



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA MECÂNICA
CURSO DE ENGENHARIA DE ENERGIAS RENOVÁVEIS

MIRELLA MARTINS CAMELO

**IRRADIÂNCIA SOLAR SINTÉTICA A PARTIR DO SENSORIAMENTO REMOTO
EM SATÉLITES GEOESTACIONÁRIOS: VALIDAÇÃO DOS SERVIÇOS SOLCAST
E CAMS RADIATION EM PETROLINA-PE, NORDESTE BRASILEIRO**

FORTALEZA

2021

MIRELLA MARTINS CAMELO

IRRADIÂNCIA SOLAR SINTÉTICA A PARTIR DO SENSORIAMENTO REMOTO EM
SATÉLITES GEOESTACIONÁRIOS: VALIDAÇÃO DOS SERVIÇOS SOLCAST E CAMS
RADIATION EM PETROLINA-PE, NORDESTE BRASILEIRO

Trabalho de Conclusão de Curso apresentado à
Universidade Federal do Ceará, como requisito
parcial à obtenção do título de Engenheiro de
Energias Renováveis

Orientador: Prof. Dr. Paulo Alexandre Costa
Rocha.

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- C189i Camelo, Mirella Martins.
 Irradiância solar sintética a partir do sensoriamento remoto em satélites geoestacionários :
 validação dos serviços Solcast e CAMS Radiation em Petrolina-PE, Nordeste brasileiro / Mirella
 Martins Camelo. – 2021.
 75 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de
 Tecnologia, Curso de Engenharia de Energias Renováveis, Fortaleza, 2021.
 Orientação: Prof. Dr. Paulo Alexandre Costa Rocha.
1. Sensoriamento remoto. 2. Dados de satélites. 3. Radiação. 4. Satélites geoestacionários.
 5. Energia solar. I. Título.

CDD 621.042

MIRELLA MARTINS CAMELO

IRRADIÂNCIA SOLAR SINTÉTICA A PARTIR DO SENSORIAMENTO REMOTO EM
SATÉLITES GEOESTACIONÁRIOS: VALIDAÇÃO DOS SERVIÇOS SOLCAST E CAMS
RADIATION EM PETROLINA-PE, NORDESTE BRASILEIRO

Trabalho de Conclusão de Curso apresentado à
Universidade Federal do Ceará, como requisito
parcial à obtenção do título de Engenheiro de
Energias Renováveis

Aprovada em: ___/___/_____.

BANCA EXAMINADORA

Prof. Dr. Paulo Alexandre Costa Rocha (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Ana Fabíola Leite Almeida
Universidade Federal do Ceará (UFC)

Prof. Dr. Paulo Cesar Marques de Carvalho
Universidade Federal do Ceará (UFC)

À minha família.

AGRADECIMENTOS

Aos professores Drs. Carlos Estêvão Rolim Fernandes, Natália Maria Cordeiro Barroso e Luis Gonzaga Rodrigues Filho, pela dedicação ao ensino que sempre tiveram.

À Eduarda, Gabryel e Luiz, meus primeiros anos de graduação foram incríveis e menos difíceis por conta da companhia de vocês.

À Alessia, Gabriela, Máyla, Valéria, Rebeca e Ruthe pela nossa longa amizade e pelo apoio em todos os momentos. À Lara, Jocélia e Pamella pela inesperada amizade que sei que durará pela vida.

Aos companheiros de RETECjr, por toda a nossa troca de conhecimento, crescimento profissional e evolução pessoal coletiva.

Ao Júnior, pelo carinho, acolhimento e por todo o suporte durante nossa jornada.

À Isabella, por ser uma das minhas maiores referências.

Aos meus pais, Jacmary e Silvio, pela criação amorosa e esforço diário na construção das possibilidades necessárias para realizarmos nossos sonhos.

Ao Prof. Dr. Paulo Alexandre Costa Rocha, pela brilhante orientação.

Aos professores participantes da banca examinadora, Profa. Dra. Ana Fabíola Leite Almeida e Prof. Dr. Paulo Cesar Marques de Carvalho pela disponibilidade e colaborações.

“Diante da vastidão do tempo e da imensidão do universo, é um imenso prazer para mim dividir um planeta e uma época com você.”
(Carl Sagan)

RESUMO

Como consequência do crescente esforço em pesquisas voltadas ao aprimoramento do uso da energia solar na matriz elétrica, há nos últimos anos uma expressiva relevância na análise e validação dos dados de irradiância dos produtos de satélite. Isto ocorre devido às vantagens relacionadas com os serviços destes produtos: o acesso simples e fácil à internet, a aquisição de baixo custo (alguns deles são serviços gratuitos) e a cobertura espacial praticamente ao longo de toda a extensão global. Por essas razões, pode-se obter dados de irradiância sem a necessidade de caros equipamentos de medição, reduzindo custos para fins acadêmicos ou comerciais. Entre os produtos de satélites geoestacionários, as radiações sintéticas imitam o comportamento estatístico de observações reais e são oriundas de modelagem computacional com base física, entretanto, possuem erros de estimativa. Nesse sentido, o presente estudo tem como objetivo validar os produtos de satélite dos serviços CAMS Radiation Service e Solcast a partir da comparação entre estimativas de Irradiância Global Horizontal e Irradiância Direta Normal (DNI) com observações reais coletadas da estação de medição terrestre localizada na cidade de Petrolina, Nordeste do Brasil, sob um clima tropical. A estação pertence ao Instituto Nacional de Pesquisas Espaciais (INPE), fazendo parte da rede SONDA (Sistema de Organização Nacional de Dados Ambientais). Além disso, este trabalho também tem como objetivo analisar as causas de possíveis variabilidades na precisão obtida, como os efeitos das estações do ano no clima local. Ademais, a verificação da precisão dos produtos de satélite é feita pelo cálculo das métricas de erro: Raiz do Erro Quadrático Médio (RMSE), Erro Médio Absoluto (MAE) e correlação R^2 , assim como a análise gráfica da distribuição dos dados. As métricas de erro são aplicadas em um banco de dados de irradiância solar para os anos de 2004-2018, com todas as resoluções temporais disponibilizadas por cada serviço.

Palavras-chave: Sensoriamento remoto; Dados de satélites; Radiação; Satélites geoestacionários; Energia solar.

ABSTRACT

Due to the increase of research efforts focused on improvements of solar energy utilization, there is in the last years an expressive relevance about the analysis and validation of the accuracy of irradiance data from satellite products. This occurs due to the advantages related to these satellite products services: the simple and easy access to the internet, the low-cost purchase (some of them are free data services) and the wide data availability for essentially the entire global territory. For these reasons, one can obtain radiation data without the need of expensive measurement equipment, thus reducing costs for academic or commercial purposes. Among the products of geostationary satellites, synthetic radiation imitates the statistical behavior of real observations and comes from computational modeling with a physical basis, however, they have estimation errors. In this sense, this study aims to validate the satellite products CAMS Radiation Service and Solcast by the comparison against the irradiance data for Global Horizontal Radiation (GHI) and Normal Direct Radiation (DNI) of a ground-based measurement station located in the city of Petrolina, Northeastern Brazil, under a tropical climate. The station belongs to the Brazilian National Institute of Spatial Research (INPE) and is part of the SONDA network (National Environmental Data Organization System). In addition, this work has also the purpose of analyzing the causes of possible variability in the accuracy obtained, such as the effects of the seasons on the local climate. Furthermore, the accuracy verification of the satellite products is done by the calculation of the error metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R^2 correlation. The error metrics are applied on a solar radiation database for the years ranging from 2004-2018, with a 60-minute temporal resolution.

Keywords: Remote sensing; Satellite data. Radiation; Geostationary satellite; Solar energy.

LISTA DE FIGURAS

Figura 1 – Evolução temática.....	30
Figura 2 – Observações de um satélite com órbita polar.....	37
Figura 3 – Posição e observação de alguns satélites geoestacionários, classificados em cor de acordo com suas resoluções temporais, vide legenda.....	37
Figura 4 – Ilustração da irradiância recebida pelo sensoriamento.....	39
Figura 5 – Fluxograma de blocos simples do processo de validação dos produtos de satélite CAMS e Solcast.....	40
Figura 6 – Localização da estação da rede SONDA em Petrolina, Pernambuco.....	42
Figura 7 – Fonte de coleta do Serviço CAMS de irradiância para ilustração de como os dados são coletados.....	43
Figura 8 – Fonte de coleta do Serviço Solcast para ilustração de como os dados do local são inseridos antes do pedido.....	44
Figura 9 – Escolha dos parâmetros no site do Serviço Solcast.....	44

LISTA DE GRÁFICOS

Gráfico 1	– Evolução do consumo de energia até 2050.....	20
Gráfico 2	– Produção científica anual.....	24
Gráfico 3	– Produção no tempo pelos autores mais relevantes.....	26
Gráfico 4	– Crescimento do tema nas principais fontes científicas.....	27
Gráfico 5	– Termos em tendência.....	29
Gráfico 6	– Dispersão com distribuição de frequências entre os dados de GHI estimados pelo CAMS-RAD vs. Medidos pela estação do INPE (60 min)..	51
Gráfico 7	– Dispersão com distribuição de frequências entre os dados de GHI estimados pelo Solcast vs. Medidos pela estação do INPE (60 min)	52
Gráfico 8	– Dispersão com distribuição de frequências entre os dados de DNI estimados pelo CAMS-RAD vs. Medidos pela estação do INPE (60 min)..	52
Gráfico 9	– Dispersão com distribuição de frequências entre os dados de DNI previstos pelo Solcast vs. Medidos pela estação do INPE (60 min)	52
Gráfico 10	– Erros de GHI e DNI estimados pelo CAMS-RAD para cada ano observado, em resolução temporal de 60 minutos.....	55
Gráfico 11	– Erros de estimação da GHI e DNI pelo Solcast para cada ano observado, em resolução temporal de 60 minutos.....	55
Gráfico 12	– CAMS-RAD: Distribuição da irradiância GHI medida dados valores estimados (à esquerda) e distribuição da GHI estimada dados valores medidos (à direita)	57

Gráfico 13 – CAMS-RAD: Distribuição da DNI medida dados valores estimados (à esquerda) e distribuição da DNI estimada dados valores medidos (à direita)	57
Gráfico 14 – Solcast: Distribuição da GHI medida dados valores estimados (à esquerda) e distribuição da GHI estimada dados valores medidos (à direita)	59
Gráfico 15 – Solcast: Distribuição da DNI medida dados valores estimados (à esquerda) e distribuição da DNI estimada dados valores medidos (à direita)	59
Gráfico 16 – Erros de estimação da GHI e DNI pelo CAMS-RAD para cada estação do ano, em resolução temporal de 60 minutos.....	61
Gráfico 17 – Série temporal da irradiância global horizontal medida (INPE) e estimada (CAMS) ao longo de 3 dias com erro associado.....	62
Gráfico 18 – Série temporal da irradiância direta normal medida (INPE) e estimada (CAMS) ao longo de 3 dias com erro associado.....	62
Gráfico 19 – Distribuição da irradiância global horizontal estimada (CAMS) e medida (INPE) ao longo das estações do ano (acima) e distribuição da GHI total (abaixo).....	63
Gráfico 20 – Distribuição da irradiância direta normal estimada (CAMS) e medida (INPE) ao longo das estações do ano (acima) e distribuição da DNI total (abaixo)	63
Gráfico 21 – Histograma 2D da irradiância global horizontal estimada pelo CAMS vs. Medida pelo INPE.....	64
Gráfico 22 – Histograma 2D da irradiância direta normal estimada pelo CAMS vs. Medida pelo INPE.....	65
Gráfico 23 – Erros de previsão da GHI e DNI pelo Solcast para cada estação do ano, em resolução temporal de 60 minutos.....	66
Gráfico 24 – Série temporal da irradiância global horizontal medida (INPE) e estimada (Solcast) ao longo de 3 dias com erro associado.....	67

Gráfico 25 – Série temporal da irradiância direta normal medida (INPE) e estimada (Solcast) ao longo de 3 dias com erro associado.....	67
Gráfico 26 – Distribuição da GHI (Solcast) e medida (INPE) ao longo das estações do ano (acima) e distribuição da GHI total (abaixo).....	68
Gráfico 27 – Distribuição da irradiância direta normal estimada (Solcast) e medida (INPE) ao longo das estações do ano (acima) e distribuição da DNI total (abaixo)	68
Gráfico 28 – Histograma 2D da irradiância global horizontal estimada pelo Solcast vs. Medida pelo INPE.....	69
Gráfico 29 – Histograma 2D da irradiância direta normal estimada pelo Solcast vs. Medida pelo INPE.....	69

LISTA DE TABELAS

Tabela 1 – Principais informações do banco de dados	23
Tabela 2 – Top 20 autores mais relevantes	24
Tabela 3 – Top 20 palavras-chave mais relevantes	28
Tabela 4 – Top 20 países com mais publicações	31
Tabela 5 – Top 20 países mais citados em publicações	32
Tabela 6 – Resumo das principais missões satelitais durante a história	38
Tabela 7 – Informações de cada dataframe montado, valores brutos correspondentes a antes da qualificação	45
Tabela 8 – CAMS-RAD: CAMS-RAD: Erros associados à estimação da GHI e DNI, para os dois grupos de dados e resolução temporal de 60 minutos	50
Tabela 9 – Solcast: Erros associados à estimação da GHI e DNI, para os dois grupos de dados e resolução temporal de 60 minutos	50
Tabela 10 – CAMS-RAD: Erros anuais de estimação da GHI e DNI, resolução temporal de 60 minutos	53
Tabela 11 – Solcast: Erros anuais de estimação da GHI e DNI, em resolução temporal de 60 minutos	54
Tabela 12 – CAMS-RAD: Erros associados à estimação da GHI e DNI por estação do ano, em resolução temporal de 60 minutos (Grupo II)	60
Tabela 13 – Solcast: Erros associados à estimação da GHI e DNI, por estação do ano, em resolução temporal de 60 minutos (Grupo II)	66
Tabela 14 – Erros associados à estimação da GHI em variadas resoluções temporais (Grupo II)	70
Tabela 15 – Erros associados à estimação da DNI em variadas resoluções temporais (Grupo II)	70

LISTA DE ABREVIATURAS E SIGLAS

ANEEL	Agência Nacional de Energia Elétrica
CAMS	Copernicus Atmosphere Monitoring Service (Serviço de monitoramento atmosférico Copernicus)
DNI	Direct Normal Irradiance (Irradiância Direta Normal)
EPE	Empresa de Pesquisa Energética
GHI	Global Horizontal Irradiance (Irradiância Global Horizontal)
INPE	Instituto Nacional de Pesquisa Espacial
MAE	Mean Absolute Error (Erro Médio Absoluto)
MME	Ministério de Minas e Energia
PNE	Plano Nacional de Energia
REN	Resolução Normativa
RMSE	Root Mean Squared Error (Raíz do Erro Quadrático Médio)
tep	Tonelada equivalente de petróleo

LISTA DE SÍMBOLOS

- ® Marca Registrada
- θ_z Ângulo zênite
- μ Unidade de medida equivalente a 10^{-6} metros

SUMÁRIO

1	INTRODUÇÃO	18
1.1	Objetivos	21
<i>1.1.1</i>	<i>Objetivo geral</i>	21
<i>1.1.2</i>	<i>Objetivos específicos</i>	21
1.2	Organização do trabalho	21
2	ANÁLISE BIBLIOMÉTRICA	23
32.1	Autores e produção científica	24
2.2	Tendência temática	28
2.3	Produção científica nos países	31
3	FUNDAMENTAÇÃO TEÓRICA	33
3.1	Radiação solar	33
3.2	Irradiância solar sintética	34
3.3	Princípios da estimação de irradiância solar por satélites	35
<i>3.3.1</i>	<i>Conceitos básicos</i>	35
<i>3.3.1.1</i>	<i>Condições do céu e ângulo zênite em agrupamento de dados</i>	35
<i>3.3.1.2</i>	<i>Resolução e cobertura dos satélites</i>	35
<i>3.3.2</i>	<i>Panorama geral</i>	36
<i>3.3.3</i>	<i>Tipos de modelos</i>	38
<i>3.3.3.1</i>	<i>Modelos estatísticos e modelos semi-empíricos</i>	39
<i>3.3.3.2</i>	<i>Modelos físicos</i>	39
4	METODOLOGIA	40
4.1	Banco de dados	41
<i>4.1.1</i>	<i>Coleta de dados</i>	41
<i>4.1.1.1</i>	<i>Coleta das observações reais</i>	41
<i>4.1.1.2</i>	<i>Coleta no CAMS</i>	42
<i>4.1.1.3</i>	<i>Coleta no Solcast</i>	43
<i>4.1.2</i>	<i>Montagem do banco de dados</i>	45
4.2	Protocolo de validação	46
<i>4.2.1</i>	<i>Métricas de erro</i>	46
<i>4.2.2</i>	<i>Análise da distribuição de dados</i>	47
5	RESULTADOS E DISCUSSÃO	49

5.1	Erros gerais da estimação de irradiância e qualificação de dados	49
5.2	Acurácia por ano.....	53
5.3	Estudo da distribuição de dados	56
5.4	Efeitos da sazonalidade	60
5.4.1	<i>CAMS-RAD</i>	60
5.4.2	<i>Solcast</i>	65
5.5	Outras resoluções temporais	70
6	CONCLUSÃO	72
	REFERÊNCIAS	73

1 INTRODUÇÃO

A incidência de radiação solar é uma das informações mais importantes para diversas áreas, como em estudos do clima, a meteorologia e suas previsões e também outras aplicações, especialmente as que tangem as energias renováveis. A “radiação solar superficial [...] é um parâmetro chave que dita as interações entre a superfície [da Terra] e a atmosfera, impulsionando processos radiativos, hidrológicos [...], e pode, portanto, afetar fortemente o clima” (HUANG *et al.*, 2019, p. 1, tradução nossa), tornando-se indispensável o conhecimento de dados de várias componentes de radiação.

Dados de radiação solar superficial podem ser adquiridos diretamente por medições terrestres através do piranômetro, um equipamento que converte radiação solar em sinais elétricos. Apesar das incertezas envolvidas no processo de medição e calibração, são considerados acurados e de alta qualidade. Todavia, equipamentos de medição resultam em dados pontuais e necessitam de manutenção frequente, não sendo adequados para aplicações em larga escala.

Dentro deste contexto, o fornecimento de dados radiométricos obtidos a partir do sensoriamento remoto realizado por satélites geoestacionários se tornou quase mandatório nas últimas décadas e são geralmente referidos como “produtos de satélite”, ou também como “irradiância sintética”, termo sugerido pelo autor Bright (2021). Os serviços de satélite “[...] oferecem uma grande cobertura geográfica, mas possuem incertezas relativamente grandes em suas estimativas de [Irradiância Global Horizontal] GHI, principalmente devido à resolução espacial e erros de modelagem” (KAMATH; SRINIVASAN, 2020, p. 45, tradução nossa).

Segundo Huang *et al.* (2019) a história da estimação de radiação solar superficial por sensoriamento remoto apresenta passos iniciais em 1960 com o lançamento do satélite TIROS-1 (Television Infrared Observation Satellites), porém, apenas a partir dos anos 70 foram realizadas as primeiras estimativas de irradiância solar. O período entre os anos 1980 e 2000 é marcado como o processo de desenvolvimento e aprimoramento dos diversos algoritmos utilizados na modelagem das estimativas. Por fim, as últimas duas décadas podem ser resumidas com o avanço tecnológico dos sensores e por métodos de modelagem mais robustos, principalmente em consequência de maiores capacidades computacionais. O artigo dos autores Huang *et al.* (2019) apresenta um maior detalhamento histórico.

Polo e Perez (2019, p. 183, 185) introduzem a ideia básica acerca do funcionamento do sensoriamento remoto em satélites

As observações de satélite são geralmente incorporadas entre modelos de previsão numérica do tempo (NWP) por meio de algoritmos de assimilação de dados, [com objetivo de] produzir a melhor estimativa do estado atmosférico e melhorar a previsão do tempo. [...] Sensores a bordo de satélites meteorológicos recebem ondas curtas de radiação (solar) ressurgentes do sistema terra-atmosfera. Isto incorpora as radiações solares refletidas pela superfície terrestre e/ou pelas nuvens e a radiação retrodispersa proveniente da interação entre a radiação solar incidente com a atmosfera. Portanto, é possível aferir a radiação solar descendente na superfície terrestre [...] [a partir do sensoriamento remoto] se algumas informações atmosféricas estiverem disponíveis [...].

Ou seja, os sensores a bordo dos satélites capturam imagens atmosféricas que ilustram as ondas de radiação ressurgentes e, então, modelos computacionais são aplicados nestas imagens produzindo dados de radiação solar superficial sintéticas, e estes são oferecidos ao público por meio de serviços, gratuitos ou não.

O estudo, a avaliação e o desenvolvimento de produtos de satélite são áreas que continuam em alta mesmo após seis décadas de publicações extensivas. Isto é justificado pelo apelo da produção de energia solar, que apresenta enorme potencial de crescimento no mundo, devido aos objetivos de desenvolvimento sustentável, à transição energética e à descarbonização de diversos setores da economia. No Brasil, a produção de energia solar foi impulsionada pela Resolução Normativa (REN) 482 da Agência Nacional de Energia Elétrica (ANEEL) em 2012, que regularizou a geração distribuída em todo o país e proporcionou incentivos fiscais para a expansão da modalidade. Outro marco da geração distribuída foi no ano de 2015, com a alteração da REN-482 pela REN-687. Em 2019 as resoluções normativas entraram novamente em revisão, porém as proposições geraram respostas negativas por parte da sociedade e profissionais da área.

De acordo com o Plano Nacional de Energia – PNE 2050 (BRASIL, 2020), elaborado pelo Ministério de Minas e Energia (MME) e a Empresa de Pesquisa Energética (EPE), é esperado que a demanda de energia saia de 300 milhões de tep (tonelada equivalente de petróleo) em 2015 e alcance cerca de 600 milhões de tep em 2050. O Gráfico 1 apresenta a evolução esperada do consumo de energia, resultando em uma taxa de crescimento médio de 2,2% ao ano no cenário da expansão.

No Anuário Estatístico de Energia Elétrica de 2020 (EMPRESA DE PESQUISA ENERGÉTICA, 2020), a geração fotovoltaica em 2015 era de 59 GWh no Brasil e, em 2019, atingiu 6.651 GWh, representando apenas 1,1% de participação da matriz energética.

Apesar da expressiva participação de fontes alternativas, o atual agravamento da crise hídrica em 2021, que apresenta baixíssimos patamares de nível nos reservatórios

hidroelétricos, com volume útil de 43,65% no dia 31 de outubro de 2021 (AGÊNCIA NACIONAL DE ÁGUAS, 2021), respalda a urgência de uma maior diversificação da matriz.

Gráfico 1 – Evolução do consumo de energia até 2050



Fonte: Brasil (2020, p. 21)

Um dos maiores desafios da atualidade é a expansão das fontes de energia renováveis intermitentes na matriz elétrica brasileira, que pode ocasionar problemas de flutuação de tensão e instabilidade. Então, devido à conjuntura atual, se torna imprescindível a obtenção em larga escala de dados radiométricos confiáveis e acurados para estudos acadêmicos de desenvolvimento da energia solar, ou para fins comerciais.

Trabalhos de validação dos produtos de satélite já foram realizados intensivamente na literatura. Apesar disto, uma análise mais detalhada no território brasileiro e em seus diversos climas agrega à área de estudo, visto o interesse nacional em seu enorme potencial fotovoltaico, principalmente quando se trata da região Nordeste. Assim, a partir disso, a motivação para a escolha desse tema é avaliar a acurácia de dados de irradiância sintética oriundos de dois dos principais serviços de satélite, a partir de um estudo de caso. A cidade escolhida para o estudo é Petrolina (09°04'08"S, 40°19'11"O, Altitude 387 m), em Pernambuco, no Nordeste brasileiro.

1.1 Objetivos

1.1.1 Objetivo geral

Este trabalho tem como objetivo geral a validação dos dados sintéticos de irradiância a partir de dois produtos de satélite, o CAMS Radiation Service (CAMS Rad) e o Solcast, em um estudo de caso para Petrolina, em Pernambuco.

1.1.2 Objetivos específicos

Como objetivos específicos:

- a) Análise bibliométrica do tema;
- b) Verificação e comparação de acurácia das estimativas feitas pelos satélites escolhidos, a partir de métricas de performance;
- c) Estudo gráfico da distribuição estatística das estimativas feitas pelos satélites escolhidos;
- d) Avaliação das estimativas feitas pelos satélites escolhidos nas quatro estações do ano.

1.2 Organização do trabalho

O presente trabalho possui seis capítulos: introdução, análise bibliométrica, fundamentação teórica, metodologia, resultados e discussão e, por fim, conclusão. O primeiro capítulo trata da introdução do tema, apontando contextualizações gerais e nacionais sobre irradiância sintética, a justificativa para a escolha do tema, assim como os objetivos do trabalho.

O segundo capítulo, a análise bibliométrica apresenta uma descrição quantitativa acerca das publicações científicas de irradiância oriunda do sensoriamento remoto de satélites, com artigos compreendidos entre o período de 1964 e 2021. Neste, há uma caracterização da produção científica anual, principais autores, palavras chaves, países e evolução temática.

No terceiro capítulo, a fundamentação teórica introduz os aspectos mais relevantes, como a definição de radiação solar e suas componentes, além do significado de irradiância sintética. Por fim, os princípios da estimação da irradiância pelo sensoriamento remoto,

incluindo conceitos básicos de termos mais utilizados na área, histórico das missões de alguns satélites meteorológicos e classificação dos modelos computacionais.

O quarto capítulo sobre a metodologia do trabalho descreve o processo de coleta dos dados de irradiâncias reais e sintéticas, de construção e qualificação do banco de dados e de validação das estimativas realizadas pelos modelos dos satélites.

Nos resultados e discussão, estão apontados os erros das estimativas por grupos, por ano, estação do ano e por resolução temporal. Além disso, é realizada a análise gráfica distributiva dos dados de irradiância.

Ademais, no último capítulo são analisadas as principais conclusões obtidas a partir da discussão dos resultados, expondo se os produtos de satélite são confiáveis e acurados para a região estudada.

2 ANÁLISE BIBLIOMÉTRICA

A análise bibliométrica proporciona ferramentas que resultam em uma descrição quantitativa das publicações científicas acerca de qualquer tema, como a investigação dos autores e referências mais relevantes, produção anual, identificação das principais revistas científicas e pontuar os termos ou palavras-chave mais utilizadas. Além disso, é possível obter mapeamentos das tendências do tópico em estudo. Sendo assim, essa visualização quantitativa fundamenta o panorama geral de um tema e é importante no reconhecimento das principais tendências globais de pesquisa ou das lacunas que precisam de maiores esforços de pesquisa.

Poucos trabalhos sobre sensoriamento remoto e produtos de satélite realizaram análise bibliométrica, o mais recente foi publicado pelos autores Scheidsteger e Haunschild (2020), porém eles apenas analisaram as principais referências no decorrer de alguns marcos temporais. Logo, o presente trabalho é ampliado com a análise dos principais autores, da produção científica anual, de tendências temáticas e estudo geográfico das publicações.

No presente trabalho foi utilizada a função *biblioshiny* do pacote *Bibliometrix* no R, sugerida por Aria e Cuccurullo (2017). A bibliometria foi coletada na base de dados Scopus, no dia seis de junho de 2021, utilizando as chaves de pesquisa “**satellite E solar E model OU modeling E irradiance OU irradiation**”. Essa coleta obteve 1377 documentos de publicações, porém, foi realizada uma filtragem para incluir apenas artigos, resultando 958 documentos para a análise.

A Tabela 1 resume as principais informações bibliométricas, como o período de abrangência, quantidade total de fontes científicas com artigos publicados, número de artigos, número de autores, de palavras-chave, entre outros. Foram 958 artigos publicados entre 1965 e 2021, totalizando 3236 autores e média de 3,38 autores por publicações sobre o tema.

Tabela 1 – Principais informações do banco de dados

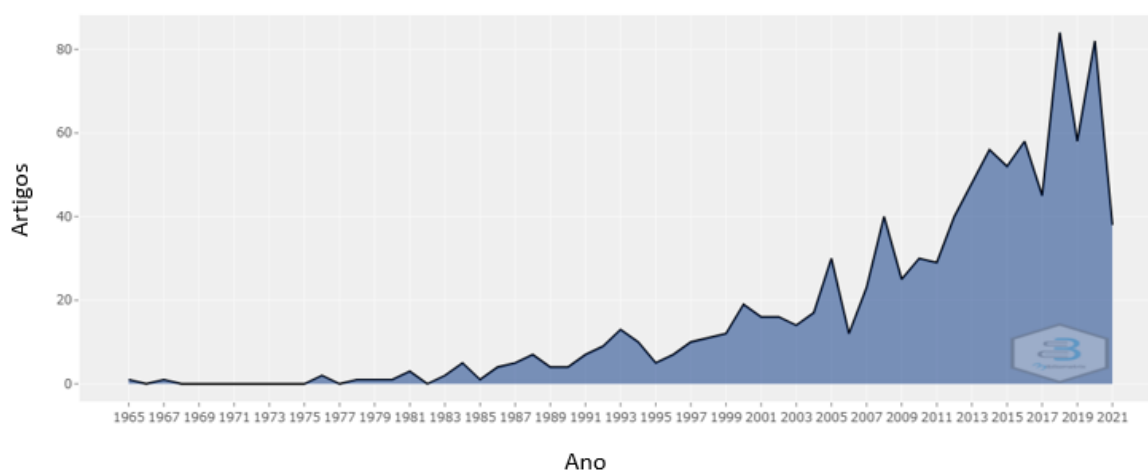
Principais informações	Resultado
Período	1965 - 2021
Fontes científicas	240
Artigos	958
Autores	3236
Palavras-chaves	1780
Publicações por autor	0,296
Autores por publicação	3,38

Fonte: Adaptado da base de dados Scopus.

2.1 Autores e produção científica

Partindo das principais informações da Tabela 1, uma análise descritiva dos 958 artigos é apresentada no Gráfico 2, que ilustra o número de artigos publicados anualmente desde 1965.

Gráfico 2 – Produção científica anual, representa o número de artigos publicados por ano entre 1965 e 2021



Fonte: Adaptado da base de dados Scopus.

É notável que, mesmo após cinco décadas, a publicação sobre produtos de satélite continua em crescimento expressivo, com mais de 80 documentos em 2018 e 2020, demonstrando que o tópico não se encontra saturado. Nesse contexto de importância do tema, a Tabela 2 apresenta os 20 autores com maior número de publicações.

Tabela 2 – Top 20 autores mais relevantes

Autor	Nº de publicações
WOODS TN	26
POLO J	15
SOLOMON SC	15
YANG D	15
KAZADZIS S	13
WALD L	12
CHAMBERLIN PC	11
PILEWSKIE P	11
GUEYMARD CA	10
TOBISKA WK	10
HAIGH JD	9

	(conclusão)
KATO S	9
MARTINS FR	9
PEREIRA EB	9
WILLSON RC	9
JANJAI S	8
LIZ	8
LIANG S	8
QIAN L	8
ROZANOV E	8

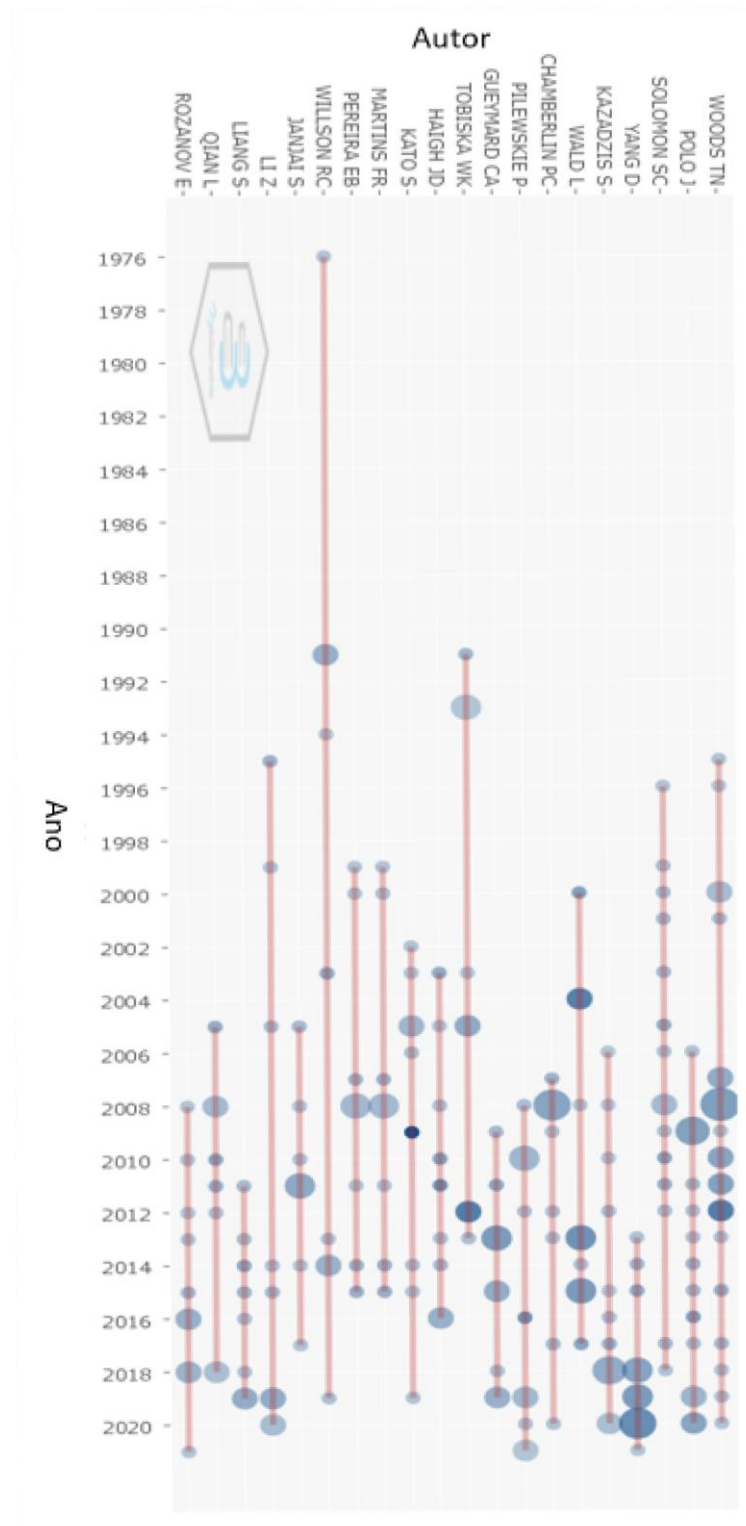
Fonte: Adaptado da base de dados Scopus.

Dentre estes, Jesús Polo, Dazhi Yang e Christian A. Gueymard são alguns dos autores de referência neste trabalho, por serem referências com pesquisas bastante atuais sobre o tema. Além disso, Fernando R. Martins e Ênio B. Pereira são autores que fazem parte diretamente do INPE e estão entre o top 20 mundiais no tema.

No caso do Gráfico 3, a produção no tempo dos 20 autores mais relevantes demonstra maior concentração de publicações a partir dos anos 2000, um período marcado pelo desenvolvimento e aprimoramento dos modelos de processamento das imagens capturadas pelos satélites geoestacionários, consequência dos avanços computacionais.

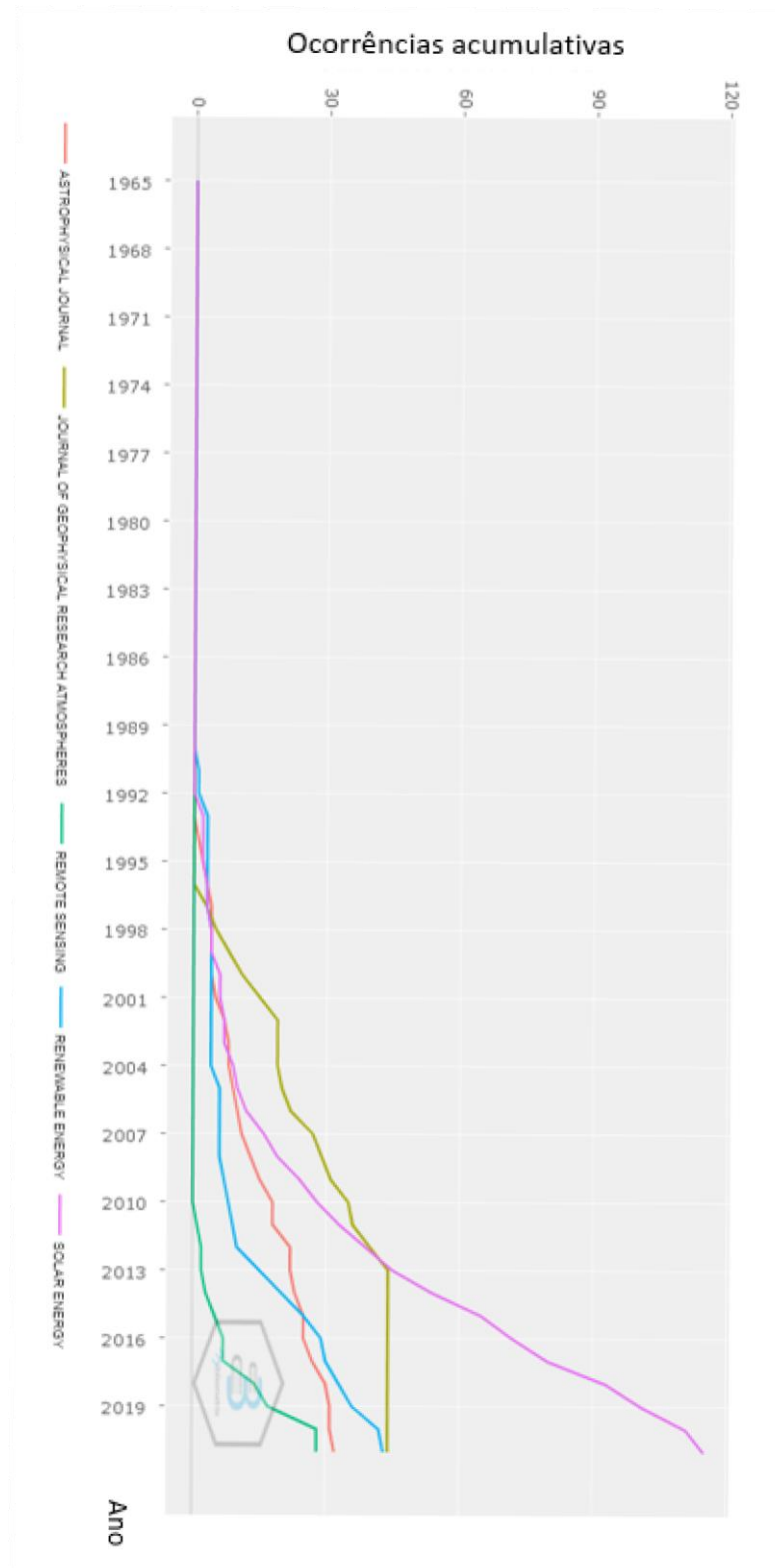
Acerca das principais fontes de publicação (Gráfico 4), desde 2013 a revista Solar Energy possui a maior expressividade sobre o tema, com 115 artigos acumulados, seguida pelo Journal of Geophysical Research Atmospheres e Renewable Energy, com 44 e 43, respectivamente.

Gráfico 3 – Produção no tempo pelos autores mais relevantes



Fonte: Adaptado da base de dados Scopus.

Gráfico 4 – Crescimento do tema nas principais fontes científicas



Fonte: Adaptado da base de dados Scopus.

2.2 Tendência temática

Após fundamentar a relevância desta temática e tomar conhecimento dos principais autores e fontes, é necessário entender quais os principais focos tomados atualmente.

No campo dos produtos de satélites, uma variedade de tópicos é estudada além da irradiância sintética, como os dados de aerossóis, de vapor de água, de ozônio, entre outros. Também há diversas abordagens, no qual o foco de uma pesquisa pode estar no desenvolvimento ou aprimoramento dos algoritmos ou na validação das estimativas. Assim, é de grande interesse analisar as palavras-chave ou termos mais utilizados e a atual tendência de pesquisa. Na Tabela 3 é apresentado o top 20 palavras-chave mais relevantes nos artigos entre 1965 e 2021, seguido por suas ocorrências.

Termos como “radiação solar” e derivados de “irradiância” e “energia solar” são as palavras-chave mais utilizadas pelos autores, consequência de estudos mais focados no avanço da irradiância sintética e aproveitamento solar.

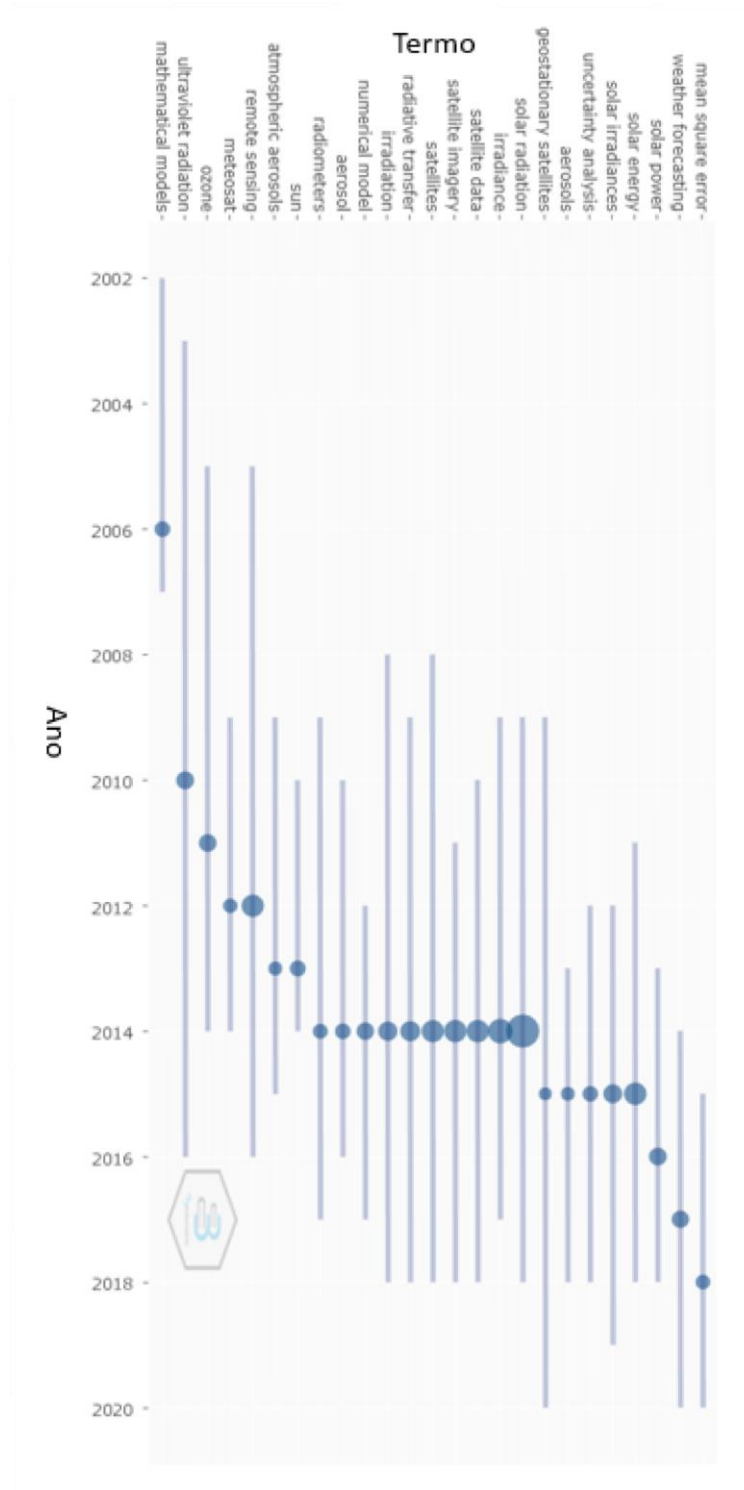
Tabela 3 – Top 20 palavras-chave mais relevantes

Palavras-chave	Tradução	Ocorrências
Solar radiation	Radiação solar	607
Irradiance	Irradiância	238
Satellite data	Dados de satélite	178
Solar energy	Energia solar	178
Satellite imagery	Imagem de satélite	177
Remote sensing	Sensoriamento remoto	174
Satellites	Satélites	171
Radiative transfer	Transferência radiativa	123
Irradiation	Irradiação	109
Solar irradiance	Irradiância solar	101
Ultraviolet radiation	Radiação ultravioleta	86
Solar power	Potência solar	85
Ozone	Ozônio	83
Numerical model	Modelo numérico	81
Weather forecasting	Previsão do tempo	79
Sun	Sol	64
Aerosol	Aerossol	62
Mathematical models	Modelos matemáticos	62
Uncertainty analysis	Análise de incertezas	61
Radiometers	Radiométricos	57

Fonte: Adaptado da base de dados Scopus.

No gráfico de uso dos principais termos na última década (Gráfico 5), entre 2018 e 2020, “erro quadrático médio”, “previsão do tempo”, “satélites geoestacionários” e “irradiação solar” possuem maior destaque.

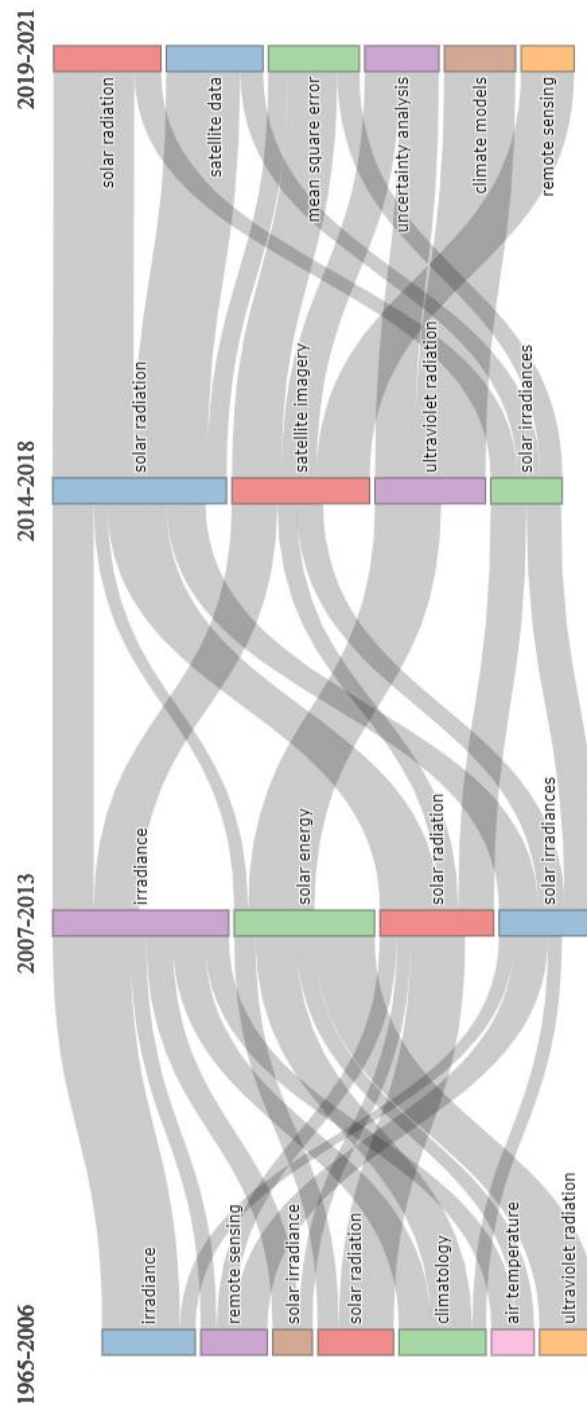
Gráfico 5 – Termos em tendência



Fonte: Adaptado da base de dados Scopus.

Outra ferramenta do *biblioshiny* é o mapa de evolução temática em fatias de tempo. No caso da Figura 1 são quatro divisões: 1965 a 2006, 2007 a 2013, 2014 a 2018 e 2019 a 2021. Temas como “dados de satélite”, “erro quadrático médio” e “análise de incerteza” em 2019-2021, que estão relacionados com “radiação solar” e “imagens de satélite” em 2014-2018 representam a evolução atual de tendências na pesquisa.

Figura 1 – Evolução temática



Fonte: Adaptado da base de dados Scopus.

2.3 Produção científica nos países

Por fim, a análise bibliométrica do presente trabalho é finalizada com a descrição da concentração geográfica de publicações no mundo. Para isto, o número de publicações sobre o tema em cada um dos 20 principais países é apresentado na Tabela 4 e, na Tabela 5, os países mais citados. O Brasil se posiciona na 11ª colocação em ambas as tabelas, totalizando 17 artigos com 354 citações no total.

Tabela 4 – Top 20 países com mais publicações, em ordem decrescente de número de artigos.

País	Artigos
USA	254
China	64
Alemanha	53
Espanha	48
França	40
Reino Unido	36
Itália	30
Austrália	25
Suíça	22
Grécia	18
Brasil	17
Canadá	16
Coreia do Sul	15
Índia	14
Japão	13
Holanda	11
Argélia	10
Singapura	10
Tailândia	8
Argentina	7

Fonte: Adaptado da base de dados Scopus.

Tabela 5 – Top 20 países mais citados em publicações

País	Citações totais	Citações médias por artigo
USA	11933	46,98
Alemanha	2210	41,70
Reino Unido	1431	39,75
Espanha	1404	29,25
França	1341	33,52
Suíça	810	36,82
Itália	738	24,60
China	707	11,05
Nova Zelândia	480	80,00
Austrália	377	15,08
Brasil	354	20,82
Canadá	331	20,69
Índia	331	23,64
Grécia	298	16,56
Japão	263	20,23
Tailândia	244	30,50
Singapura	232	23,20
Argélia	215	21,50
Chile	194	32,33
Coréia do Sul	190	12,67

Fonte: Adaptado da base de dados Scopus.

Estados Unidos, China e Alemanha são países de destaque. Na Europa, os esforços da Alemanha na transição de suas matrizes elétrica e energética é reconhecido mundialmente, sendo uma das referências em fontes alternativas. Atualmente, a China demonstra crescente interesse nas fontes renováveis, consequência da preocupação com a poluição gerada pelo seu acelerado crescimento econômico nas últimas décadas. Na América do Norte, apesar do foco nas fontes fósseis, os Estados Unidos vêm investindo mais na geração de energia renovável, para o qual um dos principais motivos é a competição econômica com a China no ramo. A geração de energia elétrica por fontes renováveis em 2017 foi de 502 TWh, 423 TWh e 203 TWh na China, EUA e Alemanha, respectivamente (EMPRESA DE PESQUISA ENERGÉTICA, 2020).

3 FUNDAMENTAÇÃO TEÓRICA

3.1 Radiação solar

A radiação e a energia solar são essenciais para a vida terrestre, tanto para sua manutenção biológica, como para as aplicações resultantes do desenvolvimento social e econômico humano. Segundo Gómes *et al.* (2018, p. 1), a energia solar tem efeitos nos “processos químicos, físicos e biológicos”, impactando o clima terrestre. Desta forma, o Sol é considerado uma das principais fontes de energia da Terra.

O transporte dessa energia acontece por meio da propagação da radiação eletromagnética emitida pela superfície solar, e ocorre em várias faixas espectrais: raios gama (γ), raios-x, ultravioleta (UV), visível, infravermelho, micro-ondas e ondas de rádio. Os autores Gómes *et al.* (2018, p. 1, 2 *apud* RESNICK; HALLIDAY; KRANE, 2008, p. 883) definem que

A radiação eletromagnética está relacionada com cargas elétricas aceleradas que irradiam energia. Em outras palavras, refere-se as variações temporais do campo elétrico (E), provocando uma variação do campo magnético (B) [...]. Estas variações se propagam radialmente [...]. A geração do campo magnético é associada com a operação de um dínamo que atua no interior do Sol. [Porém] esse complexo mecanismo ainda precisa ser melhor entendido.

Geralmente, nas engenharias e em estudos do clima é trabalhado o conceito de irradiância, que é a taxa do fluxo energético de radiação solar por unidade de área (W/m^2), ou seja, é a potência de energia solar em uma superfície. A medida da potência solar em toda a faixa espectral por metro quadrado no topo da atmosfera e perpendicular à direção do sol é chamada de Irradiância Solar Total. Ao tomar essa medida em uma distância média entre a Terra e o Sol é obtida a Constante Solar e, atualmente, esse valor é estimado em $1.361,1 \text{ Wm}^{-2}$ (GUEYMARD, 2018).

Uma superfície recebe radiação solar em todas as direções a partir da incidência direta ou difusa. A irradiância direta normal (do inglês DNI) é a irradiância recebida por um feixe direto de energia solar, enquanto a irradiância difusa (do inglês DHI ou DIF) é a incidência na superfície da energia que foi espalhada e refletida em todas as outras direções. Por fim, a irradiância global horizontal (do inglês GHI) é a soma de ambas componentes.

A componente direta varia de acordo com a presença de gases, nuvens, poluição e aerossóis, o que acaba afetando o aproveitamento de energia solar para conversão em eletricidade.

3.2 Irradiância solar sintética

Irradiância solar sintética é um tipo de irradiância que imita aspectos da irradiância solar real e é produzida por meios de modelagens computacionais, normalmente de bases físicas. Dados de irradiância solar podem ser obtidos por: (a) medições em solo terrestre com equipamentos de observação, (b) modelos numéricos e (c) sensoriamento remoto por satélite (HUANG *et al.*, 2019 apud LIANG *et al.*, 2019, tradução nossa).

O primeiro é de alta qualidade, porém pontual e necessita de manutenções frequentes. Em seguida, as simulações numéricas produzem mapas de irradiância que resultam em erros significativos na tentativa de estimar a magnitude exata ponto-a-ponto. Por fim, há os sensores a bordo de satélites que capturam imagens dos espectros de radiação resurgentes e, por modelagem computacional, produzem dados sintéticos que se assemelham estatisticamente aos valores reais (BRIGHT, 2021).

Apesar da natureza estatística das radiações solares sintéticas, Bright (2021, p. 24, tradução nossa) prossegue o trabalho descrevendo um dos erros de conceito mais comuns, no qual

O equívoco é que os modelos de irradiância solar sintética são preditivos. Com isso, quero dizer que algumas pessoas acreditam incorretamente que os modelos de irradiância solar sintética tentam derivar a medição real naquele exato momento e local, de forma que seriam precisos iguais dados de medição, como se tivessem realmente sido medidos. Este equívoco muitas vezes leva as pessoas a acreditarem que os modelos de irradiância solar sintética são equivalentes a modelos de previsão ou que sejam versões de técnicas de estimativa de recursos solares.

Os modelos de predição e de estimativas de recursos solares possuem como objetivo reproduzir exatamente os valores ponto-a-ponto da irradiância real no tempo, a única diferença seria que o primeiro estima para o futuro em diversos horizontes de tempo, e o segundo realiza estimativas entre o passado e o presente (BRIGHT, 2021).

De forma contrária, a irradiância sintética busca reproduzir o comportamento estatístico dos dados reais, representando também a variabilidade e várias outras relações de forma aceitável.

3.3 Princípios da estimacão de irradiância solar por satélites

3.3.1 Conceitos básicos

3.3.1.1 Condição do céu e ângulo zênite em agrupamento de dados

Geralmente, no estudo dos produtos de irradiância satelitais, os autores agrupam dados de acordo com as condições do céu ou altura do sol, e utilizam expressões como:

- a) “**All-sky**” (todo o céu), “**Clear-sky**” (céu claro) e “**Cloudy-sky**” (céu nublado) são abrangentemente utilizadas para designar agrupamentos de dados tomados de acordo com a condição do céu;
- b) “**Low-sun**” (sol baixo), “**high-sun**” (sol alto) e “**all-sun**” (todos os sóis), caracterizam o período do dia de um conjunto de dados. Dados de irradiância entre o entardecer e o amanhecer (maiores valores de ângulo zênite), classificam-se como “sol baixo”, enquanto o “sol alto” é marcado depois do amanhecer e antes do entardecer (menores valores de ângulo zênite). Por fim, “todos os sóis” abrange as 24 horas (todos os ângulos zênite).

3.3.1.2 Resolução e cobertura dos satélites

Existem vários tipos de resoluções que envolvem o sensoriamento remoto:

- a) Resolução espectral: é a quantidade de canais ou bandas espectrais e a faixa do intervalo de comprimento de onda do sensor. Alta resolução espectral é quando há o maior número de canais com faixas espectrais mais estreitas.
- b) Resolução espacial: capacidade de um sensor observar uma área na superfície da Terra. Alta resolução espacial se dá quando o instrumento é capaz de observar uma menor área possível.
- c) Resolução temporal: período entre cada medição de um ponto. Quanto menor for este período, o sensor alcança maior resolução temporal.

Por fim, os serviços provêm informações de cobertura temporal e espacial dos seus satélites. Cobertura temporal é a abrangência histórica da série temporal estimada e cobertura espacial é a abrangência territorial do serviço.

3.3.2 Panorama geral

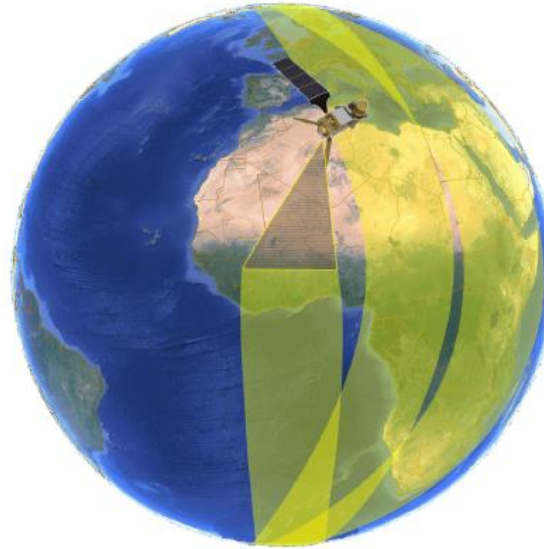
Estudos meteorológicos consideram os satélites como uma das suas principais fontes de medições. Estes são equipamentos lançados na órbita terrestre, onde há dois tipos de classificação: satélites de órbita polar e satélites geoestacionários.

Satélites polares (FIGURA 2) orbitam a Terra a uma altura de 890 km, de forma a realizar uma volta completa em 100 minutos, passando até 3 vezes por dia pelo mesmo ponto e, no geral, possuem resolução temporal de várias horas (MARTÍN, 2018). Por sua posição consideravelmente baixa, desenvolvem uma alta velocidade para manter-se em órbita. Em contraponto, satélites geoestacionários orbitam no plano equatorial a uma altura de 35.790 km e desempenham a mesma velocidade relativa da Terra. Desta forma, são capazes de observar continuamente o mesmo ponto, alcançando resoluções temporais de minutos (MARTÍN, 2018). Os diferentes satélites geoestacionários se diferenciam pela longitude que se localizam estacionados (FIGURA 3).

Segundo Polo e Perez (2019), sensores de radiação resurgente acoplados nos satélites geralmente medem três bandas espectrais: visível (0,5 a 1,1 μm), infravermelho térmico (10,5 a 12,5 μm) e infravermelho de vapor de água (5,7 a 7,1 μm). Os satélites têm uma quantidade n de canais espectrais, cada um responsável pela recepção de algum comprimento de onda, evoluindo no decorrer das gerações. O sensor MVIRI a bordo da primeira geração do Meteosat possuía apenas três canais, contra 12 canais espectrais do SEVIRI, na sua segunda geração.

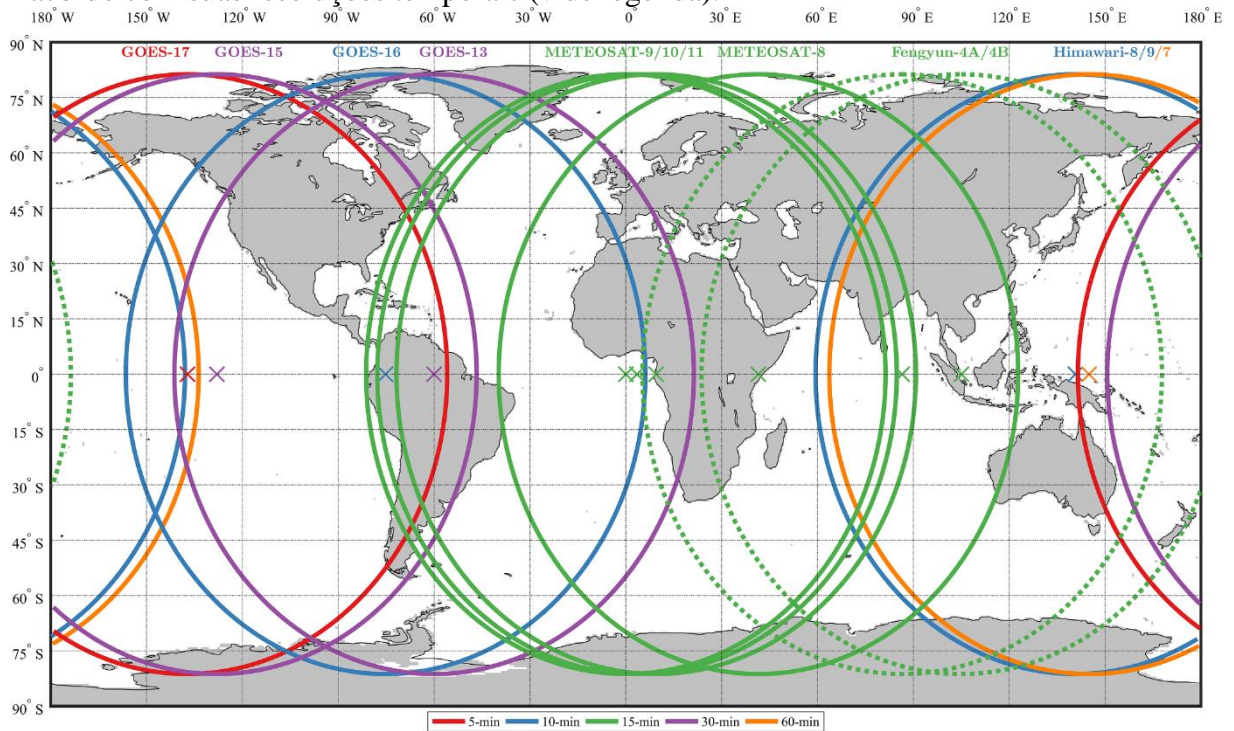
A Figura 3 apresenta a longitude de alguns satélites geoestacionários e sua cobertura espacial: GOES (13, 15, 16, 17) nas Américas, Meteosat (8, 9, 10, 11) na África e Europa, além de Fengyun (4A, 4B) e Himawari (8, 9, 7) cobrindo parte da Ásia. O GOES-17 e Himawari-8 são as gerações mais recentes nos dias atuais, caracterizados por suas altas resoluções. Segundo Bright (2019a), o primeiro realiza medições a cada 5 minutos, com 0,5 km^2 de resolução espacial. O segundo coleta as informações com resolução temporal de 10 minutos e mesma resolução espacial do GOES-17.

Figura 2 – Observações de um satélite com órbita polar



Fonte: Martín (2018, p. 84).

Figura 3 – Posição e observação de alguns satélites geostacionários, classificados em cor de acordo com suas resoluções temporais (vide legenda).



Fonte: Bright (2019^a, p. 436).

A autora Martín (2018) descreveu um breve histórico das missões satelitais, assim, a Tabela 6 foi elaborada com resumo de algumas informações mais importantes apontadas em seu trabalho.

Tabela 6 – Resumo das principais missões satelitais durante a história

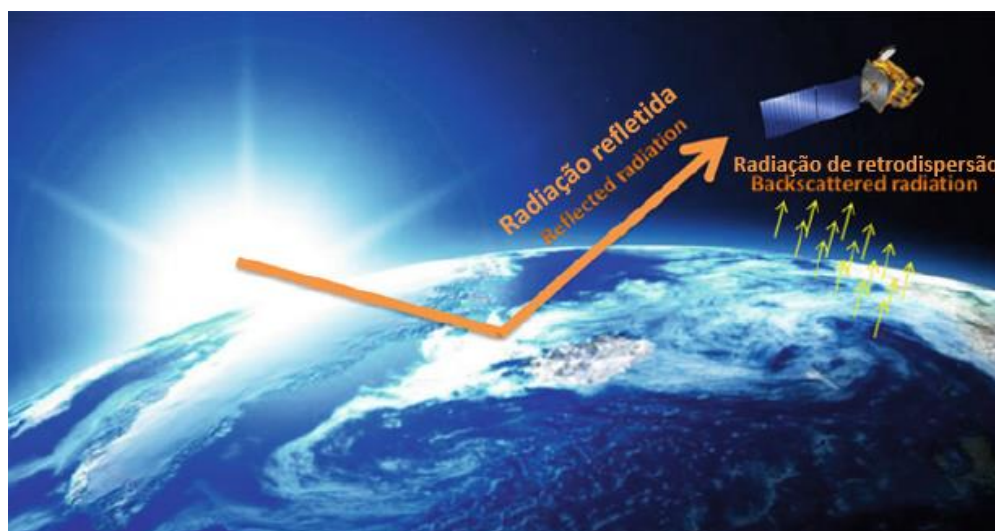
Satélite	Informações
TIROS-1 (Television Infrared Observational Satellite)	Lançado em 1960 pelos EUA, foi responsável pela primeira imagem televisiva da Terra.
Syncom-3 (Synchronous Communication Satellite)	Lançado em 1964, foi o primeiro satélite geostacionário.
TIROS-9	Um dos lançamentos experimentais da NASA, foi a nona geração do TIROS, e o primeiro satélite polar.
NIMBUS	Conjunto de 7 satélites de investigação científica da NASA.
NOAA-2	Lançado em 1972, marcou o primeiro uso de radiômetros com medição multiespectral e de alta resolução.
Meteosat	Primeira geração lançada em 1977 pela Agência Espacial Europeia. O sensor MVIRI a bordo foi considerado um avanço por possuir um canal com faixa espectral de vapor de água.
Meteosat-11	Geração mais atual, com 12 canais no sensor SEVIRI.
GOES-16 e -17 (Geostationary Operational Environmental Satellite Program)	GOES-16 teve seu lançamento em 2016 e GOES-17 em 2018. Ambos são satélites da NASA em cooperação com o NOAA (National Oceanic and Atmospheric Administration) e apresentam varredura mais rápida e com maior resoluções que as gerações anteriores do GOES.

Fonte: Adaptado de Martín (2018).

3.3.3 Tipos de modelos

Os instrumentos acoplados nos satélites responsáveis pelo sensoriamento remoto (SR) recebem a radiação solar refletida pelo topo da atmosfera e pela superfície terrestre e a radiação retrodispersa pela interação entre a radiação incidente com a atmosfera (FIGURA 4). Estatístico, semiempírico e físico são os três tipos de modelos. É então, a partir destes modelos, que pesquisadores desenvolvem, aprimoram e propõem algoritmos que estimam a irradiância sintética. As diferenças de resolução espacial, espectral, temporal entre os satélites, além dos algoritmos utilizados, afetam na acurácia dos produtos satelitais

Figura 4 – Ilustração da radiação recebida pelo sensoriamento remoto



Fonte: Adaptado de Polo e Perez (2019, p. 186).

3.3.3.1 Modelos estatísticos e modelos semi-empíricos

A ideia para o cálculo da irradiância sintética global horizontal (GHI) por modelos estatísticos foi baseada primeiramente nas relações entre a quantidade refletida e incidente de irradiância. Esta relação seria entre um parâmetro que representa a transmitância de radiação através da atmosfera e o índice de cobertura de nuvens.

No método estatístico, o índice de claridade é o parâmetro que representa a transmitância e é a proporção entre a GHI em uma superfície terrestre e a irradiância extraterrestre no topo da atmosfera. O índice de cobertura de nuvens é um parâmetro que normaliza a refletividade medida pelo sensor de satélite dentro da faixa entre menor e maior observação de irradiância (POLO; PEREZ, 2019).

Os modelos semi-empíricos utilizam o índice de céu claro como parâmetro para representação da transmitância atmosférica. Este índice é a proporção entre a irradiância horizontal global no solo e a irradiância global no solo sob condições totalmente sem nuvens (POLO; PEREZ, 2019).

3.3.3.2 Modelos físicos

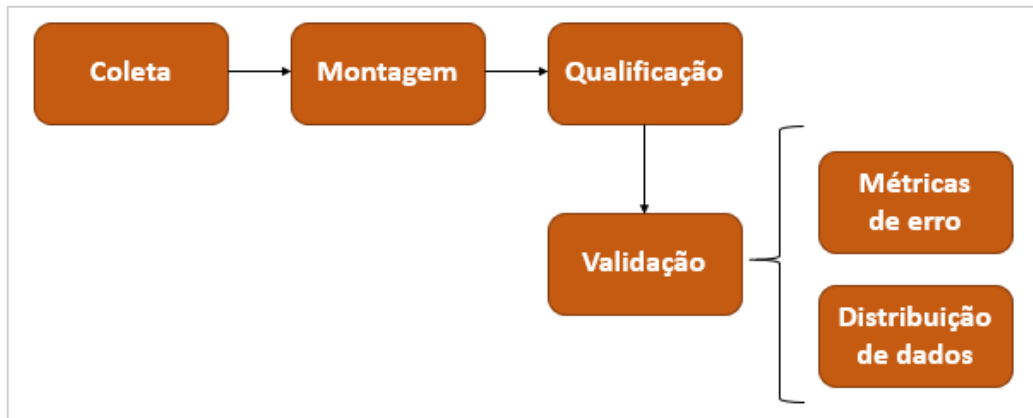
Por fim, os autores Polo e Perez (2019) definem que “os modelos físicos diferem conceitualmente dos modelos semi-empíricos, visto que resolvem os [equacionamentos dos] processos de transferência radiativa na atmosfera”.

4 METODOLOGIA

A Figura 5 ilustra um fluxograma de blocos simples que representa o processo de validação dos produtos de satélite. A metodologia deste trabalho tem como objetivo verificar se os dados de irradiância solar sintética resultantes do processamento de imagens obtidas pelo sensoriamento remoto dos serviços CAMS e Solcast conseguem representar observações reais, tomando a cidade de Petrolina-PE como estudo de caso.

Primeiramente foi realizada a coleta das observações reais, que foram medidas na estação terrestre de Petrolina (09°04'08"S, 40°19'11"O, Altitude 387 m) pela rede SONDA, seguida pela coleta dos produtos de satélite e a montagem do banco de dados. Após a montagem, a qualificação (tratamento) de dados foi aplicada no banco. Por fim, o protocolo de validação foi feito através do cálculo dos erros nas estimativas e do estudo orientado à distribuição de dados, ambos aplicados tanto no banco de dados por completo, como também em subconjuntos separados de acordo com a estação do ano.

Figura 5 – Fluxograma de blocos simples do processo de validação dos produtos de satélite CAMS e Solcast



Fonte: Elaborado pelo autor.

4.1 Banco de dados

4.1.1 Coleta de dados

A coleta dos dados de irradiância solar superficial foi realizada em três etapas: coleta de observações reais, coleta no CAMS e coleta no Solcast.

4.1.1.1 Coleta das observações reais

O projeto SONDA do Instituto Nacional de Pesquisas Espaciais (INPE), disponibiliza gratuitamente dados ambientais de suas estações de medição, tudo em planilhas mensais que possuem formato “.csv”. É importante observar que as planilhas possuem várias versões e, caso não seja observada a diferença entre elas, há a possibilidade de afetar a montagem do banco. Cada planilha de dados ambientais é padronizada, contendo 17 colunas nos arquivos até a versão 3.3, entretanto, a partir desta versão, há uma coluna a menos. Neste trabalho, coletaram-se todos os 132 arquivos compreendidos entre os anos de 2004 e 2018 da estação de Petrolina (09°04'08"S, 40°19'11"O, Altitude 387 m), localizada no estado de Pernambuco.

A base de dados radiométricos do INPE possui periodicidade de 1 minuto entre cada medição. Além disso, a própria rede SONDA realiza a qualificação dos dados, baseada nos critérios utilizados pela BSRN (Baseline Surface Radiation Network). Entretanto, após o tratamento, todos os dados são mantidos intactos, inclusive os dados incorretos e duvidosos. As únicas informações da qualificação realizada pelo INPE são códigos de verificação que sinalizam se o dado foi aprovado nas etapas do processo, sendo fornecidos em planilhas à parte. Desta forma, fica a cargo dos usuários a decisão de utilizar ou não os dados suspeitos e, no presente trabalho, não foram utilizados os arquivos de qualificação do SONDA, visto que foi escolhida outra estratégia de tratamento de dados.

A Figura 6 apresenta a localização da estação de Petrolina e o site no qual as radiações solares da estação de medição de Petrolina foram coletadas pode ser acessado pelo endereço (<http://sonda.ccst.inpe.br/basedados/petrolina.html>).

Figura 6 – Localização da estação da rede SONDA em Petrolina, Pernambuco



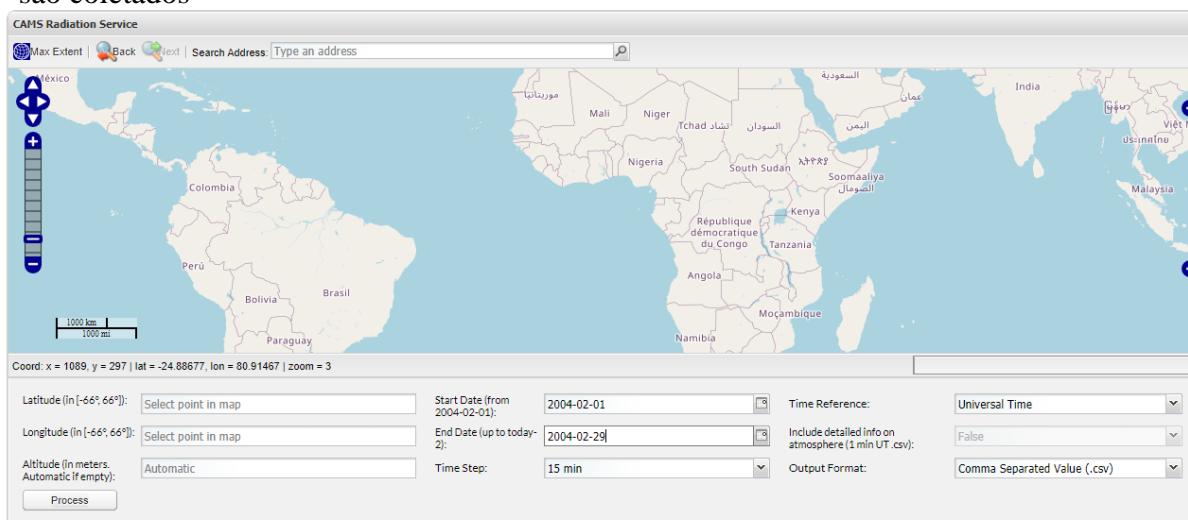
Fonte: Elaborado pelo autor.

4.1.1.2 Coleta no CAMS

Ao criar uma conta no site do serviço de irradiância CAMS (<http://www.soda-pro.com/web-services/radiation/cams-radiation-service>) é possível realizar o acesso gratuito dos dados necessários, apenas com a introdução de algumas informações de Petrolina, como a latitude, a longitude e a altitude nos campos solicitados (FIGURA 7).

Foi realizado o *download* de um arquivo no formato “.csv” em todas as resoluções temporais disponíveis (1-, 15-, 60-min, diário e mensal), com data de início em 01/02/2004, data final em 31/12/2018 e habilitando “Universal Time (UTC)” na referência de tempo, que é a mesma base do Projeto SONDA.

Figura 7 – Fonte de coleta do Serviço CAMS de Radiação para ilustração de como os dados são coletados



Fonte: Elaborado pelo autor.

4.1.1.3 Coleta no Solcast

Em consequência de o Solcast ser um serviço comercial (<https://solcast.com/historical-and-tmy/time-series/>), entrou-se em contato com um dos colaboradores do serviço, com o objetivo de solicitar crédito para acesso dos dados, o que é permitido para fins acadêmicos. De forma bastante solícita, foi oferecido crédito suficiente para o *download* de todo o banco de dados disponível para a localidade de Petrolina, incluindo todas as resoluções temporais oferecidas pelo serviço: 5-, 10-, 15-, 30-, 60-min (FIGURA 8 e FIGURA 9).

Figura 8 – Fonte de coleta do Serviço Solcast para ilustração de como os dados do local são inseridos antes do pedido

Fonte: Elaborado pelo autor.

Figura 9 – Escolha dos parâmetros no site do Serviço Solcast

Fonte: Elaborado pelo autor.

4.1.2 Montagem do banco de dados

Após a coleta dos dados, o banco foi montado com a utilização do R Studio®. Os arquivos coletados foram lidos pelo software e armazenados por meio da criação de onze *dataframes* (*estrutura básica de armazenamento de dados tabulares no R*): um do SONDA, cinco do CAMS e cinco do Solcast.

Todos os 132 arquivos de observações reais foram agrupados em apenas um *dataframe* e, como sua resolução temporal é de 1 minuto, também foram calculadas as médias móveis de 5-, 10-, 15-, 30-, 60-min, diário e mensal, para tornar possível o emparelhamento com os dados equivalentes dos satélites. No caso dos produtos de satélite, um *dataframe* foi criado para cada resolução temporal diferente, com objetivo de evitar conflito ou duplicidade de *timestamps* (coluna de identificação da hora/data de aquisição). A quantidade de dados em cada um dos onze *dataframes* criados estão apresentadas na Tabela 7.

Após o armazenamento dos dados nestes *dataframes* (Tabela 7), foi realizada a padronização das *timestamps*. As estações de medição e os produtos de satélites possuem o formato data/hora padronizados, porém diferentes entre si. Como exemplo, a rede SONDA tem colunas separadas para ano, dia do ano e minuto do dia, enquanto no CAMS essas informações são sinalizadas em uma coluna só. Desta forma, para o cruzamento de dados equivalentes (com mesma data e hora), foi necessário converter as informações de datas do tipo caractere para uma classe de data universal que o R reconhece, utilizando as funções `POSIXct` e `POSIXlt`, e possibilitando assim uma manipulação de dados mais simples e eficaz.

Tabela 7 – Informações de cada dataframe montado, valores brutos correspondentes a antes da qualificação

Dafaframe	Resolução temporal	Quantidade de dados
SONDA	1 min	7.142.400
CAMS1	1 min	7.845.120
CAMS15	15 min	558.048
CAMS60	60 min	130.752
CAMSDia	24 h	5.448
CAMSMês	30 dias	179
Solcast5	5 min	1.501.607
Solcast10	10 min	750.803
Solcast15	15 min	500.535
Solcast30	30 min	250.267
Solcast60	60 min	125.133

Fonte: Elaborado pelo autor.

O emparelhamento (ou junção) do *dataframe* SONDA com os outros dez foi feito separadamente para, então, ser realizado o tratamento de dados. O tratamento ou qualificação dos dados consistiu em dois passos: remoção dos valores ausentes (NA – Not Available, em português “não disponível”) e descarte dos dados de Sol baixo (ângulo zênite $> 85^\circ$) para estudos que envolvem resolução menor que 24 horas.

O protocolo de validação foi realizado em cada um destes 10 *dataframes* emparelhados e com dados tratados.

4.2 Protocolo de validação

O protocolo de validação consiste no cálculo dos erros de estimativas a partir de métricas de performance e na análise distributiva dos dados. O protocolo de validação foi realizado em cada um dos 10 *dataframes* emparelhados e tratados, resultantes do processo de montagem do banco de dados.

Além disso, com intuito de avaliar a variabilidade da acurácia em decorrência da mudança das estações do ano na localidade estudada, foram criados subconjuntos de dados, separando-os de acordo com as estações do ano (verão, outono, inverno e primavera) e com resolução temporal de 60 min (SONDA/CAMS60 e SONDA/Solcast60). O mesmo protocolo de validação foi aplicado nestes subconjuntos.

4.2.1 Métricas de erro

As métricas de erro fazem parte da análise com abordagem baseada na distância, ou seja, esse tipo de validação permite a comparação entre um valor estimado obtido por um modelo e uma observação real, com base na medida da distância ou do desvio entre ambos. Existem inúmeras métricas de desempenho, mas segundo Bright (2019a), algumas das mais utilizadas nos trabalhos de validação são: raiz do erro quadrático médio (RMSE), erro médio absoluto (MAE) e coeficiente de correlação (R^2). Estas são as utilizadas neste trabalho.

O RMSE, apresentado na Equação 1, pode ser interpretado como a “distância Euclidiana normalizada entre o vetor de valores estimados e o de observações” (FRIMANE; BRIGHT, 2021, p. 43). O RMSE se diferencia do MAE (Equação 2) por aplicar maior peso em erros maiores e, em ambos, quando há divisão por N, obtêm-se erros para uma única estimativa.

Ademais, na Equação 3, o coeficiente de correlação R^2 é uma medida adimensional no qual é 1 à medida que o modelo replica precisamente as observações reais e 0 se não. Segundo Kuhn e Johnson (2013, p. 95), o R^2 “pode ser interpretado como a proporção da informação nos dados que é explicada pelo modelo. Assim, um valor de 0,75 implica que o modelo pode explicar três quartos da variação no resultado”.

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

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (2)$$

$$R^2 = 1 - \left[\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \right] \quad (3)$$

Onde, \hat{y}_i são os valores estimados das observações y_i , \bar{y}_i e N é a média e o número das observações, respectivamente.

4.1.2 Análise da distribuição de dados

Na literatura ainda há debates sobre quais são as métricas ou abordagens mais adequadas para a avaliação de dados oriundos do sensoriamento remoto por satélites, geralmente ficando a cargo dos autores escolherem as que mais se adequam ao seu objetivo de estudo. Apesar disso, relevante fração da literatura utiliza o tipo de análise orientada à distribuição de dados, como em Bright (2019b), Munkhammar e Widén (2019), Peruchena *et al.* (2018), Yang e Bright (2020) e Zhang *et al.* (2018), todos estes baseados em variadas técnicas e ferramentas.

Segundo Yang e Bright (2020), respaldar a validação dos modelos de satélite apenas com abordagem orientada à medição de performance não está sendo mais recomendada nos últimos anos. Frimane e Bright (2021, p.2) justificam esta mudança ao revelar que “não é esperado que os dados de irradiância sintética se equiparem às observações ponto-a-ponto, mas ambos devem ser similares em termos de representação estatística”. Desta forma, métricas de performance são mais adequadas para algoritmos de previsão de irradiância solar, enquanto as

estimativas de radiação superficial em formato de produto de satélites (radiações sintéticas) devem ser focadas em reproduzir de maneira similar as estruturas e distribuições das observações.

Trabalhos de análise de dados possuem uma vasta gama de ferramentas visuais que auxiliam na investigação da estrutura e comportamento de dados. Estas ferramentas de visualização de dados devem fornecer ilustrações claras de acordo com o objetivo almejado. Thrun, Gehlert e Ultsch (2020) citaram e estudaram algumas diferentes abordagens utilizadas para avaliação da estrutura de dados, como os gráficos de histogramas, funções de densidade de probabilidade, *boxplots* e *ridgelines*.

Em vista disto, o presente trabalho utiliza a ferramenta visual *ridgeline* em um estudo orientado à distribuição de dados, com objetivo de complementar a discussão. O Gráfico de Ridgeline foi obtido por meio da função *geom_density_ridges*, disponível no pacote do R chamado *ggridges*. Esta função calcula as múltiplas densidades dos pontos de dados que estão mapeados no eixo-x e as agrega ao longo do eixo-y por meio de uma categorização, não tendo assim um eixo quantitativo da densidade. Por consequência, o objetivo desse tipo de gráfico é permitir a comparação qualitativa dos formatos das curvas de densidades entre os grupos ao longo do eixo y (WILKIE, 2019). Além disso, também há o uso de histogramas no decorrer do trabalho.

5 RESULTADOS E DISCUSSÃO

5.1 Erros gerais da estimação de irradiância e qualificação de dados

Satélites fornecem dados de irradiância sintética, no qual estes possuem erros de estimativa, geralmente associados à modelagem computacional utilizada para o processamento das imagens do sistema Terra-Atmosfera capturadas e às resoluções do SR. Para avaliar esse desvio de estimativas, os produtos de satélite são comparados a valores resultantes de medições terrestres (AMEEN *et al.*, 2018; BRIGHT; GUEYMARD, 2019; YANG, 2018; YANG, 2020).

Como já citado, primeiramente foram adquiridos os dados de irradiância global horizontal e irradiância direta normal de 1 minuto fornecidos pela estação de Petrolina-PE, abrangendo um período de quase 14 anos (entre 2004 e 2018). Estes dados foram cruzados com dois produtos de satélites: o CAMS Radiation e o Solcast, formando a base de dados.

Em seguida, foi realizada a avaliação do processo de qualificação dos dados ao comparar dois grupos de resultados. O Grupo I de resultados consiste nos erros obtidos pela aplicação das métricas na base de dados com todos os sóis (“all-sun”). Já o Grupo II são os mesmos resultados obtidos apenas com dados de sol alto (“high-sun”), com θ_Z menor que 85° . O Grupo I pode ser considerado de validação, uma vez que há trabalhos que usam todo o banco de dados para avaliar a acurácia dos produtos (CLARKE *et al.*, 2021; THOMAS *et al.*, 2016). Já o Grupo II se mostra preferível, pois analisa o desempenho dos modelos nos momentos em que a irradiância solar é útil (existente). Essa abordagem é utilizada geralmente e recomendada (BRIGHT, 2019a; YANG; BRIGHT, 2020; YANG, 2021) logo é o foco do presente trabalho.

A partir disto, a Tabela 8 apresenta as métricas de erro para as componentes GHI e DNI estimadas pelo produto de satélite CAMS-RAD em ambos os grupos e resolução temporal de 60 minutos. Similarmente, a Tabela 9 pontua os erros da irradiância sintética do Solcast.

Tabela 8 – CAMS-RAD: Erros associados à estimação da GHI e DNI, para os dois grupos de dados e resolução temporal de 60 minutos

Métricas CAMS-RAD (60 min)	GHI		DNI	
	Grupo I	Grupo II	Grupo I	Grupo II
RMSE (W/m ²)	125,83	181,06	219,66	313,31
MAE (W/m ²)	77,54	154,75	114,20	224,02
R ²	0,85	0,66	0,58	0,35

Fonte: Elaborado pelo autor.

Tabela 9 – Solcast: Erros associados à estimação da GHI e DNI, para os dois grupos de dados e resolução temporal de 60 minutos

Métricas Solcast (60 min)	GHI		DNI	
	Grupo I	Grupo II	Grupo I	Grupo II
RMSE (W/m ²)	57,66	82,36	173,42	250,98
MAE (W/m ²)	26,51	53,04	73,44	149,66
R ²	0,97	0,92	0,69	0,52

Fonte: Elaborado pelo autor.

Thomas *et al* (2016b) realizaram uma larga validação da GHI estimada por três produtos de satélites (HC3v4, HC3v5 e CAMS) em 42 estações do Brasil, entre elas a de Petrolina (INPE) e utilizando resolução temporal de 60 minutos, no qual obteve RMSE de 106,2 Wh/m² pelo CAMS (contra 121,83 Wm⁻² na Tabela 8). O processo de qualificação de dados pode ter originado essa diferença de resultados, uma vez que os autores utilizaram rotinas adicionais no banco de dados. Neste trabalho os autores concluíram que, entre os três produtos analisados, o CAMS possui resultados de desempenho inferior, com tendência de super estimação da GHI, que também será observada no presente estudo. Bright (2019a) obteve RMSE de 78,6 W/m² (GHI) após a validação do Solcast na estação de Petrolina (contra 82,36 Wm⁻² na Tabela 9), utilizando o Grupo II atrelado a outra metodologia de qualificação.

As estimativas de DNI horária do CAMS em Petrolina foram validadas em Salazar *et al.* (2020). O autor utilizou dados abrangendo apenas quatro anos (2010-2014) com um rígido controle de qualidade para o Grupo II, atingindo um valor RMSE de 154,3 Wm⁻² (contra 313,31 Wm⁻² na Tabela 8). Neste trabalho, o modelo CAMS teve o melhor desempenho em relação a outros quatro serviços avaliados. Estudos de validação de DNI para Solcast e que abordam a variabilidade das componentes GHI e DNI relacionados aos climas brasileiros são uma lacuna na literatura que este trabalho pretende preencher.

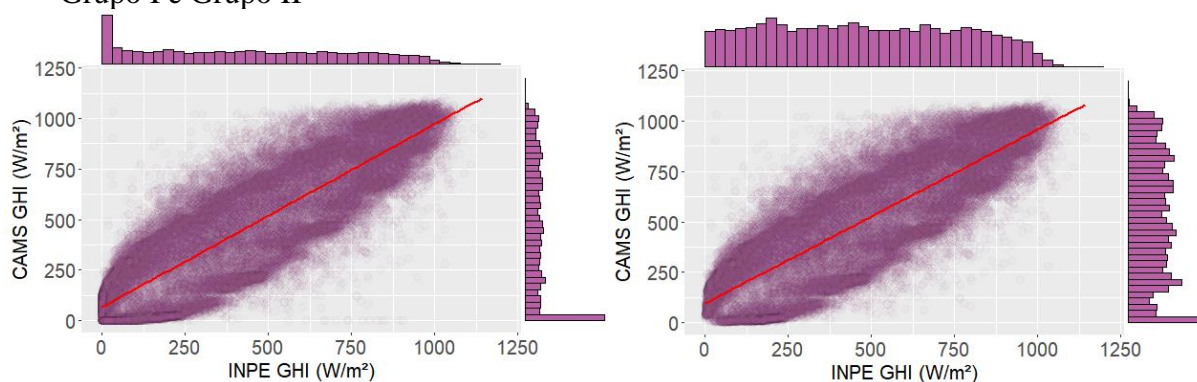
Em geral, conforme as Tabelas 8 e 9, os valores de RMSE para o Grupo II foram aproximadamente 44% maiores que os do primeiro grupo. Os resultados do erro médio absoluto (MAE) foram pelo menos 96% maiores no segundo grupo e todos os coeficientes R² se

apresentaram inferiores, comparados ao primeiro grupo. Estes resultados ilustram como os dados de radiação solar superficial correspondentes aos altos ângulos de zênite do Sol (“sol baixo”) interferem na acurácia. Altos ângulos de zênite equivalem ao período entre o entardecer e o amanhecer, no qual não há registro significativo de irradiância solar e são indicados com valores aproximadamente nulos na base de dados. Utilizar a irradiância horária durante todo o período de 24 horas desloca a distribuição de dados para valores em zero, e como consequência aumenta a correlação de dados e diminui os erros na estimação da GHI e da DNI pelos satélites analisados.

Este fenômeno de distribuição de dados relacionado ao descarte de dados com $\theta_z > 85$ (GUEYMARD; RUIZ-ARIAS, 2016) é exposto nos gráficos 6 a 9. Por este motivo, é importante ressaltar que o maior foco deste trabalho foi feito tomando o Grupo II como objeto de estudo.

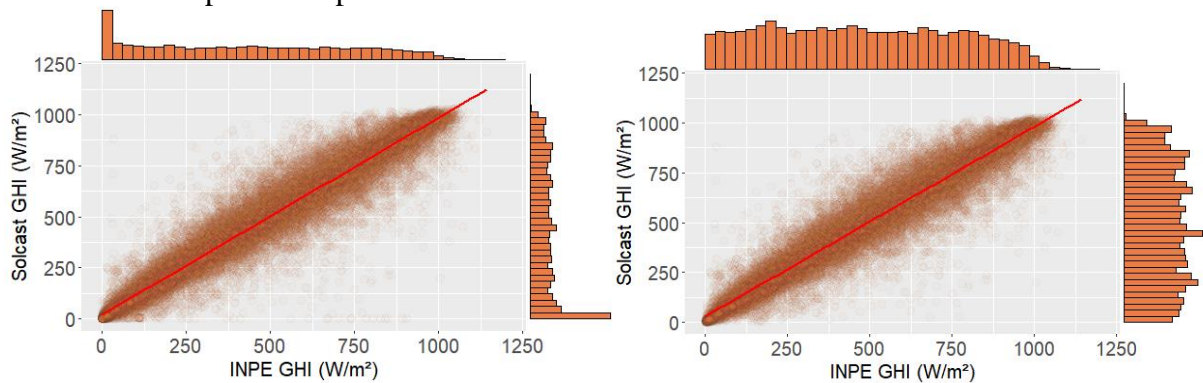
Os Gráficos 6 a 9 são gráficos de dispersão em pares de dados, utilizada para ilustrar a relação entre duas variáveis de mesma unidade - neste caso, a GHI estimada versus GHI real (Gráfico 6 para o CAMS e Gráfico 7 para o Solcast). Além da DNI estimada versus DNI real (Gráfico 8 para o CAMS e Gráfico 9 para o Solcast). Todas as figuras apresentam linhas de tendência e histogramas marginais. Ademais, cabe salientar que gráficos em cor roxa correspondem aos resultados relacionados aos produtos de satélite do Serviço CAMS-RAD e cor laranja do serviço Solcast.

Gráfico 6 – Dispersão com distribuição de frequências entre os dados de GHI estimados pelo CAMS-RAD vs. Medidos pela estação do INPE (60 min). Da esquerda para direita: Grupo I e Grupo II



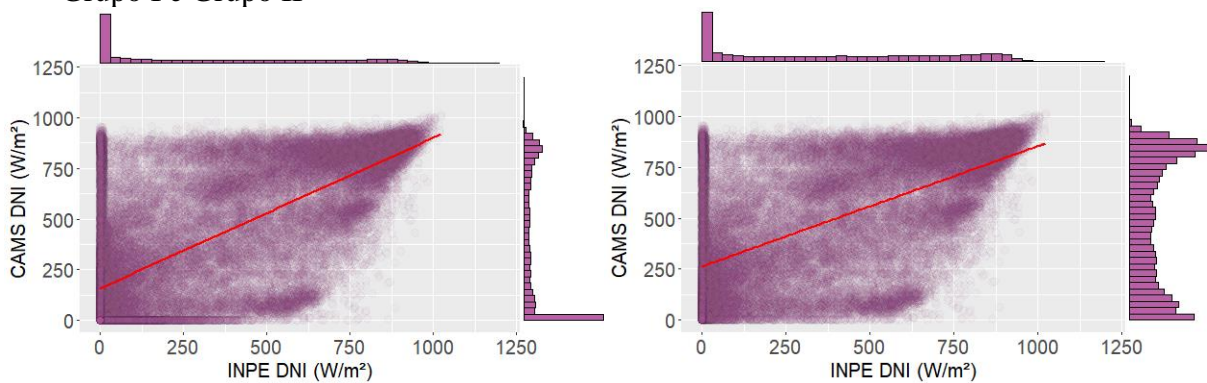
Fonte: elaborado pelo autor.

Gráfico 7 – Dispersão com distribuição de frequências entre os dados de GHI estimados pelo Solcast vs. Medidos pela estação do INPE (60 min). Da esquerda para direita: Grupo I e Grupo II



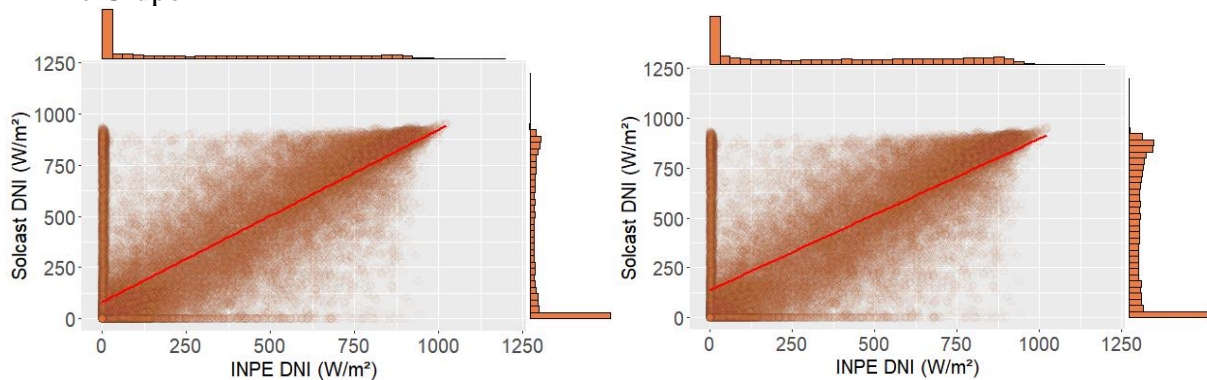
Fonte: elaborado pelo autor.

Gráfico 8 – Dispersão com distribuição de frequências entre os dados de DNI estimados pelo CAMS-RAD vs. Medidos pela estação do INPE (60 min). Da esquerda para direita: Grupo I e Grupo II



Fonte: elaborado pelo autor.

Gráfico 9 – Dispersão com distribuição de frequências entre os dados de DNI previstos pelo Solcast vs. Medidos pela estação do INPE (60 min). Da esquerda para direita: Grupo I e Grupo II



Fonte: elaborado pelo autor.

A partir da dispersão de dados ilustrada nos Gráficos 6 e 7, é possível constatar que o modelo do Solcast tem notadamente menor dispersão em comparação com o CAMS-RAD, apresentando visualmente uma acurácia mais satisfatória. Outro apontamento relevante é a tendência de superestimação do CAMS-RAD para a maioria dos valores de irradiância e, em contraste, também subestimação para menores valores de irradiância.

Com o Gráfico 9, também é verificado menor dispersão da componente DNI estimada pelo do Solcast (em comparação com o CAMS), exceto em valores nulos. Trabalhos na literatura (BRIGHT, 2019a; GUEYMARD; LUIZ-ARIAS, 2016; NONNERNMATCHER; KAUR; COIMBRA, 2014; YANG, 2020) atestam a possibilidade de os satélites estimarem dados de céu claro quando na verdade há céu nublado. A complexidade das estimativas de DNI ocorre em virtude de erros relacionados à resolução espacial correspondentes a cada satélite e erros de modelagem que se encarregam dos ajustes e correção das diferentes perspectivas entre o Sol-satélite-nuvem e Sol-nuvem-superfície. Fenômenos semelhantes podem ser observados para o modelo do CAMS-RAD.

5.2 Acurácia por ano

Este estudo utilizou observações e previsões que abrangem um espaço de tempo “tão longo quanto possível”, ou seja, com o máximo de dados disponíveis no INPE e nas plataformas dos produtos de satélite. Para a análise da precisão dos dados ao longo dos anos, as Tabelas 10 e 11 apresentam os erros de irradiância estimada, arranjados por ano de estudo, para o CAMS e Solcast, respectivamente.

Tabela 10 – CAMS-RAD: Erros anuais de estimação da GHI e DNI, resolução temporal de 60 minutos

Ano CAMS-RAD	GHI			DNI		
	RMSE	MAE	R ²	RMSE	MAE	R ²
2004	187,66	160,78	0,68	269,03	195,55	0,46
2005	-	-	-	-	-	-
2006	-	-	-	-	-	-
2007	176,79	142,48	0,69	497,47	376,21	0,064
2008	177,47	152,48	0,67	317,07	225,93	0,31
2009	183,47	156,20	0,65	337,10	245,03	0,30
2010	179,41	153,82	0,65	309,61	222,74	0,32
2011	183,42	157,73	0,64	269,71	196,40	0,41
2012	177,26	154,07	0,68	259,58	188,07	0,43
2013	176,34	151,07	0,68	234,73	172,48	0,56

(conclusão)						
2014	181,21	155,70	0,67	274,40	202,28	0,46
2015	186,63	160,48	0,65	262,68	196,26	0,46
2016	182,69	159,30	0,66	385,62	280,62	0,22
2017	173,42	149,10	0,67	239,83	178,84	0,53
2018	196,54	163,61	0,55	273,66	207,96	0,37

Fonte: Elaborado pelo autor.

Tabela 11 – Solcast: Erros anuais de estimação da GHI, em resolução temporal de 60 minutos

Ano Solcast	GHI			DNI		
	RMSE	MAE	R ²	RMSE	MAE	R ²
2007	114,77	69,63	0,85	449,66	307,70	0,11
2008	85,91	56,42	0,91	264,98	166,50	0,44
2009	80,06	53,97	0,92	278,89	172,90	0,43
2010	82,57	54,36	0,92	259,78	157,34	0,48
2011	74,62	51,03	0,93	191,34	114,46	0,68
2012	74,61	46,88	0,93	192,10	119,89	0,65
2013	73,03	49,03	0,94	153,27	99,69	0,78
2014	70,57	47,39	0,94	183,26	112,46	0,71
2015	69,98	44,90	0,94	148,70	96,96	0,79
2016	69,20	46,19	0,95	328,74	201,98	0,34
2017	76,68	53,13	0,92	155,38	103,36	0,77
2018	132,63	85,07	0,77	206,22	140,73	0,57

Fonte: Elaborado pelo autor.

A partir da Tabela 10 foi montado um gráfico de barras (GRÁFICO 10), no qual as linhas tracejadas correspondem aos erros globais de toda a extensão temporal estudada (vide Tabela 8 (Grupo II)). O mesmo foi feito com os valores da Tabela 11, da qual foi montado um gráfico de barras (GRÁFICO 11), cujas as linhas tracejadas correspondem aos resultados globais dos de toda a extensão temporal estudada apresentados na Tabela 9 (Grupo II).

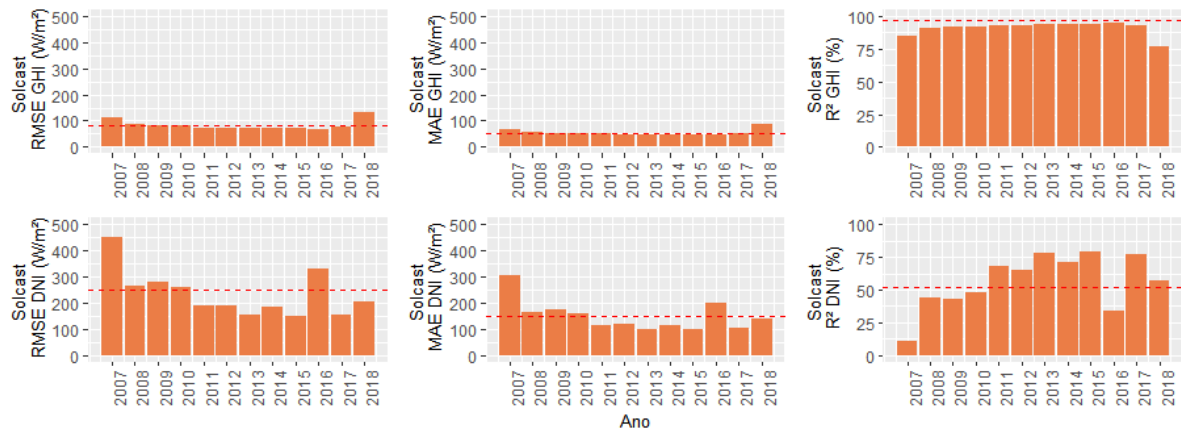
Do Gráfico 10 é possível atentar que os erros da GHI estimada pelo CAMS-RAD são precisos durante os anos, não desviando muito dos valores totais. O mesmo não ocorre para a componente DNI, que apresenta alguns desvios significativos (o que é esperado), como nos anos de 2007, 2013 e 2016. Da mesma forma, o modelo do Solcast é preciso no decorrer dos anos quando se trata da previsão da GHI e menos preciso na previsão da DNI, o que é esperado.

Gráfico 10 – Erros de GHI e DNI estimados pelo CAMS-RAD para cada ano observado, em resolução temporal de 60 minutos. Neste gráfico os resultados são comparados com valores globais da Tabela 1 (linha tracejada)



Fonte: elaborado pelo autor.

Gráfico 11 – Erros de estimação da GHI e DNI pelo Solcast para cada ano observado, em resolução temporal de 60 minutos. Neste gráfico os resultados são comparados com valores totais da Tabela 2 (linha tracejada)



Fonte: elaborado pelo autor.

De modo geral, um dos modelos utilizados pelo Solcast (REST2v5) é qualificado entre os três melhores nos termos de performance em escala global (SUN *et al.*, 2019). No presente trabalho, a acurácia de previsão da GHI pelo Solcast é maior, visto que mantém erros da ordem de 100 W/m² (RMSE) e 50 W/m² (MAE). Em comparação, o CAMS-RAD tem erros que permanecem em torno de 200 W/m² (RMSE) e 150 W/m² (MAE).

5.3 Estudo da distribuição de dados

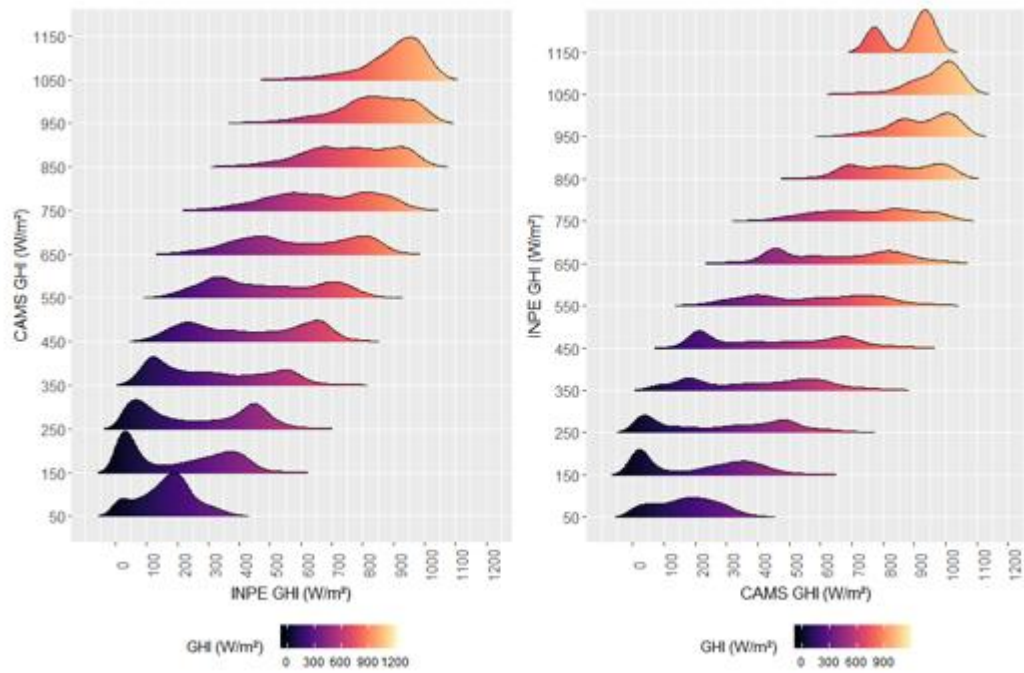
Na Seção 5.1 utilizaram-se gráficos de dispersão com histogramas marginais, sendo útil para o estudo da relação entre as variáveis observadas e estimadas. Entretanto, como existe uma imensa quantidade de pontos sobrepostos, esse tipo de gráfico acaba não demonstrando clareza suficiente e pode levar a deduções errôneas.

Seguindo a abordagem voltada a validação por meio de análise de distribuição de dados, no Gráfico 12 (à esquerda) é exibida a distribuição das médias móveis horárias das GHI medidas pela estação de Petrolina (INPE) entre 2004 e 2018 em relação aos agrupamentos de valores do produto de satélite CAMS: estimativas que possuíam valores entre zero e cem foram agrupadas na categoria de 50 (W/m^2) e, assim por diante, a cada cem Watts por metro quadrado. Cada curva dessas representa a função de densidade de probabilidade condicional das observações “o” relacionadas aos dados valores previstos “p”, $f(o/p)$. Para qualquer conjunto de estimativas é esperado que a expectativa (concentração) das observações seja próxima àquela previsão, ou seja, $E(o/p) = p$. (YANG; BRIGHT, 2020)

Analogamente, foi feita a distribuição da GHI estimada no eixo-x dadas categorias de valores medidos (Gráfico 12, à direita). Neste caso, cada curva dessas representa a função de densidade de probabilidade condicional das previsões “p” dados os valores observados “o”, $f(p/o)$. Ela indica a frequência que diferentes estimativas podem ocorrer antes de uma observação se concretizar. Para qualquer conjunto de observações é esperado que a expectativa (concentração) das previsões seja próxima àquela observação, assim, $E(p/o) = o$ (YANG; BRIGHT, 2020).

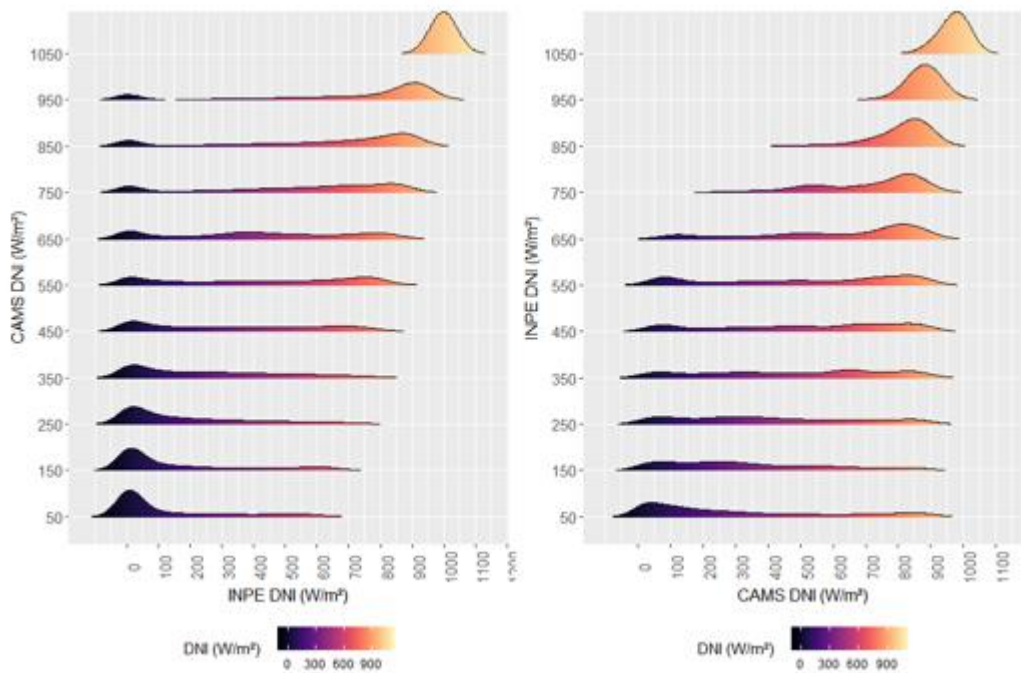
O Gráfico 13 (à esquerda) foi obtido conforme o mesmo método sob os dados de DNI do INPE em relação aos do CAMS-RAD e, à direita, distribuição da DNI estimada dados os valores medidos.

Gráfico 12 – CAMS-RAD: Distribuição da irradiância GHI medida dados os valores estimados (à esquerda) e distribuição da GHI estimada dados valores medidos (à direita)



Fonte: elaborado pelo autor.

Gráfico 13 – CAMS-RAD: Distribuição da DNI medida dados valores estimados (à esquerda) e distribuição da DNI estimada dados valores medidos (à direita)



Fonte: elaborado pelo autor.

No caso de cada curva possuir apenas um pico e o mesmo estiver posicionado em uma tendência $x = y$, será considerado um desempenho ideal para distribuição de dados de

irradiância - isto significa que à esquerda teria $E(o/p) = p$ e, à direita, $E(p/o) = o$. Em contraste, observa-se que tanto o Gráfico 12 quanto o Gráfico 13 não possuem este comportamento. A partir do Gráfico 12 (à esquerda) é notado que há simultaneamente sub estimação e super estimação para todos os grupos entre 150 e 750 W/m² de estimativas da GHI sob o modelo CAMS-RAD (dois picos diferentes em cada curva de densidade, acima e abaixo do valor esperado). À direita da mesma figura, no grupo 1150 W/m² de medição da GHI, há dois picos de densidade abaixo do valor esperado. Tudo isto leva a concluir que (a) para baixos valores de GHI o modelo possui tendência a superestimação, (b) para valores intermediários de GHI há ambas uma sub e superestimação e (c) para altos valores de GHI ele tende a subestimar.

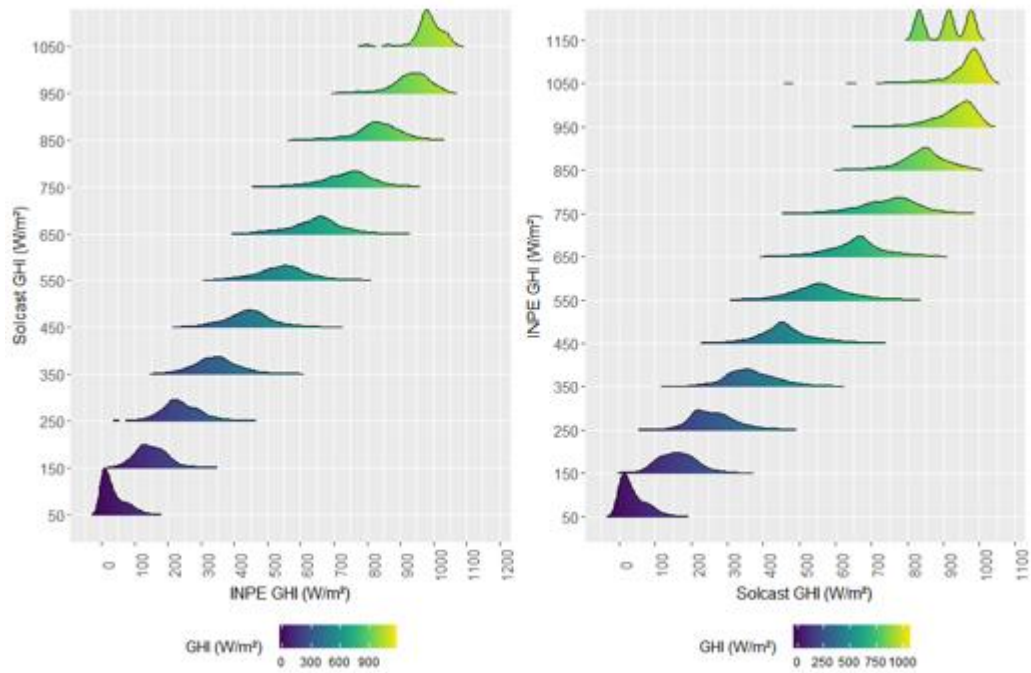
À esquerda no Gráfico 13 é ilustrado que praticamente todos os grupos de valores entre 150 e 950 W/m² de DNI estimados possuem picos aproximadamente em zero, atestando um fraco desempenho do modelo - mesmo em dias no qual o valor medido é zero de irradiância direta normal na estação de Petrolina, o modelo pode estimar valores entre 100 e 1000 W/m². Isto provavelmente ocorre por conta da sua resolução espacial, que acaba não detectando nuvens de menor tamanho, assim como também suas espessuras.

Equivalentemente aos passos de confecção dos gráficos *ridge* para o produto de satélite CAMS-RAD (Gráficos 12 e 13), foram obtidos os Gráficos 14 e 15 para o Solcast.

É observável, a partir do Gráfico 14, que o modelo do Solcast possui um satisfatório desempenho para estimação da componente GHI, uma vez que o pico da curva de densidade em cada grupo se situa aproximadamente perto dos valores esperados ($x = y$), exceto para observações entre 1100 e 1200 W/m², no qual há maior probabilidade de ocorrer estimativas de 850, 950 e 1000 W/m² (Gráfico 14, à direita).

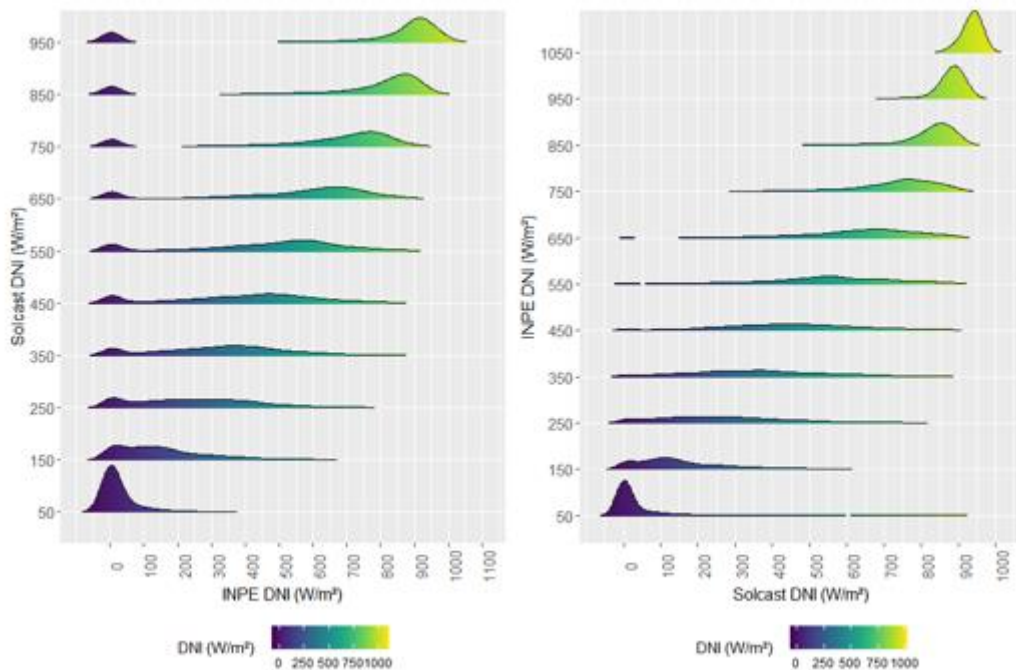
Ademais, as estimativas de DNI (Gráfico 15) possuem desvios semelhantes ao modelo do CAMS-RAD ao passo que, ainda assim, estas possuem uma performance melhor, visto que dispõem de curvas de densidade planas, mas que se posicionam próximas da tendência ideal.

Gráfico 14 – Solcast: Distribuição da GHI medida dados valores estimados (à esquerda) e distribuição da GHI estimada dados valores medidos (à direita)



Fonte: elaborado pelo autor.

Gráfico 15 – Solcast: Distribuição da DNI medida dados valores estimados (à esquerda) e distribuição da DNI estimada dados valores medidos (à direita)



Fonte: elaborado pelo autor.

5.4 Efeitos da sazonalidade

A intermitência da energia solar depende diretamente do clima local, das estações do ano, além de variar também durante o dia. Por conseguinte, a irradiância solar em uma superfície varia de acordo com o regime de nuvens, aerossóis e poluição, por exemplo. Sabe-se que a modelagem dos produtos de satélite possui erros de estimativas relacionados a condições do céu (YANG, 2020) e estas indubitavelmente são governadas pelas estações do ano de um determinado clima. Isto posto, a presente seção fica a cargo de analisar os efeitos da sazonalidade na previsão da GHI e DNI pelo CAMS e Solcast no local de estudo (Petrolina).

De acordo com Luiz *et al.* (2018), a estação chuvosa ocorre de novembro a abril (primavera e verão) e a estação seca de maio a outubro (outono e inverno), cabendo ressaltar que a estação seca tem fração mensal de cobertura por nuvens de 50% e a estação chuvosa 60%.

5.4.1 CAMS-RAD

A Tabela 12 apresenta os erros RMSE, MAE e R^2 das duas componentes solares, GHI e DNI, em cada uma das quatro estações do ano. Além disso, foi originado o Gráfico 16 a partir das informações da Tabela 12, para uma comparação visual das métricas. A linha tracejada corresponde aos erros globais de estimação do modelo para o Grupo II (Tabela 8).

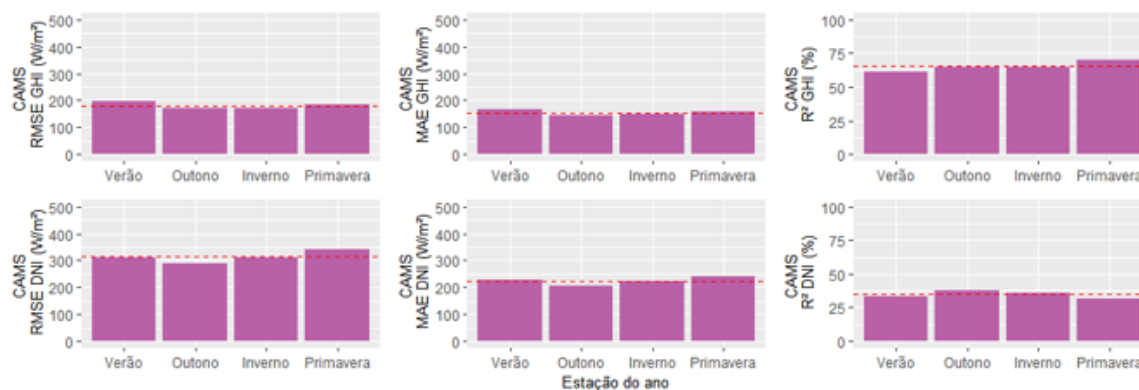
Tabela 12 – CAMS-RAD: Erros associados à estimação da GHI e DNI estação do ano, em resolução temporal de 60 minutos (Grupo II)

Ano CAMS-RAD	GHI			DNI		
	RMSE	MAE	R^2	RMSE	MAE	R^2
Verão	199,32	168,56	0,61	308,91	227,52	0,33
Outono	168,97	143,71	0,64	290,06	205,94	0,37
Inverno	170,02	147,52	0,65	310,25	222,09	0,36
Primavera	184,81	159,44	0,70	340,22	239,71	0,32

Fonte: Elaborado pelo autor.

Apesar dos valores semelhantes (Gráfico 16), os resultados de maior relevância são: (a) no verão de Petrolina encontra-se uma pior estimativa da GHI pelo modelo; (b) no outono há melhor estimativa da DNI. Porém, cabe evidenciar que os valores da Tabela 12 não possuem variações significativas dos resultados globais apresentados na Tabela 8 (representados pela linha vermelha tracejada). A seguir serão apresentadas três ferramentas de visualização de dados para verificação da estrutura destas variáveis em cada estação do ano.

Gráfico 16 – Erros de estimação da GHI e DNI pelo CAMS-RAD para cada estação do ano, em resolução temporal de 60 minutos. Neste gráfico os resultados são comparados com valores globais do Grupo II na Tabela 1 (linha tracejada)



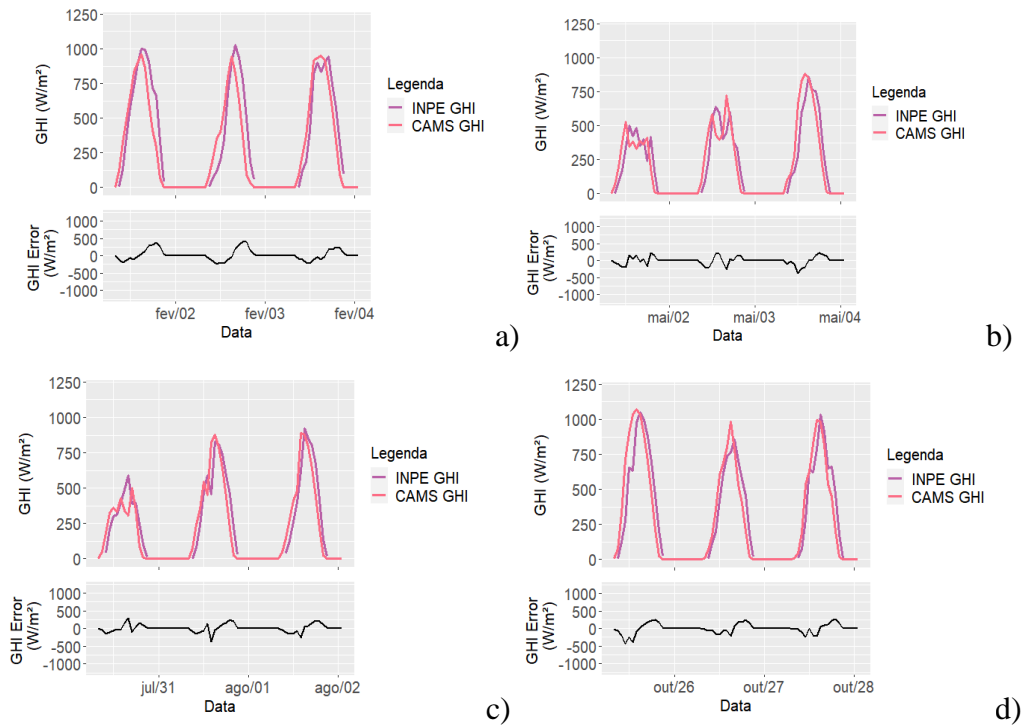
Fonte: elaborado pelo autor.

Gráficos de série temporal são bastante úteis para analisar estruturas de variáveis que mudam de comportamento ao longo do tempo, como a irradiância solar. A irradiância solar possui variações durante as 24 horas do dia, assim como também ao longo do ano, de acordo com o clima local e estações. Em virtude disso, para visualizar e comparar o comportamento da irradiância de Petrolina em cada estação do ano, foram escolhidos 5 dias sequenciais entre meados de cada estação do ano de 2017 para confecção dos gráficos de series temporais.

O Gráfico 17 exhibe no topo curvas de série temporal, uma de GHI medida (INPE-Petrolina) e outra de GHI estimada (CAMS) e, abaixo, o erro que diz respeito à diferença entre as curvas das séries temporais - tudo ao longo de três dias em meados do (a) verão, (b) outono, (c) inverno e (d) primavera. O Gráfico 18 foi confeccionado sob as mesmas premissas, só que com curvas de DNI medida e DNI estimada.

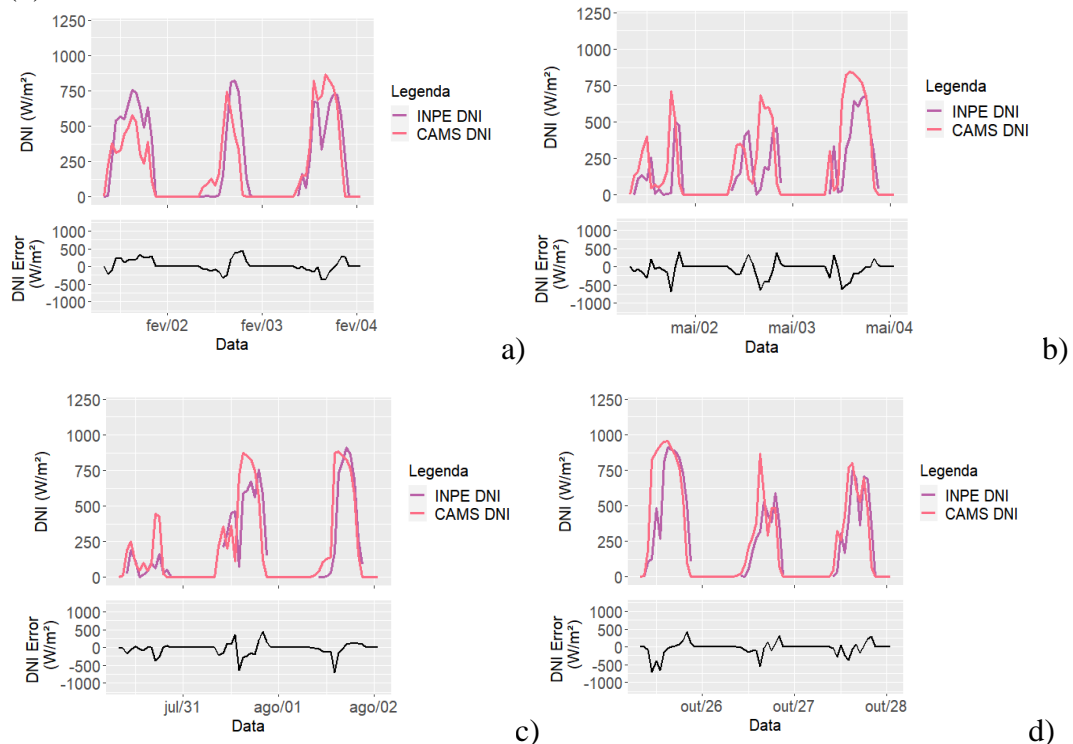
As séries temporais expostas no Gráfico 17 evidenciam uma GHI com maiores magnitudes no verão (1 a 5 de fevereiro; Figura 12(a) e na primavera (25 a 29 de outubro; Figura 12(d)). Pontos em que se manifestam erros mais acentuados possuem valores entre +500 W/m² e -500 W/m². No caso da componente DNI (Gráfico 18), os períodos selecionados de cada estação do ano não expressam padrões significativos, além dos picos de erros na metade e no final de cada dia entre +500 W/m² e -750 W/m².

Gráfico 17 – Série temporal da irradiância global horizontal medida (INPE) e estimada (CAMS) ao longo de 3 dias com erro associado. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

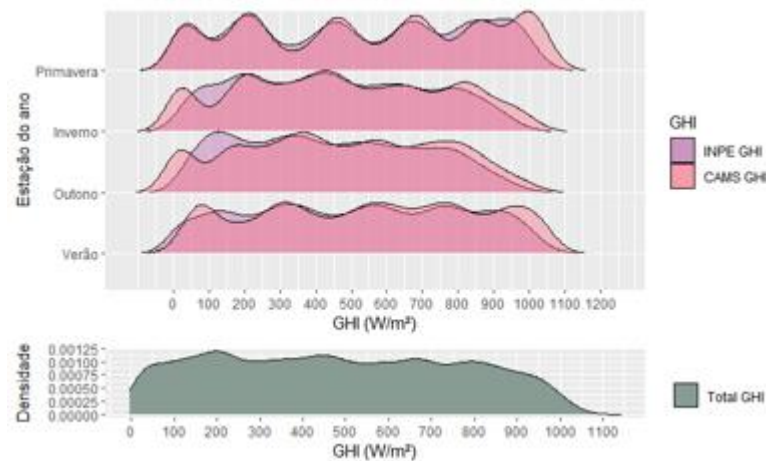
Gráfico 18 – Série temporal da irradiância direta normal medida (INPE) e estimada (CAMS) ao longo de 3 dias com erro associado. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

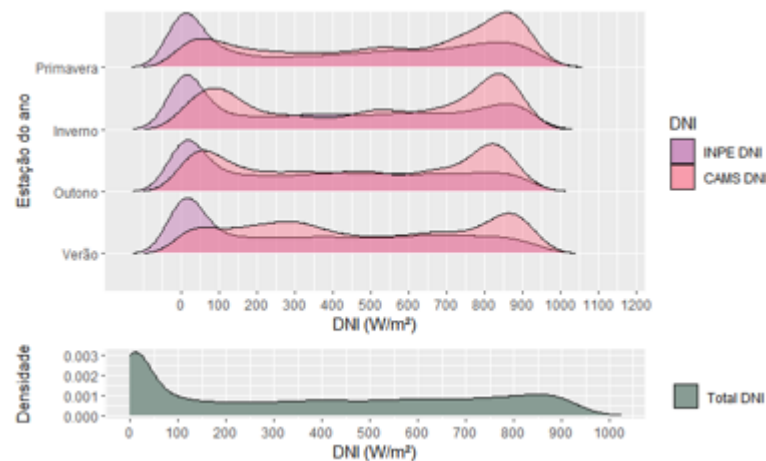
Com intuito de confrontar as distribuições de irradiância entre as estações do ano, foi montado um gráfico de Ridgeline (Gráfico 19) com ambas as densidades de GHI estimada e medida ao longo das estações do ano (superior) e distribuição da GHI total (inferior). O Gráfico 20 apresenta as mesmas informações, só que para a DNI.

Gráfico 19 – Distribuição da irradiância global horizontal estimada (CAMS) e medida (INPE) ao longo das estações do ano (acima) e distribuição da GHI total (abaixo)



Fonte: elaborado pelo autor.

Gráfico 20 – Distribuição da irradiância direta normal estimada (CAMS) e medida (INPE) ao longo das estações do ano (acima) e distribuição da DNI total (abaixo)



Fonte: elaborado pelo autor.

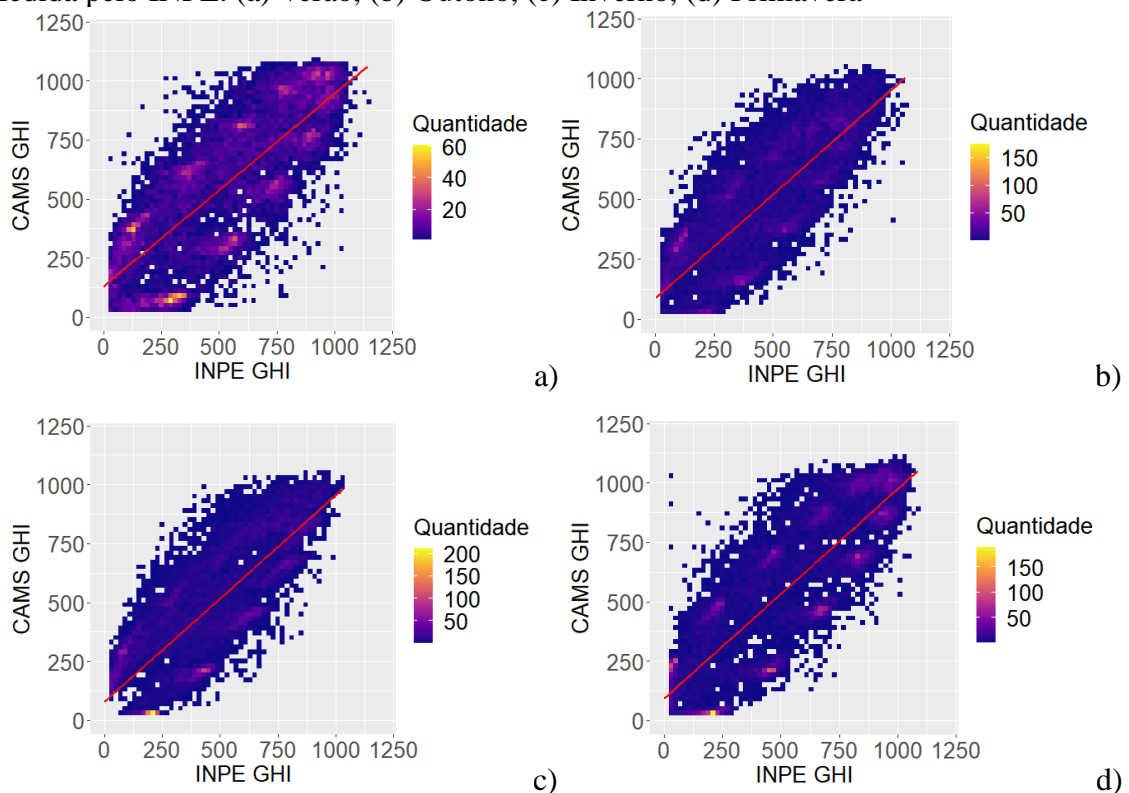
No Gráfico 19 as curvas de densidade do outono e do inverno são mais curtas e uniformes, o que talvez induza a uma menor probabilidade de erros na previsão de irradiância nestas estações. O verão e a primavera constituem majoritariamente a estação chuvosa,

expressando curvas com maior variação e, apesar de visualmente aparentar melhor ajuste entre as duas curvas, não é possível obter observações conclusivas.

Histogramas 2D são uma alternativa a diagramas de dispersão, principalmente quando há uma enorme quantidade de pontos. Um histograma 2D também mostra a relação entre duas variáveis de mesma unidade, porém ele separa a área em pequenos retângulos e a quantidade de pontos sobrepostos em cada um é representado por uma cor. O intuito desse tipo de gráfico é ilustrar melhor a concentração de dados da relação entre duas variáveis.

Deste ponto de vista, histogramas 2D da GHI estimada pelo CAMS vs. Medida pelo INPE são apresentados no Gráfico 21. E, por fim, histogramas 2D da DNI estimada pelo CAMS vs. Medida pelo INPE são apresentados no Gráfico 22.

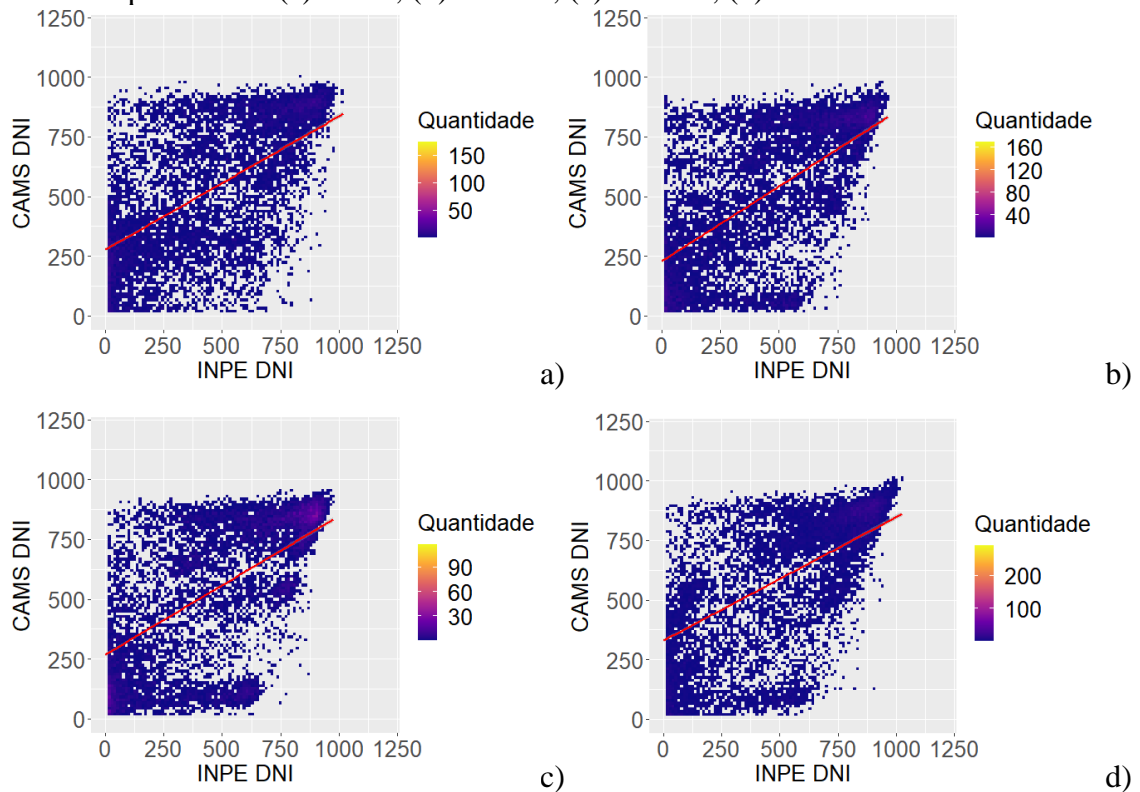
Gráfico 21 – Histograma 2D da irradiância global horizontal estimada pelo CAMS vs. Medida pelo INPE. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

Nesta última análise dos produtos de satélite CAMS Radiation, utilizando o histograma 2D, torna-se possível reunir as informações anteriores e derivar algumas conclusões. Grandes concentrações de pontos sobrepostos estão posicionadas de forma alternada acima e abaixo da linha de tendência (Gráfico 21a), isto evidencia um pior desempenho de previsão de GHI em comparação às outras estações do ano.

Gráfico 22 – Histograma 2D da irradiância direta normal estimada pelo CAMS vs. Medida pelo INPE. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

No Gráfico 22, todas as estações possuem grande concentração de pontos em ambas extremidades da linha de tendência e também possuem grande dispersão. Apesar disso, o outono (Gráfico 22b) ainda possui maior quantidade de retângulos próximos da linha de tendência que os outros três.

5.4.2 Solcast

A presente seção apresenta tabelas e gráficos análogos aos da Seção 5.4.1, porém obtidos a partir da comparação entre os dados de irradiância modelados pelo Solcast e os dados medidos da estação de Petrolina.

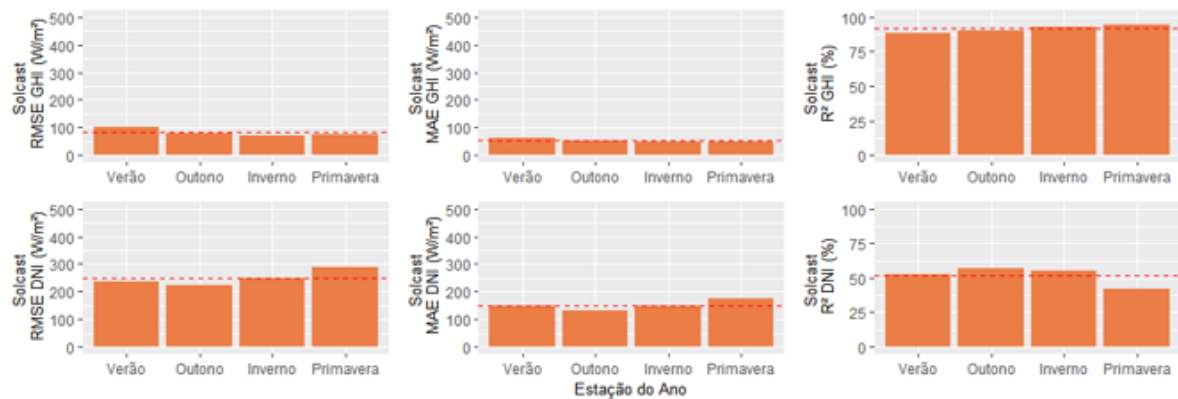
Conforme apresentado na Tabela 13 e Gráfico 23, o verão demonstra desempenho inferior nas estimativas da GHI, enquanto o outono possui melhor desempenho nas estimativas da DNI, ressaltando que o verão é a estação chuvosa e o outono é estação seca, que influem respectivamente uma menor e maior performance

Tabela 13 – Solcast: Erros associados à estimação da GHI e DNI, para cada estação do ano, em resolução temporal de 60 minutos (Grupo II)

Ano Solcast	GHI			DNI		
	RMSE	MAE	R ²	RMSE	MAE	R ²
Verão	100,00	63,38	0,89	235,86	146,47	0,53
Outono	79,30	53,11	0,90	221,61	130,66	0,57
Inverno	72,09	48,01	0,93	250,87	148,16	0,55
Primavera	75,19	47,62	0,94	290,53	173,55	0,42

Fonte: Elaborado pelo autor.

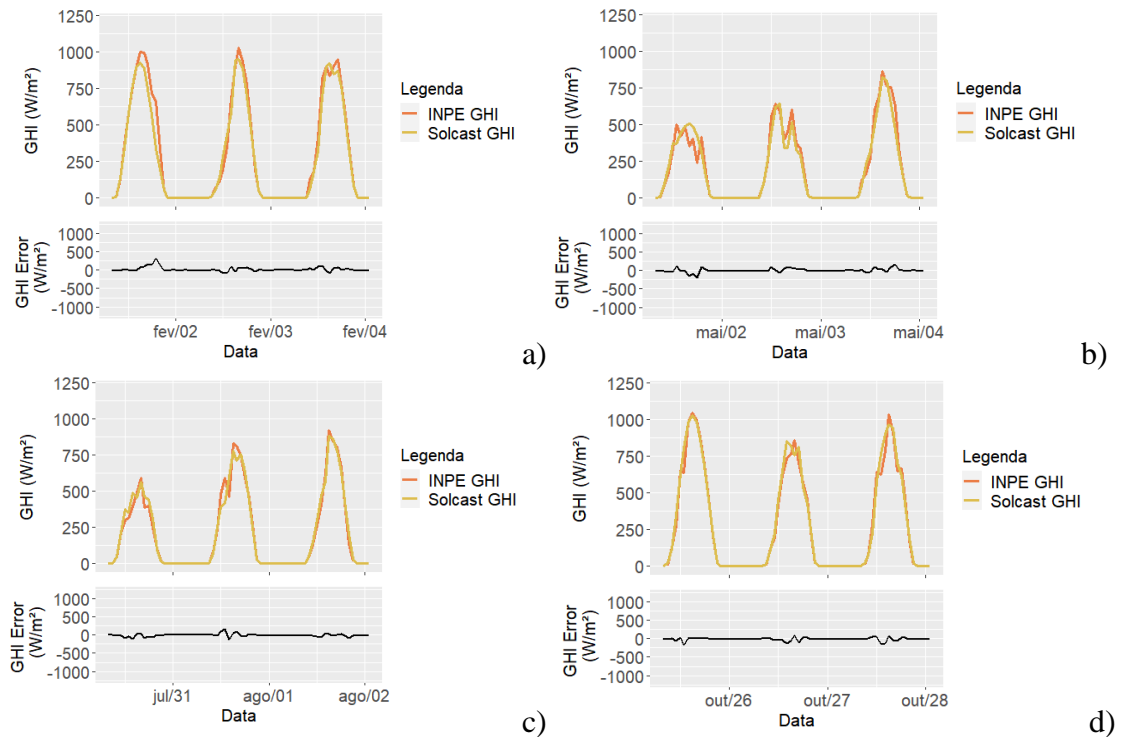
Gráfico 23 – Erros de estimação da GHI e DNI pelo Solcast para cada estação do ano, em resolução temporal de 60 minutos. Neste gráfico os resultados são comparados com valores totais da Tabela 9 (linha tracejada)



Fonte: elaborado pelo autor.

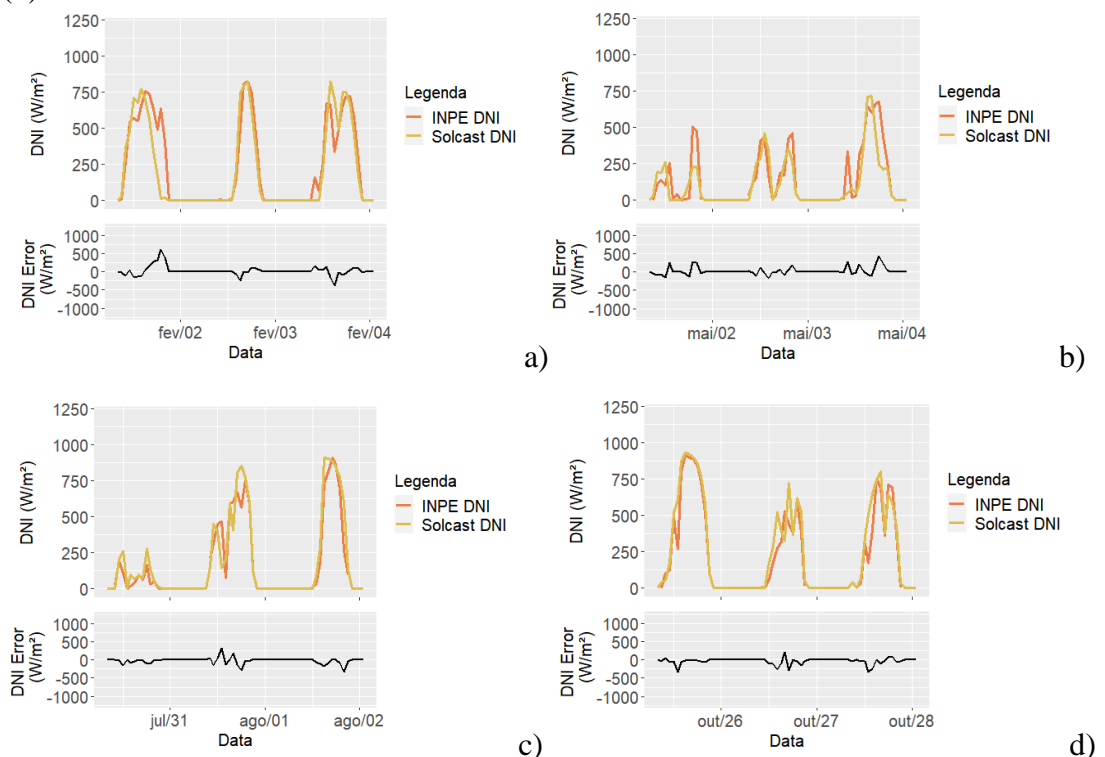
As séries temporais da GHI medida (INPE) e estimada (Solcast) são praticamente sobrepostas e estáveis nas quatro estações, com erros mantidos em valores máximos de ± 250 W/m² (Gráfico 24). No Gráfico 25 percebe-se novamente uma dificuldade de ajuste entre as curvas de medição e previsão de DNI, com erros máximos chegando a ± 750 W/m².

Gráfico 24 – Série temporal da irradiância global horizontal medida (INPE) e estimada (Solcast) ao longo de 3 dias com erro associado. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

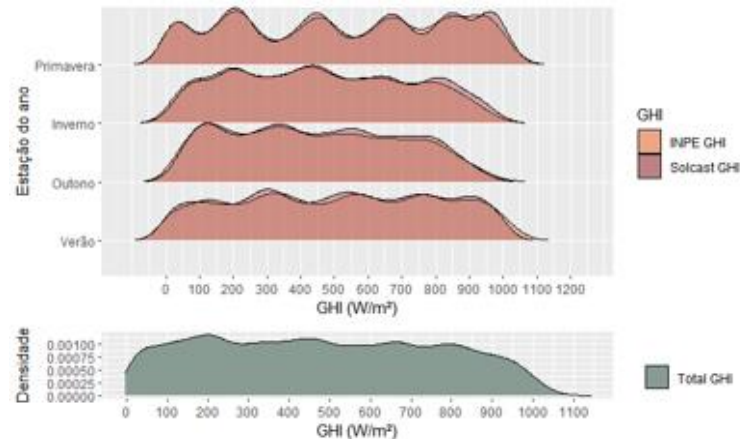
Gráfico 25 – Série temporal da irradiância direta normal medida (INPE) e estimada (Solcast) ao longo de 3 dias com erro associado. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

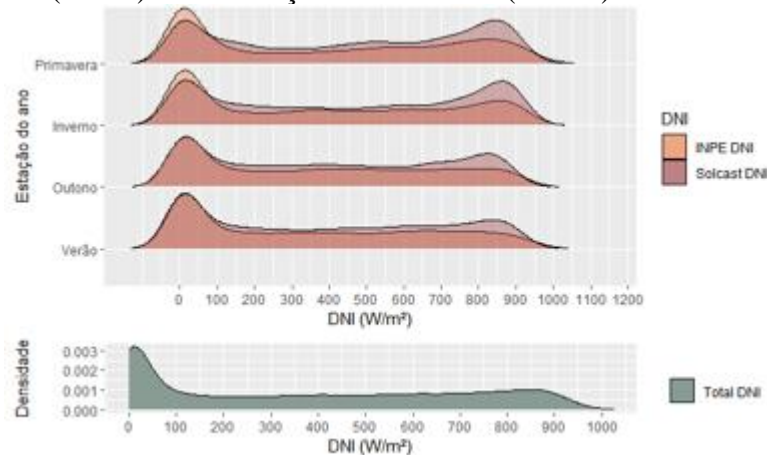
Qualitativamente, as curvas de densidade da GHI (Solcast) possuem pequenos desvios das curvas de Petrolina (Gráfico 26), tanto quanto na distribuição da DNI, vide Gráfico 27.

Gráfico 26 – Distribuição da GHI (Solcast) e medida (INPE) ao longo das estações do ano (acima) e distribuição da GHI total (abaixo)



Fonte: elaborado pelo autor.

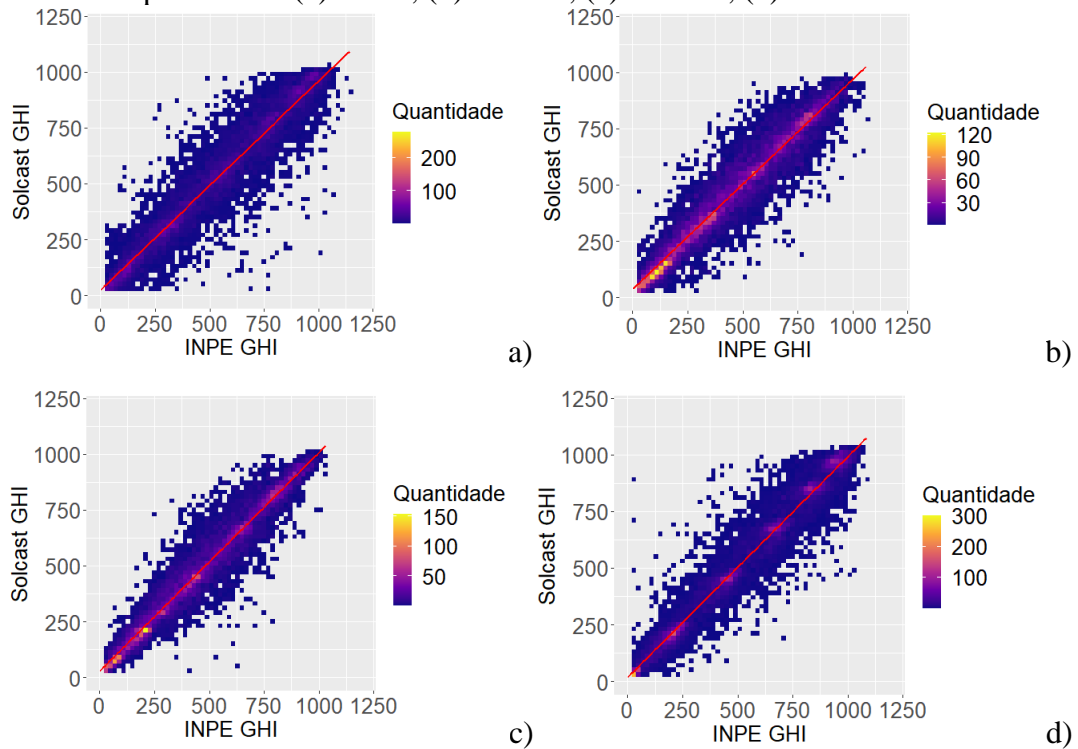
Gráfico 27 – Distribuição da irradiância direta normal estimada (Solcast) e medida (INPE) ao longo das estações do ano (acima) e distribuição da DNI total (abaixo)



Fonte: elaborado pelo autor.

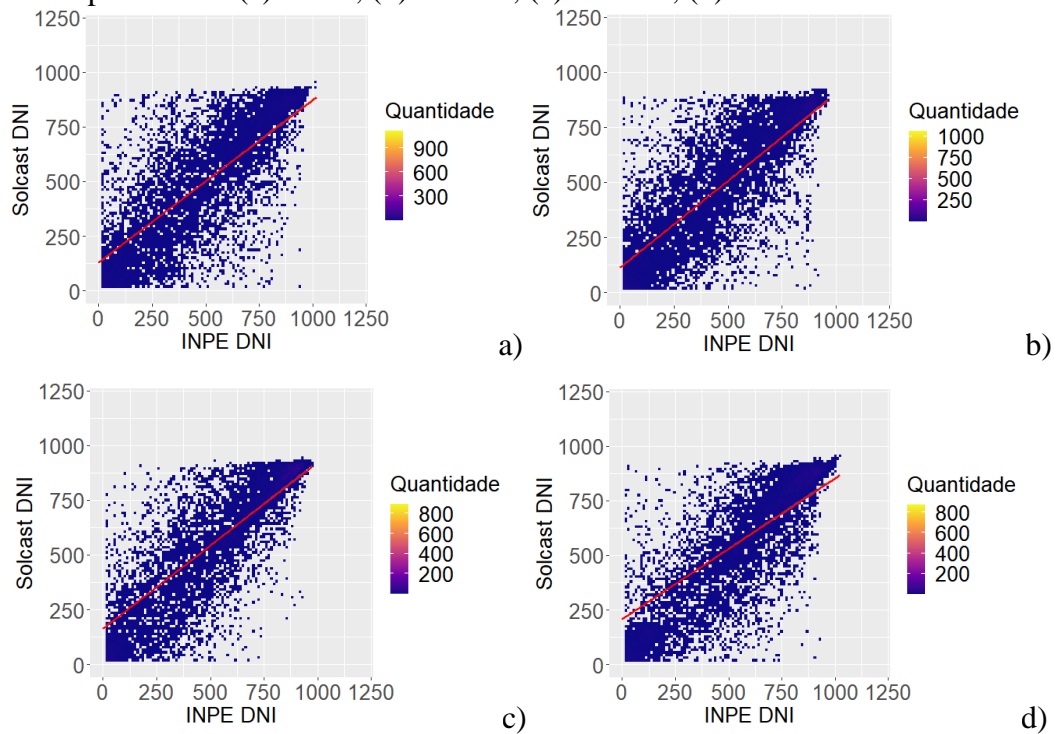
Conforme os histogramas 2D do Gráfico 28, que representam a componente GHI em cada estação, é notável uma pior concentração de pontos sobrepostos ao longo da linha de tendência no verão (Gráfico 28a) e maior concentração desses pontos na linha de tendência na estação do outono e inverno (Gráfico 28b e Gráfico 28c).

Gráfico 28 – Histograma 2D da irradiância global horizontal estimada pelo Solcast vs. Medida pelo INPE. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

Gráfico 29 – Histograma 2D da irradiância direta normal estimada pelo Solcast vs. Medida pelo INPE. (a) Verão, (b) Outono, (c) Inverno, (d) Primavera



Fonte: elaborado pelo autor.

5.5 Outras resoluções temporais

O estudo de verificação da acurácia até esta seção foi realizado em resolução temporal de 60 minutos, porém, aqui serão apresentados alguns resultados orientados às métricas de erro em várias outras resoluções temporais. O CAMS oferece resoluções de 1-, 15-, 60-min, diário e mensal, enquanto resoluções de 5-, 10-, 15-, 30- e 60-min são disponibilizados pelo Solcast. A Tabela 14 apresenta os erros obtidos em todas estas resoluções temporais oferecidas pelos satélites na estimativa de GHI. Igualmente, o mesmo conjunto de resultados é apresentado na Tabela 15, porém para a DNI.

Tabela 14 – Erros associados à estimação da GHI em variadas resoluções temporais (Grupo II). Valores com um traço correspondem a resoluções não disponibilizadas pelos serviços

Métricas	CAMS			Solcast		
	RMSE	MAE	R ²	RMSE	MAE	R ²
1 min	162,22	103,95	0,75	-	-	-
5 min	-	-	-	130,61	84,39	0,81
10 min	-	-	-	117,32	76,11	0,84
15 min	129,4	91,2	0,83	110,02	71,30	0,86
30 min	-	-	-	93,14	61,87	0,89
Diário	51,37	39,74	0,38	-	-	-
Mensal	38,98	32,66	0,34	-	-	-

Fonte: Elaborado pelo autor.

Tabela 15 – Erros associados à estimação da DNI em variadas resoluções temporais (Grupo II). Valores com um traço correspondem a resoluções não disponibilizadas pelos serviços

Métricas	CAMS			Solcast		
	RMSE	MAE	R ²	RMSE	MAE	R ²
1 min	310,09	201,93	0,44	-	-	-
5 min	-	-	-	282,14	175,20	0,45
10 min	-	-	-	272,64	168,35	0,47
15 min	301,3	197,3	0,43	277,36	170,85	0,46
30 min	-	-	-	253,76	154,54	0,51
Diário	133,96	103,19	0,15	-	-	-
Mensal	90,13	69,30	0,0085	-	-	-

Fonte: Elaborado pelo autor.

Com RMSE de respectivamente 130,61 Wm⁻² e 93,14 Wm⁻², em 5 e 30 minutos, no caso do Solcast, contra 310,09 Wm⁻² e 90,13 Wm⁻², nos valores do CAMS de 1 minuto e

mensais, é observável que o erro diminui concomitantemente ao aumento da resolução temporal. Consequentemente, as médias móveis com maior número de dados ponderam a distância entre estimativa e observações reais. Por fim, valores diários e mensais apresentaram variações menores e resultaram em um R^2 significativamente abaixo da média.

6 CONCLUSÕES

Primeiramente, a análise bibliométrica destacou a relevância das pesquisas voltadas à irradiância obtida a partir do sensoriamento remoto em satélites, que continua ascendente nas últimas duas décadas. Além disso, o Brasil, apesar de se posicionar no top 20 países que mais focam esforços no tópico, tem como premissa a urgência de uma transição da matriz elétrica. Desta forma, é esperado maiores investimentos no potencial de aproveitamento solar e um consequente crescimento de pesquisas que envolvem irradiações sintéticas ou sensoriamento remoto.

Constatou-se também que os produtos de satélite derivados do serviço Solcast possuem desempenho superior ao CAMS-RAD para a região de Petrolina, no qual o primeiro tem métricas de erro significativamente menores para as estimações de GHI (RMSE de 82,36 Wm^{-2} contra 181,06 Wm^{-2}) e DNI (RMSE de 250,98 Wm^{-2} contra 313,31 Wm^{-2}), respectivamente, e distribuições de irradiância sintética que representam as observações reais de forma mais acurada.

Ademais, por conta da fração mensal de cobertura por nuvens nas estações secas (outono) e chuvosas (verão), pode explicar o motivo do outono possuir melhores estimativas e o verão apresenta as piores estimativas em ambos os produtos, consequência de maior presença de nuvens. Apesar disso, a diferença entre as estimativas de cada estação não é relevante.

REFERÊNCIAS

- AMEEN, Bikhtiyar et al. Validation of Hourly Global Horizontal Irradiance for Two Satellite-Derived Datasets in Northeast Iraq. **Remote Sensing**, v. 10, n. 10, p. 1651, 2018.
- AGÊNCIA NACIONAL DE ÁGUAS (BRASIL). Sistema Interligado Nacional. Brasília, 2021. Disponível em: <https://www.ana.gov.br/sar/sin>. Acesso em: 31 ago. 2021.
- ARIA, M; CUCCURULLO, C. bibliometrix: An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**. v. 11, 959-975, 2017. doi:10.1016/j.joi.2017.08.007
- BRASIL, Ministério de Minas e Energia; Empresa de Pesquisa Energética. **Plano Nacional de Energia 2050**. Brasília, DF, 2020.
- BRIGHT, J. M. Solcast: Validation of a satellite-derived solar irradiance dataset. **Solar Energy**, v. 189, p. 435–449, 2019a. doi:10.1016/j.solener.2019.07.086
- BRIGHT, J. M. The impact of globally diverse GHI training data: Evaluation through application of a simple Markov chain downscaling methodology. **Journal of Renewable and Sustainable Energy**, v. 11, n. 2, 2019b. doi:10.1063/1.5085236
- BRIGHT, J. M. “Introduction To Synthetic Solar Irradiance”. Synthetic Solar Irradiance: Modeling Solar Data [AIP Publishing (online), Melville, New York, 2021], Disponível em: https://doi.org/10.1063/9780735421820_001
- BRIGHT, J. M; GUEYMARD, C. A. Climate-specific and global validation of MODIS Aqua and Terra aerosol optical depth at 452 AERONET stations. **Solar Energy**, v. 183, p. 594–605, 2019. doi:10.1016/j.solener.2019.03.043
- CLARKE, D. E et al. Which Reanalysis Dataset Should We Use for Renewable Energy Analysis in Ireland?. **Atmosphere**, v. 12, n. 5, p. 624, 2021. <https://doi.org/10.3390/atmos12050624>
- EMPRESA DE PESQUISA ENERGÉTICA. **Anuário estatístico de energia elétrica 2020**, 2020. Disponível em: <<https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/anuario-estatistico-de-energia-eletrica>>. Acesso em: 2 jul. 2021
- FRIMANE, A; BRIGHT, J. M. “Validation of Synthetic Solar Irradiance Data” em Synthetic Solar Irradiance: Modeling Solar Data [AIP Publishing (online), Melville, New York, 2021], Disponível em: https://doi.org/10.1063/9780735421820_004
- GÓMEZ, J. M. R et al. A irradiância solar: conceitos básicos. **Revista Brasileira de Ensino de Física**, v. 40, n. 3, 2018. doi:10.1590/1806-9126-rbef-2017-0342

GUEYMARD, C. A; RUIZ-ARIAS, J. A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. **Solar Energy**, v. 128, p. 1–30, 2016. doi:10.1016/j.solener.2015.10.010

GUEYMARD, C. A. A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. **Solar Energy**, v. 168, p. 2–9, 2018. doi: 10.1016/j.solener.2018.04.001

HUANG, G et al. Estimating surface solar irradiance from satellites: Past, present, and future perspectives. **Remote Sensing of Environment**. v. 233, p. 111371, 2019. doi: 10.1016/j.rse.2019.111371

INEICHEN, P. Long Term Satellite Global, Beam and Diffuse Irradiance Validation. **Energy Procedia**, v. 48, p. 1586–1596, 2014. doi:10.1016/j.egypro.2014.02.179

INEICHEN, P. Long term HelioClim-3 global, beam and diffuse irradiance validation. 2016. Disponível em: <https://archive-ouverte.unige.ch/unige:81915>

KAMATH, H.; SRINIVASAN. J. Validation of global irradiance derived from INSAT-3D over India. **Solar Energy**. v. 202, p. 45-54. doi:10.1016/j.solener.2020.03.084

KUHN, Max; JOHNSON, Kjell. **Applied Predictive Modeling**. Edição 1. Nova York: Springer Science, Business Media 2013, doi: 10.1007/978-1-4614-6849-3

LIANG, S. et al. Remote sensing of earth's energy budget: synthesis and review. **International Journal of Digital Earth**, v. 12, n. 7, p. 737-780, 2019
DOI: 10.1080/17538947.2019.1597189

LUIZ, E. W et al. Analysis of intra-day solar irradiance variability in different Brazilian climate zones. **Solar Energy**, v. 167, p. 210–219, 2018. doi:10.1016/j.solener.2018.04.005

MARTÍN, C. M. Satélites meteorológicos. Física del caos en la predicción meteorológica. Madrid: Agencia Estatal de Meteorología, p. 83-102, 2018. doi:10.31978/014-18-009-X.07

MUNKHAMMAR, J; WIDÉN, J. A spatiotemporal Markov-chain mixture distribution model of the clear-sky index. **Solar Energy**, v. 179, p. 398–409, 2019. doi:10.1016/j.solener.2018.12.064

NONNENMACHER, L; KAUR, A; COIMBRA, C. F. M. Verification of the SUNY direct normal irradiance model with ground measurements. **Solar Energy**, v. 99, 246–258, 2014. doi:10.1016/j.solener.2013.11.010

PEREZ, R et al. Producing satellite-derived irradiances in complex arid terrain. **Solar Energy**, v. 77, n. 4, 367–371, 2004 doi:10.1016/j.solener.2003.12.016

POLO, J.; PEREZ, R. Solar Radiation Modeling from Satellite Imagery. **Solar Resources Mapping**, p. 183-197, 2019.

PERUCHENA, C. F et al. High frequency generation of coupled GHI and DNI based on clustered Dynamic Paths. **Solar Energy**, v. 159, p. 453–457, 2018. doi:10.1016/j.solener.2017.11.024

RESNICK, Robert; HALLIDAY, David; KRANE, Kenneth S. **Física**. México: Grupo Editorial Patria, 2008.

RODRÍGUEZ-BENÍTEZ, F. J et al. A short-term solar radiation forecasting system for the Iberian Peninsula. Part 1: Models description and performance assessment. **Solar Energy**, v. 195, p. 396–412, 2020. doi:10.1016/j.solener.2019.11.028

RUIZ-ARIAS, J. A; GUEYMARD, C. A. Worldwide inter-comparison of clear-sky solar radiation models: Consensus-based review of direct and global irradiance components simulated at the earth surface. **Solar Energy**, v. 168, p. 10–29, 2018. doi:10.1016/j.solener.2018.02.008

SALAZAR, G. et al Solar irradiance time series derived from high-quality measurements, satellite-based models, and reanalyses at a near-equatorial site in Brazil. **Renewable and Sustainable Energy Reviews**. v. 117, p. 109478, 2020. doi:10.1016/j.rser.2019.109478

SCHEIDSTEGER, T.; HAUNSCHILD, R. Telling the story of solar energy meteorology into the satellite era by applying (co-citation) reference publication year spectroscopy. **Scientometrics**. v. 125, p. 1159-1177, 2020. doi: 10.1007/s11192-020-03597-0

SUN, X et al. Worldwide performance assessment of 75 global clear-sky irradiance models using Principal Component Analysis. **Renewable and Sustainable Energy Reviews**, v. 111, p. 550–570, 2019. doi:10.1016/j.rser.2019.04.006

THOMAS, C et al. Validation of three satellite-derived databases of surface solar radiation using measurements performed at 42 stations in Brazil. **Adv. Sci. Res**, v. 13, p. 81–86, 2016a. <https://doi.org/10.5194/asr-13-81-2016>

THOMAS, C et al. Validation of CAMS Radiation (2004 - J-1) against the measurements of 11 INPE stations, 2016b Disponível em: <<http://www.soda-pro.com/help/cams-radiation-service/validation/brazil-inpe>>. Acesso em: 20 fev. 2021.

THRUN, M.C; GEHLERT, T; ULTSCH, A. Analyzing the fine structure of distributions. **PLoS ONE**, v. 15, n. 10, p. e0238835, 2020. <https://doi.org/10.1371/journal.pone.0238835>. Disponível em: <<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0238835#references>>. Acesso em: 11 mai. 2021

VERNAY, C; PITAVAL, S; BLANC, P. Review of Satellite-based Surface Solar Irradiation Databases for the Engineering, the Financing and the Operating of Photovoltaic Systems. **Energy Procedia**, v. 57, p. 1383–1391, 2014. doi:10.1016/j.egypro.2014.10.129

WILKE, C. O. Fundamentals of data visualization: a primer on making informative and compelling figures, O'Reilly Media, 2019. Disponível em <<https://clauswilke.com/dataviz/>>. Acesso em 11 mai. 2021.

YANG, D. A correct validation of the National Solar Radiation Data Base (NSRDB). **Renewable and Sustainable Energy Reviews**, v. 97, p. 152–155, 2018. doi:10.1016/j.rser.2018.08.023

YANG, D. Ensemble model output statistics as a probabilistic site-adaptation tool for satellite-derived and reanalysis solar irradiance. **Journal of Renewable and Sustainable Energy**, v. 12, n. 1, 2020. doi:10.1063/1.5134731

YANG, D; BRIGHT, J. M. Worldwide validation of 8 satellite-derived and reanalysis solar radiation products: A preliminary evaluation and overall metrics for hourly data over 27 years. **Solar Energy**, v. 210, p. 3-19, 2020. doi:10.1016/j.solener.2020.04.016

YANG, D. Validation of the 5-min irradiance from the National Solar Radiation Database (NSRDB). **J. Renewable Sustainable Energy**, 13, p. 016101, 2021; doi: 10.1063/5.0030992

QU, Z et al. Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method, **Meteorol. Z.**, v. 26, p. 33-57, doi: 10.1127/metz/2016/0781,2017.

ZHANG, W. et al A stochastic downscaling approach for generating high-frequency solar irradiance scenarios. **Solar Energy**, v. 176, p. 370–379, 2018. doi:10.1016/j.solener.2018.10.019