futuros, estão:

- Investigar o desempenho do melhor modelo em diferentes contextos sazonais, como estação seca, estação chuvosa e períodos de transição;
- Aplicar a metodologia em outros lugares, com climas diferentes;
- Incluir tipos de nuvens e taxas de rampa como variáveis exógenas na modelagem, verificando o impacto na qualidade da previsão;
- Realizar a previsão da variável exógena (CF), a princípio usando a mesma metodologia usada para prever apenas GHI, respeitando os limites de validade temporal impostos pela dinâmica das nuvens;

- Aplicar intervalos de previsão (PIs) às estimativas de GHI e à detecção de sobreirradiância, de modo a quantificar incertezas e reduzir alarmes falsos;
- Investigar a ocorrência e os efeitos radiativos de múltiplas camadas de nuvem a partir do produto ARSCL e de classificações obtidas por ASIs;
- Investigar o impacto de diferentes valores de limiar na definição de eventos de rampa a fim de aprimorar a avaliação e a detecção dessas variações rápidas de irradiância, conforme destacado por estudos recentes como [Nouri et al. \(2024\)](#).

REFERÊNCIAS BIBLIOGRÁFICAS

- ACKERMAN, A. S.; TOON, O. B.; STEVENS, D. E.; HEYMSFIELD, A. J.; RAMANATHAN, V.; WELTON, E. J. Reduction of tropical cloudiness by soot. **Science**, v. 288, n. 5468, p. 1042–1047, maio 2000. ISSN 1095-9203. Disponível em: <<<http://dx.doi.org/10.1126/science.288.5468.1042>>>. 10
- AKIMA, H. A new method of interpolation and smooth curve fitting based on local procedures. **Journal of the ACM**, v. 17, n. 4, p. 589–602, out. 1970. ISSN 1557-735X. Disponível em: <<<http://dx.doi.org/10.1145/321607.321609>>>. 42
- ALBRECHT, B. A. Aerosols, cloud microphysics, and fractional cloudiness. **Science**, v. 245, n. 4923, p. 1227–1230, set. 1989. ISSN 1095-9203. Disponível em: <<<http://dx.doi.org/10.1126/science.245.4923.1227>>>. 10
- ALMEIDA, M. P.; ZILLES, R.; LORENZO, E. Extreme overirradiance events in São Paulo, Brazil. **Solar Energy**, v. 110, p. 168–173, dez. 2014. ISSN 0038-092X. Disponível em: <<<http://dx.doi.org/10.1016/j.solener.2014.09.012>>>. 29
- ALVES, M. A. S.; MACHADO, L. A. T.; PRASAD, G. S. S. D. Estudo da variabilidade da cobertura de nuvens altas na Amazônia central. **Acta Amazonica**, v. 37, n. 1, p. 71–79, 2007. ISSN 0044-5967. Disponível em: <<<http://dx.doi.org/10.1590/S0044-59672007000100008>>>. 62
- ANDREAS, A.; DOORAGHI, M.; HABTE, A.; KUTCHENREITER, M.; REDA, I.; SENGUPTA, M. **Solar Infrared Radiation Station (SIRS), Sky Radiation (SKYRAD), Ground Radiation (GNDRAD), and Broadband Radiometer Station (BRS) Instrument Handbook**. National Renewable Energy Laboratory (NREL) - U.S. Department of Energy (DOE), 2018. Disponível em: <<<http://dx.doi.org/10.2172/1432706>>>. 35
- ANTONANZAS, J.; OSORIO, N.; ESCOBAR, R.; URRACA, R.; PISON, F. M. de; ANTONANZAS-TORRES, F. Review of photovoltaic power forecasting. **Solar Energy**, v. 136, p. 78–111, out. 2016. Disponível em: <<<https://doi.org/10.1016/j.solener.2016.06.069>>>. 23, 37
- ARIAS, P. A.; FU, R.; HOYOS, C. D.; LI, W.; ZHOU, L. Changes in cloudiness over the Amazon rainforests during the last two decades: diagnostic and potential causes. **Climate Dynamics**, v. 37, n. 5-6, p. 1151–1164, set. 2010. Disponível em: <<<https://doi.org/10.1007/s00382-010-0903-2>>>. 16, 61
- ARM. **GoAmazon2014/5: photos from the GoAmazon campaign**. 2014. Disponível em: <<<https://www.flickr.com/photos/armgov/12841803134/in/album-72157628390811069>>>. Acesso em: 31 jul. 2025. 37
- ASSOCIAÇÃO BRASILEIRA DE ENERGIA SOLAR FOTOVOLTAICA (ABSOLAR). **Infográfico**. 2025. Disponível em: <<<https://www.absolar.org.br/mercado/infografico/>>>. 1

BALFOUR, J. **Introduction to photovoltaic installations**. Jones & Bartlett Learning, 2011. ISBN 9781449625788. Disponível em:

<<<https://books.google.com.br/books?id=Lje9wAEACAAJ>>>. 27, 84

BARHMI, K.; HEYNEN, C.; GOLROODBARI, S.; SARK, W. van. A review of solar forecasting techniques and the role of artificial intelligence. **Solar**, v. 4, n. 1, p. 99–135, fev. 2024. ISSN 2673-9941. Disponível em:

<<<http://dx.doi.org/10.3390/solar4010005>>>. 23

BECK, H. E.; ZIMMERMANN, N. E.; MCVICAR, T. R.; VERGOPOLAN, N.; BERG, A.; WOOD, E. F. Present and future köppen-geiger climate classification maps at 1-km resolution. **Scientific Data**, v. 5, n. 1, October 2018. ISSN 2052-4463. 14

BOTTINO, M. J.; NOBRE, P.; GIAROLLA, E.; JUNIOR, M. B. da S.; CAPISTRANO, V. B.; MALAGUTTI, M.; TAMAOKI, J. N.; OLIVEIRA, B. F. A. de; NOBRE, C. A. Amazon savannization and climate change are projected to increase dry season length and temperature extremes over Brazil. **Scientific Reports**, v. 14, n. 1, mar. 2024. ISSN 2045-2322. 15

BRAGA, M.; OLIVEIRA, A. K. V. de; BURNHAM, L.; DITTMANN, S.; GOTTSCHALG, R.; BETTS, T.; RODRIGUEZ-GALLEGOS, C. D.; REINDL, T.; RÜTHER, R. Solar over-irradiance events: preliminary results from a global study. In: IEEE PHOTOVOLTAIC SPECIALISTS CONFERENCE, 47., 2020. **Proceedings...** IEEE, 2020. p. 2764–2770. Disponível em:

<<<http://dx.doi.org/10.1109/PVSC45281.2020.9300868>>>. 30, 78

BROOMHEAD, D. S.; LOWE, D. Multivariable functional interpolation and adaptive networks. **Complex Systems**, v. 2, 1988. Disponível em:

<<<https://api.semanticscholar.org/CorpusID:3686496>>>. 19

BROWNLEE, J. **Python for time series forecasting**. [S.l.]: Machine Learning Mastery, 2020. 43

BURLEYSON, C. D.; LONG, C. N.; COMSTOCK, J. M. Quantifying diurnal cloud radiative effects by cloud type in the tropical western pacific. **Journal of Applied Meteorology and Climatology**, v. 54, n. 6, p. 1297–1312, jun. 2015. ISSN 1558-8432. Disponível em:

<<<http://dx.doi.org/10.1175/JAMC-D-14-0288.1>>>. 41, 42

CARNEIRO, R. G.; BORGES, C. K.; SANTOS, C. A.; OLIVEIRA, G. de; STACHLEWSKA, I. S.; FISCH, G.; SANTOS, C. A. dos. Energy balance closure and evapotranspiration hysteresis in central amazon under contrasting conditions during the goamazon project in 2014 and 2015. **Journal of South American Earth Sciences**, v. 132, p. 104686, dez. 2023. ISSN 0895-9811. Disponível em:

<<<http://dx.doi.org/10.1016/j.jsames.2023.104686>>>. 10, 11, 12, 14

CARNEIRO, R. G.; FISCH, G. Observational analysis of the daily cycle of the planetary boundary layer in the central Amazon during a non-El Niño year and El

Niño year (GoAmazon project 2014/5). **Atmospheric Chemistry and Physics**, v. 20, n. 9, p. 5547–5558, maio 2020. Disponível em: <<<https://doi.org/10.5194/acp-20-5547-2020>>>. 10, 83, 87

CECCHINI, M. A.; MACHADO, L. A.; COMSTOCK, J. M.; MEI, F.; WANG, J.; FAN, J.; TOMLINSON, J. M.; SCHMID, B.; ALBRECHT, R.; MARTIN, S. T.; ARTAXO, P. Impacts of the manaus pollution plume on the microphysical properties of amazonian warm-phase clouds in the wet season. **Atmospheric Chemistry and Physics**, v. 16, n. 11, p. 7029–7041, 2016. 16

CHEN, S.; LI, P.; BRADY, D.; LEHMAN, B. The impact of irradiance time behaviors on inverter sizing and design. In: WORKSHOP ON CONTROL AND MODELING FOR POWER ELECTRONICS, 12., 2010. **Proceedings... IEEE**, 2010. p. 1–5. Disponível em: <<<http://dx.doi.org/10.1109/COMPEL.2010.5562387>>>. 26

CHENG, H.-Y.; YU, C.-C. All-sky image processing for solar energy applications. In: INTERNATIONAL CONFERENCE ON INTERNET MULTIMEDIA COMPUTING AND SERVICE, 2014. **Proceedings... ACM**, 2014. p. 185–188. Disponível em: <<<http://dx.doi.org/10.1145/2632856.2632909>>>. 24, 37

CHU, Y.; LI, M.; COIMBRA, C. F.; FENG, D.; WANG, H. Intra-hour irradiance forecasting techniques for solar power integration: A review. **iScience**, v. 24, n. 10, p. 103136, oct 2021. Disponível em: <<<https://doi.org/10.1016/j.isci.2021.103136>>>. 23

CLIMATE PREDICTION CENTER, NCEP, NOAA. **Oceanic Niño Index (ONI) Version 5**. 2025. Disponível em: <<https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php>>. Acesso em: 3 jul. 2025. 61, 71

DAVIES, R. The effect of finite geometry on the three-dimensional transfer of solar irradiance in clouds. **Journal of the Atmospheric Sciences**, v. 35, n. 9, p. 1712–1725, set. 1978. ISSN 1520-0469. Disponível em: <<[http://dx.doi.org/10.1175/1520-0469\(1978\)035<1712:TEOFGO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1978)035<1712:TEOFGO>2.0.CO;2)>>. 27

DAVIS, J. M.; COX, S. K.; MCKEE, T. B. Design and verification of a cloud field optical simulator. **Journal of Climate and Applied Meteorology**, v. 22, n. 5, p. 947–958, maio 1983. ISSN 0733-3021. Disponível em: <<[http://dx.doi.org/10.1175/1520-0450\(1983\)022<0947:DAVOAC>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1983)022<0947:DAVOAC>2.0.CO;2)>>. 27

DIAS, M. A. F. S.; DIAS, P. L. S.; LONGO, M.; FITZJARRALD, D. R.; DENNING, A. S. River breeze circulation in eastern amazonia: observations and modelling results. **Theoretical and Applied Climatology**, v. 78, n. 1–3, abr. 2004. ISSN 1434-4483. Disponível em: <<<http://dx.doi.org/10.1007/s00704-004-0047-6>>>. 16

EMCK, P.; RICHTER, M. An upper threshold of enhanced global shortwave irradiance in the troposphere derived from field measurements in tropical mountains. **Journal of Applied Meteorology and Climatology**, v. 47, n. 11, p. 2828–2845, nov. 2008. ISSN 1558-8424. Disponível em: <<<http://dx.doi.org/10.1175/2008JAMC1861.1>>>. 28

FEISTER, U.; CABROL, N.; HÄDER, D. Uv irradiance enhancements by scattering of solar radiation from clouds. **Atmosphere**, v. 6, n. 8, p. 1211–1228, ago. 2015. ISSN 2073-4433. Disponível em: <<<http://dx.doi.org/10.3390/atmos6081211>>>. 30

FERNANDES, A. D. F. T. **Técnicas de aprendizado de máquina para a predição de eventos extremos de sobreirradiância em Natal-RN**. 79 f.: il. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, Natal, RN, 2023. 30

FERREIRA, P. M.; GOMES, J. M.; MARTINS, I. A. C.; RUANO, A. E. A neural network based intelligent predictive sensor for cloudiness, solar radiation and air temperature. **Sensors**, v. 12, n. 11, p. 15750–15777, nov. 2012. ISSN 1424-8220. Disponível em: <<<http://dx.doi.org/10.3390/s121115750>>>. 23

FERREIRA, P. M.; RUANO, A. E. Evolutionary multiobjective neural network models identification: evolving task-optimised models. In: _____. **New advances in intelligent signal processing**. Berlin: Springer, 2011. cap. 2, p. 21–53. ISBN 9783642117398. Disponível em: <<http://dx.doi.org/10.1007/978-3-642-11739-8_2>>. 19

FISCH, G.; MARENGO, J. A.; NOBRE, C. A. Uma revisão geral sobre o clima da Amazônia. **Acta Amazonica**, v. 28, n. 2, p. 101–101, jun. 1998. ISSN 0044-5967. 14

FLYNN, D. **Total Sky Imager (TSISKYCOVER), 2014-01-01 to 2015-12-01, ARM Mobile Facility (MAO), Manacapuru, Amazonas, Brazil; AMF1 (M1)**. [S.l.]: [s.n.], 2014. 34

FRANCO, M. A.; TAVARES, P. H.; RIZZO, L. V.; MORAIS, F. G.; PALÁCIOS, R.; ARTAXO, P. Aeronet sun photometer as a didactic tool for understanding aerosol refractive index in the atmosphere: a case study for central Amazon. **Revista Brasileira de Ensino de Física**, v. 47, 2025. ISSN 1806-1117. Disponível em: <<<http://dx.doi.org/10.1590/1806-9126-RBEF-2024-0341>>>. 39

FRITSCH, F. N.; BUTLAND, J. A method for constructing local monotone piecewise cubic interpolants. **SIAM Journal on Scientific and Statistical Computing**, Society for Industrial and Applied Mathematics (SIAM), v. 5, n. 2, p. 300–304, jun. 1984. ISSN 2168-3417. Disponível em: <<<http://dx.doi.org/10.1137/0905021>>>. 42

GIANGRANDE, S. E.; BISCARO, T. S.; PETERS, J. M. Seasonal controls on isolated convective storm drafts, precipitation intensity, and life cycle as observed

during goamazon2014/5. **Atmospheric Chemistry and Physics**, v. 23, n. 9, p. 5297–5316, maio 2023. ISSN 1680-7324. Disponível em:

<<<http://dx.doi.org/10.5194/acp-23-5297-2023>>>. 10, 12, 13

GIANGRANDE, S. E.; FENG, Z.; JENSEN, M. P.; COMSTOCK, J. M.; JOHNSON, K. L.; TOTO, T.; WANG, M.; BURLEYSON, C.; BHARADWAJ, N.; MEI, F.; MACHADO, L. A. T.; MANZI, A. O.; XIE, S.; TANG, S.; DIAS, M. A. F. S.; SOUZA, R. A. F. de; SCHUMACHER, C.; MARTIN, S. T. Cloud characteristics, thermodynamic controls and radiative impacts during the observations and modeling of the Green Ocean Amazon (GoAmazon2014/5) experiment. **Atmospheric Chemistry and Physics**, v. 17, n. 23, p. 14519–14541, dez. 2017. Disponível em:

<<<https://doi.org/10.5194/acp-17-14519-2017>>>. 57, 82, 83

GIBLER, G.; SIVARAMAN, C.; GREGORY, L.; MA, L.; WAGENER, R. **Sunphotometer (CSPHOTAODV3), 2013-12-20 to 2015-11-30, ARM Mobile Facility (MAO), Manacapuru, Amazonas, Brazil; AMF1 (M1)**. [S.l.]: [s.n.], 2014. 38

GONÇALVES, L. J. M.; COELHO, S. M. S. C.; KUBOTA, P. Y.; SOUZA, D. C. Interaction between cloud–radiation, atmospheric dynamics and thermodynamics based on observational data from goamazon 2014/15 and a cloud-resolving model. **Atmospheric Chemistry and Physics**, v. 22, n. 23, p. 15509–15526, dez. 2022. ISSN 1680-7324. Disponível em:

<<<http://dx.doi.org/10.5194/acp-22-15509-2022>>>. 10, 12, 13

GUEYMARD, C. A. Cloud and albedo enhancement impacts on solar irradiance using high-frequency measurements from thermopile and photodiode radiometers. part 1: impacts on global horizontal irradiance. **Solar Energy**, v. 153, p. 755–765, set. 2017. Disponível em: <<<https://doi.org/10.1016/j.solener.2017.05.004>>>. 27

HAAREN, R. van; MORJARIA, M.; FTHENAKIS, V. Empirical assessment of short-term variability from utility-scale solar pv plants. **Progress in Photovoltaics: Research and Applications**, v. 22, n. 5, p. 548–559, nov. 2012. ISSN 1099-159X. Disponível em: <<<http://dx.doi.org/10.1002/pip.2302>>>. 4

HANSEN, C. W.; STEIN, J. S.; ELLIS, A. **Statistical criteria for characterizing irradiance time series**. Albuquerque, NM, USA: [s.n.], 2010. 4

HENDRIKX, N.; BARHMI, K.; VISSER, L.; BRUIN, T. de; PÓ, M.; SALAH, A.; SARK, W. van. All sky imaging-based short-term solar irradiance forecasting with long short-term memory networks. **Solar Energy**, v. 272, p. 112463, abr. 2024. ISSN 0038-092X. Disponível em:

<<<http://dx.doi.org/10.1016/j.solener.2024.112463>>>. 24

HENKES, A.; FISCH, G.; MACHADO, L. A. T.; CHABOUREAU, J.-P. Morning boundary layer conditions for shallow to deep convective cloud evolution during the dry season in the central Amazon. **Atmospheric Chemistry and Physics**, v. 21, n. 17, p. 13207–13225, set. 2021. ISSN 1680-7324. Disponível em:

<<<http://dx.doi.org/10.5194/acp-21-13207-2021>>>. 10, 11

HENKES, A. F. **Evolução de nuvens rasas à convecção profunda na estação seca na Amazônia central: observação e modelagem.** Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2021. 14

HOBBS, P. V. **Aerosol-cloud-climate interactions.** [S.l.]: Academic Press, 1993. 10

HOLBEN, B.; ECK, T.; SLUTSKER, I.; TANRÉ, D.; BUIS, J.; SETZER, A.; VERMOTE, E.; REAGAN, J.; KAUFMAN, Y.; NAKAJIMA, T.; LAVENU, F.; JANKOWIAK, I.; SMIRNOV, A. Aeronet—a federated instrument network and data archive for aerosol characterization. **Remote Sensing of Environment**, v. 66, n. 1, p. 1–16, out. 1998. ISSN 0034-4257. Disponível em: <<[http://dx.doi.org/10.1016/S0034-4257\(98\)00031-5](http://dx.doi.org/10.1016/S0034-4257(98)00031-5)>>. 38, 39

HOLTON, J. R. **An introduction to dynamic meteorology.** 4.. ed. San Diego: Academic Press, 2004. ISBN 9780123540157. 10, 16

HOREL, J. D.; HAHMANN, A. N.; GEISLER, J. E. An investigation of the annual cycle of convective activity over the tropical americas. **Journal of Climate**, v. 2, n. 11, p. 1388–1403, nov. 1989. ISSN 1520-0442. Disponível em: <<[http://dx.doi.org/10.1175/1520-0442\(1989\)002<1388:AIOTAC>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1989)002<1388:AIOTAC>2.0.CO;2)>>. 69

HORNER, N.; BRAGA, M.; MANTELLI, S. L.; RÜTHER, R. eventos de sobreirradiação em sistemas fotovoltaicos em Florianópolis – SC. **Revista Brasileira de Energia Solar**, v. 14, n. 1, p. 11–20, nov. 2023. ISSN 2178-9606. Disponível em: <<<http://dx.doi.org/10.59627/rbens.2023v14i1.419>>>. 26

HOYOS-GÓMEZ, L. S.; RUIZ-MUÑOZ, J. F.; RUIZ-MENDOZA, B. J. Short-term forecasting of global solar irradiance in tropical environments with incomplete data. **Applied Energy**, v. 307, p. 118192, fev. 2022. ISSN 0306-2619. Disponível em: <<<http://dx.doi.org/10.1016/j.apenergy.2021.118192>>>. 23

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice.** 3. ed. Melbourne, Australia: OTexts, 2021. Disponível em: <<<https://otexts.com/fpp3/>>>. 52

INEICHEN, P.; PEREZ, R. A new airmass independent formulation for the linke turbidity coefficient. **Solar Energy**, v. 73, n. 3, p. 151–157, 2002. ISSN 0038-092X. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0038092X02000452>>>. 44

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Programa Queimadas – banco de dados de queimadas.** 2025. Disponível em: <<<https://terrabrasilis.dpi.inpe.br/queimadas/bdqueimadas>>>. Acesso em: jul. 2025. 63

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). The earth's energy budget, climate feedbacks and climate sensitivity. In: **Climate Change 2021 – The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change**. Cambridge University Press, 2023. p. 923–1054. Disponível em: <<<https://doi.org/10.1017/9781009157896.009>>>. 10

INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC). **IEC 61215-1:2021 - Terrestrial photovoltaic (PV) modules – design qualification and type approval – Part 1: test requirements**. [s.n.], 2021. Disponível em: <<<https://webstore.iec.ch/publication/61173>>>. 28

JIMÉNEZ-MUÑOZ, J. C.; MATTAR, C.; BARICHIVICH, J.; SANTAMARÍA-ARTIGAS, A.; TAKAHASHI, K.; MALHI, Y.; SOBRINO, J. A.; SCHRIER, G. v. d. Record-breaking warming and extreme drought in the Amazon rainforest during the course of El Niño 2015–2016. **Scientific Reports**, v. 6, n. 1, set. 2016. ISSN 2045-2322. Disponível em: <<<http://dx.doi.org/10.1038/srep33130>>>. 15

JOHNSON, K.; JENSEN, M. **ARM: ARSCL: multiple outputs from first Clothiaux algorithms on Vaisala or Belfort ceilometers, Micropulse lidar, and MMCR**. 1996. Disponível em: <<<https://www.osti.gov/servlets/purl/1027282/>>>. 41

JONES, C.; HOREL, J. A note on the upper level divergence field over South America during the summer season. **Revista Brasileira de Meteorologia**, v. 5, n. 2, p. 411–416, 1990. 15

KAUFMAN, Y. J.; NAKAJIMA, T. Effect of Amazon smoke on cloud microphysics and albedo-analysis from satellite imagery. **Journal of Applied Meteorology**, v. 32, n. 4, p. 729–744, abr. 1993. ISSN 1520-0450. Disponível em: <<[http://dx.doi.org/10.1175/1520-0450\(1993\)032<0729:EOASOC>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1993)032<0729:EOASOC>2.0.CO;2)>>. 10

KHAIR, A.; PUTRANTO, L. M.; ARIANANDA, D. D. Time series analysis for solar irradiance forecasting: a systematic review. In: INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING, 15., 2025. **Proceedings... IEEE**, 2025. p. 1–6. Disponível em: <<<http://dx.doi.org/10.1109/ICEENG64546.2025.11031366>>>. 23

KHOSRAVANI, H.; RUANO, A.; FERREIRA, P. A convex hull-based data selection method for data driven models. **Applied Soft Computing**, v. 47, p. 515–533, out. 2016. ISSN 1568-4946. Disponível em: <<<http://dx.doi.org/10.1016/j.asoc.2016.06.014>>>. 17, 123

KUHN, P.; NOURI, B.; WILBERT, S.; PRAHL, C.; KOZONEK, N.; SCHMIDT, T.; YASSER, Z.; RAMIREZ, L.; ZARZALEJO, L.; MEYER, A.; VUILLEUMIER, L.; HEINEMANN, D.; BLANC, P.; PITZ-PAAL, R. Validation of an all-sky imager-based nowcasting system for industrial pv plants. **Progress in**

Photovoltaics: Research and Applications, v. 26, n. 8, p. 608–621, 2018.
Disponível em: <<<https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.2968>>>. 2

KUMARI, P.; TOSHNIWAL, D. Deep learning models for solar irradiance forecasting: a comprehensive review. **Journal of Cleaner Production**, v. 318, p. 128566, oct 2021. Disponível em:
<<<https://doi.org/10.1016/j.jclepro.2021.128566>>>. 23

KWOK, T.-Y.; YEUNG, D.-Y. Constructive algorithms for structure learning in feedforward neural networks for regression problems. **IEEE Transactions on Neural Networks**, v. 8, n. 3, p. 630–645, maio 1997. ISSN 1941-0093. 19

KYROUAC, J.; SHI, Y.; TUFTEDAL, M. **Surface Meteorological Instrumentation (MET), 2014-01-01 to 2015-12-01, ARM Mobile Facility (MAO), Manacapuru, Amazonas, Brazil; AMF1 (M1)**. [S.l.]: [s.n.], 2014. 34

LAING, A. G.; EVANS, J. L. Challenges of tropical weather forecasting. In: **Introduction to Tropical Meteorology**. 2. ed. [S.l.]: COMET Program, University Corporation for Atmospheric Research (UCAR), 2011. cap. 9.1. 14

LAOUAFI, A.; LAOUAFI, F.; BOUKELIA, T. E. An adaptive hybrid ensemble with pattern similarity analysis and error correction for short-term load forecasting. **Applied Energy**, v. 322, p. 119525, set. 2022. ISSN 0306-2619.
Disponível em: <<<http://dx.doi.org/10.1016/j.apenergy.2022.119525>>>. 3

LAOUALI, I.; RUANO, A.; RUANO, M. d. G.; BENNANI, S. D.; FADILI, H. E. Non-intrusive load monitoring of household devices using a hybrid deep learning model through convex hull-based data selection. **Energies**, v. 15, n. 3, p. 1215, fev. 2022. ISSN 1996-1073. Disponível em:
<<<http://dx.doi.org/10.3390/en15031215>>>. 18

LAVE, M.; KLEISSL, J.; ELLIS, A.; MEJIA, F. Simulated pv power plant variability: impact of utility-imposed ramp limitations in puerto rico. In: **IEEE PHOTOVOLTAIC SPECIALISTS CONFERENCE**, 39., 2013. **Proceedings...** IEEE, 2013. p. 1817–1821. Disponível em:
<<<http://dx.doi.org/10.1109/PVSC.2013.6744495>>>. 3

LEBSOCK, M. D.; STEPHENS, G. L.; KUMMEROW, C. Multisensor satellite observations of aerosol effects on warm clouds. **Journal of Geophysical Research: Atmospheres**, v. 113, n. D15, ago. 2008. ISSN 0148-0227. Disponível em: <<<http://dx.doi.org/10.1029/2008jd009876>>>. 10

LEHMANN, E. L.; ROMANO, J. P. **Testing statistical hypotheses**. 4. ed. Cham, Switzerland: Springer, 2022. (Springer Texts in Statistics). ISBN 978-3-030-70577-0. Disponível em:
<<<https://link.springer.com/book/10.1007/978-3-030-70578-7>>>. 55

LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. **Quarterly of Applied Mathematics**, v. 2, n. 2, p. 164–168, 1944. ISSN 1552-4485. 21

LIN, F.; ZHANG, Y.; WANG, J. Recent advances in intra-hour solar forecasting: a review of ground-based sky image methods. **International Journal of Forecasting**, v. 39, n. 1, p. 244–265, jan. 2023. ISSN 0169-2070. Disponível em: <<<http://dx.doi.org/10.1016/j.ijforecast.2021.11.002>>>. 23

LOGOTHETIS, S.-A.; SALAMALIKIS, V.; WILBERT, S.; REMUND, J.; ZARZALEJO, L. F.; XIE, Y.; NOURI, B.; NTAVELIS, E.; NOU, J.; HENDRIKX, N.; VISSER, L.; SENGUPTA, M.; PÓ, M.; CHAUVIN, R.; GRIEU, S.; BLUM, N.; SARK, W. van; KAZANTZIDIS, A. Benchmarking of solar irradiance nowcast performance derived from all-sky imagers. **Renewable Energy**, v. 199, p. 246–261, nov. 2022. ISSN 0960-1481. Disponível em: <<<http://dx.doi.org/10.1016/j.renene.2022.08.127>>>. 5, 24, 37, 52

LONG, C. N.; SLATER, D. W.; TOOMAN, T. **Total Sky Imager (TSI) model 880 status and testing results**. Washington, D.C.: [s.n.], 2001. 36 p. Disponível em: <<<http://www.arm.gov/docs/documents/techreports/index.html>>>. 36

LÓPEZ-CUESTA, M.; ALER-MUR, R.; GALVÁN-LEÓN, I. M.; RODRÍGUEZ-BENÍTEZ, F. J.; POZO-VÁZQUEZ, A. D. Improving solar radiation nowcasts by blending data-driven, satellite-images-based and all-sky-imagers-based models using machine learning techniques. **Remote Sensing**, v. 15, n. 9, p. 2328, abr. 2023. ISSN 2072-4292. Disponível em: <<<http://dx.doi.org/10.3390/rs15092328>>>. 25, 37

MACEDO, A. d. S. **Variabilidade temporal e espacial da radiação solar na região de Manaus - AM durante o experimento GoAmazon 2014/15**. Dissertação (Mestrado) — Instituto Nacional de Pesquisas Espaciais, 2016. Disponível em: <<<http://urlib.net/8JMKD3MGP3W34P/3MFB9BE>>>. 35, 61

MACEDO, A. d. S.; FISCH, G. Variabilidade temporal da radiação solar durante o experimento goamazon 2014/15. **Revista Brasileira de Meteorologia**, v. 33, n. 2, p. 353–365, jun. 2018. ISSN 0102-7786. Disponível em: <<<http://dx.doi.org/10.1590/0102-7786332017>>>. 17, 59, 60, 63

MACHADO, L. A. T.; CALHEIROS, A. J. P.; BISCARO, T.; GIANGRANDE, S.; DIAS, M. A. F. S.; CECCHINI, M. A.; ALBRECHT, R.; ANDREAE, M. O.; ARAUJO, W. F.; ARTAXO, P.; BORRMANN, S.; BRAGA, R.; BURLEYSON, C.; EICHHOLZ, C. W.; FAN, J.; FENG, Z.; FISCH, G. F.; JENSEN, M. P.; MARTIN, S. T.; PÖSCHL, U.; PÖHLKER, C.; PÖHLKER, M. L.; RIBAUD, J.-F.; ROSENFELD, D.; SARAIVA, J. M. B.; SCHUMACHER, C.; THALMAN, R.; WALTER, D.; WENDISCH, M. Overview: precipitation characteristics and sensitivities to environmental conditions during GoAmazon2014/5 and ACRIDICON-CHUVA. **Atmospheric Chemistry and Physics**, v. 18, n. 9, p. 6461–6482, maio 2018. Disponível em: <<<https://doi.org/10.5194/acp-18-6461-2018>>>. 10, 16, 69, 74

MACHADO, L. A. T.; LAURENT, H.; DESSAY, N.; MIRANDA, I. Seasonal and diurnal variability of convection over the Amazonia: a comparison of different vegetation types and large scale forcing. **Theoretical and Applied Climatology**, v. 78, n. 1–3, abr. 2004. ISSN 1434-4483. 61

MACHADO, L. A. T.; LAURENT, H.; LIMA, A. A. Diurnal march of the convection observed during trmm-wetamc/lba. **Journal of Geophysical Research: Atmospheres**, v. 107, n. D20, set. 2002. ISSN 0148-0227. Disponível em: <<<http://dx.doi.org/10.1029/2001JD000338>>>. 17

MAHESH, B. Machine learning algorithms - a review. **International Journal of Science and Research (IJSR)**, v. 9, n. 1, p. 381–386, jan. 2020. ISSN 2319-7064. Disponível em: <<<http://dx.doi.org/10.21275/ART20203995>>>. 22

MARENGO, J. A.; ESPINOZA, J. C. Extreme seasonal droughts and floods in amazonia: causes, trends and impacts. **International Journal of Climatology**, v. 36, n. 3, p. 1033–1050, jul. 2015. ISSN 1097-0088. Disponível em: <<<http://dx.doi.org/10.1002/joc.4420>>>. 15, 61

MARENGO, J. A.; FISCH, G. F.; ALVES, L. M.; SOUSA, N. V.; FU, R.; ZHUANG, Y. Meteorological context of the onset and end of the rainy season in central amazonia during the goamazon2014/5. **Atmospheric Chemistry and Physics**, v. 17, n. 12, p. 7671–7681, jun. 2017. ISSN 1680-7324. 15

MARQUARDT, D. W. An algorithm for least-squares estimation of nonlinear parameters. **Journal of the Society for Industrial and Applied Mathematics**, v. 11, n. 2, p. 431–441, jun. 1963. ISSN 2168-3484. 21

MARTIN, S. T.; ARTAXO, P.; MACHADO, L. A. T.; MANZI, A. O.; SOUZA, R. A. F.; SCHUMACHER, C.; WANG, J.; ANDREAE, M. O.; BARBOSA, H. M. J.; FAN, J.; FISCH, G.; GOLDSTEIN, A. H.; GUENTHER, A.; JIMENEZ, J. L.; PöSCHL, U.; DIAS, M. A. S.; SMITH, J. N.; WENDISCH, M. Introduction: observations and modeling of the Green Ocean Amazon (GoAmazon2014/5). **Atmospheric Chemistry and Physics**, v. 16, n. 8, p. 4785–4797, abr. 2016. 17, 33

MARTINS, B. J.; CERENTINI, A.; MANTELLI, S. L.; CHAVES, T. Z. L.; BRANCO, N. M.; WANGENHEIM, A. von; RÜTHER, R.; ARRAIS, J. M. Systematic review of nowcasting approaches for solar energy production based upon ground-based cloud imaging. **Solar Energy Advances**, v. 2, p. 100019, 2022. Disponível em: <<<https://doi.org/10.1016/j.seja.2022.100019>>>. 23

MARTINS, G.; MANTELLI, S.; RÜTHER, R. Evaluating the performance of radiometers for solar overirradiance events. **Solar Energy**, v. 231, p. 47–56, jan. 2022. ISSN 0038-092X. Disponível em: <<<http://dx.doi.org/10.1016/j.solener.2021.11.050>>>. 2, 79

MARTINS, V. S.; NOVO, E. M.; LYAPUSTIN, A.; ARAGÃO, L. E.; FREITAS, S. R.; BARBOSA, C. C. Seasonal and interannual assessment of cloud cover and

atmospheric constituents across the amazon (2000–2015): insights for remote sensing and climate analysis. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 145, p. 309–327, nov. 2018. ISSN 0924-2716. Disponível em: <<<http://dx.doi.org/10.1016/j.isprsjprs.2018.05.013>>>. 40

MARZOUQ, M.; FADILI, H. E.; ZENKOUAR, K.; LAKHLIAI, Z.; AMOUZG, M. Short term solar irradiance forecasting via a novel evolutionary multi-model framework and performance assessment for sites with no solar irradiance data. **Renewable Energy**, v. 157, p. 214–231, set. 2020. ISSN 0960-1481. Disponível em: <<<http://dx.doi.org/10.1016/j.renene.2020.04.133>>>. 25

MAYTA, V. C.; AMBRIZZI, T.; ESPINOZA, J. C.; DIAS, P. L. S. The role of the Madden-Julian oscillation on the Amazon basin intraseasonal rainfall variability. **International Journal of Climatology**, v. 39, n. 1, p. 343–360, ago. 2018. Disponível em: <<<https://doi.org/10.1002/joc.5810>>>. 15

MAYTA, V. C.; KILADIS, G. N.; DIAS, J.; DIAS, P. L. S.; GEHNE, M. Convectively coupled kelvin waves over tropical South America. **Journal of Climate**, p. 1–52, maio 2021. Disponível em: <<<https://doi.org/10.1175/jcli-d-20-0662.1>>>. 15

MCFARLANE, S. A.; LONG, C. N.; FLAHERTY, J. A climatology of surface cloud radiative effects at the arm tropical western pacific sites. **Journal of Applied Meteorology and Climatology**, Boston MA, USA, v. 52, n. 4, p. 996 – 1013, 2013. Disponível em: <<<https://journals.ametsoc.org/view/journals/apme/52/4/jamc-d-12-0189.1.xml>>>. 41, 42

MCKEE, T. B.; COX, S. K. Scattering of visible radiation by finite clouds. **Journal of the Atmospheric Sciences**, v. 31, n. 7, p. 1885–1892, out. 1974. ISSN 1520-0469. Disponível em: <<[http://dx.doi.org/10.1175/1520-0469\(1974\)031<1885:SOVRBF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1974)031<1885:SOVRBF>2.0.CO;2)>>. 27

MOL, W.; HEERWAARDEN, C. van. Mechanisms of surface solar irradiance variability under broken clouds. **Atmospheric Chemistry and Physics**, v. 25, n. 8, p. 4419–4441, abr. 2025. ISSN 1680-7324. Disponível em: <<<http://dx.doi.org/10.5194/acp-25-4419-2025>>>. 7, 8, 69

MOLION, L. C. B. Climatologia dinâmica da região amazônica: mecanismos de precipitação. **Revista Brasileira de Meteorologia**, v. 2, n. 1, p. 107–117, 1987. 15

MONTEITH, J. L. (Ed.). **Vegetation and the atmosphere. Volume 2: case studies**. London: Academic Press, 1976. 439 p. 30

MORETTIN, P. A.; TOLOI, C. G. **Séries Temporais**. [S.l.]: Edgard Blücher, 2006. 43

- MORRIS, V. **ARM: Total Sky Imager (TSI): fractional sky coverage**. Atmospheric Radiation Measurement (ARM) Archive, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (US), 2000. Disponível em: <<<https://www.osti.gov/servlets/purl/1025308/>>>. 37
- NASCIMENTO, L. R. do; SOUZA VIANA, T. de; CAMPOS, R. A.; RÚTHER, R. Extreme solar overirradiance events: occurrence and impacts on utility-scale photovoltaic power plants in Brazil. **Solar Energy**, v. 186, p. 370–381, 2019. ISSN 0038-092X. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0038092X19304530>>>. 28, 29
- NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). **Sun and climate research area (Research Area 156)**. 2025. Página web. Disponível em: <<<https://science.gsfc.nasa.gov/earth/climate/researchareas/156>>>. Acesso em: 28 jul. 2025. 5
- NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA). **Radiation monitoring instruments – pyranometer**. 2024. Disponível em: <<<https://gml.noaa.gov/grad/antuv/Instruments.jsp>>>. Acesso em: 30 jul. 2025. 36
- NATSIS, A. N.; BAIS, A.; MELETI, C. Analysis of cloud enhancement events in a 30-year record of global solar irradiance at Thessaloniki, Greece. **Atmospheric Research**, v. 312, p. 107731, dez. 2024. ISSN 0169-8095. Disponível em: <<<http://dx.doi.org/10.1016/j.atmosres.2024.107731>>>. 28
- NOURI, B.; BLUM, N.; WILBERT, S.; ZARZALEJO, L. F. A hybrid solar irradiance nowcasting approach: combining all sky imager systems and persistence irradiance models for increased accuracy. **Solar RRL**, v. 6, n. 5, ago. 2021. ISSN 2367-198X. Disponível em: <<<http://dx.doi.org/10.1002/solr.202100442>>>. 24, 37
- NOURI, B.; FABEL, Y.; BLUM, N.; SCHNAUS, D.; ZARZALEJO, L. F.; KAZANTZIDIS, A.; WILBERT, S. Ramp rate metric suitable for solar forecasting. **Solar RRL**, v. 8, n. 24, out. 2024. ISSN 2367-198X. Disponível em: <<<http://dx.doi.org/10.1002/solr.202400468>>>. 5, 93, 103
- NOURI, B.; WILBERT, S.; BLUM, N.; FABEL, Y.; LORENZ, E.; HAMMER, A.; SCHMIDT, T.; ZARZALEJO, L. F.; PITZ-PAAL, R. Probabilistic solar nowcasting based on all-sky imagers. **Solar Energy**, v. 253, p. 285–307, mar. 2023. ISSN 0038-092X. Disponível em: <<<http://dx.doi.org/10.1016/j.solener.2023.01.060>>>. 24
- NOVALYNX CORPORATION. **260-2500 tipping bucket rain gauge and 260-2500E electric heated rain and snow gauge**. 2024. Disponível em: <<<https://novalynx.com/store/pc/260-2500-Tipping-Bucket-Rain-Gauge-and-260-2500E-Electric-Heated-Rain-and-Snow-Gauge-p210.htm>>>. Acesso em: 30 jul. 2025. 40

ÖKTEM, R.; ROMPS, D. M.; VARBLE, A. C. No warm-phase invigoration of convection detected during goamazon. **Journal of the Atmospheric Sciences**, v. 80, n. 10, p. 2345–2364, out. 2023. ISSN 1520-0469. Disponível em: <<<http://dx.doi.org/10.1175/JAS-D-22-0241.1>>>. 10, 82

PANISSET, J. S.; LIBONATI, R.; GOUVEIA, C. M. P.; MACHADO-SILVA, F.; FRANÇA, D. A.; FRANÇA, J. R. A.; PERES, L. F. Contrasting patterns of the extreme drought episodes of 2005, 2010 and 2015 in the Amazon basin. **International Journal of Climatology**, v. 38, n. 2, p. 1096–1104, 2018. Disponível em: <<<https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.5224>>>. 15

PEREIRA, E.; MARTINS, F.; GONÇALVES, A.; COSTA, R.; LIMA, F.; RÜTHER, R.; ABREU, S.; TIEPOLO, G.; PEREIRA, S.; SOUZA, J. **Atlas brasileiro de energia solar**. Universidade Federal de São Paulo, 2017. ISBN 9788517000898. Disponível em: <<<http://dx.doi.org/10.34024/978851700089>>>. 1

PFISTER, G.; MCKENZIE, R. L.; LILEY, J. B.; THOMAS, A.; FORGAN, B. W.; LONG, C. N. Cloud coverage based on all-sky imaging and its impact on surface solar irradiance. **Journal of Applied Meteorology**, Boston MA, USA, v. 42, n. 10, p. 1421 – 1434, 2003. 26, 27, 28

PIACENTINI, R. D.; SALUM, G. M.; FRAIDENRAICH, N.; TIBA, C. Extreme total solar irradiance due to cloud enhancement at sea level of the ne atlantic coast of brazil. **Renewable Energy**, v. 36, n. 1, p. 409–412, jan. 2011. ISSN 0960-1481. Disponível em: <<<http://dx.doi.org/10.1016/j.renene.2010.06.009>>>. 29, 71

RAZALI, N. M.; WAH, Y. B. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. **Journal of Statistical Modeling and Analytics**, v. 2, n. 1, p. 21–33, 2011. 54

REN21. **Renewables Global Status report**. 2025. Disponível em: <<<https://www.ren21.net/reports/global-status-report/>>>. 1

REYNOLDS, D. W.; MCKEE, T. B.; DANIELSON, K. S. Effects of cloud size and cloud particles on satellite-observed reflected brightness. **Journal of the Atmospheric Sciences**, v. 35, n. 1, p. 160–164, jan. 1978. ISSN 1520-0469. Disponível em: <<[http://dx.doi.org/10.1175/1520-0469\(1978\)035<0160:EOCSAC>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1978)035<0160:EOCSAC>2.0.CO;2)>>. 27

RICHARDSON, W.; CAÑADILLAS, D.; MONCADA, A.; GUERRERO-LEMUS, R.; SHEPHARD, L.; VEGA-AVILA, R.; KRISHNASWAMI, H. Validation of all-sky imager technology and solar irradiance forecasting at three locations: Nrel, san antonio, texas, and the canary islands, spain. **Applied Sciences**, v. 9, n. 4, p. 684, fev. 2019. ISSN 2076-3417. Disponível em: <<<http://dx.doi.org/10.3390/app9040684>>>. 23

ROCHA, V. R. d.; FISCH, G.; COSTA, R. S.; RUANO, A. Anomaly detection in all-sky images: an approach using robust ensemble modeling of cloud cover fraction and prediction bounds. **Engineering Applications of Artificial Intelligence**, v. 143, p. 110003, mar. 2025. ISSN 0952-1976. Disponível em: <<<http://dx.doi.org/10.1016/j.engappai.2025.110003>>>. 34, 54, 65

ROCHA, V. R. d.; YAMASOE, M. A. Estudo da variabilidade espacial e temporal da profundidade óptica do aerossol obtida com o modis sobre a região amazônica. **Revista Brasileira de Meteorologia**, v. 28, n. 2, p. 210–220, jun. 2013. ISSN 0102-7786. Disponível em: <<<http://dx.doi.org/10.1590/S0102-77862013000200010>>>. 50

ROCHA, V. R. da. **Análise de propriedades de nuvens em função da profundidade óptica do aerossol a partir de produtos derivados pelo MODIS da região amazônica durante a estação seca**. Dissertação (Mestrado) — Universidade de São Paulo, São Paulo, 2011. 10

ROCHA, V. R. da; COSTA, R. S.; MARTINS, F. R.; GONÇALVES, A. R.; PEREIRA, E. B. Variability index of solar resource based on data from surface and satellite. **Renewable Energy**, v. 201, p. 354–378, dez. 2022. ISSN 0960-1481. Disponível em: <<<http://dx.doi.org/10.1016/j.renene.2022.10.093>>>. 4

RODERICK, M. L.; FARQUHAR, G. D.; BERRY, S. L.; NOBLE, I. R. On the direct effect of clouds and atmospheric particles on the productivity and structure of vegetation. **Oecologia**, v. 129, n. 1, p. 21–30, set. 2001. ISSN 1432-1939. Disponível em: <<<http://dx.doi.org/10.1007/s004420100760>>>. 30

ROPELEWSKI, C. F.; HALPERT, M. S. Global and regional scale precipitation patterns associated with the el niño/southern oscillation. **Monthly Weather Review**, v. 115, n. 8, p. 1606–1626, ago. 1987. ISSN 1520-0493. Disponível em: <<[http://dx.doi.org/10.1175/1520-0493\(1987\)115<1606:GARSPP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2)>>. 15

ROSSOW, W. B.; SCHIFFER, R. A. Advances in understanding clouds from isccp. **Bulletin of the American Meteorological Society**, v. 80, n. 11, p. 2261–2287, nov. 1999. ISSN 1520-0477. Disponível em: <<[http://dx.doi.org/10.1175/1520-0477\(1999\)080<2261:AIUCFI>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2)>>. 6

RUANO, M. d. G.; RUANO, A. A multi-step ensemble approach for energy community day-ahead net load point and probabilistic forecasting. **Energies**, v. 17, n. 3, p. 696, jan. 2024. ISSN 1996-1073. 46

RÜTHER, R.; NASCIMENTO, L. R. d.; CAMPOS, R. A. Performance assessment issues in utility-scale photovoltaics in warm and sunny climates. **Renewable Energy and Environmental Sustainability**, v. 2, p. 35, 2017. ISSN 2493-9439. Disponível em: <<<http://dx.doi.org/10.1051/rees/2017035>>>. 28

SÁ, G. A. G. de; FONTES, C. H.; EMBIRUÇU, M. A new method for building single feedforward neural network models for multivariate static regression

problems: a combined weight initialization and constructive algorithm. **Evolutionary Intelligence**, v. 17, n. 2, p. 1221–1233, dez. 2022. ISSN 1864-5917. 19

SALATI, E.; MARQUES, J. Climatology of the amazon region. In: SIOLI, H. (Ed.). **The Amazon: limnology and landscape ecology of a mighty tropical river and its basin**. Dordrecht: Dr. W. Junk Publishers, 1984. p. 85–126. Total book: 763 p. 59

SAMU, R.; BHUJUN, S. G.; CALAIS, M.; SHAFIULLAH, G.; MOGHBEL, M.; SHOEB, M. A.; NOURI, B. Solar irradiance nowcasting system trial and evaluation for islanded microgrid control purposes. **Energies**, v. 15, n. 17, p. 6100, ago. 2022. ISSN 1996-1073. Disponível em: <<<http://dx.doi.org/10.3390/en15176100>>>. 25

SANTOS, M. J.; MEDVIGY, D.; DIAS, M. A. F. S.; FREITAS, E. D.; KIM, H. Seasonal flooding causes intensification of the river breeze in the central Amazon. **Journal of Geophysical Research: Atmospheres**, v. 124, n. 10, p. 5178–5197, maio 2019. Disponível em: <<<https://doi.org/10.1029/2018jd029439>>>. 16

SATYAMURTY, P.; COSTA, C. P. W. da; MANZI, A. O. Moisture source for the Amazon basin: a study of contrasting years. **Theoretical and Applied Climatology**, v. 111, n. 1–2, p. 195–209, maio 2012. ISSN 1434-4483. Disponível em: <<<http://dx.doi.org/10.1007/s00704-012-0637-7>>>. 17

SEGAL, M.; DAVIS, J. The impact of deep cumulus reflection on the ground-level global irradiance. **Journal of Applied Meteorology**, v. 31, n. 2, p. 217–222, fev. 1992. ISSN 1520-0450. Disponível em: <<[http://dx.doi.org/10.1175/1520-0450\(1992\)031<0217:TIODCR>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1992)031<0217:TIODCR>2.0.CO;2)>>. 27

SENA, E. T.; DIAS, M. A. F. S.; CARVALHO, L. M. V.; DIAS, P. L. S. Reduced wet-season length detected by satellite retrievals of cloudiness over Brazilian Amazonia: a new methodology. **Journal of Climate**, v. 31, n. 24, p. 9941–9964, dez. 2018. Disponível em: <<<https://doi.org/10.1175/jcli-d-17-0702.1>>>. 16, 61, 62

SENGUPTA, M.; HABTE, A.; ANDREAS, A.; REDA, I.; JAKER, S.; XIE, Y.; YANG, J.; GOTSEFF, P.; KUTCHENREITER, M.; SHI, Y. **Sky Radiometers on Stand for Downwelling Radiation (SKYRAD60S), 2014-01-01 to 2015-12-01, ARM Mobile Facility (MAO), Manacapuru, Amazonas, Brazil; AMF1 (M1)**. [S.l.]: [s.n.], 2014. 34

SETZER, A. W.; SANO, E. E. Cloud cover and cloud types derived from goes-8 satellite data over South America. **Revista Brasileira de Meteorologia**, v. 13, n. 1, p. 35–44, 1998. 41

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, v. 52, n. 3/4, p. 591–611, 1965. 54

SHIRAZI, E.; GORDON, I.; REINDERS, A.; CATTHOOR, F. Sky images for short-term solar irradiance forecast: a comparative study of linear machine learning models. **IEEE Journal of Photovoltaics**, v. 14, n. 4, p. 691–698, jul. 2024. ISSN 2156-3403. Disponível em: <<<http://dx.doi.org/10.1109/JPHOTOV.2024.3398365>>>. 23

SOUSA, A. F. G. de; FURTADO, H. C. M.; MACEDO, W. N.; MENESES, A. A. d. M. Analysis of artificial neural network point forecasting models and prediction intervals for solar irradiance estimation. **American Journal of Engineering and Applied Sciences**, v. 13, n. 3, p. 347–357, mar. 2020. ISSN 1941-7020. Disponível em: <<<http://dx.doi.org/10.3844/ajeassp.2020.347.357>>>. 25, 26

SOUZA, V. R. S. d.; SOUSA, A. R. d.; EMILIAVACA, S. d. A. S.; SILVA, B. d. N. e.; OLIVEIRA, J. V. d.; SILVA, L. S. d. Influência da resolução temporal na análise de eventos de sobreirradiância na cidade de Sousa – Paraíba. In: CONGRESSO BRASILEIRO DE ENERGIA SOLAR, 2024. **Anais...** ABES, 2024. Disponível em: <<<http://dx.doi.org/10.59627/cbens.2024.2494>>>. 28

STRAUB, N.; HERZBERG, W.; DITTMANN, A.; LORENZ, E. Blending of a novel all sky imager model with persistence and a satellite based model for high-resolution irradiance nowcasting. **Solar Energy**, v. 269, p. 112319, fev. 2024. ISSN 0038-092X. Disponível em: <<<http://dx.doi.org/10.1016/j.solener.2024.112319>>>. 24

TAPAKIS, R.; CHARALAMBIDES, A. Enhanced values of global irradiance due to the presence of clouds in eastern mediterranean. **Renewable Energy**, v. 62, p. 459–467, fev. 2014. ISSN 0960-1481. Disponível em: <<<http://dx.doi.org/10.1016/j.renene.2013.08.001>>>. 28

THUILLIER, G.; PERRIN, J.-M.; KECKHUT, P. L.; HUPPERT, F. Local enhanced solar irradiance on the ground generated by cirrus: measurements and interpretation. **Journal of Applied Remote Sensing**, v. 7, n. 1, p. 073543, 2013. Disponível em: <<<https://doi.org/10.1117/1.JRS.7.073543>>>. 27

TIAN, Y.; ZHANG, Y.; KLEIN, S. A.; SCHUMACHER, C. Interpreting the diurnal cycle of clouds and precipitation in the arm goamazon observations: shallow to deep convection transition. **Journal of Geophysical Research: Atmospheres**, v. 126, n. 5, mar. 2021. ISSN 2169-8996. Disponível em: <<<http://dx.doi.org/10.1029/2020JD033766>>>. 10, 11, 12, 13

TIBA, C.; LEAL, S. d. S. Enhancement of uv radiation by cloud effect in ne of Brazil. **International Journal of Photoenergy**, v. 2017, p. 1–9, 2017. ISSN 1687-529X. Disponível em: <<<http://dx.doi.org/10.1155/2017/8107435>>>. 30

TOLEDO, N.; MOULATLET, G.; GAONA, G.; VALENCIA, B.; HIRATA, R.; CONICELLI, B. Dynamics of meteorological and hydrological drought: the impact of groundwater and El Niño events on forest fires in the Amazon. **Science of The Total Environment**, v. 954, p. 176612, dez. 2024. ISSN 0048-9697. Disponível em: <<<http://dx.doi.org/10.1016/j.scitotenv.2024.176612>>>. 15

TWOMEY, S. The influence of pollution on the shortwave albedo of clouds. **Journal of the Atmospheric Sciences**, v. 34, n. 7, p. 1149–1152, jul. 1977. ISSN 1520-0469. Disponível em: <<[http://dx.doi.org/10.1175/1520-0469\(1977\)034<1149:TIOPOP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1977)034<1149:TIOPOP>2.0.CO;2)>>. 10

UNITED STATES NAVAL OBSERVATORY. **Earth's Seasons – Equinoxes, Solstices, Perihelion, and Aphelion**. 2025. Disponível em: <<https://aa.usno.navy.mil/data/Earth_Seasons>>. Acesso em: 27 jul. 2025. 59

VATS, D. S. Solar irradiance prediction in the amazon basin using machine learning: a sustainable approach for renewable energy expansion. **European Economic Letters (EEL)**, v. 14, n. 2, p. 3718–3730, jul. 2024. Disponível em: <<<https://www.eelet.org.uk/index.php/journal/article/view/1739>>>. 25, 26

VIRTANEN, P. e. a. Scipy 1.0: fundamental algorithms for scientific computing in python. **Nature Methods**, v. 17, n. 3, p. 261–272, 2020. Disponível em: <<https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize_scalar.html>>. 53

WORLD METEOROLOGICAL ORGANIZATION (WMO). **International Cloud Atlas**. 2017. Disponível em: <<<https://cloudatlas.wmo.int/en/home.html>>>. 6, 16

YAMASOE, M. A.; CORRÊA, M. d. P. **Processos radiativos na atmosfera: fundamentos**. São Paulo, Brasil: Oficina de Textos, 2016. ISBN 978-85-7975-229-2. 5, 59

YORDANOV, G. H.; SAETRE, T. O.; MIDTGÅRD, O.-M. Extreme overirradiance events in norway: 1.6 suns measured close to 60°n. **Solar Energy**, v. 115, p. 68–73, maio 2015. ISSN 0038-092X. Disponível em: <<<http://dx.doi.org/10.1016/j.solener.2015.02.020>>>. 28

ZEHNER, M.; HARTMANN, M.; WEIZENBECK, J.; GRATZL, T.; WEIGL, T.; MAYER, B.; WIRTH, G.; KRAWCZYNSKI, M.; BETTS, T.; GOTTSCHALG, R.; HAMMER, A.; GIESLER, B.; BECKER, G.; MAYER, O. Systematic analysis of meteorological irradiation effects. In: PHOTOVOLTAIC SOLAR ENERGY CONFERENCE, 25., 2010. **Proceedings...** Valencia, Spain, 2010. p. 6–10. 27

ZUO, H.-M.; QIU, J.; LI, F.-F. Ultra-short-term forecasting of global horizontal irradiance (GHI) integrating all-sky images and historical sequences. **Journal of Renewable and Sustainable Energy**, v. 15, n. 5, p. 053701, 09 2023. ISSN 1941-7012. Disponível em: <<<https://doi.org/10.1063/5.0163759>>>. 5, 24, 52

APÊNDICE A - ALGORITMO DE CONSTRUÇÃO DO APPROXHULL

Algorithm 1 Algoritmo de Construção do ApproxHull. Fonte: adaptado de [Khosravani et al. \(2016\)](#)

```
1: Escalar cada dimensão do conjunto de dados para o intervalo  $[-1, 1]$ .
2: Identificar as amostras de valor máximo e mínimo em cada dimensão.
3: Definir essas amostras como os vértices iniciais do envoltório convexo.
4: while os critérios de parada não forem atendidos do
5:     Gerar uma população de  $k$  facetas baseada nos vértices atuais do envoltório.
6:     for cada faceta  $F$  composta por  $d$  pontos em  $D$  dimensões do
7:         if  $F$  não for uma matriz válida (posto completo1) then
8:             Substituí-la por outra combinação até encontrar uma faceta válida.
9:         end if
10:        (Opcional) Introduza pequenas perturbações aleatórias (método de joggling).
11:    end for
12:    for cada face do
13:        Identifique os pontos mais distantes como novos vértices do casco convexo.
14:        Obtenha a equação do hiperplano da faceta a partir do vetor normal e do deslocamento.
15:        Calcule a distância de cada amostra ao hiperplano.
16:        Selecione a amostra com maior distância como o ponto mais distante.
17:    end for
18:    Atualize o casco convexo atual adicionando os novos vértices encontrados.
19:    Defina  $d_c$  como a distância máxima dos pontos mais distantes ao casco convexo atual.
20:    Verifique os critérios de parada:
21:    if nenhum novo vértice for encontrado then
22:        Termine o algoritmo.
23:    else if a diferença entre o máximo e o mínimo de  $d_c$  nas últimas  $w$  iterações for menor que um limiar (valor padrão: 0,1) then
24:        Termine o algoritmo.
25:    end if
26: end while
```

A motivação para o segundo critério de parada reside na necessidade de evitar a seleção de novos vértices que não contribuam significativamente para o formato do envoltório convexo, otimizando assim o esforço computacional.