



**INSTITUTO LATINO-AMERICANO DE
TECNOLOGIA, INFRAESTRUTURA E
TERRITÓRIO (ILATIT)**

ENGENHARIA DE MATERIAIS

**APRENDIZADO PROFUNDO APLICADO À PREDIÇÃO DA IRRADIÂNCIA SOLAR
POR MEIO DO MÉTODO HÍBRIDO DE PREDIÇÃO**

GUSTAVO CAMPOI DE SOUZA

Foz do Iguaçu
2025



**INSTITUTO LATINO-AMERICANO DE
TECNOLOGIA, INFRAESTRUTURA E
TERRITÓRIO (ILATIT)**

ENGENHARIA DE MATERIAIS

**APRENDIZADO PROFUNDO APLICADO À PREDIÇÃO DA IRRADIÂNCIA SOLAR
POR MEIO DO MÉTODO HÍBRIDO DE PREDIÇÃO**

GUSTAVO CAMPOI DE SOUZA

Trabalho de Conclusão de Curso apresentado ao Instituto Latino-Americano de Tecnologia, Infraestrutura e Território da Universidade Federal da Integração Latino-Americana, como requisito parcial à obtenção do título de Bacharel em Engenharia de Materiais.

Orientador: Prof. Dr. Joylan Nunes Maciel
Co-orientadora: Prof. Dra. Priscila Lemes

Foz do Iguaçu
2025

GUSTAVO CAMPOI DE SOUZA

**APRENDIZADO PROFUNDO APLICADO À PREDIÇÃO DA IRRADIÂNCIA SOLAR
POR MEIO DO MÉTODO HÍBRIDO DE PREDIÇÃO**

Trabalho de Conclusão de Curso apresentado ao Instituto Latino-Americano de Tecnologia, Infraestrutura e Território da Universidade Federal da Integração Latino-Americana, como requisito parcial à obtenção do título de Bacharel em Engenharia de Materiais.

BANCA EXAMINADORA

Orientador: Prof. Dr. Joylan Nunes Maciel
UNILA

Coorientadora: Profa. Dra. Priscila Lemes
UNILA

Profa. Dra. Liliane Cristina Battirola
UNILA

Prof. Dr. Willian Zalewski
UNILA

Foz do Iguaçu, 13 de Março de 2025.

Dedico este trabalho à minha mãe, cujo apoio e amor incondicionais foram essenciais em toda a minha jornada.

AGRADECIMENTOS

Agradeço, em primeiro lugar, aos meus pais e à minha irmã pelo apoio incondicional e por sempre acreditarem em mim. O amor e a dedicação de vocês foram fundamentais em minha trajetória.

Ao meu professor orientador, Joylan Nunes Maciel, pelos puxões de orelha, pelo apoio incondicional e, acima de tudo, pela amizade. Sem dúvida, sua orientação foi fundamental nessa trajetória. Mais do que um amigo, considero-o um verdadeiro irmão mais velho.

À minha professora coorientadora, Priscila Lemes, pela constante orientação, pelo carinho e pela amizade, que, sem dúvida, levarei para a vida.

Aos professores da banca, pelas valiosas orientações e contribuições para o aprimoramento deste trabalho.

Aos meus colegas de curso, que estiveram presentes durante toda essa jornada. Sem dúvida, foi um prazer conhecê-los e compartilhar esse caminho com vocês.

E, por fim, à minha namorada, que esteve ao meu lado nos melhores e piores momentos ao longo dessa trajetória. Seu amor e carinho tornaram esse percurso mais leve e especial.

*A vida não é sobre quão duro você é capaz de
bater, mas sobre quão duro você é capaz de
apanhar e continuar indo em frente.*

Rocky Balboa

RESUMO

A predição da irradiância solar é essencial para otimizar sistemas fotovoltaicos e mitigar a intermitência da geração solar na rede elétrica. Este estudo avalia e compara a acurácia do Método Híbrido de Predição (MHP) utilizando Redes Neurais Artificiais (RNA) *MultiLayer Perceptron* (MLP) e duas arquiteturas de Aprendizado Profundo (*Deep Learning*): Redes Neurais Recorrentes (RNN) e *Long Short-Term Memory* (LSTM). O objetivo é verificar se modelos que capturam dependências temporais (RNN e LSTM) proporcionam maior acurácia preditiva em comparação à RNA convencional do MHP. Os experimentos foram realizados com dados de irradiância solar e imagens all-sky (180°) coletados entre 2014 e 2016, considerando três horizontes de predição (1, 5 e 15 minutos). Foram aplicadas validação cruzada de 3-folds e otimização Bayesiana para ajuste de hiperparâmetros. A acurácia foi avaliada pelas métricas RMSE, MAE e R². Os resultados indicam que a RNA obteve o menor RMSE médio geral (70,42), enquanto RNN apresentou o maior erro médio (75,5). Apesar da pequena vantagem da RNA, o teste estatístico de Friedman revelou que as diferenças de desempenho entre os modelos avaliados não são estatisticamente significativas ($p > 0,05$). As principais contribuições do estudo incluem: (i) demonstrar que modelos mais simples (RNA) podem alcançar acurácia semelhante a arquiteturas mais robustas (RNN e LSTM), evitando custos computacionais elevados; (ii) reforçar a importância do pré-processamento de imagens *all-sky* do MHP, que pode reduzir a necessidade de redes recorrentes; e (iii) fornecer um protocolo experimental estruturado para avaliação de modelos de predição de irradiância solar. Estudos futuros podem explorar modelos híbridos, diferentes horizontes temporais e abordagens baseadas em *Transformers* para aprimorar a eficácia do MHP.

Palavras-chave: método híbrido de predição; energia solar fotovoltaica; inteligência artificial; acurácia de predição.

RESUMEN

La predicción de la irradiancia solar es esencial para optimizar los sistemas fotovoltaicos y mitigar la intermitencia de la generación solar en la red eléctrica. Este estudio evalúa y compara la precisión del Método Híbrido de Predicción (MHP) utilizando Redes Neuronales Artificiales (RNA) MultiLayer Perceptron (MLP) y dos arquitecturas de Aprendizaje Profundo (Deep Learning): Redes Neuronales Recurrentes (RNN) y Long Short-Term Memory (LSTM). El objetivo es verificar si los modelos que capturan dependencias temporales (RNN y LSTM) proporcionan una mayor precisión predictiva en comparación con la RNA convencional del MHP. Los experimentos se realizaron con datos de irradiancia solar e imágenes all-sky (180°) recopilados entre 2014 y 2016, considerando tres horizontes de predicción (1, 5 y 15 minutos). Se aplicaron validación cruzada de 3 particiones (3-folds) y optimización Bayesiana para el ajuste de hiperparámetros. La precisión se evaluó mediante las métricas RMSE, MAE y R^2 . Los resultados indican que la RNA obtuvo el RMSE promedio más bajo (70,42), mientras que RNN presentó el mayor error promedio (75,5). A pesar de la ligera ventaja de la RNA, la prueba estadística de Friedman reveló que las diferencias de rendimiento entre los modelos no son estadísticamente significativas ($p > 0,05$). Las principales contribuciones del estudio incluyen: (i) demostrar que modelos más simples (RNA) pueden lograr una precisión similar a arquitecturas más complejas (RNN y LSTM), evitando altos costos computacionales; (ii) reforzar la importancia del preprocesamiento de imágenes all-sky en el MHP, lo que puede reducir la necesidad de redes recurrentes; y (iii) proporcionar un protocolo experimental estructurado para la evaluación de modelos de predicción de irradiancia solar. Estudios futuros pueden explorar modelos híbridos, diferentes horizontes temporales y enfoques basados en Transformers para mejorar la eficacia del MHP.

Palabras clave: método híbrido de predicción; energía solar fotovoltaica; inteligencia artificial; precisión de predicción.

ABSTRACT

The solar irradiance prediction is essential for optimizing photovoltaic systems and mitigating the intermittency of solar generation in the power grid. This study evaluates and compares the accuracy of the Hybrid Prediction Method (MHP) using Artificial Neural Networks (RNA) MultiLayer Perceptron and two Deep Learning architectures: Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM). The objective is to verify whether models that capture temporal dependencies (RNN and LSTM) provide greater predictive accuracy compared to the conventional ANN of the HPM. Experiments were conducted using solar irradiance data and all-sky images (180°) collected between 2014 and 2016, considering three prediction horizons (1, 5, and 15 minutes). A 3-fold cross-validation and Bayesian optimization were applied for hyperparameter tuning. Accuracy was assessed using the RMSE, MAE, and R^2 metrics. The results indicate that RNA obtained the lowest overall average RMSE (70.42), while RNN presented the highest average error (75.5). Despite ANN's slight advantage, the Friedman statistical test revealed that the performance differences among the models are not statistically significant ($p > 0.05$). The main contributions of the study include: (i) demonstrating that simpler models (RNA) can achieve similar accuracy to more complex architectures (RNN and LSTM), avoiding high computational costs; (ii) reinforcing the importance of the all-sky image preprocessing in the HPM, which can reduce the need for recurrent networks; and (iii) providing a structured experimental protocol for evaluating solar irradiance prediction models. Future studies may explore hybrid models, different temporal horizons, and Transformer-based approaches to enhance the effectiveness of the MHP.

Key words: hybrid prediction method; photovoltaic solar energy; artificial intelligence; prediction accuracy.

LISTA DE ILUSTRAÇÕES

Figura 1 – Células de Silício.	16
Figura 2 – Estrutura de um cristal de silício.....	17
Figura 3 – Estrutura genérica ABX ₃ de perovskitas.	18
Figura 4 – Estrutura de uma placa solar fotovoltaica.....	19
Figura 5 – Método Híbrido de Predição (MHP).....	22
Figura 6 – Estrutura de uma Rede Neural Artificial MultiLayer Perceptron.	24
Figura 7 – Célula e estrutura de uma Recurrent Neural Network (RNN).....	25
Figura 8 – Célula e estrutura de uma Long Short-Term Network (LSTM).	26
Figura 9 – Exemplo do conjunto de medidas e irradiância solar no dataset do MHP...	28
Figura 10 – Delineamento experimental.	29
Figura 11 – Acurácia média dos modelos por horizonte de predição.	34
Figura 12 – Acurácia média geral dos modelos e composição proporcional dos folds.	35
Figura 13 – Variabilidade da acurácia dos modelos.	35

LISTA DE TABELAS

Tabela 1 – Detalhamento do dataset utilizado neste estudo.....	27
Tabela 2 – Divisão do dataset utilizando validação cruzada de 3 partições (folds).	30
Tabela 3 – Melhores Hiperparâmetros Obtidos com Bayesian Search Optimization....	31
Tabela 4 – Acurácias dos Modelos Preditivos	33
Tabela 5 – Comparações estatísticas de RMSE.....	37

LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina
AP	Aprendizado Profundo
ARIMA	<i>Autoregressive Integrated Moving Average</i>
CdTe	Telureto de Cádmio
CIGS	Cobre-Índio-Gálio-Selênio
CPV	<i>Concentrator Photovoltaics</i>
DL	<i>Deep Learning</i>
FTO	Óxido de índio dopado com estanho
GaAs	<i>Arsenieto de Gálio</i>
GHI	Irradiância solar global horizontal
GRU	<i>Gated Recurrent Units</i>
IA	Inteligência Artificial
ITO	Óxido de índio dopado com estanho
LightGBM	<i>Light Gradient Boosting Machine</i>
LSTM	<i>Long Short-Term Network</i>
MAE	<i>Mean Absolute Error</i>
MHP	Método Híbrido de Predição
ML	<i>Machine Learning</i>
MLP	<i>MultiLayer Perceptron</i>
OPV	Células solares fotovoltaicas orgânicas
PERC	<i>Passivated Emitter and Rear Cell</i>
PGESF	Predição da Geração de Energia Solar Fotovoltaica
PI	Processamento de Imagens
R ²	Coeficiente de Determinação
ReLU	<i>Rectified Linear Unit</i>
RMSE	<i>Root Mean Squared Error</i>
RNA	Rede Neural Artificial
RNN	<i>Recurrent Neural Network</i>
SVR	<i>Support Vector Regression</i>
Tanh	Tangente hiperbólica
TiO ₂	Dióxido de titânio
UNILA	Universidade Federal da Integração Latino-Americana

SUMÁRIO

1. INTRODUÇÃO	13
2. EMBASAMENTO TEÓRICO.....	16
2.1. ENERGIA SOLAR FOTOVOLTAICA.....	16
2.2. PREDIÇÃO DA GERAÇÃO DE ENERGIA SOLAR FOTOVOLTAICA (PGESF)	20
2.3. MÉTODO HÍBRIDO DE PREDIÇÃO (MHP)	21
2.4. MODELOS PREDITIVOS DE INTELIGÊNCIA ARTIFICIAL	23
2.4.1. REDES NEURAS ARTIFICIAIS.....	23
2.4.2. REDES NEURAS RECORRENTES	24
3. MATERIAL E MÉTODO	27
3.1. BANCO DE DADOS.....	27
3.2. DELINEAMENTO EXPERIMENTAL	28
3.3. FERRAMENTAS E TECNOLOGIAS	32
4. RESULTADOS E DISCUSSÃO.....	33
4.1. ACURÁCIA DE PREDIÇÃO	33
4.2. ANÁLISES DESCRITIVA E EXPLORATÓRIA.....	34
4.3. VALIDAÇÃO ESTATÍSTICA	36
5. CONCLUSÃO	38
6. REFERÊNCIAS.....	40
APÊNDICE	45

1. INTRODUÇÃO

Estudos com energias limpas têm se expandido significativamente devido à crescente demanda global por fontes de energia elétrica, sendo uma prioridade a descarbonização do setor elétrico e a promoção da sustentabilidade ambiental (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b). Entre as fontes renováveis, a energia solar fotovoltaica vem se destacando pela sua capacidade de geração e diminuição dos custos, apresentando um rápido crescimento no Brasil (ABSOLAR, 2022) e em países como China, Estados Unidos da América, Alemanha e Espanha (SOLAR GYCX, 2023). Constantemente, novas tecnologias e materiais de painéis fotovoltaicos têm sido pesquisados e desenvolvidos, aumentando a eficiência e minimizando o custo destes sistemas (RENEWABLE ENERGY AGENCY, 2020). No entanto, o processo de geração de energia a partir da fonte solar fotovoltaica é intermitente, pois, depende da irradiância solar que varia conforme as condições atmosféricas, tais como os ventos, precipitação, temperatura e, principalmente, a cobertura de nuvens que absorvem e refletem irradiância solar em diferentes níveis e direções (LAPPALAINEN; VALKEALAHTI, 2017; WANG, F. et al., 2015).

No Brasil, o crescimento dos sistemas fotovoltaicos conectados à rede tem gerado preocupações quanto à estabilidade e segurança do sistema elétrico, devido à natureza intermitente da energia solar (SHIVASHANKAR et al., 2016). Uma das abordagens utilizadas em estratégias que permitem mitigar essa intermitência, e a consequente insegurança e instabilidade da rede energética, baseia-se no uso de técnicas de Predição da Geração de Energia Solar Fotovoltaica (PGESF). A PGESF tornou-se um campo de pesquisa relevante na literatura (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2021), sendo utilizada também para aprimorar a gestão energética e o armazenamento energético *off-grid*, especialmente em sistemas que integram baterias para suportar a geração intermitente (ELLIS et al., 2012).

Eficiência dos sistemas fotovoltaicos não depende apenas da predição da irradiância, mas também das tecnologias empregadas nas células solares. As células fotovoltaicas operam com base na absorção de fótons e na conversão de energia solar em eletricidade, sendo impactadas por fatores como a irradiância incidente e a temperatura. No caso das células de silício, a tecnologia PERC aumenta a absorção de luz e reduz perdas por recombinação, melhorando a eficiência do sistema (CREMASCO et al.,

2021). Já para as perovskitas, a modulação do *bandgap* por substituição de cátions e haletos permite otimizar a absorção de luz e o transporte de cargas, favorecendo a conversão fotovoltaica (PEREIRA et al., 2024). Além disso, estudos demonstram que a cobertura de nuvens, analisada por imagens *all-sky* (180°), influencia diretamente a irradiância solar recebida, o que pode impactar a geração fotovoltaica ao alterar os níveis de corrente gerada pelos painéis (SILVA; BATISTA, 2020). Dessa forma, a integração de dados meteorológicos e ópticos não apenas permite uma modelagem mais precisa da geração de energia solar, como também contribui para a implementação de sistemas mais eficientes e resilientes.

Um aumento expressivo do número de trabalhos na PGESF pode ser observado na literatura científica, com ênfase a partir de 2017 (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2021), principalmente com Inteligência Artificial por meio de Aprendizado de Máquina (*Machine Learning* – ML) e Aprendizado Profundo (*Deep Learning* - DL) (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2006; KUMAR et al., 2020; KUMARI; TOSHNIWAL, 2021; OBANDO; CARVAJAL; PINEDA AGUDELO, 2019; VOYANT et al., 2017; WANG, H. et al., 2019; YING et al., 2023). O aumento do poder de processamento computacional favorece o uso do Aprendizado Profundo com Imagens *all-sky* (CHOW et al., 2011; DEV et al., 2019; FENG; ZHANG, 2020; RAJAGUKGUK; KAMIL; LEE, 2021; WANG, F. et al., 2020; YANG et al., 2021; ZHANG et al., 2018; ZHEN et al., 2022).

Historicamente, as Redes Neurais Artificiais (RNAs) (RUSSELL; NORVIG, 2016) constituem o modelo de ML mais aplicado na PGESF, devido a sua habilidade de capturar padrões temporais em dados complexos e não-lineares (JAILANI et al., 2023; LIM et al., 2022). Uma linha de pesquisa da Universidade Federal da Integração Latino-Americana (UNILA) tem desenvolvido estudos científicos sobre a PGESF de curto prazo (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b; MACIEL, Joylan Nunes et al., 2021; SERRANO ARDILA et al., 2022; WENTZ et al., 2022).

A partir dessas pesquisas, o estudo de (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b) propôs e avaliou um novo Método Híbrido de Predição (MHP) da irradiância solar global, que aplica conjuntamente técnicas de Processamento de Imagens (PI) e Aprendizado de Máquina. O MHP foi desenvolvido a partir de um *dataset* de dados históricos padronizado com imagens *all-sky*, informações meteorológicas e irradiância solar, com qualidade controlada e amostras coletadas a cada minuto durante o período completo de 3 anos (2014 a 2016) (PEDRO; LARSON; COIMBRA, 2019). Em

seu estudo, (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b) menciona a necessidade de se avaliar a otimização do MHP com o uso de algoritmos mais robustos de Aprendizado Profundo (AP).

Neste contexto, este estudo tem como objetivo avaliar e comparar a acurácia de predição da irradiância solar do MHP, que utiliza Rede Neural Artificial (RNA) *MultiLayer Perceptron*, a partir de duas arquiteturas distintas de redes neurais artificiais recorrentes: i) *Recurrent Neural Network* (RNN) e *Long Short-Term Memory* (LSTM), as quais são modelos de AP que conseguem trabalhar com dependências temporais (RUSSELL; NORVIG, 2016). Portanto, neste estudo busca-se identificar qual das três arquiteturas (RNA, RNN ou LSTM) provê melhor acurácia de predição no MHP, considerando os horizontes de curto prazo de 1, 5 e 15 minutos. Nesse sentido, a hipótese a ser investigada neste estudo é: o uso de redes RNN e LSTM, por reterem informações ao longo do tempo, proporcionam acurácia de predição superior à RNA originalmente adotada no MHP?

Portanto, a seção 2 descreve o embasamento teórico com conceitos e trabalhos científicos relacionados à PGESF. A seção 3 detalha os materiais e o método utilizado na avaliação experimental. A seção 4 exhibe os resultados, análises e a discussão relativa à hipótese da pesquisa. Por fim, a seção 5 descreve as conclusões e trabalhos futuros.

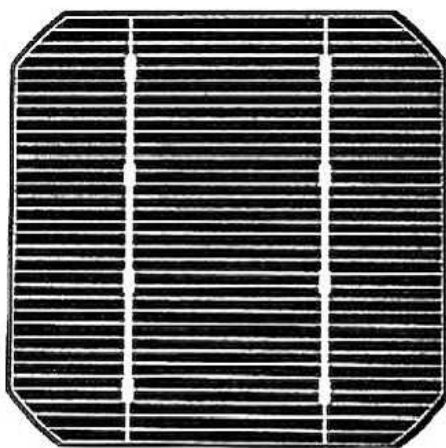
2. EMBASAMENTO TEÓRICO

Esta seção apresenta a contextualização teórica e os trabalhos científicos relacionados a este estudo. Inicialmente são apresentados os conceitos sobre os painéis fotovoltaicos e a geração da energia solar fotovoltaica. Posteriormente, são apresentados os trabalhos relacionados, incluindo o detalhamento do Método Híbrido de Predição – MHP (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b), que foi utilizado como base nas avaliações.

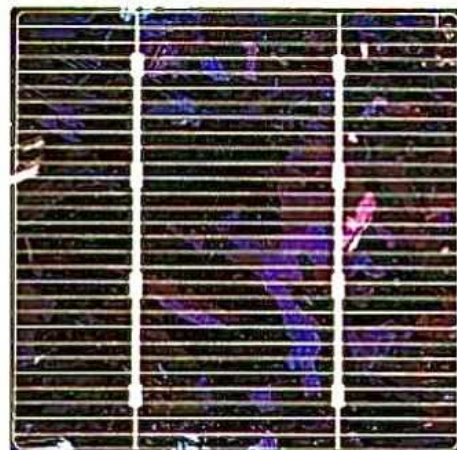
2.1. Energia Solar Fotovoltaica

A energia solar fotovoltaica fundamenta-se no efeito fotovoltaico, que consiste na conversão direta da luz solar em eletricidade. Este processo ocorre quando os fótons da luz excitam elétrons em materiais semicondutores, permitindo que eles se movam livremente, gerando uma corrente elétrica. Este princípio é a base das diversas tecnologias fotovoltaicas existentes, como as células de silício cristalino (Figura 1) e as de filmes finos (BRYKALSKI M, 2023; FERNANDES C, 2017).

Figura 1 – Células de Silício.



a) Célula de Silício Monocristalina



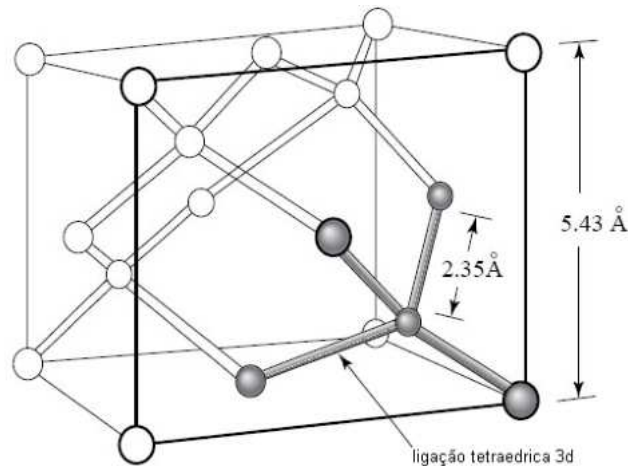
b) Célula de Silício Policristalina

FONTE: (RIBEIRO R, 2016).

As células de silício cristalino, predominantes no mercado, destacam-se pela alta eficiência e durabilidade (CREMASCO et al., 2021). As células monocristalinas podem atingir eficiências de até 24,4%, enquanto as policristalinas (Figura 1) alcançam até 22,3% em laboratório (CREMASCO et al., 2021). Além disso, as células de filmes finos, como as de dissênio de cobre-índio-gálio (CIGS) e telureto de cádmio (CdTe),

são alternativas que oferecem vantagens como flexibilidade e melhor resposta térmica, embora apresentem eficiências ligeiramente inferiores (CREMASCO et al., 2021). Em sua forma pura, o silício apresenta uma estrutura metálica com coloração cinza e brilho característico, composta por cristais (Figura 2) (BRAGA, 2025).

Figura 2 – Estrutura de um cristal de silício

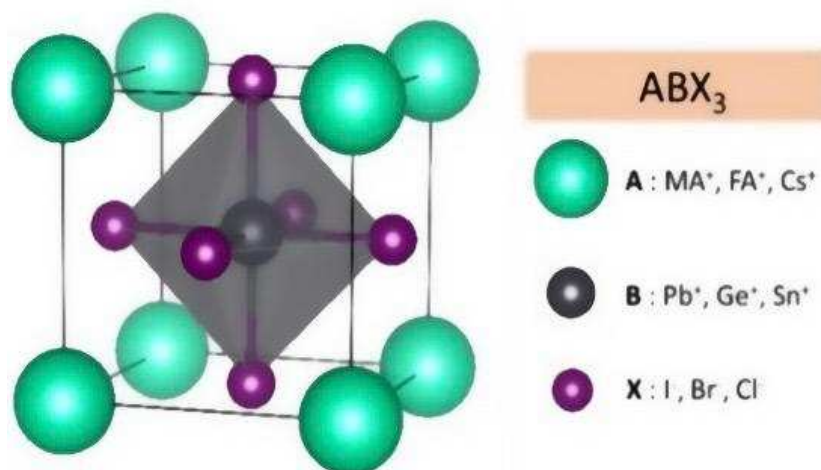


FONTE: (BRAGA, 2025).

O avanço das tecnologias fotovoltaicas engloba módulos de terceira geração, como os sistemas CPV (Concentrator Photovoltaics), que empregam células baseadas em compostos do grupo III-V, como o Arsenieto de Gálio (GaAs), conhecidas por sua alta eficiência, são amplamente empregadas em aplicações aeroespaciais e satélites, devido à sua capacidade de desempenho superior em condições extremas (DIAS P, 2015).

Por outro lado, as células de perovskita emergem como uma alternativa promissora, combinando eficiência crescente com custos de produção significativamente mais baixos. Essas células demonstraram rápida evolução em termos de desempenho, atingindo eficiências superiores a 25% em um curto período de desenvolvimento, graças à sua estrutura cristalina versátil e processos de fabricação simplificados (Pereira et al., 2024). Além disso, a flexibilidade na composição química das perovskitas permite a modulação de propriedades ópticas e eletrônicas, ampliando suas aplicações potenciais em diferentes contextos, desde dispositivos portáteis até integrações em edifícios (CREMASCO et al., 2021; PEREIRA et al., 2024). A estrutura apresentada na Figura 3 ilustra a organização da perovskita, onde os átomos B^{2+} estão coordenados com os átomos X^- , formando um arranjo octaédrico. Esse modelo segue a fórmula geral ABX_3 , característica das perovskitas (PEREIRA et al., 2024).

Figura 3 – Estrutura genérica ABX₃ de perovskitas.



FONTE: (PEREIRA et al., 2024).

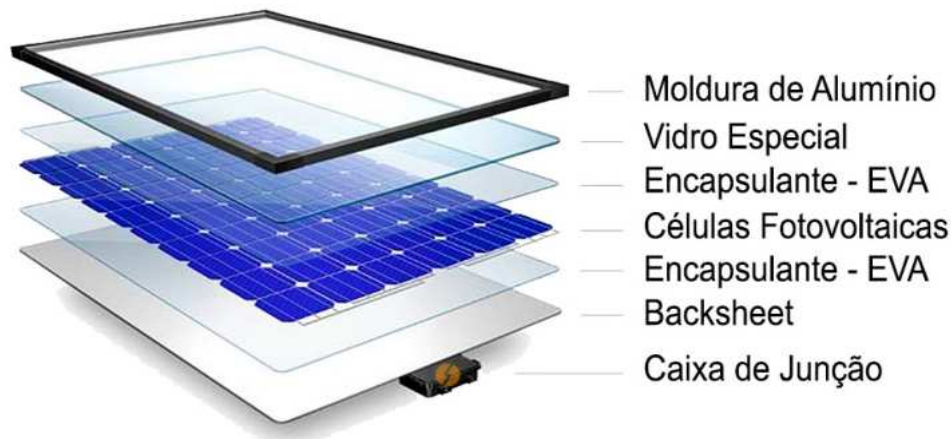
Apesar dos desafios relacionados à estabilidade e à degradação em condições ambientais adversas, as pesquisas continuam avançando para superar essas limitações, tornando as células de perovskita uma opção viável para a expansão do uso de tecnologias fotovoltaicas em escala global (PEREIRA et al., 2024).

Outro aspecto em destaque é a pesquisa contínua por novas configurações geométricas que possam maximizar a absorção de energia e reduzir perdas. No contexto das tecnologias cristalinas, destaca-se a arquitetura PERC (*Passivated Emitter and Rear Cell*), que eleva a eficiência das células de silício, especialmente as monocristalinas, ao aumentar a absorção de luz na região posterior e reduzir perdas por recombinação (CREMASCO et al., 2021). Para células orgânicas (OPV), a estrutura multicamada — composta por eletrodos transparentes (FTO/ITO), camadas orgânicas doadoras e receptoras de elétrons, e um eletrodo metálico — é apresentada como uma estratégia para otimizar a captação de fótons (CREMASCO et al., 2021). Já nas células sensibilizadas por corante, a configuração mesoporosa de TiO₂ é descrita como fundamental para ampliar a área superficial de fixação do corante, melhorando a geração de fotoelétrons (CREMASCO et al., 2021).

Para otimizar a captura de luz e a durabilidade das células fotovoltaicas na geração de energia, elas são organizadas em módulos que incluem camadas de vidro e alumínio, garantindo proteção e estabilidade. O vidro frontal permite a passagem da luz solar, enquanto o silício atua como o principal material responsável pela geração de eletricidade, com contatos metálicos de prata ou cobre conduzindo a eletricidade gerada (DIAS P, 2015; PEREIRA et al., 2024). A Figura 4 esboça a estrutura de uma placa solar

fotovoltaica, composta por vidro temperado especial, células fotovoltaicas encapsuladas em camadas de EVA, um *backsheet* protetor e uma moldura de alumínio que garante resistência. A caixa de junção na parte traseira abriga os diodos de *bypass* e possibilita a conexão elétrica segura do módulo (PORTAL SOLAR, 2025).

Figura 4 – Estrutura de uma placa solar fotovoltaica.



FONTE: (PORTAL SOLAR, 2025).

A reciclagem de placas solares tem ganhado destaque como solução para reduzir impactos ambientais. O vidro fotovoltaico, componente principal, pode ser reaproveitado para produzir espumas vítreas usadas em isolamento térmico e acústico, aumentando o ciclo de vida desses materiais (BRYKALSKI M, 2023; CREMASCO et al., 2021). Este reaproveitamento está alinhado com as demandas por soluções sustentáveis em energia renovável.

A estrutura cristalina e a composição química são determinantes para a eficiência das células fotovoltaicas. Em perovskitas (ABX_3), a substituição de cátions (A^+ , B^{2+}) e haletos (X^-), aliada ao fator de tolerância de Goldschmidt, permite modular o *bandgap* e prever arranjos estruturais otimizados para absorção de luz e transporte de cargas (PEREIRA et al., 2024). A engenharia de composição, como em perovskitas mistas ($MA_{0.8}FA_{0.2}PbI_3$), e nanoestruturas (2D/3D) equilibram eficiência e estabilidade ambiental. No silício cristalino (m-Si/p-Si), a uniformidade estrutural minimiza perdas em fronteiras de grãos, enquanto filmes finos (CIGS/CdTe) exploram alta absorção óptica (CREMASCO et al., 2021).

Por fim, as placas solares fotovoltaicas representam uma tecnologia-chave na transição para fontes de energia limpa. O desenvolvimento de novos materiais, como as células de perovskita, e o foco na reciclagem de componentes reforçam seu

potencial de sustentabilidade e expansão futura (CREMASCO et al., 2021; PEREIRA et al., 2024).

2.2. Predição da Geração de Energia Solar Fotovoltaica (PGESF)

A variabilidade inerente à geração de energia solar fotovoltaica, causada por fatores climáticos incontrolláveis como nuvens, condições atmosféricas e sazonalidade, representa um desafio significativo para a integração desta fonte à rede elétrica. Nesse contexto, a Predição da Geração de Energia Solar Fotovoltaica (PGESF) tornou-se uma área relevante de pesquisa (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2021).

Os modelos de previsão amplamente utilizados incluem métodos estatísticos, como *Autoregressive Integrated Moving Average* (ARIMA) e técnicas baseadas em aprendizado de máquina, como *Support Vector Regression* (SVR), Redes Neurais Artificiais (RNA) e Redes Neurais Recorrentes (RNN). Cada um desses métodos possui características específicas para lidar com dados históricos (URRACA et al., 2016) e irradiância solar (URRACA et al., 2016).

Embora as Redes Neurais Artificiais seja o modelo de inteligência artificial mais aplicado na PGESF (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2021), métodos mais recentes e robustos, baseados em Redes Neurais Recorrentes (RNN) e *Long Short Term Memory* (LSTM) (KONSTANTINOU; PERATIKOU; CHARALAMBIDES, 2021; WENTZ, 2021; YU; CAO; ZHU, 2019), têm mostrado desempenho promissor em diversas aplicações, devido à sua capacidade de capturar dependências de longo prazo em séries temporais (RUSSELL; NORVIG, 2016).

A previsão da geração fotovoltaica pode ser classificada em duas abordagens principais: predição direta e predição indireta. A predição direta se concentra na estimativa da energia gerada pelos sistemas fotovoltaicos, utilizando dados históricos, como variáveis meteorológicas. Na predição indireta são utilizadas variáveis intermediárias como a irradiância solar para simular o desempenho dos sistemas fotovoltaicos (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b).

Os dados de entrada utilizados nos modelos de predição variam de acordo com a abordagem escolhida (KONSTANTINOU; PERATIKOU; CHARALAMBIDES, 2021; WENTZ, 2021). Informações meteorológicas, como irradiância solar, índice de claridade, temperatura ambiente, umidade relativa e velocidade do vento, são comuns, sendo a irradiância solar global horizontal (GHI) a variável mais importante para estimar a geração de energia. Já o uso de imagens do céu permite capturar padrões

atmosféricos instantâneos, como a cobertura e o movimento de nuvens (MACIEL, J. N., 2022). Variáveis temporais, como hora do dia, data e sazonalidade, também são frequentemente incluídas, permitindo que os modelos capturem oscilações sazonais e diárias da irradiância solar (MOSAVI et al., 2019).

Os horizontes de previsão referem-se ao intervalo de tempo para o qual a geração de energia é estimada. No curto prazo, previsões de minutos a horas são utilizadas para ajustes operacionais, como o balanceamento e mitigação da variabilidade da potência da rede em conjunto com sistemas de baterias (LAPPALAINEN; VALKEALAHTI, 2017).

As principais categorias de modelos da PGESF são os modelos estatísticos, físicos e de inteligência artificial, que se destacam-se pela capacidade de aprender padrões não lineares em dados, os quais apresentam resultados mais robustos em cenários de alta variabilidade climática (TINA et al., 2021). Por fim, os métodos híbridos integram modelos estatísticos, físicos e de aprendizado de máquina, combinando suas vantagens (MACIEL, J. N., 2022).

A etapa de avaliação da acurácia dos modelos preditivos é essencial para validar seu desempenho. Existem diversas métricas que podem ser utilizadas na análise de acurácia dos modelos. A métrica de *Root Mean Squared Error* (RMSE) representa a diferença média entre os valores previstos e reais, sendo a mais utilizada na PGESF (MACIEL, 2021). Além da RMSE, este estudo adotou o *Mean Absolute Error* (MAE) e o coeficiente de determinação (R^2) em suas análises (GUEYMARD, 2014).

2.3. Método Híbrido de Predição (MHP)

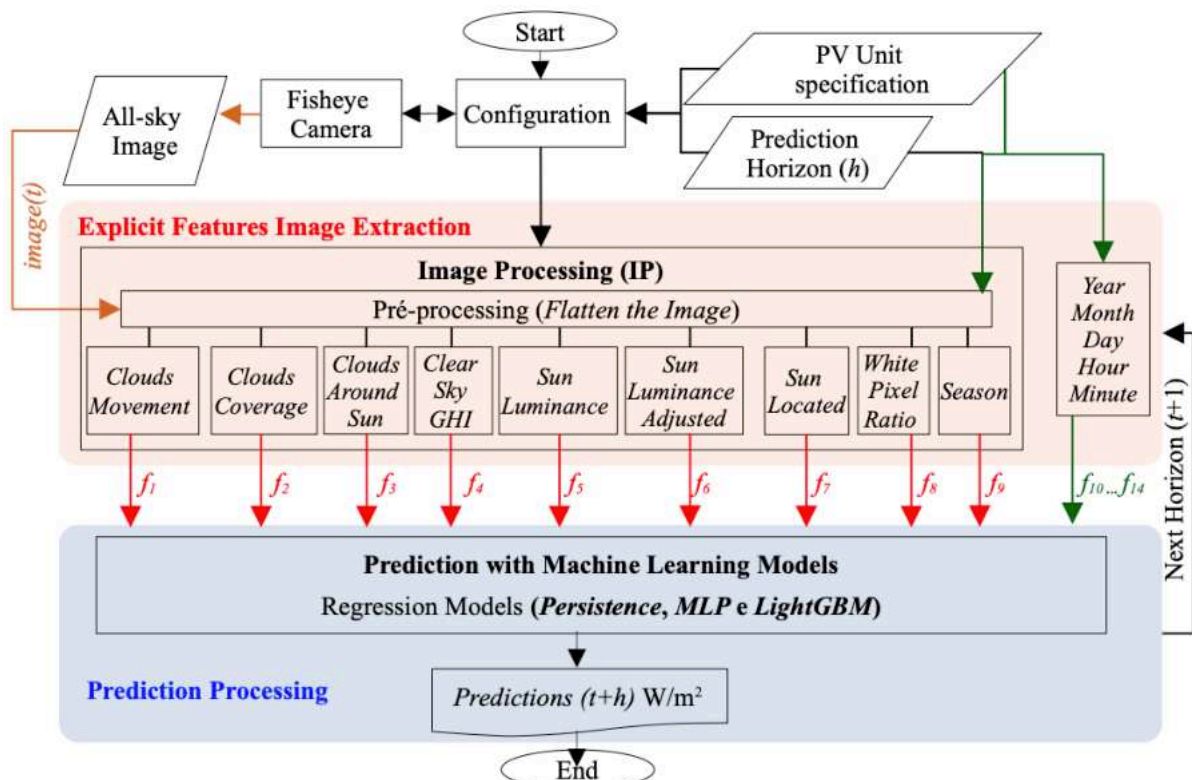
O Método Híbrido de Predição (MHP) foi proposto (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b) e adota uma abordagem original que combina técnicas de aprendizado de máquina com processamento de imagens para prever a irradiância solar global. Essa abordagem utiliza um conjunto de características extraídas de imagens totais do céu (all-sky de 180°) como entrada para modelos de aprendizado de máquina, como a RNA *MultiLayer Perceptron* (MLP) e *Light Gradient Boosting Machine* (LightGBM), realizar as previsões de curto prazo.

Conforme exibido na Figura 5, a etapa inicial do MHP consiste na coleta de dados de imagens do céu capturadas em intervalos regulares e informações de geo-

localização (latitude, longitude e altitude). Essas imagens fornecem informações detalhadas sobre a cobertura de nuvens e outras condições atmosféricas que afetam a irradiância solar (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b).

Posteriormente, o processamento de cada imagem coletada é realizado para extrair características relevantes das condições atmosféricas por meio de um conjunto de técnicas de processamento de imagens. Estas medidas compreendem a: movimento das nuvens (*Clouds Movement*), cobertura das nuvens (*Clouds Coverage*), nuvem em torno do sol (*Clouds Around Sun*), luminância do sol (*Sun Luminance*), luminância do sol ajustada (*Sun Luminance Adjusted*), irradiância de céu limpo (*Clear Sky GHI*), localização do sol (*Sun Located*) e taxa de pixels brancos (*White Pixel Ratio*). Estas medidas são combinadas com um conjunto de informações de geolocalização e temporais, sendo então fornecidas como entrada para um modelo preditivo de Aprendizado de Máquina baseado em RNA no MHP (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b), conforme exibido na Figura 5.

Figura 5 – Método Híbrido de Predição (MHP).



Fonte: (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024).

Em seu estudo de (MACIEL, J. N., 2022) analisou a acurácia das predições do MHP com as métricas de *Root Mean Squared Error* (RMSE), o *Mean Absolute Error* (MAE) e o Coeficiente de Determinação (R^2). Além disso, foi adotada a validação

cruzada de 2 partições (*3-folds*), em que cada partição correspondeu ao período de um ano completo. Os resultados indicaram que o MHP com algoritmo de predição baseado em RNA, possui precisão superior ao modelo de referência (*Persistence*) e acurácia competitiva se comparada aos estudos atuais da literatura que utilizam técnicas mais robustas de aprendizado profundo. A escalabilidade do MHP também é um fator positivo, permitindo sua aplicação em diferentes horizontes e regiões, tornando-o uma ferramenta promissora para a otimização de sistemas fotovoltaicos e gestão redes elétricas (MACIEL, J. N., 2022).

2.4. Modelos Preditivos de Inteligência Artificial

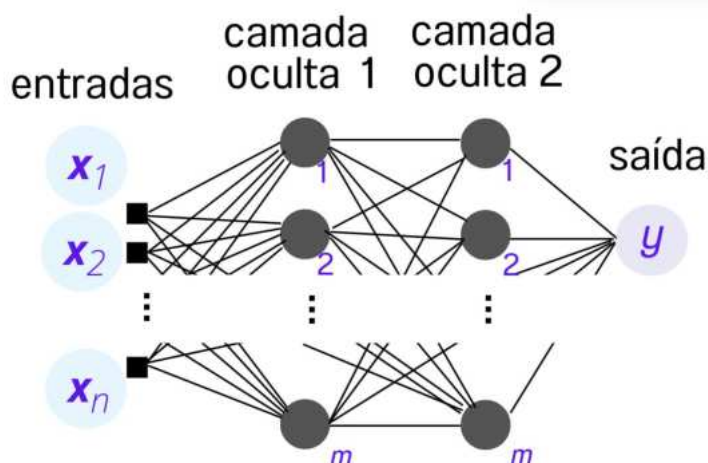
2.4.1. Redes Neurais Artificiais

A Rede Neural Artificial (RNA) *MultiLayer-Perceptron* é uma arquitetura de aprendizado supervisionado amplamente utilizada em problemas de regressão e classificação. Essa arquitetura é composta por uma rede de camadas totalmente conectadas, onde cada nó (neurônio) em uma camada está conectado a todos os nós da camada. A RNA é caracterizada por sua capacidade de modelar relações não lineares, o que a torna adequada para tarefas como previsão de irradiância solar (RUSSELL; NORVIG, 2016).

A estrutura da RNA é dividida em três componentes principais. A camada de entrada recebe os dados brutos. As camadas ocultas são compostas por neurônios que realizam transformações nos dados por meio de funções de ativação, como ReLU (*Rectified Linear Unit*) ou *Sigmoid*, criando representações mais abstratas e permitindo que a rede capture relações não lineares. Por fim, a camada de saída fornece o resultado previsto, que pode ser um valor contínuo em problemas de regressão ou uma classe em tarefas de classificação (RUSSELL; NORVIG, 2016).

Cada conexão entre os neurônios possui pesos ajustáveis, que são otimizados durante o processo de treinamento para minimizar uma função de erro (RUSSELL; NORVIG, 2016). Esse processo de treinamento iterativo é repetido em várias épocas até que a convergência seja alcançada e o modelo apresente um bom desempenho nas previsões. A Figura 6 exibe o diagrama de uma RNA, em que os neurônios estão organizados em camadas interconectadas, permitindo o processamento de informações por meio da propagação de sinais e ajuste de pesos sinápticos durante o treinamento (SINKOVICS, 2016).

Figura 6 – Estrutura de uma Rede Neural Artificial.



FONTE: Adaptado de (RUSSELL; NORVIG, 2016).

Apesar de sua flexibilidade e robustez, a RNA apresenta algumas limitações. Uma delas é a falta de memória temporal, o que impede sua aplicação em problemas que dependem de relações ao longo do tempo. Além disso, a RNA é sensível ao sobreajuste (*overfitting*), especialmente quando aplicada a conjuntos de dados pequenos ou com muitos parâmetros. Para superar esse problema, são utilizadas técnicas de regularização como *dropout*, que realiza o desligamento aleatório de neurônios e ajuda a melhorar a prevenir o *overfitting* (RUSSELL; NORVIG, 2016).

2.4.2. Redes Neurais Recorrentes

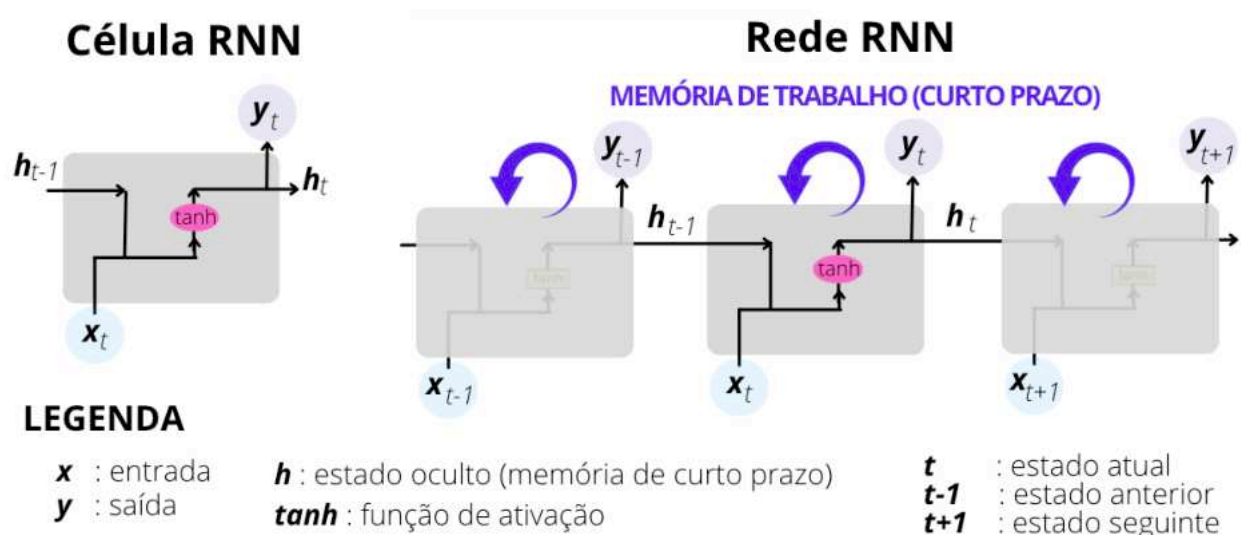
Muitos problemas empregam dados relacionados a um período temporal. Esses problemas podem ter seus dados classificados como Séries Temporais, que são sequências de dados indexados ao longo do tempo, onde a ordem das observações é essencial para a análise e predição. Diferentemente de problemas não temporais, onde cada amostra pode ser tratada independentemente, em séries temporais há dependência temporal, exigindo modelagem que capture padrões sazonais, tendências e correlações temporais. Métodos clássicos incluem ARIMA, modelos de suavização exponencial e regressão temporal, enquanto abordagens modernas utilizam redes neurais recorrentes (RNNs), LSTMs e *Transformers* (KIM et al., 2024).

As Redes Neurais Recorrentes (RNN) representam uma classe de arquiteturas de aprendizado profundo projetadas para processar dados sequenciais, como séries temporais, texto e reconhecimento de fala. Diferentemente das RNAs, que tratam cada entrada de forma independente, as RNNs possuem conexões recorrentes que permitem a retenção de informações de estados anteriores (curto prazo), possibilitando a

modelagem de dependências temporais. Essa característica é especialmente vantajosa para tarefas que exigem contexto ao longo do tempo, como tradução automática e geração de texto (YU, Yong et al., 2019).

A Figura 7 ilustra uma célula e o fluxo de informações na arquitetura de uma RNN tradicional. As RNNs são modelos projetados para processar dados sequenciais, mantendo um estado oculto h_t que captura dependências temporais ao longo do tempo. A cada instante t , a célula recorrente recebe uma entrada x_t e atualiza seu estado oculto h_t com base no estado anterior h_{t-1} e na transformação não linear da função de ativação, frequentemente a tangente hiperbólica (\tanh). Esse estado oculto h_t é utilizado para calcular a saída y_t , que pode representar previsões e/ou classificações.

Figura 7 – Célula e estrutura de uma *Recurrent Neural Network* (RNN).



FONTE: o autor, (2025).

No diagrama sequencial da Figura 7, observa-se que a rede mantém dependências temporais ao longo do tempo, permitindo a modelagem de dados sequenciais, como séries temporais e processamento de linguagem natural. Embora a aplicação de RNNs tem crescido em diversas áreas, impulsionada pelo avanço da computação e pelo acesso a grandes volumes de dados, o treinamento de RNNs pode ser computacionalmente custoso devido à necessidade de processar sequências inteiras de dados antes de gerar previsões (YU, Yong et al., 2019). Além disso, as RNNs tradicionais enfrentam desafios como o desvanecimento do gradiente (*vanishing gradient*), dificultando o aprendizado de dependências de longo prazo (YU, Yong et al., 2019).

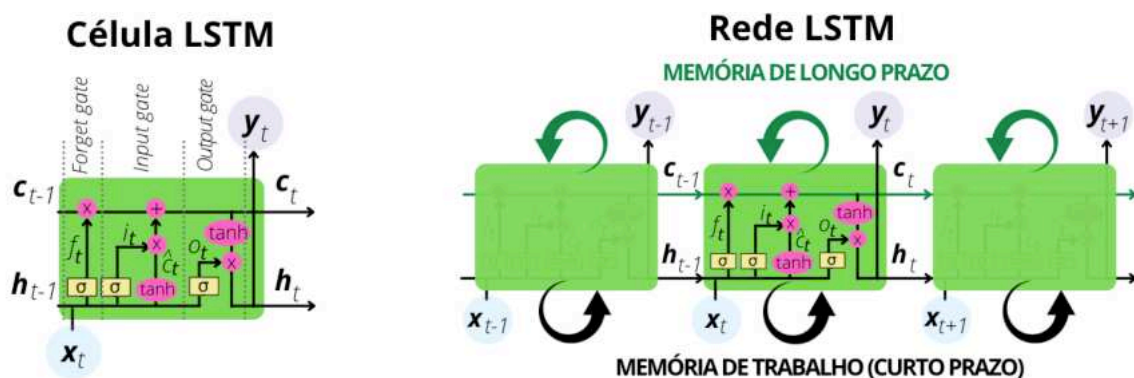
Em 1997, (HOCHREITER; SCHMIDHUBER, 1997), propuseram uma arquitetura de rede neural, denominada *Long Short-Term Memory* (LSTM), para superar a

limitação de captura de dependências de longo prazo em dados sequenciais (*vanishing gradient*) das RNN tradicionais, introduzindo um mecanismo especializado para armazenar informações relevantes ao longo do tempo.

Conforme a Figura 8, a célula LSTM mantém dois estados distintos: o estado da célula c_t , que armazena informações de longo prazo, e o estado oculto h_t , responsável pela memória de curto prazo e pela produção da saída. O fluxo de informações dentro da célula LSTM é controlado por três portas principais: a *forget gate* f_t , que decide quais informações do estado da célula anterior c_{t-1} serão descartadas; a *input gate* i_t , que regula a adição de novas informações \hat{c}_t , ao estado da célula; e a *output gate* o_t , que determina quais informações do estado da célula c_t serão utilizadas para calcular o estado oculto h_t e, conseqüentemente, a saída y_t . As funções de ativação sigmoide (σ) e tangente hiperbólica *tanh* são empregadas para controlar o fluxo e a transformação das informações (HOCHREITER; SCHMIDHUBER, 1997).

A estrutura recorrente da LSTM permite a propagação eficiente de informações ao longo do tempo, mitigando o problema do desvanecimento do gradiente e facilitando o aprendizado de dependências de longo prazo (HOCHREITER; SCHMIDHUBER, 1997). A capacidade de capturar dependências de curto e longo prazo em séries temporais permitem sua aplicação na previsão da irradiância solar utilizando variáveis meteorológicas (WENTZ, 2021). Todavia, as LSTMs apresentam maior complexidade computacional devido à presença de múltiplas portas de controle e operações internas, resultando em tempos de treinamento mais longos e maior poder computacional.

Figura 8 – Célula e estrutura de uma *Long Short-Term Network* (LSTM).



LEGENDA

x : entrada	y : saída	t : estado atual
h : estado oculto (memória de curto prazo)	σ, \tanh : funções de ativação	$t-1$: estado anterior
c : estado da célula (memória de longo prazo)		$t+1$: estado seguinte

Fonte: Adaptado de (WENTZ, 2021).

3. MATERIAL E MÉTODO

Este capítulo apresenta o material e método empregados no desenvolvimento deste trabalho, abordando a descrição do banco de dados utilizado e provido pelo Método Híbrido de Predição (MHP), descrito na Seção 2.3, bem como o delineamento experimental projetado para realizar as avaliações de acurácia dos modelos de RNA (MHP), RNN e LSTM avaliados no MHP para a predição da irradiância solar global. Por fim, são detalhadas as ferramentas e tecnologias adotadas.

3.1. Banco de Dados

O bando de dados (*dataset*) utilizado neste estudo está disponível em (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024a) e contempla um conjunto de medidas (características) extraídas com o processamento de imagens *all-sky*, conforme detalhado na Tabela 1, onde são disponibilizados aproximadamente 775 mil registros referentes ao ano de 2014, 2015 e 2016, com amostragens para 1, 5, 10, 15, 30 e 60 minutos.

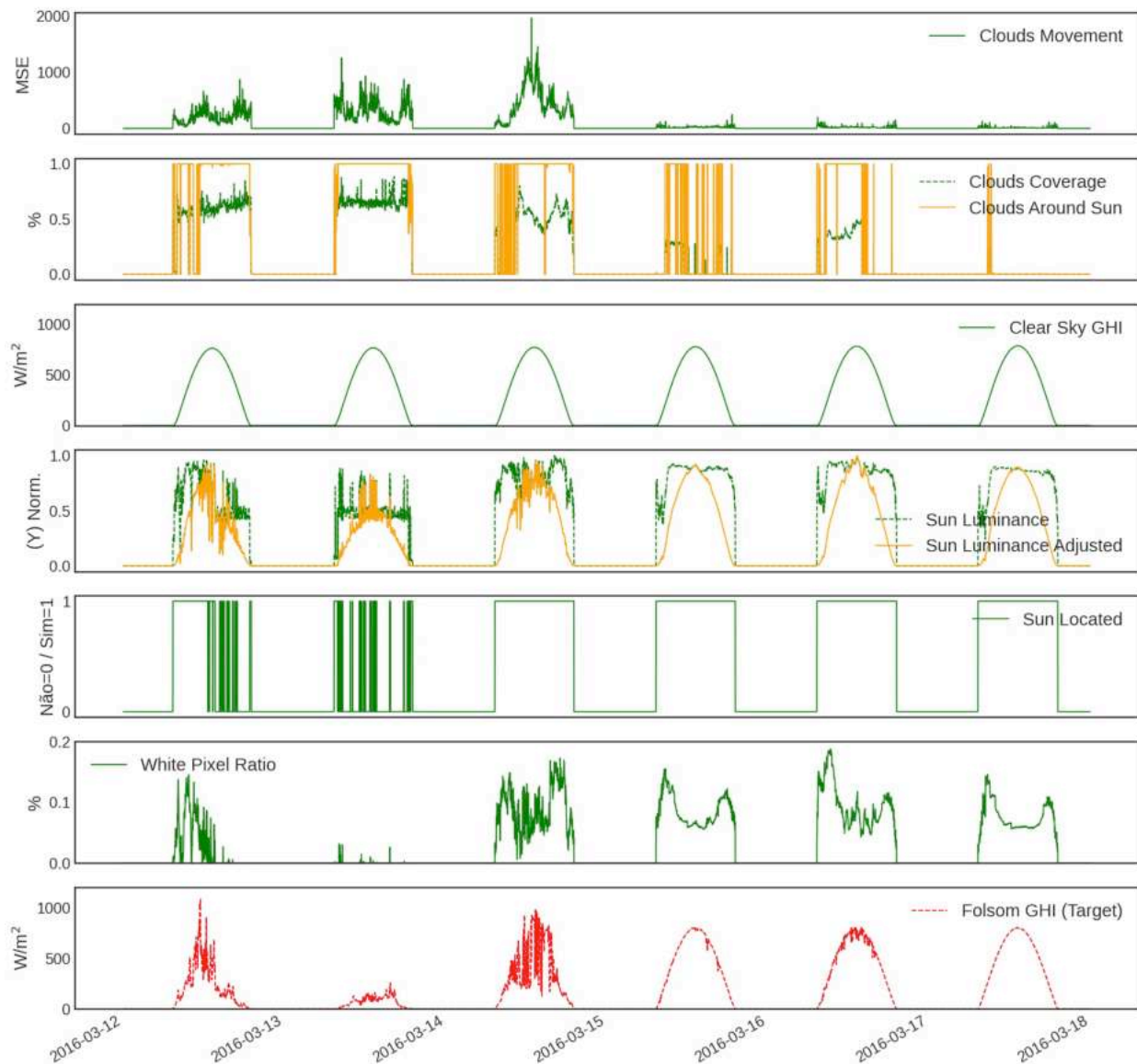
Tabela 1 – Detalhamento do *dataset* utilizado neste estudo.

Variáveis	Tipo	Horizonte (minutos)	Amostras			Total
			2014	2015	2016	
Year, Month, Day, Hour, Minute e Season	Temporal	1	250.579	263.023	261.961	775.563
		5	250.574	263.018	261.956	775.548
Clouds Movement, Clouds Coverage, Clouds Around Sun, Clear Sky GHI, Sun Luminance, Sun Luminance Adjusted, Sun Located e White Pixel Ratio	Medidas das imagens <i>all-sky</i>	10	250.569	263.013	261.951	775.533
		15	250.564	263.008	261.946	775.518
Global Horizontal Irradiance (GHI)	Irradiância Solar	30	250.549	262.993	261.931	775.473
		60	250.519	262.963	261.901	775.383

Fonte: Adaptado de (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024a).

Já a Figura 9 exibe um exemplo, para seis dias completos, demonstrando o comportamento do conjunto de medidas, extraídas das imagens *all-sky* e disponíveis no *dataset*, onde a variável *Folsom GHI* é aquela considerada alvo (a ser predita), e as demais correspondem a medidas extraídas com processamento de imagens *all-sky*, proposto no MHP (Figura 5).

Figura 9 – Exemplo do conjunto de medidas e irradiância solar no *dataset* do MHP.



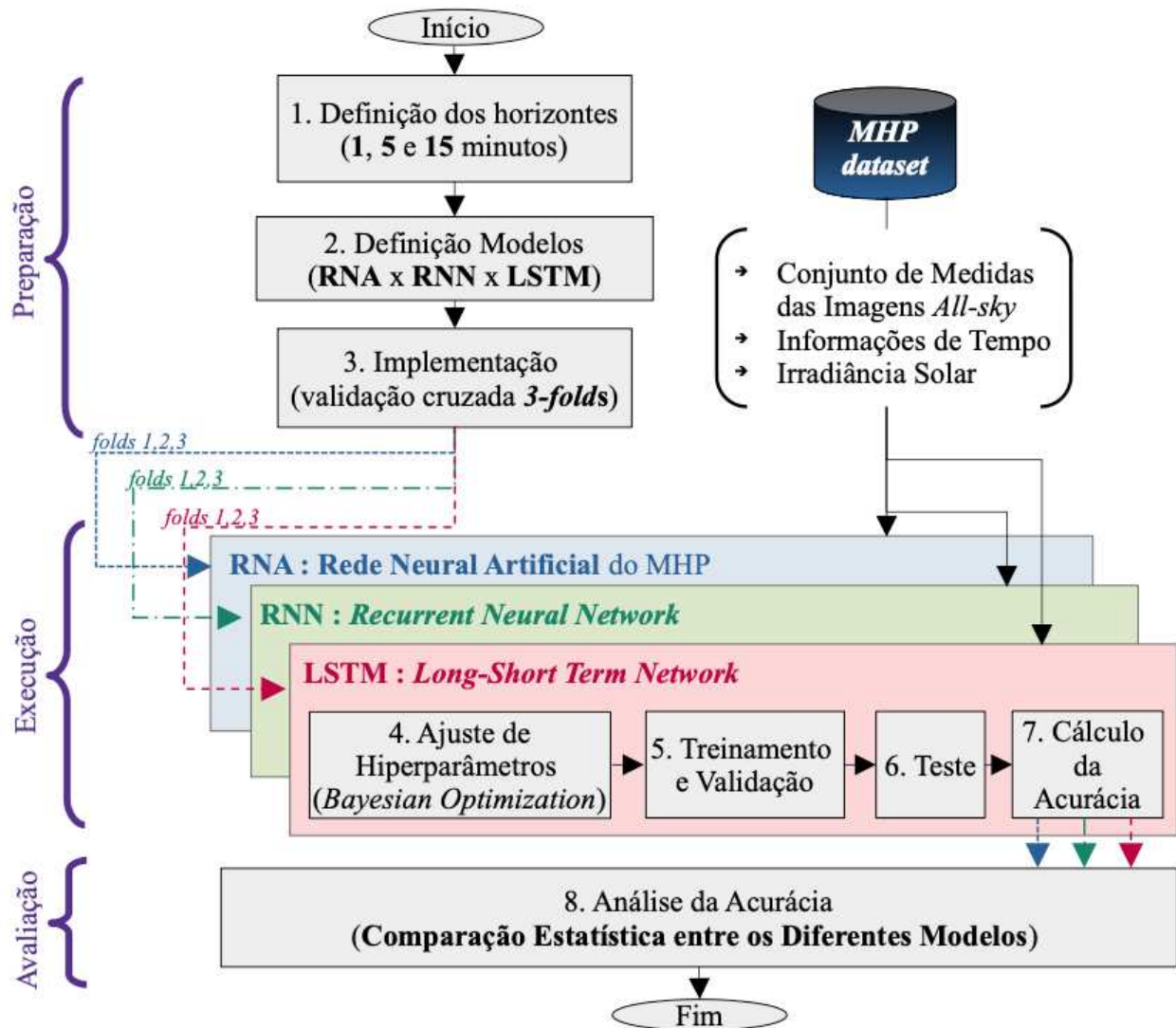
Fonte: (MACIEL, J. N., 2022).

3.2. Delineamento Experimental

O delineamento experimental deste estudo foi projetado para avaliar a acurácia de diferentes modelos de aprendizado de máquina e aprendizado profundo, especificamente com Redes Neurais Recorrentes (RNNs) tradicionais e com Memória de Curto e Longo Prazo (LSTM), na previsão da irradiância solar de curto do MHP, que adota Rede Neural Artificial (RNA) do tipo *MultiLayer Perceptron*. O objetivo principal é comparar a acurácia de previsão entre estes modelos, em diferentes horizontes de previsão, tendo como entrada (*input*) o conjunto de medidas extraídas das imagens *all-sky* (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b).

A Figura 10 ilustra o fluxo geral do método experimental concebido e utilizado para permitir uma comparação, justa e homogênea, entre os algoritmos de previsão avaliados (RNA, RNN e LSTM). O método possui a etapa inicial de preparação e configuração, seguido da execução e avaliação dos resultados.

Figura 10 – Delineamento experimental.



Fonte: o autor, (2025).

Portanto, o presente estudo emprega um método sistemático e estruturado contemplando as seguintes etapas:

1. **Definição dos Horizontes de Previsão:** estabeleceu-se três horizontes de previsão correspondentes a 1, 5 e 15 minutos, permitindo avaliar a capacidade preditiva dos modelos em diferentes escalas temporais;
2. **Definição dos Modelos:** considerou-se três arquiteturas de redes

neurais para análise neste estudo:

- a. **RNA**: modelo de RNA *MultiLayer Perceptron* adotado no MHP;
- b. **RNN**: modelo de rede neural recorrente tradicional com memória de curto prazo;
- c. **LSTM**: modelo rede neural recorrente com memória de curto e longo prazos.

Cabe ressaltar que as arquiteturas RNN e LSTM foram escolhidas por suas capacidades de capturar padrões espaciais e temporais nos dados, diferente da RNA *MultiLayer Perceptron* adotada no MHP;

3. **Implementação com Validação Cruzada (3-Folds)**: desenvolveu-se os modelos e a preparação dos dados para uso da validação cruzada com 3 partições (*3-folds*), buscando a minimização de vieses nos dados de treinamento e teste. Ou seja, o *dataset* foi particionado em três *folds* onde cada um corresponde a um ano completo. Esta estratégia foi adotada para permitir que o modelo captasse a sazonalidade anual da irradiância solar de forma mais eficaz. Além disso, uma janela (*time_step*) com valor 2 (dois) valores foram usados as redes recorrentes (RNN e LSTM).

Tabela 2 – Divisão do dataset utilizando validação cruzada de 3 partições (*folds*).

Partição	2014	2015	2016
<i>fold-1</i>	Teste	Treino	Treino
<i>fold-2</i>	Treino	Teste	Treino
<i>fold-3</i>	Treino	Treino	Teste

Os dados de validação correspondem a 20% (4,8 meses) da partição de treino.

Fonte: Adaptado de (MACIEL, J. N.; LEDESMA; ANDO JUNIOR, 2024b).

4. **Ajuste de Hiperparâmetros**: submeteu-se cada modelo ao processo de ajuste fino (*fine tuning*) para identificação dos melhores hiperparâmetros para obtenção da melhor acurácia preditiva. Foi aplicada a abordagem *Bayesian Optimization* que otimização a acurácia por meio da inferência probabilística bayesiana (SHAHRIARI et al., 2016). Visando promover a reprodutibilidade deste estudo, a Tabela 2 exhibe hiperparâmetros analisados e aqueles selecionados para cada modelo preditivo. Cabe ressaltar que os hiperparâmetros da RNA são oriundos do modelo preditivo adotado no MHP (MACIEL, J. N., 2022).

Tabela 3 – Melhores Hiperparâmetros Obtidos com *Bayesian Search Optimization*.

Hiperparâmetros Avaliados e Selecionados (*) para cada Modelo Preditivo	
RNA (MHP)	<p>Número de camadas ocultas = [1, 2, 3*] <i>Batch Size</i> = [32, 64, 128*] <i>Learn Rate</i> = [0,01; 0,001*; 0,0001] <i>Epochs</i> = 200* Função de ativação = [tanh, ReLU*] Neurônios na camada de entrada – <i>input</i> (<i>x</i>) = 12 Neurônios na camada de entrada – <i>output</i> (<i>y</i>) = 1 Estrutura da rede: {<i>x</i> + camadas_ocultas + <i>y</i>} = {12 + 30 + 210 + 150 + 1} Total de neurônios da rede = 402</p>
RNN	<p>Número de camadas ocultas = [1, 2, 3*] <i>Batch Size</i> [16 ... 128] = 101* <i>Learn Rate</i> = [0,01; 0,001*; 0,0001] <i>Epochs</i> = 200* <i>Dropout</i> = [0%, 10%, 15%, 20%*, 30%, 40%, 50%] Função de ativação = [tanh, ReLU*] Neurônios na camada de entrada – <i>input</i> (<i>x</i>) = 12 Neurônios na camada oculta 1 [16 ... 256] = 60* Neurônios na camada oculta 2 [32 ... 256] = 138* Neurônios na camada oculta 3 [32 ... 256] = 154* Neurônios na camada de entrada – <i>output</i> (<i>y</i>) = 1 Estrutura da rede: {<i>x</i> + camadas_ocultas + <i>y</i>} = {12 + 60 + dropout + 138 + dropout + 154 + 1} Total de neurônios da rede = 365</p>
LSTM	<p>Número de camadas ocultas = [1, 2, 3*] <i>Batch Size</i> [16 ... 128] = 64* <i>Learn Rate</i> = [0,01; 0,001; 0,0001*] <i>Epochs</i> = 200* <i>Dropout</i> = [0%*, 10%, 15%, 20%, 30%, 40%, 50%] Função de ativação = [tanh, ReLU*] Neurônios na camada de entrada – <i>input</i> (<i>x</i>) = 12 Neurônios na camada oculta 1 [16 ... 256] = 128* Neurônios na camada oculta 2 [32 ... 256] = 32* Neurônios na camada oculta 3 [32 ... 256] = 96* Neurônios na camada de entrada – <i>output</i> (<i>y</i>) = 1 Estrutura da rede: {<i>x</i> + camadas_ocultas + <i>y</i>} = {12 + 128 + 32 + 96 + 1} Total de neurônios da rede = 269</p>

Fonte: o autor, (2025).

- 5. Treinamento e Validação:** treinou-se os modelos utilizando validação cruzada com 3 partições (*3-folds*) e validação correspondente a 20% dos dados de treinamento;
- 6. Teste:** após o treinamento, avaliou-se os modelos com um conjunto de teste independente e diferente para cada *fold*, permitindo verificar sua capacidade de generalização. A acurácia final de teste de cada modelo preditivo é computada com a média aritmética dos três *folds*;
- 7. Cálculo da Acurácia:** calculou-se as métricas estatísticas de acurácia (MAE, MSE, RMSE e R²) para mensurar qualidade das

predições dos modelos;

8. **Análise Comparativa da Acurácia:** realizou-se uma análise descritiva e comparação estatística da acurácia média (dos *3-folds*) entre cada modelo preditivo avaliado (RNA, RNN e LSTM).

3.3. Ferramentas e Tecnologias

As execuções experimentais realizadas neste estudo exigiram o uso de diversas ferramentas e tecnologias gratuitas. Dentre as principais tecnologias utilizadas citam-se a linguagem de programação Python (MINING, 2019). As bibliotecas *scikit-learn* e *TensorFlow/Keras* (DÜRR; SICK; MURINA, 2020), que disponibilizam pacotes que facilitam a programação dos modelos de RNA, CNN e LSTM, e o framework de otimização de hiperparâmetros Optuna (AKIBA et al., 2019).

Em relação ao hardware utilizado no desenvolvimento dos modelos de empregou-se o ambiente colaborativo da Google, denominado *Google Colaboratory* (BISONG, 2019). Este ambiente disponibiliza gratuitamente para uso as máquinas virtuais com placas gráficas Nvidia Tesla T4 com 16 GB de memória. Os resultados e as análises estatísticas foram executados com auxílio do software estatístico JASP (JASP TEAM, 2021), assim como o editor de Planilhas do Google.

4. RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados e discussões referentes à acurácia dos modelos RNA, RNN e LSTM empregados na predição da irradiância solar do Método Híbrido de Predição (MHP). Após a execução experimental as acurácias de MSE, RMSE, MAE e R^2 foram coletadas e sintetizadas para os horizontes de predição de 1, 5 e 15 minutos. Posteriormente, conduziu-se a análise descritiva e estatística para interpretação dos resultados e validação da hipótese deste estudo (página 15), respectivamente.

4.1. Acurácia de Predição

Os três modelos preditivos RNA (Rede Neural Artificial), RNN (*Recurrent Neural Network*) e LSTM (*Long Short-Term Memory*) foram avaliados conforme o delineamento experimental exibido na Figura 10 (seção 3.2). A Tabela 4 detalha o desempenho dos modelos nas três partições (*folds*), considerando os horizontes de predição de 1, 5 e 15 minutos.

Tabela 4 – Acurácias dos Modelos Preditivos

	Modelo	Horizonte	MSE	RMSE	MAE	R^2
Fold-1	RNA	1	4968,55	70,49	37,31	0,944
		5	5467,11	73,94	39,74	0,938
		15	6033,54	77,68	41,57	0,932
	RNN	1	5079,14	71,27	38,93	0,942
		5	5452,44	73,84	39,80	0,938
		15	6555,32	80,96	47,99	0,926
	LSTM	1	5131,47	71,63	37,59	0,942
		5	5533,86	74,39	40,32	0,937
		15	6047,06	77,76	42,01	0,931
Fold-2	RNA	1	4858,86	69,71	39,44	0,943
		5	5143,49	71,72	39,26	0,940
		15	5842,28	76,43	42,20	0,932
	RNN	1	5411,15	73,56	44,85	0,937
		5	5939,19	77,07	49,05	0,931
		15	6200,65	78,74	45,48	0,928
	LSTM	1	4862,57	69,73	37,74	0,943
		5	5140,57	71,70	39,34	0,940
		15	5920,23	76,94	41,92	0,931
Fold-3	RNA	1	4566,97	67,58	36,93	0,948
		5	4699,10	68,55	36,75	0,947
		15	5729,01	75,69	41,10	0,935
	RNN	1	4816,24	69,40	39,05	0,945
		5	5609,04	74,89	45,57	0,936
		15	6427,71	80,17	48,79	0,927

LSTM	1	4716,31	68,68	36,55	0,947
	5	5032,05	70,94	37,12	0,943
	15	6515,21	80,72	47,60	0,926

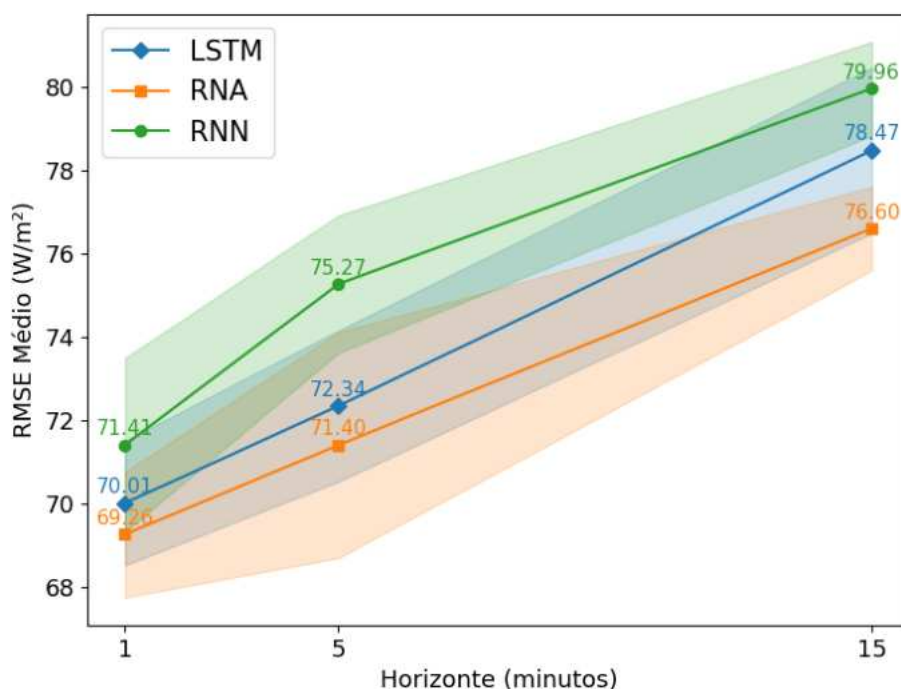
Fonte: o autor, (2025).

A análise dos resultados em cada *fold*, por meio da métrica de RMSE em termos absolutos, evidencia a existência de variações de acurácia entre os *fold*s avaliados. Estes resultados evidenciam a existência de variabilidade anual climática e sazonal de cada *fold*, o qual representa um ano completo.

4.2. Análises Descritiva e Exploratória

A partir dos resultados da Tabela 4, apresenta-se a análise descritiva e exploratória estruturada em três eixos principais: as acurácias relativas aos horizontes de predição, as arquiteturas dos modelos (RNA, RNN e LSTM) avaliados e as implicações técnicas para sistemas fotovoltaicos. Na Figura 11 é demonstrado o RMSE médio e a respectiva variabilidade (desvio padrão) entre os modelos ANN, RNN e LSTM para os horizontes de 1, 5 e 15 minutos. O RMSE tende a aumentar conforme o horizonte se amplia, conforme já observado na literatura (MACIEL, 2021). De modo geral, a RNN demonstra ter RMSE médio geral mais elevado, enquanto ANN e LSTM obtiveram desempenhos próximos, com melhor a acurácia média para a ANN nos três horizontes.

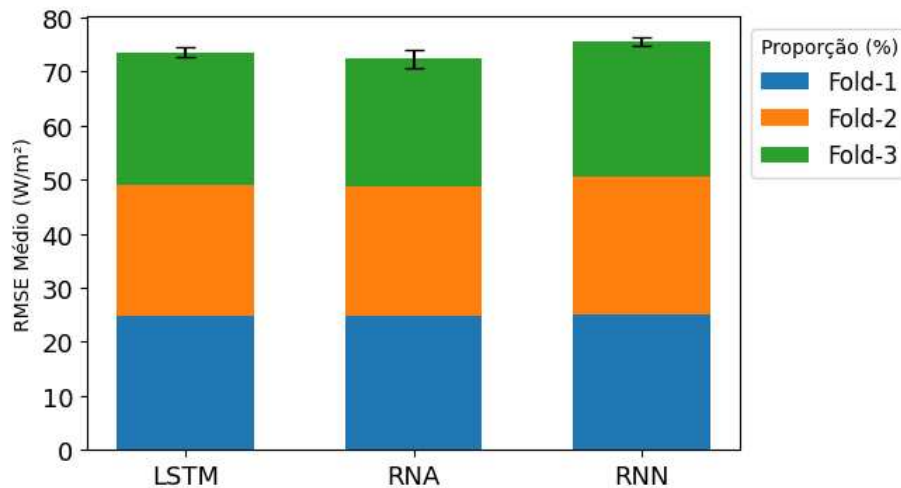
Figura 11 – Acurácia média dos modelos por horizonte de predição.



Fonte: o autor, (2025).

Considerando o desvio padrão exibidos na Figura 11, observa-se que as acurácias dos modelos são praticamente sobrepostas, com exceção para o horizonte de 15 minutos, existindo ainda uma maior variabilidade para o horizonte de 5 minutos. Na Figura 12 é exibido o desempenho médio geral de cada modelo em todos os horizontes. A menor acurácia média geral de RMSE pertence ao modelo RNA (70,42), enquanto o RNN exibiu o maior RMSE médio (75,5).

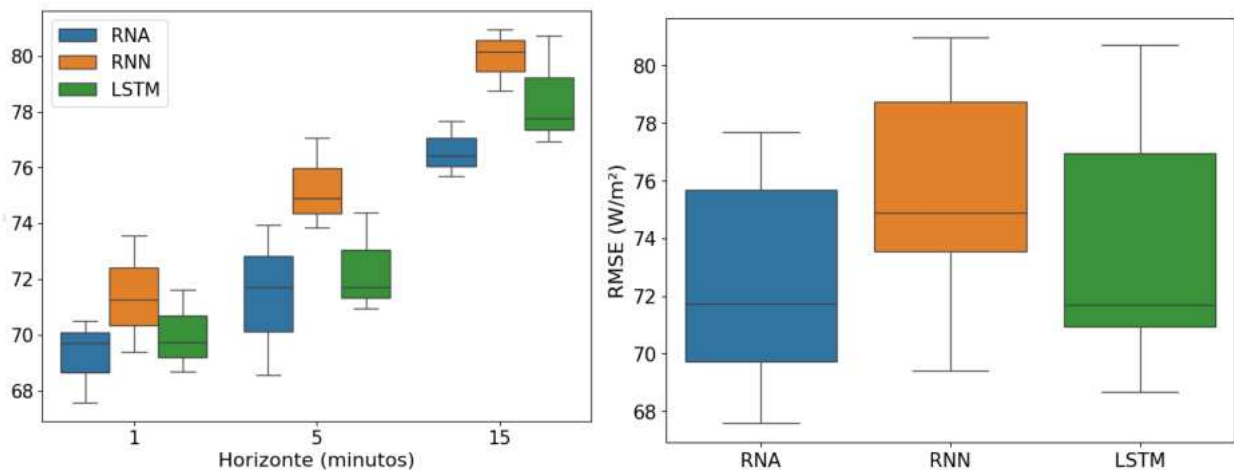
Figura 12 – Acurácia média geral dos modelos e composição proporcional dos folds.



Fonte: o autor, (2025).

A superioridade da RNA pode ser atribuída à sua capacidade de mapear relações imediatas entre variáveis de entrada sem a necessidade de mecanismos temporais complexos, seja de curto ou longo prazos. O desempenho inferior da RNN em relação a LSTM, sendo ambas redes recorrentes, demonstra que a habilidade de memorizar sequências de longo prazo não produz melhoria de desempenho no MHP, embora a LSTM tenha apresentado acurácias mais próximas ao modelo de RNA.

Figura 13 – Variabilidade da acurácia dos modelos.



Fonte: o autor, (2025).

Na Figura 13 apresenta-se a variabilidade da acurácia geral e em cada horizonte dos modelos analisados. Percebe-se um comportamento não regular nos três horizontes, assim como uma menor variabilidade da RNA, apesar do grande desvio padrão observado no horizonte de 5 minutos. A proximidade numérica entre LSTM e RNA sugere que, em cenários sem dependências temporais críticas, a simplicidade da RNA pode ser vantajosa, especialmente em sistemas com restrições computacionais. Por outro lado, nestes cenários avaliados, a maior complexidade computacional da RNN justifica sua não aplicação no MHP.

A similaridade visual entre os modelos em horizontes de 1 e 5 minutos contrasta com a divergência progressiva em horizontes longos, reforçando a necessidade de uma análise estatística formal para validar a significância das diferenças observadas. Desse modo, a próxima seção apresenta uma análise estatística para validar a comparação de desempenho dos modelos preditivos.

4.3. Validação Estatística

Neste estudo foi adotado na análise estatística de significância o valor de 5% ($\alpha = 0,05$) para validação das diferenças nas acurácias observadas. A validação estatística considerou os três modelos preditivos (RNA, RNN e LSTM) avaliados em três horizontes de predição (1, 5 e 15 minutos), utilizando a técnica de validação cruzada com três partições (*3-folds*).

Fundamentando-se nestas características, o Teste não paramétrico de Friedman (MORETTIN; BUSSAB, 2017) foi aplicado como método estatístico, pois, os dados foram obtidos a partir do processo de validação cruzada com 3-folds, sendo apropriado para situações em que se deseja comparar múltiplos métodos sobre amostras pareadas, como é o caso deste estudo, em que os valores de RMSE de cada modelo foram avaliados sob as mesmas partições dos dados em cada *fold*. Além disso, o Teste de Friedman não exige que os dados sigam a distribuição normal, tornando-o mais robusto em contextos de pequenos tamanhos amostrais, como o presente caso, em que há apenas três observações (*folds*) por modelo em cada horizonte de previsão (MORETTIN; BUSSAB, 2017).

Definiu-se como Hipótese Nula (H_0) que a acurácia de predição (RMSE) entre os pares de modelos é estatisticamente igual ($\alpha = 0,05$). Na Tabela 5 é exibido o resultado do teste de Friedman para acurácia de RMSE. Em todas as comparações, a

estatística de Friedman revelou que não houveram diferenças estatisticamente significativas entre os modelos em nenhum dos horizontes avaliados ($p > 0,05$).

Tabela 5 – Comparações estatísticas de RMSE.

Horizonte (minutos)	Estatística de Friedman (χ^2)	p-valor
1	4,67	0,097
5	0,67	0,717
15	4,67	0,097

Fonte: o autor, (2025).

Nas comparações do horizonte de 5 minutos, os modelos apresentando as menores variações entre si ($\chi^2=0,67$ e p-valor = 0,717). Desse modo, não há evidências estatísticas suficientes para afirmar que um dos modelos apresenta desempenho significativamente superior aos demais em termos de RMSE nos horizontes analisados.

Portanto, com base nestes resultados, apesar das diferenças das acurácias dos modelos em termos absolutos, pode-se afirmar que não há evidências estatísticas suficientes para rejeitar a hipótese nula (H_0), ou seja, isto indica que as variações observadas entre os modelos podem ser atribuídas ao acaso e não refletem diferenças reais de desempenho. Sendo assim, a hipótese investigada neste estudo, considerando o método, os horizontes e modelos preditivos analisados, pode ser rejeitada. Em outras palavras, os mecanismos de captura de dependências de curto ou longo prazos em dados sequenciais não contribuíram para o aumento da acurácia dos modelos preditivos analisados no MHP.

5. CONCLUSÃO

Os resultados deste estudo mostram que, embora redes neurais recorrentes (RNN) e de memória de curto e longo prazo (LSTM) sejam reconhecidas na literatura por capturar dependências temporais, não apresentaram melhorias estatisticamente significativas em relação à rede neural tradicional (RNA) *MultiLayer Perceptron* adotada no Método Híbrido de Predição (MHP). A comparação englobou horizontes de 1, 5 e 15 minutos, contemplando otimização de hiperparâmetros e validação cruzada com três partições anuais, buscando conferir rigor e minimizar vieses nos dados.

De modo geral, a RNA alcançou ligeiramente melhores valores médios de RMSE e MAE, mas sem diferença estatisticamente relevante ($p > 0,05$) em relação à RNN e LSTM; portanto, não se evidenciou vantagem no uso de mecanismos de memória sequencial para o conjunto de variáveis de entrada empregado. Isto permite concluir que a eficácia do MHP pode ser atribuída principalmente ao pré-processamento robusto baseado em imagens *all-sky*.

A contribuição principal deste estudo reside na análise sistemática que confronta diferentes arquiteturas de redes neurais para a predição de irradiância solar de curto prazo no MHP, oferecendo evidências de que, em determinados cenários de geração de dados (principalmente quando se dispõe de variáveis que capturam aspectos críticos da atmosfera, como a cobertura de nuvens), a adoção de modelos mais simples e menos custosos computacionalmente pode ser suficiente. Além disso, os resultados reforçam a importância de avaliar estatisticamente as diferenças de desempenho entre modelos, pois, em muitas situações, a aparente superioridade de uma arquitetura pode não se sustentar diante de análises inferenciais.

Quanto às limitações citam-se validação cruzada com apenas 3 *folds* e a análise restrita a 3 curtos horizontes que poderiam ser investigados em trabalhos futuros. Além disso, sugere-se: (i) investigar modelos híbridos, combinando diferentes arquiteturas para otimização preditiva, por exemplo, aliando CNNs ao processamento sequencial com *Transformers*; (ii) ampliar as janelas de entrada para capturar padrões temporais mais longos; e (iii) avaliar novos horizontes de predição, explorando o impacto do tempo de antecipação na acurácia.

Os avanços deste estudo podem aprimorar a aplicação de IA na gestão da energia renovável e, desta forma, a presente pesquisa contribui para a seleção criteriosa de arquiteturas para a PGESEF e reforça a relevância de um pré-processamento adequado, assim como para a evolução das técnicas de predição, fornecendo subsídios para futuras pesquisas e para o aprimoramento de sistemas inteligentes aplicados à gestão de energia renovável.

Por fim, destaca-se que o presente estudo se originou e integra uma linha de pesquisa em que o autor participou dos seguintes trabalhos científicos relacionados ao tema (**Apêndice**):

- 1) Maciel, J. N.; **Souza, Gustavo Campoi et al.** Optimizing Solar Irradiance Prediction: Feature Selection for All-Sky Image Processing Using a Hybrid Prediction Method. *IEEE Latin America Transactions*, [s. l.], v. 23, n. 1, p. 50–57, 2025. Disponível em: <https://doi.org/10.1109/TLA.2025.10810404>
- 2) **Souza, Gustavo Campoi et al.** Avaliação de métricas de imagens all-sky com redes neurais artificiais aplicadas à predição da irradiância solar. In: *Anais do XIII Encontro Anual de Iniciação Científica (EICTI)*. Foz do Iguaçu, Brasil. 2024.
- 3) **Souza, Gustavo Campoi**; Maciel, J. N. et al. Influência das medidas de processamento de imagem na predição da irradiância solar com redes neurais artificiais. In: *5a Congreso de Ingenierías y Ciencias Aplicadas de Las Tres Fronteras (MEC3F)*, 2024. Foz do Iguaçu, Brasil. 2024.
- 4) Bastos, M; **Souza, Gustavo Campoi**; Maciel, J. N. et al. Algoritmos de inteligência artificial aplicados ao método híbrido de predição da irradiância solar e energia fotovoltaica. In: *5o Congreso de Ingenierías y Ciencias Aplicadas de las Tres Fronteras (MEC3F)*, 2024.

6. REFERÊNCIAS

- ABSOLAR. **Energia Solar Fotovoltaica no Brasil - Infográfico ABSOLAR**. Disponível em: <<https://www.absolar.org.br/mercado/infografico>>. Acesso em: 19 jul. 2022.
- AKIBA, T. et al. Optuna: A Next-generation Hyperparameter Optimization Framework. 25 jul. 2019, New York, NY, USA: ACM, 25 jul. 2019. p. 2623–2631. Disponível em: <<https://dl.acm.org/doi/10.1145/3292500.3330701>>.
- BISONG, E. Google Colaboratory. **Building Machine Learning and Deep Learning Models on Google Cloud Platform**. Berkeley, CA: Apress, 2019. p. 59–64. Disponível em: <http://link.springer.com/10.1007/978-1-4842-4470-8_7>.
- BRAGA, Newton C. **Silício**. 2025. Disponível em: https://www.newtoncbraga.com.br/index.php?option=com_content&view=article&id=809:silicio&catid=36:almanaque&Itemid=55. Acesso em: 29 mar. 2025.
- BRYKALSKI M. **Produção de Espumas Vítreas Utilizando Resíduo de Vidro de Placas Fotovoltaicas**. Pelotas: [s.n.], 2023.
- CHOW, C. W. et al. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. **Solar Energy**, 2011. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0038092X11002982>>.
- CREMASCO, N. P. et al. **Estudo de Diferentes Tecnologias de Células Fotovoltaicas**. **Cient. Inov. Tecnol.** [S.l.: s.n.], 2021. Disponível em: <<https://periodicos.utfpr.edu.br/recitv>>.
- DEEPSEEK. **DeepSeek** [Assistente inteligente]. 2023. Disponível em: <https://www.deepseek.com>. Acesso em: 18 fev. 2025.
- DEV, S. et al. Estimating Solar Irradiance Using Sky Imagers. 11 out. 2019. Disponível em: <<http://arxiv.org/abs/1910.04981>>.
- DIAS P, V. H. **Caracterização e Reciclagem de Materiais de Módulos Fotovoltaicos**. Porto Alegre: [s.n.], 2015.
- DÜRR, O.; SICK, B.; MURINA, E. **Probabilistic Deep Learning: with Python, Keras and Tensorflow Probability**. [S.l.]: Manning Publications, 2020. Disponível em: <<https://books.google.com.br/books?id=-bYCEAAAQBAJ>>.
- ELLIS, A. et al. PV output smoothing with energy storage. **Conference Record of the IEEE Photovoltaic Specialists Conference**, p. 1523–1528, 2012.
- FEI, W. et al. Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. **Energy and Buildings**, v. 86, p. 427–438, jan. 2015. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0378778814008226>>.
- FENG, C.; ZHANG, J. SolarNet: A sky image-based deep convolutional neural network for intra-hour solar forecasting. **Solar Energy**, v. 204, p. 71–78, jul. 2020. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X20303285>>. Acesso em: 14 dez. 2023.

FERNANDES C. **Célula Fotovoltaica ZnO@ZnS Core@Shell, Obtida a Partir do Método Hidrotérmico Assistido por Microondas.** Pelotas: [s.n.], 2017.

GUEYMARD, C. A. A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. **Renewable and Sustainable Energy Reviews**, v. 39, p. 1024–1034, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.rser.2014.07.117>>.

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1 nov. 1997. Disponível em: <<https://direct.mit.edu/neco/article/9/8/1735-1780/6109>>.

JAILANI, N. L. M. et al. **Investigating the Power of LSTM-Based Models in Solar Energy Forecasting. Processes.** [S.l.]: Multidisciplinary Digital Publishing Institute (MDPI). , 1 maio 2023

JASP TEAM. **JASP (Version 0.16)[Computer software].** Disponível em: <<https://jasp-stats.org/>>.

KIM, J. et al. A Comprehensive Survey of Time Series Forecasting: Architectural Diversity and Open Challenges. 24 out. 2024. Disponível em: <<http://arxiv.org/abs/2411.05793>>.

KONSTANTINOU, M.; PERATIKOU, S.; CHARALAMBIDES, A. G. Solar Photovoltaic Forecasting of Power Output Using LSTM Networks. **Atmosphere**, v. 12, n. 1, p. 124, 18 jan. 2021. Disponível em: <<https://www.mdpi.com/2073-4433/12/1/124>>.

KOTSIANTIS, S. B.; ZAHARAKIS, I. D.; PINTELAS, P. E. Machine learning: A review of classification and combining techniques. **Artificial Intelligence Review**, v. 26, n. 3, p. 159–190, 2006.

KUMAR, D. S. et al. Solar irradiance resource and forecasting: a comprehensive review. **IET Renewable Power Generation**, v. 14, n. 10, p. 1641–1656, 2020.

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

LAPPALAINEN, K.; VALKEALAHTI, S. Output power variation of different PV array configurations during irradiance transitions caused by moving clouds. **Applied Energy**, v. 190, p. 902–910, mar. 2017. Disponível em: <<http://dx.doi.org/10.1016/j.apenergy.2017.01.013>>.

LIM, S. C. et al. Solar Power Forecasting Using CNN-LSTM Hybrid Model. **Energies**, v. 15, n. 21, 1 nov. 2022.

MACIEL, J. N. L. J. J. G. A. J. O. H. Forecasting Solar Power Output Generation: A Systematic Review with the Proknow-C. **IEEE Latin America Transactions**, v. 19, p. 612–624, 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9448544/>>. Acesso em: 3 fev. 2025.

MACIEL, J. N. **Método Híbrido de Predição da Irradiância Solar Com Processamento de Imagens e Inteligência Artificial Aplicável a Geração de Energia Solar Fotovoltaica.** 2022. 1–146 f. Tese de Doutorado–Universidade Federal da Integração Latino-Americana (UNILA), Foz do Iguaçu - PR, 2022. Disponível em: <<https://dspace.unila.edu.br/handle/123456789/7135>>. Acesso em: 8 fev. 2023.

MACIEL, J. N.; LEDESMA, J. J. G.; ANDO JUNIOR, O. H. Dataset for Machine Learning: Explicit All-Sky Image Features to Enhance Solar Irradiance Prediction. **Data**, v. 9, n. 10, p. 113, 29 set. 2024a. Disponível em: <<https://www.mdpi.com/2306-5729/9/10/113>>.

MACIEL, J. N.; LEDESMA, J. J. G.; ANDO JUNIOR, O. H. Forecasting Solar Power Output Generation: A Systematic Review with the Proknow-C. **IEEE Latin America Transactions**, v. 19, n. 4, p. 612–624, abr. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9448544/>>.

MACIEL, J. N.; LEDESMA, J. J. G.; ANDO JUNIOR, O. H. Hybrid prediction method of solar irradiance applied to short-term photovoltaic energy generation. **Renewable and Sustainable Energy Reviews**, v. 192, p. 114185, mar. 2024b. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1364032123010432>>.

MACIEL, Joylan Nunes et al. Analysis of Artificial Neural Networks for Forecasting Photovoltaic Energy Generation with Solar Irradiance. **Brazilian Archives of Biology and Technology**, v. 64, n. spe, 2021. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-89132021000200212&tlng=en>.

MINING, E. **Python Machine Learning: Understand Python Libraries (Keras, NumPy, Sci-kit-Lear, TensorFlow) for Implementing Machine Learning Models in Order to Build Intelligent Systems**. [S.l.]: Amazon Digital Services LLC - KDP Print US, 2019. Disponível em: <<https://books.google.com.br/books?id=qqQdzAEACAAJ>>.

MORETTIN, P. A.; BUSSAB, W. de O. **Estatística básica**. 9ª edição ed. São Paulo: Saraivauni, 2017.

MOSAVI, A. et al. State of the Art of Machine Learning Models in Energy Systems, a Systematic Review. **Energies 2019, Vol. 12, Page 1301**, v. 12, n. 7, p. 1301, 4 abr. 2019. Disponível em: <<https://www.mdpi.com/1996-1073/12/7/1301/htm>>. Acesso em: 13 abr. 2022.

OBANDO, E. D.; CARVAJAL, S. X.; PINEDA AGUDELO, J. Solar Radiation Prediction Using Machine Learning Techniques: A Review. **IEEE Latin America Transactions**, v. 17, n. 04, p. 684–697, abr. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8891934/>>.

OPENAI. **ChatGPT** (modelo GPT-4). 2025. Disponível em: <https://openai.com/>. Acesso em: 7 fev. 2025.

PEDRO, H. T. C.; LARSON, D. P.; COIMBRA, C. F. M. A Comprehensive Dataset for the Accelerated Development and Benchmarking of Solar Forecasting Methods. **Journal of Renewable and Sustainable Energy**, v. 11, n. 3, p. 036102, maio 2019. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.5094494>>.

PEREIRA, G. N. et al. **Célula Fotovoltaica de Perovskita: Uma Revisão do Estado da Arte**. [S.l.: s.n.], 2024.

PORTAL SOLAR. **Portal Solar**. 2025. Disponível em: <https://www.portalsolar.com.br/>. Acesso em: 5 fev. 2025.

RAJAGUKGUK, R. A.; KAMIL, R.; LEE, H.-J. A Deep Learning Model to Forecast Solar Irradiance Using a Sky Camera. **Applied Sciences**, v. 11, n. 11, p. 5049, 29 maio 2021. Disponível em: <<https://www.mdpi.com/2076-3417/11/11/5049>>.

RENEWABLE ENERGY AGENCY, I. **Renewable power generation costs in 2019**. [S.l.: s.n.], 2020. Disponível em: <www.irena.org>.

RIBEIRO R. **Estudo de caso: Dimensionamento de um Sistema Fotovoltaico Residencial**. [S.l.: s.n.], 2016.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 3rd edition ed. United Kingdom: Pearson Education, 2016.

SERRANO ARDILA, V. M. et al. Fuzzy Time Series Methods Applied to (In)Direct Short-Term Photovoltaic Power Forecasting. **Energies**, v. 15, n. 3, p. 845, 24 jan. 2022. Disponível em: <<https://www.mdpi.com/1996-1073/15/3/845>>.

SHAHRIARI, B. et al. **Taking the human out of the loop: A review of Bayesian optimization. Proceedings of the IEEE**. [S.l.]: Institute of Electrical and Electronics Engineers Inc. , 1 jan. 2016

SHIVASHANKAR, S. et al. Mitigating methods of power fluctuation of photovoltaic (PV) sources - A review. **Renewable and Sustainable Energy Reviews**, v. 59, p. 1170–1184, 2016. Disponível em: <<https://www2.scopus.com/inward/record.uri?eid=2-s2.0-84961305114&doi=10.1016%2Fj.rser.2016.01.059&partnerID=40&md5=91a7393c7d8d2f166b6f58f12fd10213>>.

SILVA, T. M.; BATISTA, Y. N. **Estimativa de Irradiância Solar Utilizando Redes Neurais Artificiais e Imagens do Céu**. Pesqueira: [s.n.], 15 ago. 2020.

SINKOVICS, J. G. **RNA/DNA and Cancer**. M.D. Anderson Hospital Comprehensive Cancer Center, The University of Texas, Houston, TX, United States: Springer International Publishing, 2016. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84969199789&doi=10.1007%2F978-3-319-22279-0&partnerID=40&md5=242bac36af4240b764e8c4bad6f25a1c>>.

SOLAR GYCX. **EU Market Outlook For Solar Power - 2023 - 2027**. Jiangxi, China: [s.n.], 2023. Disponível em: <<https://gycxsolar.com/Solar-Global-Market-Outlook/>>. Acesso em: 12 jun. 2024.

TINA, G. M. et al. A State-of-Art-Review on Machine-Learning Based Methods for PV. **Applied Sciences**, v. 11, n. 16, p. 7550, 17 ago. 2021. Disponível em: <<https://www.mdpi.com/2076-3417/11/16/7550>>.

URRACA, R. et al. Smart baseline models for solar irradiation forecasting. **Energy Conversion and Management**, v. 108, p. 539–548, jan. 2016.

VOYANT, C. et al. Machine Learning Methods for Solar Radiation Forecasting: A Review. **Renewable Energy**, v. 105, p. 569–582, 2017. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85008622769&doi=10.1016%2Fj.renene.2016.12.095&partnerID=40&md5=4c1d53f9e15b4f37b0bf62a15dcc01e6>>.

WANG, F. et al. A minutely solar irradiance forecasting method based on real-time sky image-irradiance mapping model. **Energy Conversion and Management**, v. 220, p. 113075, 15 set. 2020. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0196890420306191>>.

WANG, H. et al. A review of deep learning for renewable energy forecasting. **Energy Conversion and Management**, v. 198, p. 111799, out. 2019. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0196890419307812>>.

WENTZ, V.H et al. Solar Irradiance Forecasting to Short-Term PV Power: Accuracy Comparison of ANN and LSTM Models. **Energies**, v. 15, n. 7, p. 2457, 27 mar. 2022. Disponível em: <<https://www.mdpi.com/1996-1073/15/7/2457>>.

WENTZ, Victor Hugo. **Avaliação de Modelos de Inteligência Artificial Aplicados à Predição de Irradiância Solar**. 2021. 1–105 f. Trabalho de Conclusão de Curso–Universidade Federal da Integração Latino-Americana, 2021.

YANG, H. et al. 3d-cnn-based sky image feature extraction for short-term global horizontal irradiance forecasting. **Water (Switzerland)**, v. 13, n. 13, 1 jul. 2021.

YING, C. et al. Deep learning for renewable energy forecasting: A taxonomy, and systematic literature review. **Journal of Cleaner Production**, v. 384, p. 135414, 15 jan. 2023. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0959652622049885>>.

YU, Yong et al. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. **Neural Computation**, v. 31, n. 7, p. 1235–1270, jul. 2019. Disponível em: <<https://direct.mit.edu/neco/article/31/7/1235-1270/8500>>.

YU, Yunjun; CAO, J.; ZHU, J. An LSTM Short-Term Solar Irradiance Forecasting Under Complicated Weather Conditions. **IEEE Access**, v. 7, p. 145651–145666, 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8864021/>>.

ZHANG, R. et al. Forecast of Solar Energy Production - A Deep Learning Approach. nov. 2018, [S.l.]: IEEE, nov. 2018. p. 73–82. Disponível em: <<https://ieeexplore.ieee.org/document/8588777/>>.

ZHEN, Z. et al. Ultra-short-term irradiance forecasting model based on ground-based cloud image and deep learning algorithm. **IET Renewable Power Generation**, v. 16, n. 12, p. 2604–2616, 6 set. 2022. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1049/rpg2.12280>>.

Optimizing Solar Irradiance Prediction: Feature Selection for All-Sky Image Processing using a Hybrid Prediction Method

Joylan Nunes Maciel ¹, Gustavo Campoi de Souza ¹, Willian Zalewski ¹, Jorge Javier Gimenez Ledesma ¹, and Oswaldo Hideo Ando Junior ¹

Abstract—The forecasting of solar irradiance is crucial for photovoltaic solar energy generation, as production is subject to intermittency due to climatic conditions, such as cloud cover, wind and, temperature. Based on the Hybrid Prediction Method (HPM), this study investigated the influence of a set of all-sky image processing features on the HPM's Artificial Neural Network prediction accuracy. Using correlation-based attribute selection, three predictive models with different input feature sets were evaluated. The results show that, when considering all horizons together and paired, the Medium set of 6 features achieves prediction accuracy statistically similar to the Complete set with 9 features, reducing the computational time (14.4%) and model input dimensionality (33.3%). However, when comparing individual horizons, the Complete set outperforms the Medium set at 5- and 15-minute horizon, while maintain similar accuracy at the 1-minute horizon. The Reduced set, with three features, consistently underperformed. This study provides news insights into the optimization of solar irradiance forecasting using HPM, contributing to advances in photovoltaic energy forecasting.

Link to graphical and video abstracts, and to code: <https://latam.ieceer9.org/index.php/transactions/article/view/9222>

Index Terms— all-sky image, solar irradiance prediction, artificial neural network, Hybrid Prediction Method.

I. INTRODUCTION

THE growing energy demand has been increasingly met by renewable sources like wind and photovoltaics, with significant investments in solar energy installations, particularly in China, Japan, the United States, Brazil, Germany, and India [1], [2]. Photovoltaic solar energy adoption has spurred research into methods for predicting solar irradiance, the key factor in energy production [3]. The primary challenge in solar energy generation is the intermittency caused by atmospheric variations, especially cloud cover [4], [5]. Since 2017, there has been a surge in publications on photovoltaic energy prediction [5], mainly using Machine Learning (ML) and Deep Learning (DL) techniques [4], [6], [7], [8], [9]. Advances in computational

The associate editor coordinating the review of this manuscript and approving it for publication was Ruth Aguilar (Corresponding author: Joylan Nunes Maciel).

J. N. Maciel, G. C. de Souza, W. Zalewski, and J. J. D. Ledesma are with Federal University of Latin American Integration, Foz do Iguaçu, Brazil (e-mails: joylan.maciell@unila.edu.br, gustavocampoi17@gmail.com, willian.zalewski@unila.edu.br, and jorge.ledesma@unila.edu.br).

O. H. Ando Junior is with Federal Rural University of Pernambuco, Cabo de Santo Agostinho, Brazil (e-mail: oswaldo.ando@ufrpe.br).

power and DL models have further promoted prediction studies using all-sky images [10], [11], [12], [13], [14], [15].

A research group at the Federal University of Latin American Integration (UNILA) has developed scientific studies on the short-term Prediction of Solar Photovoltaic Energy Generation (PSPEG) [16], [17], [18], [19], [20], [21]. Based on this research, the study by [21] proposed and evaluated a new hybrid approach to predicting solar irradiance, applying Image Processing (IP) [22] and Machine Learning techniques [23]. This approach is referred to as Hybrid Prediction Method (HPM). The HPM was developed from a dataset containing historical all-sky image data (180°, meteorological, and solar irradiance information, created with controlled quality and samples collected every minute over a complete period of 3 years (2014 to 2016) [24].

Based on the HPM, the aim of this research is to investigate the set of features (characteristics) extracted from all-sky images, proposed in the HPM, in order to verify the influence of different combinations of these features in predicting short-term solar irradiance (1, 5 and 15 minutes), using the Artificial Neural Network (ANN) model [23] adopted in the HPM. The hypothesis to be investigated is that an optimized set of these features can provide predictions with similar accuracy to the use of all features proposed in the HPM, thus reducing complexity and computational processing.

A theoretical literature review on concepts and works related to PSPEG, is presented in Section 2. The material and methods area described in Section 3. Section 4 presents and discusses the experimental results. The conclusion, contributions and, limitations are outlined in Section 5.

II. THEORETICAL BACKGROUND

The growing demand for clean, so-called renewable energy is driving investment in sources such as photovoltaic solar energy. However, the intermittency of energy generation, caused mainly by variations in cloud cover [25], requires accurate methods of predicting solar irradiance [26]. In this context, Prediction of Solar Photovoltaic Energy Generation (PSPEG) plays a crucial role in optimizing photovoltaic system operations. Its enables adjustments in load distribution to maintain grid stability, facilitate real-time management through the activation of backup resources, promotes early trading in the energy market, and allows for more reliably battery sizing, thereby enhancing both operational and economic performance of photovoltaic systems [27], [28].

Artigo 2



EICTI

XIII ENCONTRO ANUAL DE INICIAÇÃO CIENTÍFICA E
IX ENCONTRO DE INICIAÇÃO EM DESENVOLVIMENTO
TECNOLÓGICO E INOVAÇÃO



AVALIAÇÃO DE MÉTRICAS DE IMAGENS ALL-SKY COM REDES NEURAIS ARTIFICIAIS APLICADAS À PREDIÇÃO DA IRRADIÂNCIA SOLAR

Gustavo Campoi de Souza^{1,2*}, Joylan Nunes Maciel^{2,3}, Willian Zalewski², Jorge Javier Gimenez Ledesma² e Oswaldo Hideo Ando Junior²

1) Bolsista ITI-UNILA, Engenharia de Materiais, ILATIT, UNILA.

2) Pesquisador colaborador, Laboratório de Computação (LACA), ILATIT, UNILA.

3) Orientador, Universidade Federal da Integração Latino Americana, ILATIT, UNILA.

*E-mail de contato: gc.souza.2020@aluno.unila.edu.br

1. RESUMO

A crescente demanda por energia elétrica impulsiona a busca por fontes renováveis e sustentáveis, como a energia solar fotovoltaica, que tem se expandido significativamente na matriz elétrica brasileira (Maciel *et al.*, 2021). No entanto, essa fonte enfrenta intermitências na geração devido a condições climáticas adversas, o que afeta negativamente a estabilidade do fornecimento de energia. Para mitigar essas intermitências, esta pesquisa analisou o Método Híbrido de Predição (MHP), que combina técnicas de Processamento de Imagens (PI) all-sky e Aprendizado de Máquina com Rede Neural Artificial (RNA) para prever a irradiância solar de curto prazo, com base na Irradiância Solar Global (GHI) (Maciel *et al.*, 2024). No MHP, as imagens all-sky (180°) são processadas para se extrair um conjunto de nove métricas (características) usadas como entrada para uma Rede Neural Artificial (RNA). A presente pesquisa avaliou a influência destas métricas na acurácia preditiva do MHP, nos horizontes de 1, 5 e 15 minuto. As nove métricas (*Clouds Movement, Clouds Around Sun, Sun Luminance Adjusted, Clear Sky GHI, Clouds Coverage, Sun Luminance, Season, Sun Located* e *White Pixel Ratio*) foram agrupadas por correlação (Pearson) com a variável GHI (alvo), resultando em três conjuntos: (i) Completo (CC) com todas as 9 medidas; (ii) Médio (CM) com 6 medidas (*Clouds Movement, Clouds Around Sun, Sun Luminance Adjusted, Clear Sky GHI, Clouds Coverage, Sun Luminance*) e; (iii) Reduzido (CR) com 3 medidas (*Clouds Movement, Clouds Around Sun, Sun Luminance Adjusted*). Os modelos preditivos com RNAs utilizando estes conjuntos de entrada foram desenvolvidos, treinados e suas acurácias foram analisadas. O dataset utilizado contém imagens e valores de GHI de um período completo de 3 anos (2014 a 2016), coletados a cada minuto. Durante o treinamento, a validação cruzada de partições (*3-fold cross validation*) foi aplicada, sendo cada *fold* correspondente a um ano completo. A métrica de acurácia adotada foi o *Root Mean Squared Error* (RMSE), que é medida mais utilizada na área (Maciel *et al.*, 2024). Os resultados, exibidos no Quadro 1, mostraram que o modelo com CR, com 3 medidas, apresentou os maiores erros e dispersão. Por outro lado, a similaridade visual dos resultados, entre os modelos com CM e CC, motivou a aplicação do teste estatístico estatística de ANOVA de uma via, seguida do pós-teste



INFLUÊNCIA DAS MEDIDAS DE PROCESSAMENTO DE IMAGEM NA PREDIÇÃO DA IRRADIÂNCIA SOLAR COM REDES NEURAIS ARTIFICIAIS

G. C. de Souza^{1*}, M. Bastos¹, W. Zalewski¹, J. J. G. Ledesma^{1,2}, O. H. Ando Junior^{2,4} e J. N. Maciel^{1,2,3}

¹Universidade Federal da Integração Latino Americana (UNILA).

²Programa de Pós-Graduação Interdisciplinar em Energia e Sustentabilidade (PPGIES/UNILA).

³Programa de Pós-Graduação em Física Aplicada (PPGFISA/UNILA).

⁴Universidade Federal Rural de Pernambuco (UFRPE).

*gustavocampoi17@gmail.com

Palavras-chave: irradiância solar, rede neural artificial, acurácia de previsão.

INTRODUÇÃO

A demanda crescente por energia elétrica impulsiona a busca por fontes renováveis e sustentáveis, as quais viabilizam o suprimento da demanda energética e contribuem para a sustentabilidade ambiental (Maciel et al., 2021). No Brasil, a energia solar fotovoltaica se destaca por seu crescimento expressivo na matriz elétrica brasileira, alcançando 42 Gigawatts de capacidade instalada em 2024 (ABSOLAR, 2024). No entanto, as condições climáticas adversas, como nebulosidades e ventos, afetam negativamente a geração de energia solar, ocasionando intermitências na geração de energia elétrica.

Os métodos de Predição da Geração de Energia Solar Fotovoltaica (PGESF) se apresentam como ferramentas para mitigar essa intermitência (Carneiro et al., 2022), permitindo o uso de técnicas para contorná-las. O Método Híbrido de Predição (MHP), proposto em (Maciel et al., 2024), combina técnicas de Processamento de Imagens (PI) e Aprendizado de Máquina (AM), para prever a geração de energia solar fotovoltaica, de curto prazo, com base na irradiância solar global (GHI).

No MHP, as imagens *all-sky* (180°) são processadas para extrair um conjunto de 8 (oito) medidas (características) que servem como entrada para uma Rede Neural Artificial (RNA). Visando aprimorar a acurácia preditiva do MHP,

no horizonte de 1 minuto, esta pesquisa tem como objetivo avaliar experimentalmente quais conjuntos de medidas, extraídas no PI do MHP, apresentam maior influência e eficiência na acurácia preditiva do MHP. Portanto, este estudo avalia e apresenta as análises de acurácia de RNAs construídas com três diferentes conjuntos de de medidas usadas como entrada (*inputs*) de dados.

MATERIAIS E MÉTODO

A análise experimental deste estudo aplicou o Método Híbrido de Predição (MHP) para extrair, das imagens *all-sky*, oito (8) medidas (m_1 a m_8) de características do céu, com técnicas de Processamento de Imagens (PI) (Figura 1).

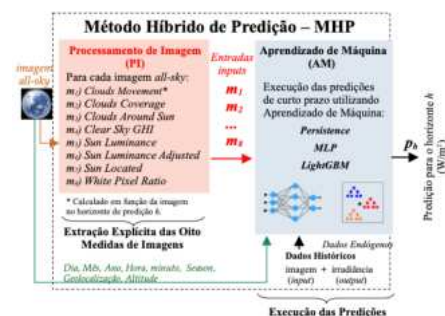


Figura 1. Método Híbrido de Predição. Fonte: adaptado de (Maciel, 2022).



ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL APLICADOS AO MÉTODO HÍBRIDO DE PREDIÇÃO DA IRRADIÂNCIA SOLAR E ENERGIA FOTOVOLTAICA

M. Bastos^{1*}, G. C. de Souza¹, W. Zalewski¹, J. J. G. Ledesma^{1,2}, O. H. Ando Junior^{2,4} e J. N. Maciel^{1,2,3}

¹Universidade Federal da Integração Latino Americana (UNILA).

²Programa de Pós-Graduação Interdisciplinar em Energia e Sustentabilidade (PPGIES/UNILA).

³Programa de Pós-Graduação em Física Aplicada (PPGFISA/UNILA).

⁴Universidade Federal Rural de Pernambuco (UFRPE).

*mariellebastos11@gmail.com

Palavras-chave: energia fotovoltaica, aprendizado de máquina, acurácia de predição.

INTRODUÇÃO

O crescimento populacional e a demanda crescente por energia elétrica impulsionam a busca por fontes renováveis e sustentáveis. Nesse contexto, a inclusão de fontes complementares ajuda a suprir a demanda energética, contribuindo para a sustentabilidade ambiental (Maciel et al., 2021). No Brasil, a energia solar fotovoltaica se destaca com um crescimento significativo na matriz elétrica brasileira, atingindo 42 GW em 2024 (ABSOLAR, 2024). Todavia, as condições meteorológicas, como nuvens e ventos, influenciam negativamente e causam intermitências na geração da energia elétrica.

Os métodos de Predição da Geração de Energia Solar Fotovoltaica (PGESF) surgem como ferramenta para mitigar essa intermitência e otimizar a geração de energia (Carneiro et al., 2022). Tais métodos possibilitam a aplicação de técnicas para contornar as intermitências e otimizar o controle na geração e despacho da energia futura.

Um estudo recente (Maciel et al., 2024) propôs o Método Híbrido de Predição (MHP), que combina conjuntamente técnicas de Processamento de Imagens (PI) e Aprendizado de Máquina (AM), para prever a geração de energia solar fotovoltaica utilizando a irradiância solar. No MHP, as imagens *all-sky* (180°) são processadas

para se extrair um conjunto de medidas (características) que servem como entrada para dois modelos de AM: uma Rede Neural Artificial (RNA) e um modelo LightGBM, sendo que a RNA apresentou o melhor desempenho na acurácia de predição (Maciel et al., 2024).

Neste contexto, para aperfeiçoar a acurácia preditiva do MHP, no horizonte de 1 minuto, esta pesquisa tem por objetivo avaliar experimentalmente a aplicação de três modelos de AM alternativos no MHP: *K-Nearest Neighbor* (KNN), *Random Forest* (RF) e *AdaBoost Regressor* (ADA) (Shalev-Shwartz & Shai Ben-David, 2014).

MATERIAIS E MÉTODO

A análise experimental deste estudo foi baseada na abordagem do Método Híbrido de Predição (MHP), o qual processa imagens *all-sky* e extrai um conjunto de medidas (características do céu) com técnicas de Processamento de Imagens (PI). Esse conjunto foi então utilizado como entrada para um modelo de AM, baseado em RNA *MultiLayer Perceptron* (Russell & Norvig, 2016).

Após o treinamento com dados históricos, a RNA executou as predições de curto prazo para horizontes entre 1 e 60 minutos (**Figura 1**).