

31 - 05 | 2025

GESTÃO DA ESTIMATIVA DE DISPONIBILIDADE DE ENERGIA SOLAR UTILIZANDO MODELOS DE APRENDIZADO DE MÁQUINA

Management of solar energy availability estimation using machine learning models

Gestión de la estimación de la disponibilidad de energía solar mediante modelos de aprendizaje automático

Fernando V. Mucomole⁸⁴ Carlos A.S. Silva⁸⁵ Lourenço L. Magaia⁸⁶

¹ CS-OGET—Center of Excellence of Studies in Oil and Gas Engineering and Technology, Faculty of Engineering, Eduardo Mondlane University, Mozambique Avenue Km 1.5, Maputo 257, Mozambique

² Department of Mechanical Engineering, Instituto Superior Técnico, University of Lisbon, 1600-214 Lisbon, Portugal.

³ Department of Mathematics and Informatics, Faculty of Science, Eduardo Mondlane University, Main Campus No. 3453, Maputo 257, Mozambique

Autor para correspondência: miguelbungo1108@yahoo.com

Data de recepção: 17-01-2025

Data de aceitação: 02-03-2025

Como citar este artigo: Mucomole, F.V.; Silva, C.A.S.; Magaia, L.L. (2025). Gestão da estimativa de disponibilidade de energia solar utilizando modelos de aprendizado de máquina. *ALBA – ISFIC Research and Science Journal*, 2(7), pp. 823 – 834. <https://alba.ac.mz/index.php/alba/issue/view/10>.

RESUMO

Devido a factores de diversa natureza a energia a sua chegada a superfície da terra pode apresentar-se intermitente, propiciando um obstáculo para o seu aproveitamento. Motivada pela necessidade de maior rentabilização de energia solar, primordialmente para produção de energia fotovoltaica. Foi feita a gestão da estimativa de disponibilidade de energia solar em curta escala de medição, para conhecer o real comportamento da energia solar, na

localidade de Pomene. Foi colhido a energia solar no ano de 2014, durante 12 meses, com período diário e intervalo de medição de 10 minutos. Usando os modelos de aprendizado por máquina nomeadamente random forest, regressive k-riking, e artificial neural networks foi predita a energia solar. A gestão da energia solar predita mostra a predominância de dias de sol pleno com classificação de intermediário, e claro, potência do ar existência de energia para aproveitamento, ademais concluiu-se que o modelo random

forest apresenta menor erro de predição da potência com os modelos de APM para maior análise e mais eficiência de produção energética.

Palavras-chave: gestão, disponibilidade, energia solar, estimativa, aprendizado por máquina.

ABSTRACT

Due to factors of different nature, energy upon arrival at the surface of the earth may appear intermittent, creating an obstacle to its use. Motivated by the need for greater profitability of solar energy, primarily for the production of photovoltaic energy. An estimate of the availability of solar energy was managed on a short measurement scale, to understand the real behavior of solar energy in the town of Pomene. Solar energy was collected in 2014, for 12 months, with a daily period and a 10-minute measurement interval. Using machine learning models such as random forest, regressive Kriking, and artificial neural networks, solar energy was predicted. The predicted energy management shows the predominance of full sun with intermediate classification, and of course, the power of the existence of energy for use, in addition, it was concluded that the random forest model presents lower prediction error and enhances the APM models. for greater analysis, more energy production efficiency.

Keywords: management, availability, solar energy, estimation, machine learning.

RESUMEN

Debido a factores de diversa índole, la energía al llegar a la superficie terrestre puede presentar intermitencia, lo que supone un obstáculo para su aprovechamiento. Motivados por la necesidad de una mayor rentabilidad de la energía solar, principalmente para la producción de energía fotovoltaica. Se realizó una estimación de la disponibilidad de energía solar en una escala de medición corta, para comprender el comportamiento real de la energía solar en la localidad de Pomene. La energía solar se recogió en 2014, durante 12 meses, con un periodo diario y un intervalo de medición de 10 minutos. Utilizando modelos de aprendizaje automático como el

bosque aleatorio, el Kriking regresivo y las redes neuronales artificiales, se predijo la energía solar. La gestión energética prevista muestra el predominio del sol pleno con clasificación intermedia y, por supuesto, la potencia de la existencia de energía para su uso. Además, se concluyó que el modelo de bosque aleatorio presenta un menor error de predicción y mejora los modelos APM para un mayor análisis y una mayor eficiencia en la producción de energía.

Palabras clave: gestión, disponibilidad, energía solar, estimación, aprendizaje automático

1 INTRODUÇÃO

O crescimento das sociedades está ligado com a necessidade de maior demanda de consumo de energia. Nos anos actuais os combustíveis fósseis constituem cerca de 80% da energia usada pelo homem para suas necessidades. Contudo estes poluem consideravelmente a natureza, colocando em perigo a vida do homem, sem considerar que o seu ritmo de consumo pelo homem é maior em relação a sua demanda de regeneração com o tempo, isto é são esgotáveis. A adopção de outras formas de energias tem sido uma estratégia usada no âmbito dos objectivos sustentáveis, cujo essência primordial é atingir uma taxativa electrificação por voltada de 2030, através da massificação de fontes que sejam menos nocivas ao ambiente até por volta de 2050. A aposta no uso das energias renováveis tais conhecidas energias verdes, com enfoque a energia solar fotovoltaica (FV) vem revelando-se uma alternativa por este constituir uma forma de energia limpa, sustentável, eficiente, e sem oscilações podendo ser usado em qualquer parte do mundo onde há disponibilidade da energia solar, em detrimento a outras formas de energias renováveis que contribuem para o

devastação ecológica, desflorestamento, etc (Leirvik, T., & Yuan, M. 2021).

O Sol é a fonte primária de energia na terra e garante a existência de vários processos biológicos que garantem a vida na terra e a sustentabilidade. Contudo este é o principal recurso de excitação de fonte de motor de geração das várias formas de energias solar, com destaque para a tecnologia solar fotovoltaica que usa a energia solar horizontal para geração de electricidade em uma usina solar. O recurso solar a sua chegada a superfície da terra pode apresentar-se distribuída de maneira desigual devido a vários factores como: a presença de aerossóis na atmosfera, gases nocivos, bolhas de água, poeiras entre outros que dispersão a radiação solar a sua chegada á superfície da terra, dividindo o feixe de radiação em radiação directa (a que chega a terra sem ser dispersa) e difusa (o que chega a terra após ser dispersa). A gestão da variabilidade da energia solar é um parâmetro bastante importante a medida que influencia a geração FV que pode ser afectada pelas flutuações da energia solar. O estabelecimento da gestão da variabilidade é um parâmetro para melhor dimensionamento de sistemas fotovoltaicos, o que significaria a existência de um instrumento padronizado de apuramento dos aparatos do sistema FV para a óptima instalação e monitoramento deste, que resultaria no aumento de tempo de vida em relação ás avarias, sobrecargas, ruptura que veem sofrendo os sistemas já estabelecido. Isto surge devido ao fraco estabelecimento de normas de acordo com a gestão da radiação no local em que se pretende instalar em qualquer parte do mundo. Aliado a isto a extrapolação da energia solar nos arredores de uma estação de medição, contribuiria consideravelmente

para conhecer a disponibilidade de potencial em locais de difícil alcance onde se pode aproveitar para instalar as centrais FVs, reduzindo a devastação ambiental e economia de espaço, em vez de aproveitamento de áreas recicláveis como terras aráveis, bem como superfícies de águas. O óptimo dimensionamento dos sistemas FV com recurso a radiação solar gerada a longo prazo e em curta escala de tempo, reduziria consideravelmente o custo de acesso a esta tecnologia e tornando-a acessível a energia a todas populações, com destaque as sociedades rurais que constituem 81% da população mundial sem acesso a electrificação. Uma ainda maior viabilidade e relevância no dimensionamento e gestão de potência a saída de uma usina solar revela-se na estimativa futura da energia solar com base no recurso medido e ou extrapolado, isto permite ao projectista monitorar e simular a geração de energia com cenário ao futuro (Mucomole et. al., 2024).

Vários modelos de estimativa de energia solar foram propostos em estudos recentes, como o caso em Leirvik, T., & Yuan, M. (2021) que usaram o modelo Random Forest (RF) e encontraram menor margem de erro MAE de $10,2 \text{ Wm}^{-2}$, Puga-Gil et al., (2022) usaram o RF, e artificial neural network (ANN), e encontram um erro MAPE no valor entre 3,9 e 13,8% e Sridharan, (2023) faz a estimativa baseada em modelo de rede neural de regressão generalizada de energia solar global, erro é de 3,55%, 4,64% e 5,49%, entre outros.

A presente pesquisa baseou-se em apresentar a gestão da estimativa da disponibilidade do recurso de energia solar a longo prazo e em curta escala de medição. Tomando como amostra fonte

de dados medido por piranômetro epyly instalado na província de Inhambane na estação de Pomente, medida no ano 2014, com período de um dia compreendido de Janeiro á Dezembro e um intervalo de medição de 10 minutos, o tempo suficiente para perceber os reais factores que podem afectar a variabilidade e quantificar as flutuações de energia solar na sua chegada á superfície da terra. A pesquisa baseou-se na essência de como a gestão de estimativa de energia solar usando modelos de APM pode influenciar no planeamento de cenário de aproveitamento de energia e consequentemente melhor gestão de impactos ambientais. Com tudo, já se havia relatado a hipótese a comprovar de que o rápido desenvolvimento recente da mineração de dados, abordagens alternativas de estimativa por modelos de APM caso de: RF, ANN, OK, são melhores para reduzir o erro de estimativa em detrimento de avaliação de amostra de dados brutos que carregam consigo vários constrangimentos introduzidos por diversos factores, por outro lado, ha que considerar que tornaram-se mais populares e mais fáceis de operar. Na presente pesquisa fazer se a gestão com auxílio ao modelo RF, ANN, OK para estimar a variabilidade de longo prazo na irradiância solar da superfície e para quantificar o recurso solar e projetar os piores cenários de produção de energia em tais sistemas.

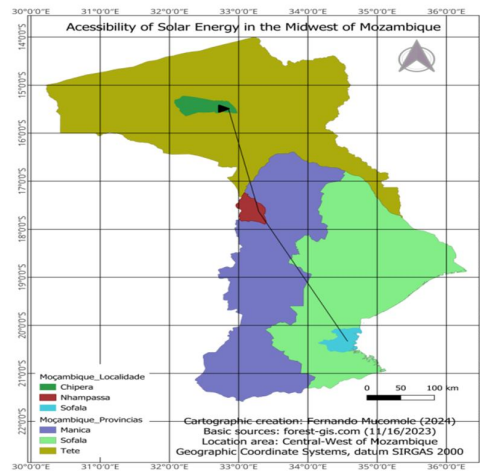
A presente pesquisa contribui no estabelecimento de técnicas aprimoradas

de estimativa de energia solar usando APM com enfoque para: RF, ANN, OK, que reduzem consideravelmente o erro de predição e melhor gestão preditiva da potência FV á saída de uma usina solar, podendo a descrição metodologica aqui usada ser usada para estimar e gerir o recurso de energia solar FV em qualquer parte do mundo.

2 MATERIAIS E MÉTODOS

2.1. COLECTA E PROCESSAMENTO DE DADOS

Os dados de GHI foram colhidos na estação FV estabelecida na localidade de Pomene na Província de Inhambane, situada na região Sul de Moçambique, localizada entre os paralelos 22°56'44,26" S de latitude Sul e entre os meridianos 35°35'35,52" E de longitude Leste como se pode ver na *Figura 1(a)*. Os dados usados no processamento foram colhidos durante o ano 2014, a partir de 1 de Janeiro de 2014 á 31 de Dezembro do mesmo ano, constituindo uma medição a longo prazo, uma amplitude de um dia e com um intervalo de medição de 10 min. Os dados foram colectado usando um piranômetro do tipo empply para medição de radiação global apresentado na *Figura 1(b)* e um piranômetro com anel apresentado na *Figura 1(c)* que mede a radiação global difusa, estes dois dispositivos encontravam-se associados formando uma única plataforma de medição durante a realização do projecto.



(b)



(c)

Figura 1 – (a) Região de realização das medições em Pomene; (b) Piranômetro epyly e (c) Piranômetro com anel.

A amostra de dados colhida foi tratada de modo a aferir-se a regularidade das medições e a eventuais incidentes de falta de medição que significativamente poderiam dificultar a

gestão da amostra, contudo foram estabelecido baseadamente em existência de dias aceitáveis, não aceitáveis e dias não aplicáveis como se pode ver na tabela 1.

Tabela 1 – Comportamento de radiação solar na estação de Pomene em Inhambane.

Ano: 2014	Dias aceitáveis		Dias não aceitáveis	
Mês	Nr.	Identificação	Nr.	Identificação
Janeiro	18	2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 15, 16, 20, 23, 27, 28, 29, 30	13	1, 6, 7, 14, 17, 18, 19, 21, 22, 24, 25, 26, 31
Fevereiro	24	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26, 27, 28	4	7, 13, 19, 20,
Março	26	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 30, 31	5	10, 13, 21, 28, 29
Abril	27	1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30	3	4, 5, 25
Maio	30	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 20, 31	1	22
Junho	21	1, 2, 3, 4, 5, 7, 8, 9, 10, 12, 13, 15, 16, 18, 19, 23, 24, 26, 27, 29, 30	9	6, 11, 14, 17, 20, 21, 22, 25, 28
Julho	24	1, 3, 4, 5, 6, 7, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31	7	2, 8, 9, 10, 11, 20, 30
Agosto	28	1, 2, 3, 4, 5, 6, 7, 8, 10, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31	3	9, 11, 16,
Setembro	27	1, 3, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 22, 23, 24, 25, 26, 27, 28, 29, 30	3	2, 20, 21
Outubro	28	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 14, 15, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31	3	10, 16, 17
Novembro	23	1, 2, 4, 6, 7, 8, 9, 10, 11, 12, 13, 15, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	7	3, 5, 14, 16, 17, 18, 28
Dezembro	13	2, 2, 3, 4, 5, 9, 10, 12, 19, 20, 21, 22, 23, 24, 25, 26, 27, 31	18	1, 6, 7, 11, 13, 14, 15, 16, 17, 18, 28, 29, 30
Total		271		63

As medições são ótimas não havendo dias não aceitáveis estimadas em suma de 81% de dias aceitáveis e 19% de dias não aceitáveis, como se mostra na *Figura 2(a)*, o que motiva a melhor eficiência de gestão da potência á saída do gerador FV com maior probabilidade

de alcance de real variabilidade predita de disponibilidade da energia solar usando a amostra de GHI. De modo analisar-se os períodos de Sol pleno compreendidas em 12 horas na região na estao durante o ano 2014, onde contactou-se possuir menor radiação no nascer e

por do solar sendo máxima ao meio dia. Contudo os dias variam havendo dias em que a radiação experimental se aproxima da radiação teórica que constituem bons dias para nossa estimativa pois apresentam menores erros de medição

(ausência estatística de outliers), e dias com radiação muito maior e ou menor que a radiação teórica que contudo apresentam, maior erro de medição (presença estatística de outliers), como se pode ver na *Figura 2(b)*.

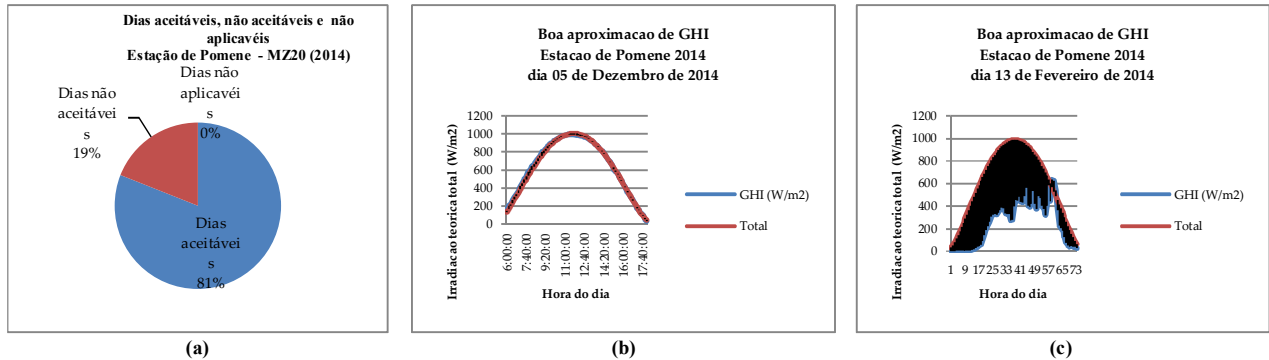


Figura 2 – (a) Quantificação de dias aceitáveis e não aceitáveis **(b)** Aproximação de trajetória radiação global no período de interesse e **(c)** Má aproximação da trajetória de radiação no período de interesse.

Durante as medições, cada instrumento foi submetido a manutenções regulares sazonalmente, para transferência de dados, troca de bateria, limpeza e realinhamento, neste processo foram registrados períodos de limpeza e orientação de orientação, e exceto na seleção de dados como períodos de interferência e turbulência.

2.2 METODOLOGIA

Apos a extração de amostra de GHI, apartir do datta Loger e disponibilizados para verificação, foi investigado o erro amostral ligado principalmente a erros de escala, e erros intrumentais normalmente fornecidos pelo fabricante que pode favorecer melhor a gestão da frota FV futura usando a relação:

$$e = \sqrt{e_a + e_e + e_c} \quad (\text{eq. 3})$$

onde: e_a é o erros aleatório, dado por $e_a = \frac{E.S}{\sqrt{n}}$, (S , é o desvio tipico da amostra,

E e um parametro que depende do nível de confiança ou confiabilidade (P_k eleito), ademais, e_e é o erro de escala e e_c é o erro de calibração ligada ao nível de calibração do instrumento e é fornecido pelo fabricante.

Recorrendo a modelos de machine learning foi predita a GHI. Usando primeiro o aplicação do **RF**, onde foram definidas três camadas de vizinhanças com escopo a saber k_1, k_2 e k_3 . Usamos a notação $\bar{z}(N_0^k)$ para a média de z_i na vizinhança, $\bar{z}(N_0^k)$ e N_0^k para a contrapartida média ponderada. A média ponderada é então dada por:

$$\bar{z}(N_0^k) = \sum_{z_i \in N_0^k} \lambda_i z_i \quad (\text{eq. 3})$$

Definindo o conjunto de dados espaciais $Z = \{z_1, z_2, z_3, \dots, z_n\}$ onde z_i é o valor da radiação z no local i , $N_0^k = \{z_i \in Z : d(0, i) \leq d(0, j_k), \forall i \neq 0\}$ a vizinhança da estação o , onde $d(0, i)$ indica o distância entre a estação o e a estação i , $d(0, j_k)$ denota a distância entre a estação o e sua

k –ésima estação mais próxima j_k (Saraçlı, 2008).

Para capturar a dispersão dos valores, usamos a variancia dos valores dentro desta vizinhança,

$$\sigma^2(N_0^k) = \frac{1}{|N_0^k|} \sum_{z_i \in N_0^k} (z_i - \bar{z}(N_0^k))^2 \quad (\text{eq. 3})$$

$\sigma(N_0^k)$ mede a volatilidade do SSR dentro da vizinhança N_0^k . Para suavizar os dados, introduziu-se a Seasonal Moving Average, $RAD_{t-12}, RAD_{t-24}, RAD_{t-36}, \dots$ onde, por exemplo, RAD_{t-12} é o RAD observado no 12 meses antes do mês t dada pela equação (2),

$$RAD_{t_mayK} = \frac{\sum_{i=1}^K RAD_{t-i \times s}}{K} \quad (\text{eq. 3})$$

Foi usado o modelo **Regression Kriging (RK)**, que combina tendências previstas de regressão com resíduos previstos de Krigagem dado por,

$$\hat{z}(x_0) = \hat{m}(x_0) + \hat{e}(x_0) = \sum_{k=0}^p \hat{\beta}_k q_k(x_0) + \sum_{i=1}^m \hat{w}_i e(x_i) \quad (\text{eq. 3})$$

onde $\hat{m}(x_0)$ é a tendência (variáveis climáticas e geográficas), e

$\hat{e}(x_0)$ é o resíduo, $\hat{\beta}_k, k = 0, 1, \dots, p$ são os coeficientes de regressão estimados, p é o número de variáveis auxiliares, $q_0(x_0) = 1$ é um termo constante e $\hat{w}_i, i = 1, \dots, m$ são os pesos de Krigagem (Saraçlı, 2008).

Foi usado o modelo ANN considerando uma gama de entrada para cada neurônio j , de u_1 , a u_n , multiplicada pelos pesos w_{ij} e coletada no neurônio je dados pela equação (1) (Tucan & Saracti, 2021) (Saraçlı, 2008).

$$S_j = \sum_{i=1}^n w_{ij} u_j + w_{0j} \quad (\text{eq. 3})$$

onde w_{0j} , chama-se viés, e está incluído. Uma função de ativação é aplicada ao valor S_j dada pela equação (Dawson e Wilby, 1998). Esta função é dada por $f(x) = \frac{1}{1+e^{-x}}$ (Saraçlı, 2008).

3. RESULTADOS

3.1. ESTIMATIVA DA ENERGIA SOLAR GLOBAL

Com tudo a comparação entre a energia solar medida, a radiação predita usando o modelo RF apresenta-se também próxima da radiação original, seguida pela radiação estimada pelo modelo RK e por fim a ainda uma muito boa estimativa se pode mostrar com o modelo ANN.

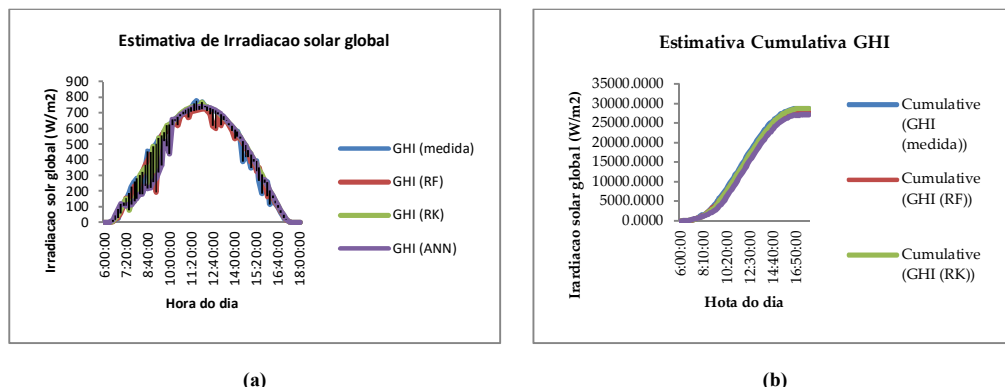


Figura 3 – Estimativa temporal de (a) irradiação Solar Global, (b) Irradiação solar global cumulativa

3.2. ESTIMATIVA CUMULATIVA GHI

Uma análise cumulativa de GHI, mostra uma aproximação dos valores preditos com um comportamento evolutivo dos parametros atmosféricos e meteorológicos que influenciam a radiação solar á chegada a superfície da terra. Contudo uma aproximação representado sempre melhor estimativa apresenta o modelo ANN estado sempre por debaixo da radiação medida por radiômetros tradicionais.

3.3. ESTIMATIVA DE INCREMENTOS DE GHI

O curso diário com que se descreve a GHI apresenta-se na *Figura 5*. O diapasão de todos modelos incluindo os valores medidos, apresenta incrementos dos valores de GHI no intervalo -150.3 a 169.8. contudo os valores estimado pelos métodos considerados encontra-se dentro da faixa de variação dos valores originais de incremento de GHI.

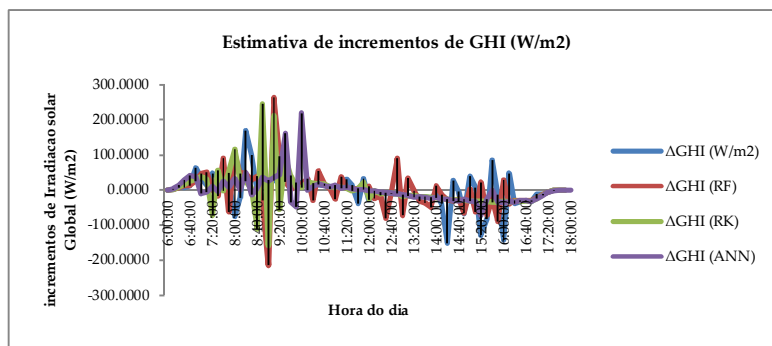


Figura 5 – Estimativa de incrementos de irradiacao solar global (W/m2)

Este indicador descreve um a boa predição, com maior proximidade da radiação estimada pelos modelos ANN, seguido por RF e por fim pelo RK. Este indicador auxilia na gestão do desempenho da saída do gerador FV, descrevendo predizendo se num futuro próximo, poderia energia solar ser afectado por incrementos através da passagem, da nuvem em um sentido de diminuição e ou acrescimo. Este indicativo pode e consideravelmente afectar o dimensionameto e longanimidade dos componentes do aparato solar. A medida que uma baixa de radiação por passagem de nuvem e outro influenciador reduzem consideravelmente a energia previamente estabelecida e pondo em causa o funcionamento deste esemble.

3.4. ESTIMATIVA DE TIPOS DE DIAS

A amostra de dados originalmente medidos tanto como os preditos possui uma distribuição normal como a maioria dos processos que se sucedem na natureza. Esta amostra foi classificada em classes enquadrada em quartis de medição. Para melhor gestão e economia de radiação solar foram classificados em dias nublado os dias com muito baixa radiação perfazendo-se no primeiro quadrante Q1 (isto e com valores $GHI < Q1$, para cada dia), em dias claros os dia com valor medio de radiação próximo a da constante solar ou muito elevado e situados no último quartil Q3 ($GHI > Q3$), acima de tudo foi classificada em dias céu intermediário os dias com características mistas de dias claros e nublado, isto é com valores

moderados de energia solar localizados entre o primeiro e o último quantil com tendência central para o valor médio (isto é $Q1 < GHI < Q3$).

A *Figura 6*, descreve que a classificação geral da amostra mostra com dados medidos in situ, mostra-se uma maior predominância de dias de céu

intermediário, sendo para os dias aceitáveis classificados em 24% de dias de céu claro, 24% de dias de céu claro e 52% de dias de céu claro, e a análise em termo de dias não aceitáveis mostrou-se 27% de dias de céu claro, 23% de dias de céu nublado e 50% de dias de céu intermédio ao longo de todo ano.

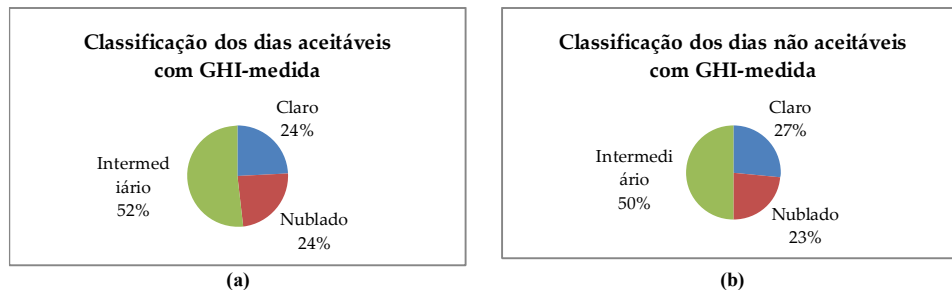


Figura 6 – Quantificação dos dias (a) dias aceitáveis e (b) dias não aceitáveis

Por outro lado aas margem de valores preditos, tanto em RF, RK e em ANN, mostram uma tendência central apresentando em dias aceitáveis 26% de dias de céu claro, 25% de dias de céu nuublado e 50% em dias de céu

intermédio, sendo para dias de não aceitáveis verificada uma margem de 29% de dia de cáu claro, 25% de dias de céu nublado e 46% de dias de céu intermedio.

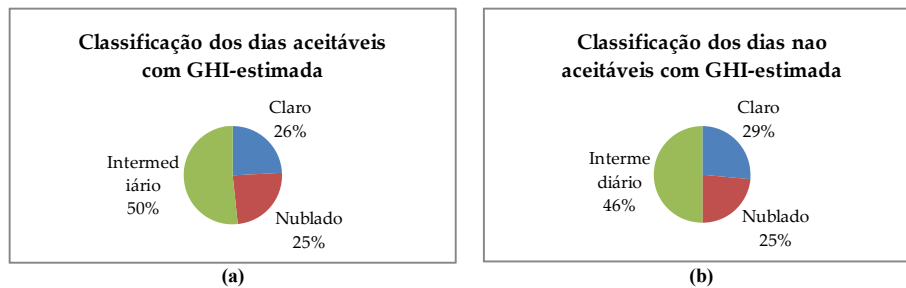


Figura 7 – Quantificação dos dias da amostra predita por RF (a) dias aceitáveis e (b) dias não aceitáveis

Entretanto a gestão apresentada propicia a gestão da energia solar como potência para aproveitamento em uma margem de mais de 81% e 80%, com dias de predominância de radiação solar plena durante a vigência do ano olhando a amostra actual assim do ponto de vista da amostra predita.

Contudo a gestão da energia predita mostra uma tendência central muito próxima aos valores anteriores, acompanhando paulatinamente a alteração de parâmetros meteorológicos, atmosféricos entre outros relatados que influenciam a energia solar a superfície da terra. A Tabela 2 a seguir apresenta a magnitude da dimensão do modelo mais adequado para estimativa e gestão da disponibilidade da energia solar.

Table 2 – Estimativa de erro de medição de GHI na estação de Pomene, Província de Inhambane em Moçambique

	Medida	Modelos		
		RF	RK	ANN
MAE	0.52	0.51	0.53	0.54
MAPE (%)	0.51	0.50	0.52	0.53
RMSE	0.45	0.42	0.59	0.58
R^2	0.50	0.49	0.51	0.53

O modelo RF, com uma estimativa de erro RMSE, em cerca de 0.42 comparativamente a outros modelos que exibem maior erro de estimativa, conforme se pode ver na *Figura 2*.

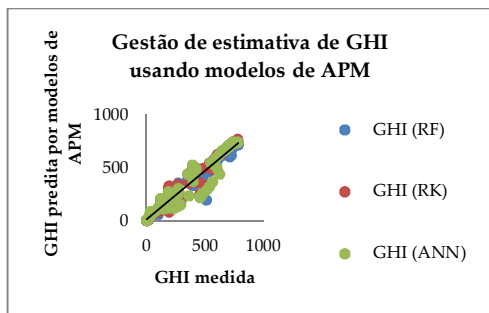


Figura 8 – Comparação entre a gestão de estimativa da disponibilidade de energia solar em curta escala de medição.

4. DISCUSSÃO

O cenário de viabilidade energética está diretamente relacionado com a modelação usando parâmetros que influenciam diretamente os valores estimados, para melhor conhecimento e gestão da frota de energia para produção FV, os valores preditos aqui mostram sempre uma margem de erro por RMSE não superior a 0.45 com destaque para o modelo RF e os outros dois modelos estudados, além de que a construção dos parâmetros introduzidos como a temperatura, variáveis atmosféricas, clima são determinísticos para a saída pois determinam os parâmetros das variáveis de modelos de APM. Contudo há a massificação de eficiência de gestão de energia predita havendo maior incidência de radiação solar em quase todo ano com maior destaque para dias de céu intermédio, estes resultados concordam com os resultados de validação cruzada em Leirvik, T., & Yuan, M. (2021) mostra que o RF tem um erro médio absoluto (MAE) de 10,2 Wm^{-2} e um desvio padrão de 1,5 Wm^{-2} , por outro lado, o MAE médio menor em relação aos métodos de interpolação convencionais, isso destaca os benefícios do uso do aprendizado de máquina na pesquisa ambiental, entre outros modelos aqui avaliados.

A gestão de disponibilidade de energia solar mostra maior predominância de fluxo de energia solar para aproveitamento solar, porém com maioria percentagem de dias de céu intermédio na estimativa de 50 e 52% para valores medidos e 48 e 50% para valores preditos, contudo esta predominância e uma boa estimativa a medida que exibe condições de pleno sol

para ótimo funcionamento, contudo há que referenciar que na maioria das regiões tropicais com mesma aproximações, ocorre o problema primordial de flutuações que provocam oscilações de energia solar como sendo os dias de céu intermediário os mais potências a flutuações de energias solar, uma combinação com a frota predita atendendo os parâmetros de conjugação de variáveis de calibração eliminam em grande margem as possibilidades de oscilação da saída em uma usina solar, resultados no mesmo contexto são exibidos em Mucomole et al., (2023) onde relata a variabilidade no Sul da região de Moçambique uma região adjacente e descreve-a com potenciais dias de céu intermediário, e em Mucomole et al., (2024) que mostra que a energia solar apresenta incrementos tendenciosos com as mudanças atmosféricas que se podem verificar no modelo predito de energia solar apresentado na *Figura 8*.

5. CONCLUSÕES

A gestão prévia e a estimativa de disponibilidade de energia solar e constituem as ferramentas-chaves para o ótimo funcionamento de diversos sistemas pre-estabelecidos para aproveitamento de energia solar, caso de aproveitamento solar fotovoltaico que é directamente influenciado pelo seu motor gerador de electricidade, radiação ou energia solar, no processamento e gestão da disponibilidade da energia solar, pode concluir-se que:

- ❖ A energia solar gerenciona na estação de Pomene apresenta-se maioritariamente com maior quantidade de plena radiação solar perfazendo-se com maiores dias

intermédios e claro, com percentagem de 52% e 50% para dias aceitáveis e não aceitáveis. Adicionalmente verifica-se 24 e 23% de dias nublados e, 24 e 27% de dias claros. Contudo há uma aproximação dos valores preditos perfazendo também maior percentagem de dias intermédios em 50 e 46% de dias intermédios em dias aceitáveis e não aceitáveis, contudo verificam-se 24 e 27% de dias claros e 23 e 27 de dias de céu nublado;

- ❖ A gestão de estimativa por modelo RF, apresenta menor erro de predição por RMSE, na margem de 0.45 em relação a outros modelos, isto pode influenciar em maior valia o dimensionamento prévio de sistemas solares conhecendo o seu real prévio processo de funcionamento;
- ❖ Os valores preditos decrescem com menor erro de exatidão da radiação futura para funcionamento de uma frota fotovoltaica, potenciando menores articulações que podem ser devastadoras para o ambiente e o ecossistema e contribuindo para maiores estratégias de eficiência de alcance de electrificação solar fotovoltaica, o que também pode reduzir consideravelmente o custo do watt-pico.

6. Tarefas futuras: A mesma extrapolação e gestão da estimativa da disponibilidade da energia solar pode ser alcançada usando outros modelos caso de ordinary kriging, multiregressão, entre outros que não foram possíveis apresentar aqui, principalmente por não constituírem objectivos desta pesquisa,

porém torna-se determinístico para o futuro uma comparação do erro de estimativa dos diversos métodos.

7. Patentes: Este artigo de investigação surge como uma das primeiras obrigações na obtenção do grau de Doutor no Curso de Doutoramento em Ciências e Tecnologia da Energia, da Universidade Eduardo Mondlane, que o autor principal está a frequentar.

8. Financiamento: Esta investigação é financiada pelo CS-OGET, da Faculdade de Engenharia da Universidade Eduardo Mondlane no âmbito da investigação de doutoramento.

Mathematics, 10(24), Artigo 24. <https://doi.org/10.3390/math10244746>
Sridharan, M. (2023). Generalized Regression Neural Network Model Based Estimation of Global Solar Energy Using Meteorological Parameters. *Annals of Data Science*, 10(4), 1107–1125. <https://doi.org/10.1007/s40745-020-00319-4>

REFERÊNCIAS BIBLIOGRÁFICAS

A Machine Learning Technique for Spatial Interpolation of Solar Radiation Observations. ([s.d.]). Recuperado 22 de março de 2024, de <https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2020EA001527>

Mucomole, F. V., Silva, C. A. S., & Magaia, L. L. (2023). Temporal Variability of Solar Energy Availability in the Conditions of the Southern Region of Mozambique. *American Journal of Energy and Natural Resources*, 2(1), Artigo 1. <https://doi.org/10.54536/ajenr.v2i1.1311>

Mucomole, F. V., Silva, C. S. A., & Magaia, L. L. (2024). Quantifying the Variability of Solar Energy Fluctuations at High-Frequencies through Short-Scale Measurements in the East-Channel of Mozambique Conditions. *American Journal of Energy and Natural Resources*, 3(1), Artigo 1. <https://doi.org/10.54536/ajenr.v3i1.2569>

Puga-Gil, D., Astray, G., Barreiro, E., Gálvez, J. F., & Mejuto, J. C. (2022). Global Solar Irradiation Modelling and Prediction Using Machine Learning Models for Their Potential Use in Renewable Energy Applications.

Mucomole, F. V.; Silva, C. A. S.; Magaia, L. L. (2025). *Gestão da estimativa de disponibilidade de energia solar utilizando modelos de aprendizado de máquina*