

LEONARDO ASSIS GASPAR

**O impacto da sujidade em módulos fotovoltaicos
no desempenho de usinas solares**

**Uberlândia, MG
2025**

LEONARDO ASSIS GASPAR

O impacto da sujidade em módulos solares para desempenho de usinas fotovoltaicas

Trabalho de Conclusão de Curso da
Engenharia de Controle e Automação da
Universidade Federal de Uberlândia - UFU -
Câmpus Santa Mônica, como requisito para
a obtenção do título de Graduação em
Engenharia de Controle e Automação

Universidade Federal de Uberlândia – UFU
Faculdade de Engenharia Elétrica - FEELT

Orientadora Prof.^a Dra. Gabriela Vieira Lima

Uberlândia, MG

2025

AGRADECIMENTOS

Gostaria de agradecer primeiramente aos meus pais, Heloisa e Rogério que sempre foram extremamente acolhedores e me deram apoio durante toda minha vida e em especial durante a graduação.

À minha noiva, Fernanda, que sempre esteve ao meu lado nos bons e maus momentos, oferecendo apoio emocional e sempre acreditando no meu potencial.

Aos amigos de graduação que me acompanharam nessa jornada e tornaram o dia a dia na faculdade mais leve.

À minha orientadora, Gabriela Vieira Lima, por ter me guiado de forma tão profissional e paciente durante todo o processo de idealização e produção deste trabalho.

Agradeço, por fim, à todos os colaboradores e demais alunos da Universidade Federal de Uberlândia com os quais eu tive contato durante minha formação nesta instituição. Acredito que nada é possível através do trabalho individual de ninguém, mas sim do esforço coletivo de todos que passam por este lugar.

RESUMO

A energia fotovoltaica tem se consolidado como uma das principais fontes renováveis no cenário energético mundial, destacando-se pela sustentabilidade e pela ampla aplicabilidade. Contudo, a eficiência dos módulos fotovoltaicos pode ser significativamente impactada por fatores ambientais, dentre os quais se destaca a sujidade. Neste sentido, este trabalho tem como objetivo analisar os efeitos da sujidade, como poeira, terra e resíduos orgânicos, sobre o desempenho de módulos em usinas fotovoltaicas, considerando as perdas na geração de energia elétrica e os desafios operacionais decorrentes desse fenômeno. Assim, foram analisados dados operacionais de 42 usinas, comparando-se a geração de energia nos 30 dias anteriores e posteriores ao procedimento de limpeza. Os resultados indicaram um ganho médio absoluto de 11,51% e um ganho médio relativo de 12,78% na geração, confirmando a relação entre a limpeza e a recuperação de desempenho. O estudo fornece subsídios técnicos para otimizar ciclos de manutenção e apoiar a tomada de decisão no setor fotovoltaico.

Palavras-chaves: Energia fotovoltaica; Sujidade; Eficiência, Manutenção preventiva.

ABSTRACT

Photovoltaic energy has consolidated itself as one of the main renewable sources in the global energy scenario, standing out for its sustainability and wide applicability. However, the efficiency of photovoltaic modules can be significantly impacted by environmental factors, among which soiling is a major concern. In this context, this study aims to analyze the effects of soiling, such as dust, dirt, and organic residues, on the performance of modules in photovoltaic plants, considering the losses in electricity generation and the operational challenges arising from this phenomenon. Thus, operational data from 42 power plants were analyzed, comparing energy generation in the 30 days before and after the cleaning procedure. The results indicated an average absolute gain of 11.51% and an average relative gain of 12.78% in generation, confirming the relationship between cleaning and performance recovery. The study provides technical subsidies to optimize maintenance cycles and support decision-making in the photovoltaic sector.

Keywords: Photovoltaic energy; Soiling; Efficiency; Preventive maintenance.

LISTA DE FIGURAS

Figura 1 - Efeito fotovoltaico.....	15
Figura 2 - Tipos de células fotovoltaicas.	15
Figura 3 - Diferença entre módulos limpos e sujos.	18
Figura 4 - Sujidade causando <i>hotspots</i> em módulo fotovoltaico.	19
Figura 5 - Limpeza automatizada de módulos fotovoltaicos.	20
Figura 6 - Diagrama do processo ETL.....	23
Figura 7 - Representação da conversão de arquivos Excel para DataFrames em memória.	24
Figura 8 - Fluxo de unificação e preparação da base de dados de desempenho.	25
Figura 9 - Processo de <i>Feature Engineering</i> para enriquecimento dos dados.....	26
Figura 10 - Divisão dos períodos de análise.	28
Figura 11 - Gráfico de dispersão da usina Capim Branco III.	30
Figura 12 - Gráfico <i>boxplot</i> da usina Capim Branco III.	31
Figura 13 - <i>Boxplot</i> das variações relativas de eficiência após limpeza.	35
Figura 14 - Dispersão da eficiência antes e depois da limpeza para a usina Araguari.	37
Figura 15 - <i>Boxplot</i> comparativo antes vs. depois da limpeza para a usina Araguari.	38
Figura 16 - Dispersão da eficiência antes e depois da limpeza para a usina Lagoa Formosa II... 39	
Figura 17 - <i>Boxplot</i> comparativo antes vs. depois da limpeza para a usina Lagoa Formosa II.... 40	
Figura 18 - Dispersão da eficiência antes e depois da limpeza para a usina Poconé III..... 41	
Figura 19 - <i>Boxplot</i> comparativo antes vs. depois da limpeza para a usina Poconé III..... 42	
Figura 20 - Resultado da análise preditiva para a usina Araguari.	43
Figura 21 - Resultado da análise preditiva para a usina Lagoa Formosa II..... 44	
Figura 22 - Resultado da análise preditiva para a usina Poconé III..... 45	

LISTA DE TABELAS

Tabela 1 - Resumo das estatísticas de ganho de eficiência após a limpeza.	34
--	----

SUMÁRIO

1	INTRODUÇÃO.....	11
1.1	JUSTIFICATIVAS	12
1.2	OBJETIVOS	13
1.2.1	<i>Objetivos Específicos</i>	<i>13</i>
2	REFERENCIAL TEÓRICO.....	14
2.1	MÓDULOS FOTOVOLTAICOS.....	14
2.1.1	<i>Células Fotovoltaicas</i>	<i>14</i>
2.2	IRRADIÂNCIA E IRRADIAÇÃO	16
2.3	EFICIÊNCIA DE GERAÇÃO	16
2.4	IMPACTO DA SUJIDADE NOS MÓDULOS FOTOVOLTAICOS	17
2.5	ESTRATÉGIAS DE LIMPEZA E MANUTENÇÃO	19
2.6	IMPORTÂNCIA DA MANUTENÇÃO PARA O SETOR FOTOVOLTAICO	21
3	METODOLOGIA.....	22
3.1	FERRAMENTA COMPUTACIONAIS UTILIZADAS	22
3.2	AQUISIÇÃO E DISPOSIÇÃO DE DADOS	22
3.3	TRATAMENTO E ORGANIZAÇÃO DE DADOS: UMA ABORDAGEM ETL	23
3.3.1	<i>Extração (Extract).....</i>	<i>23</i>
3.3.2	<i>Transformação (Transform)</i>	<i>24</i>
3.3.3	<i>Carga (Load)</i>	<i>27</i>
3.4	ANÁLISE ESTATÍSTICA E VALIDAÇÃO DE DADOS.....	27
3.4.1	<i>Definição da Janela de Análise</i>	<i>28</i>
3.4.2	<i>Cálculo das Métricas de Desempenho e Análise Agregada</i>	<i>29</i>
3.4.3	<i>Visualização de Dados.....</i>	<i>30</i>
3.4.4	<i>Validação Estatística com Teste de Hipótese</i>	<i>31</i>
3.4.5	<i>Análise Preditiva com Regressão Linear.....</i>	<i>32</i>
4	RESULTADOS	34
4.1	ANÁLISE AGREGADA DO IMPACTO DA LIMPEZA	34
4.2	VALIDAÇÃO ESTATÍSTICA DO RESULTADO	35
4.3	ANÁLISE DE DESEMPENHO POR USINA: ESTUDOS DE CASO	36
4.3.1	<i>Estudo de Caso I: Arguari – O Exemplo Típico</i>	<i>36</i>

4.3.2	<i>Estudo de Caso II: Lagoa Formosa 2 – O Outlier de Baixo Impacto</i>	38
4.3.3	<i>Estudo de Caso III: Poconé III – O Exemplo de Alta Variabilidade</i>	40
4.4	ANÁLISE PREDITIVA: DESAFIOS E LIMITAÇÕES DO MODELO	42
5	CONCLUSÃO	46
	REFERÊNCIAS BIBLIOGRÁFICAS	47

1 INTRODUÇÃO

O conceito de energia fotovoltaica tem suas origens no século XIX, quando o físico francês Alexandre Edmond Becquerel descobriu, em 1839, o efeito fotovoltaico. Contudo, foi somente em meados de 1950, com o desenvolvimento das primeiras células solares de silício, que a energia solar fotovoltaica começou a ser utilizada de forma mais prática. Inicialmente, a tecnologia foi aplicada em satélites espaciais, devido ao fato de ser uma solução eficiente e independente para a geração de energia. A partir da década de 1970, com a crise do petróleo, a energia fotovoltaica passou a ser vista como uma alternativa viável para a geração de energia renovável, e, desde então, o avanço tecnológico tem permitido que os sistemas fotovoltaicos se tornem cada vez mais eficientes e acessíveis, estabelecendo sua presença no mercado global de energia (PORTAL SOLAR, 2025b).

Em termos de energia sustentável, o Brasil é referência, de modo que em 2023 o país bateu a marca de 83,79% de sua matriz energética composta por fontes renováveis, sendo que a energia fotovoltaica corresponde a segunda maior fonte do território nacional (BRASIL, 2023; PORTAL SOLAR, 2025). Além disso, um dos principais facilitadores desse progresso em direção a um futuro mais limpo é a área denominada Cinturão Solar, que, localizada ao longo de 14 estados brasileiros, possui o maior índice de irradiação do país, o que a torna extremamente favorável à produção fotovoltaica (ALBA ENERGIA, 2023).

Neste contexto, o surgimento de usinas fotovoltaicas foi inevitável, de modo que já são mais de 18 mil construídas no Brasil. Dessa forma, essas usinas representam uma fonte lucrativa, oferecendo retornos financeiros consistentes a partir da venda de energia limpa. Em 2023, o Brasil atingiu um recorde histórico, adicionando mais de 12 GW de capacidade instalada, o que gerou investimentos superiores a R\$ 60 bilhões no setor (BRASIL, 2023). Para o consumidor comum, as usinas proporcionam uma alternativa sustentável e relativamente acessível, possibilitando reduções de até 95% nas contas de energia elétrica (ALBA ENERGIA, 2023). Esses fatores, fazem com que as usinas fotovoltaicas sejam um pilar essencial no avanço da transição energética global, promovendo benefícios ambientais e sociais, enquanto impulsionam o desenvolvimento econômico (PORTAL SOLAR, 2025).

Um componente fundamental para o funcionamento de uma planta solar são os módulos fotovoltaicos, que consistem no agrupamento de diversas células fotovoltaicas, que, por sua vez, são responsáveis por converter a radiação eletromagnética, proveniente da luz solar, em energia elétrica (CARNEIRO, 2010; FARIA, 2000). Para reduzir a proximidade com o solo em usinas fotovoltaicas do tipo terrestre e otimizar a captação da luz solar, os módulos são

instalados em estruturas chamadas mesas, que podem ficar em inclinações de máxima eficiência ou serem móveis, em caso de sistemas com rastreadores que acompanham a posição do Sol ao longo do dia.

Dessa forma, tendo em vista que o bom funcionamento dos módulos fotovoltaicos é essencial para a produção energética eficiente, o problema da sujeira, como poeira, terra e resíduos orgânicos, torna-se ainda mais relevante, pois seu acúmulo pode comprometer significativamente a operação das usinas fotovoltaicas. (SARVER; AL-QARAGHULI; KAZMERSKI, 2013). Assim, esse trabalho analisará como diferentes níveis de sujeira impactam a eficiência de usinas fotovoltaicas, assim como possíveis estratégias de mitigação desse fenômeno, de modo a contribuir para o desenvolvimento de soluções mais sustentáveis e confiáveis para utilização na operação dessas plantas energéticas.

1.1 Justificativas

A energia solar fotovoltaica tem se consolidado como uma das principais alternativas para a geração de energia limpa e renovável, contribuindo significativamente para a diversificação da matriz energética global. Tal transição se mostra cada vez mais viável, visto que o planeta Terra recebe cerca de $1,8 * 10^{11}$ MW de energia por segundo na forma de raios solares. No entanto, para que usinas fotovoltaicas operem com máxima eficiência, é essencial que os módulos fotovoltaicos estejam em condições ideais de funcionamento. Além de fatores como degradação, irradiância e até mesmo fumaça, um dos elementos que mais compromete o desempenho desses sistemas é o acúmulo de sujeira, incluindo poeira, poluição e matéria orgânica, que reduz a capacidade de absorção da luz solar e, consequentemente, a geração de energia (RASHID *et al.*, 2023).

Estudos indicam que, dependendo da região e das condições ambientais, a sujeira acumulada nos módulos pode causar perdas de eficiência de até 40%, sendo regiões áridas, como desertos e afins, as áreas mais afetadas. Essas perdas não afetam apenas a produtividade das usinas, mas também aumentam os custos operacionais, uma vez que a limpeza frequente dos módulos é necessária para manter a eficiência do sistema (SARVER; AL-QARAGHULI; KAZMERSKI, 2013).

Diante desse cenário, entender o impacto da sujeira no desempenho das usinas fotovoltaicas torna-se crucial para a otimização desses sistemas. Assim, este trabalho busca analisar como o acúmulo de sujeira afeta a eficiência dos módulos fotovoltaicos e propor soluções para mitigar esse problema. Além disso, a pesquisa visa fornecer embasamento para a

elaboração de estratégias de manutenção mais eficientes, reduzindo custos e maximizando a geração de energia.

Por fim, a relevância deste estudo também se justifica pelo crescimento acelerado da energia solar no Brasil e no mundo, tendo em vista que, com o aumento da demanda por fontes renováveis, é fundamental garantir que as usinas fotovoltaicas operem com o máximo de eficiência possível, minimizando perdas e garantindo a viabilidade econômica e ambiental desses projetos. Portanto, investigar o impacto da sujeira nos módulos fotovoltaicos não apenas contribui para o avanço do conhecimento científico, mas também tem aplicações práticas diretas no setor energético.

1.2 Objetivos

Este trabalho tem como objetivo analisar o impacto da sujeira acumulada nos módulos fotovoltaicos sobre a eficiência da geração de energia em usinas solares, quantificando as perdas de desempenho associadas por meio da análise de dados operacionais.

1.2.1 Objetivos Específicos

- Realizar o levantamento dos dados das usinas que serão utilizados no projeto.
- Preparar os dados para análise, de modo que os dados a serem utilizados estejam organizados de maneira padrão e centralizada em um único banco de dados.
- Implementar um script em *Python* para leitura e organização dos dados de eficiência de geração e datas de limpeza dos módulos fotovoltaicos, utilizando bibliotecas como Pandas e NumPy.
- Interpretar os resultados da análise estatística, verificando a existência de correlação entre os períodos sem limpeza e a variação na eficiência dos módulos.

2 REFERENCIAL TEÓRICO

Nesta seção, são apresentados os principais conceitos que embasam o tema deste trabalho, abordando desde o funcionamento dos módulos fotovoltaicos até os impactos da sujeira e as estratégias de manutenção aplicadas no setor.

2.1 Módulos Fotovoltaicos

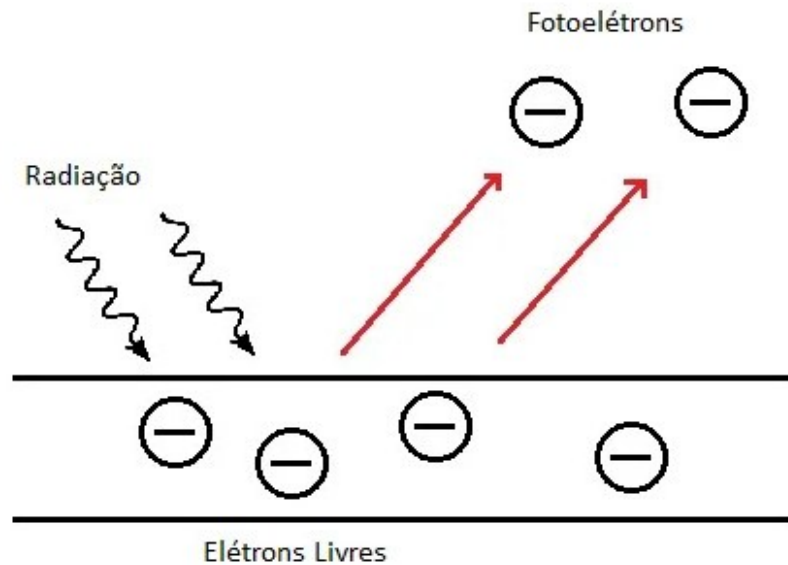
Os módulos fotovoltaicos são componentes essenciais nos sistemas de geração de energia solar, responsáveis por converter a radiação solar em eletricidade por meio do efeito fotovoltaico. Cada módulo é constituído por várias células fotovoltaicas encapsuladas em estruturas de vidro e polímeros, protegendo-as contra fatores ambientais como umidade, poeira e variações de temperatura (FERNANDES COELHO e VALDÉS SERRA, 2018).

A eficiência desses módulos depende de fatores como orientação e inclinação adequadas em relação ao Sol, ausência de sombreamento e manutenção apropriada. Dessa forma, em usinas fotovoltaicas, os módulos podem ser instalados em estruturas fixas ou em sistemas de rastreamento solar que acompanham a trajetória solar diária, otimizando a captação de luz, (ROSSI, LEITE e RAMPINELLI, 2014). Além disso, estratégias como a inclinação otimizada dos módulos auxiliam na autolimpeza, reduzindo o acúmulo de sujeira e aumentando a eficiência do sistema (OLIVEIRA *et al.*, 2022).

2.1.1 Células Fotovoltaicas

As células fotovoltaicas são os elementos fundamentais dos módulos solares e operam com base no efeito fotovoltaico, descoberto por Alexandre Edmond Becquerel em 1839. Esse efeito ocorre quando a luz solar incide sobre materiais semicondutores, geralmente silício, liberando elétrons e gerando uma corrente elétrica contínua (LIMA *et al.*, 2019). Uma representação simples e didática deste fenômeno pode ser vista na Figura 1.

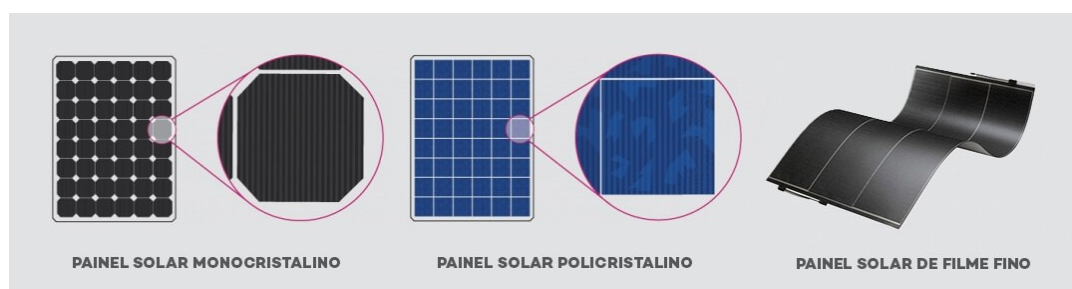
Figura 1 - Efeito fotovoltaico.



Fonte: Minha Casa Solar, 2025.

Além disso, existem diferentes formas de se fabricar uma célula fotovoltaica, como as células monocristalinas, que são fabricadas a partir de um único cristal de silício, apresentando alta eficiência energética devido à organização uniforme dos átomos, permitindo maior mobilidade dos elétrons. Em contrapartida, as células policristalinas são produzidas a partir de múltiplos cristais, tornando seu processo de fabricação mais econômico, embora sua eficiência seja ligeiramente inferior (FERNANDES COELHO e VALDÉS SERRA, 2018). Já as células de silício amorfo, pertencentes à categoria de filmes finos, possuem uma estrutura desorganizada, resultando em flexibilidade e menor custo, porém com menor rendimento energético em comparação às tecnologias cristalinas (GOÉS e TANIMOTO, 2021). Algumas dessas formas podem ser vistas na Figura 2 - Tipos de células fotovoltaicas.

Figura 2 - Tipos de células fotovoltaicas.



Fonte: Neosolar, 2025.

Com isso, a eficiência das células fotovoltaicas está diretamente relacionada a fatores

como a qualidade do material semicondutor, a temperatura de operação e a incidência de sujeira sobre sua superfície (OLIVEIRA *et al.*, 2022).

2.2 Irradiância e irradiação

Para que o efeito fotovoltaico aconteça, é necessário que exista irradiação e, consequentemente, irradiância. Tais conceitos se referem à quantidade de energia solar que atinge uma determinada superfície, mas diferem quanto à dimensão física envolvida (LIMA *et al.*, 2019).

Irradiância (medida em W/m^2) é uma grandeza instantânea que representa a potência da radiação solar recebida por unidade de área, ou seja, é a taxa de fluxo de energia solar que incide sobre uma superfície em um dado instante. Essa medida é particularmente importante para avaliar, em tempo real, as condições de insolação em uma usina fotovoltaica.

Já a irradiação (medida em Wh/m^2 ou kWh/m^2) refere-se à energia solar acumulada sobre uma superfície durante um intervalo de tempo específico, que pode ser uma hora, um mês ou até um ano. Assim, uma forma de obter-se a irradiação H em um determinado período, como um dia, é realizar a integral da irradiância $E(t)$ medida ao longo desse espaço de tempo, como mostra a equação 2.1 (RABELO, 2022).

$$H = \int E(t) dt \quad (2.1)$$

Dessa forma, a irradiação se consolida como a grandeza mais utilizada nas análises de desempenho de usinas fotovoltaicas, uma vez que permite avaliar a energia solar total disponível ao longo de um período específico. É com base na comparação entre a irradiação simulada, utilizada nas etapas de projeto e dimensionamento do sistema, e a irradiação realmente medida em campo que se torna possível julgar se a geração de energia da usina foi satisfatória ou não (RABELO, 2022).

2.3 Eficiência de geração

A eficiência de geração é uma métrica amplamente utilizada na avaliação do desempenho operacional de usinas fotovoltaicas. Ela permite quantificar o quão próxima a produção real de energia está daquilo que era esperado em condições ideais de projeto. A análise da eficiência torna-se ainda mais relevante ao considerar que, em condições reais, diversos fatores, como

acúmulo de sujeira nos módulos, falhas de equipamentos, sombreamentos e variações climáticas, impactam diretamente a geração de energia (DURISCH *et al.*, 2007).

Dessa forma, é possível calcular a eficiência de geração diária E por meio da equação 2.2.

$$E = \left(\frac{\text{Energia Gerada}}{\text{Irradiação Realizada}} \right) \times \left(\frac{\text{Irradiação Simulada}}{\text{Energia Simulada}} \right) \quad (2.2)$$

É fundamental destacar a natureza distinta das variáveis que compõem a Equação 2.2. Os termos Energia Simulada e Irradiação Simulada são oriundos de simulações computacionais, geralmente realizadas na fase de projeto da usina. Eles representam a expectativa de desempenho em condições ideais para um determinado mês e, portanto, são valores de referência que permanecem fixos para todos os dias dentro daquele mesmo mês. Em contrapartida, os termos Energia Gerada e Irradiação Realizada são as medições efetivas, coletadas diariamente pelos sensores da usina, refletindo a performance real da planta sob as condições operacionais e climáticas daquele dia específico. Dessa forma, a métrica E é, de fato, uma eficiência diária, pois compara o desempenho real de um dia com a expectativa de projeto, ao mesmo tempo que é normalizada pela insolação real daquele mesmo dia.

É importante ressaltar que, devido à sua natureza normalizada, esta métrica de eficiência pode, em certas condições, resultar em valores superiores a 100%. Isso não representa uma violação das leis da física, mas indica que a usina, em um determinado dia, operou com um desempenho superior ao previsto na simulação de referência. Fatores como temperaturas mais baixas que as previstas no modelo (o que aumenta a eficiência dos semicondutores), condições de céu mais limpo que a média histórica ou mesmo erros inerentes aos sensores de medição podem levar a esse 'superdesempenho' em relação à linha de base simulada.

2.4 Impacto da Sujidade nos Módulos Fotovoltaicos

A sujidade acumulada na superfície dos módulos fotovoltaicos é um fator que pode impactar negativamente o desempenho desses sistemas, pois a deposição de poeira, poluentes atmosféricos, resíduos orgânicos e outras partículas reduz a quantidade de radiação solar que atinge as células fotovoltaicas, diminuindo a eficiência na conversão de energia solar em elétrica. Estudos indicam que a sujidade pode causar perdas de eficiência significativas, variando conforme a localização geográfica, condições climáticas e ambientais (ARAÚJO, MARQUES DE CARVALHO e DUPONT, 2019).

Assim, a intensidade do impacto da sujeira depende de diversos fatores, incluindo a taxa de deposição de partículas, a inclinação dos módulos e a frequência de eventos naturais de limpeza, como chuvas. Sob essa ótica, em regiões áridas ou com baixa precipitação, a acumulação de sujeira tende a ser mais pronunciada, exigindo estratégias de manutenção mais rigorosas. Por exemplo, um estudo realizado em Fortaleza-CE avaliou os efeitos da sujeira no desempenho de uma planta fotovoltaica de 3,90 kWp ao longo de 13 meses, utilizando a métrica *Performance Ratio* (PR) ou índice de desempenho. Nesse estudo, os resultados demonstraram que a sujeira impacta diretamente o PR, evidenciando a necessidade de práticas de limpeza adequadas para manter a eficiência do sistema (ARAÚJO, MARQUES DE CARVALHO e DUPONT, 2019). Na Figura 3 - Diferença entre módulos limpos e sujos. é possível notar a nítida discrepância entre um módulo limpo e um módulo sujo.

Figura 3 - Diferença entre módulos limpos e sujos.



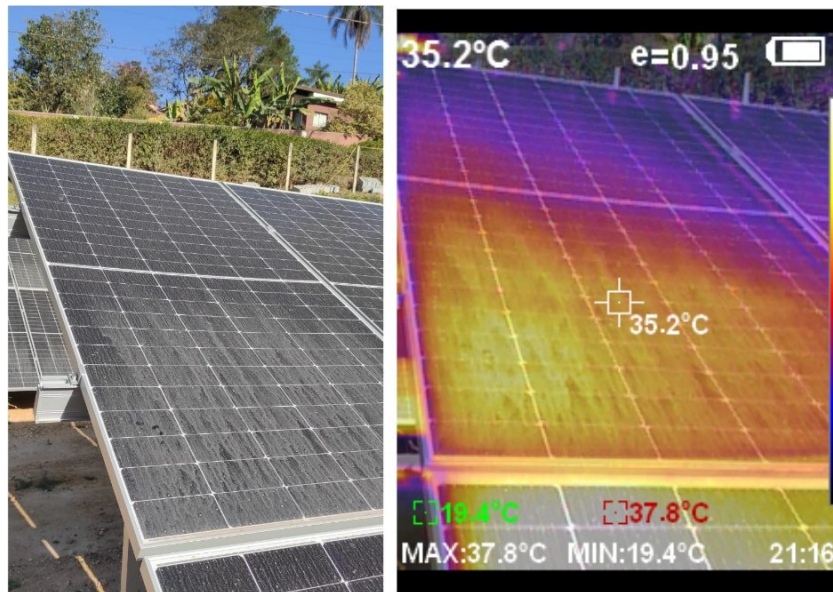
Fonte: PVClean, 2025.

Além da perda de eficiência, a sujeira pode provocar aquecimento desigual na superfície do módulo, um fenômeno conhecido como *hotspot* (ponto quente), que pode danificar permanentemente as células fotovoltaicas (RASHID *et al.*, 2023). Isso ocorre porque as áreas com maior concentração de detritos absorvem mais radiação térmica e dissipam menos calor, criando um diferencial de temperatura em relação às áreas mais limpas do mesmo painel.

A Figura 4 ilustra esse efeito através de uma imagem termográfica. É possível observar

que as áreas com maior acúmulo de sujeidade (manchas mais claras na imagem térmica) atingem temperaturas superiores a 35.2°C, enquanto áreas adjacentes no mesmo módulo, porém mais limpas, operam em temperaturas significativamente menores. Essa variação térmica em um mesmo componente é a causa direta do estresse que leva a falhas e à redução da vida útil do sistema.

Figura 4 - Sujidade causando *hotspots* em módulo fotovoltaico.



Fonte: Canal Solar, 2025.

Como é possível notar, em locais onde há maior acúmulo de partículas e detritos, a temperatura dos módulos tende a se elevar de forma significativa. Esse aumento térmico não é apenas um sintoma da sujeidade, mas também um fator que acelera a degradação dos materiais constituintes do módulo, reduz a eficiência de conversão e pode comprometer sua vida útil. Dessa forma, para mitigar tais impactos, torna-se essencial a implementação de estratégias de limpeza e manutenção regulares. A frequência e o método de limpeza devem ser adaptados às condições específicas de cada instalação, considerando fatores como localização, clima e tipo de sujeira predominante. Além disso, a adoção de práticas de manutenção adequadas assegura não apenas maior eficiência energética dos módulos, mas também contribui para sua durabilidade e para a viabilidade econômica dos sistemas fotovoltaicos (JUNIOR, 2021).

2.5 Estratégias de Limpeza e Manutenção

Diversas estratégias de limpeza e manutenção têm sido desenvolvidas com o objetivo de

mitigar os impactos da sujeira sobre os módulos fotovoltaicos, como a limpeza manual úmida que é amplamente utilizada em sistemas menores ou onde o acúmulo de poeira é moderado. Essa técnica consiste na aplicação de água e escovas macias, sendo eficaz para remover resíduos superficiais sem danificar a superfície dos módulos. Em contrapartida, regiões com escassez hídrica demandam alternativas mais sustentáveis, como a limpeza a seco, que utiliza ar comprimido ou panos eletrostáticos para a remoção de poeira (JUNIOR, 2021).

Avanços tecnológicos também vêm permitindo a automação desses processos, principalmente em grandes usinas solares. Sistemas automatizados de limpeza, como robôs autônomos, vêm sendo integrados em instalações de grande porte, otimizando a higienização constante dos módulos sem a necessidade de intervenção manual. Alguns desses sistemas não apenas removem a sujeira, mas também aproveitam a energia térmica captada pelos módulos para o aquecimento de água, agregando um valor adicional à operação (DE CAMPOS, 2017; RAMPINELLI; KRENZINGER; ROMERO, 2013). Na Figura 5, é possível observar um robô desenvolvido especificamente para a limpeza de módulos fotovoltaicos, ilustrando a aplicação prática dessa tecnologia.

Figura 5 - Limpeza automatizada de módulos fotovoltaicos.



Fonte: Canal Solar, 2025.

Além disso, a implementação de práticas preventivas de manutenção, como inspeções

regulares e monitoramento de falhas, é crucial para garantir o desempenho energético dos sistemas. Essas práticas auxiliam na identificação precoce de falhas, como *hotspots* ou microfissuras, que podem ser agravadas pela sujeira acumulada (CORREA; NASCIMENTO; NEVES, 2023).

A escolha da estratégia de limpeza deve considerar fatores específicos de cada instalação, como localização, tipo de sujidade predominante e condições climáticas. Em regiões desérticas ou com grande quantidade de poeira em suspensão, sistemas de ar comprimido ou robôs sem uso de água são mais indicados. Já em áreas industriais, onde a sujeira pode conter resíduos oleosos ou salinos, o uso de água pressurizada ou agentes de limpeza específicos pode ser mais eficaz (ARAÚJO; MARQUES DE CARVALHO; DUPONT, 2019; JUNIOR, 2021). Assim, a adoção de um plano de manutenção ajustado às condições ambientais é fundamental para a longevidade e viabilidade econômica dos sistemas fotovoltaicos.

2.6 Importância da Manutenção Para o Setor Fotovoltaico

A manutenção de sistemas fotovoltaicos desempenha um papel essencial para garantir que a geração de energia ocorra com eficiência ao longo da vida útil dos módulos. Em particular, a limpeza periódica dos painéis é fundamental para prevenir perdas de desempenho associadas à acumulação de poeira, fuligem e outras formas de sujidade. Essas partículas atuam como uma barreira à radiação solar, impedindo que a luz atinja as células fotovoltaicas de maneira plena e causando reduções significativas na eficiência dos módulos, comprometendo diretamente a produção energética da planta (ARAÚJO; MARQUES DE CARVALHO; DUPONT, 2019).

Além dos impactos diretos na eficiência energética, a ausência de manutenção adequada pode comprometer a durabilidade e a confiabilidade do sistema como um todo. Negligenciar a manutenção preventiva pode resultar em falhas permanentes nos módulos, além de danos aos demais componentes, como cabos e estruturas de suporte. Em sistemas de grande porte, como usinas solares, essas perdas são ainda mais significativas do ponto de vista econômico (DE CAMPOS, 2017). Estudos reforçam que a redução de desempenho causada pela sujeira pode atingir até 30%, dependendo das condições ambientais. Dessa forma, a implementação de estratégias de manutenção contínuas, adaptadas ao tipo de instalação, clima, local e tipo de sujidade predominante é indispensável para garantir o retorno do investimento e a sustentabilidade do sistema fotovoltaico ao longo do tempo (MANI; PILLAI, 2010).

3 METODOLOGIA

Este capítulo descreve os procedimentos adotados para analisar o impacto da sujidade acumulada nos módulos fotovoltaicos sobre a eficiência de geração de energia em usinas solares. A metodologia foi estruturada para quantificar as variações de desempenho antes e depois da execução de limpezas programadas, utilizando dados operacionais reais de várias usinas fotovoltaicas. Para isso, foram aplicadas etapas de extração, tratamento e organização dos dados, seguidas de análises estatísticas que permitiram avaliar a significância dos ganhos de eficiência observados.

3.1 Ferramenta computacionais utilizadas

A implementação da metodologia de análise de dados deste trabalho foi desenvolvida utilizando a linguagem de programação *Python*. A escolha desta tecnologia foi motivada por três fatores principais. Primeiramente, seu robusto ecossistema de bibliotecas de código aberto, que oferece ferramentas especializadas para cada etapa do projeto; com destaque para a biblioteca *Pandas*, utilizada para a manipulação e estruturação dos dados; *NumPy*, para operações numéricas; e as bibliotecas *Matplotlib* e *Seaborn*, para a geração de visualizações gráficas. Em segundo lugar, a sintaxe clara e legível do *Python* facilita o desenvolvimento e a manutenção do código. Por fim, a possibilidade de utilizar o ambiente de desenvolvimento em formato *Jupyter Notebook* (.ipynb) foi decisiva. Este formato permite um desenvolvimento iterativo e uma documentação integrada, onde células de código executável são intercaladas com texto formatado e visualizações. A capacidade de executar blocos de código de forma independente agiliza os ciclos de teste, depuração e exploração dos dados, garantindo a clareza e a reprodutibilidade da análise apresentada.

3.2 Aquisição e Disposição de dados

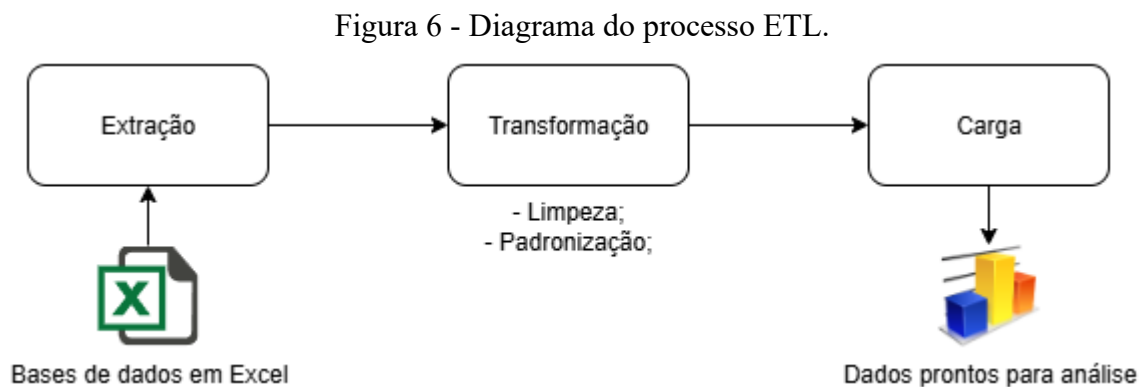
Os dados utilizados neste trabalho foram fornecidos pela empresa (re)energisa, com a devida autorização para uso acadêmico. Foram empregadas duas bases principais. A primeira corresponde ao banco de dados central da empresa, que reúne informações operacionais diárias das usinas, incluindo energia gerada, irradiação registrada, metas de geração, expectativas de irradiação realizada, eficiência de geração diária, entre outros indicadores relevantes para o monitoramento do desempenho. A segunda base utilizada foi o banco de ordens de serviço, no qual estão registradas as manutenções e intervenções realizadas nas usinas, com destaque para

as ordens de serviço referentes à limpeza dos módulos fotovoltaicos, contendo as datas em que cada atividade foi executada.

Essas duas fontes de informação foram essenciais para a análise proposta, uma vez que permitiram relacionar eventos de manutenção (limpeza de módulos) com o desempenho operacional das usinas, possibilitando a avaliação do impacto dessas intervenções sobre a eficiência de geração.

3.3 Tratamento e Organização de dados: Uma Abordagem ETL

Para garantir a qualidade, a consistência e a confiabilidade das análises, o processo de preparação dos dados deste trabalho foi estruturado seguindo o paradigma ETL (*Extract, Transform, Load*). Essa abordagem, amplamente consolidada na área de engenharia e ciência de dados, consiste em três fases distintas: a extração dos dados brutos de suas fontes, a transformação desses dados para limpá-los e adequá-los ao objetivo do estudo, e, por fim, o carregamento em um formato final, pronto para a análise estatística (SEENIVASAN, 2021). Cada uma dessas fases é mostrada no diagrama da Figura 6.



Fonte: Autoria própria.

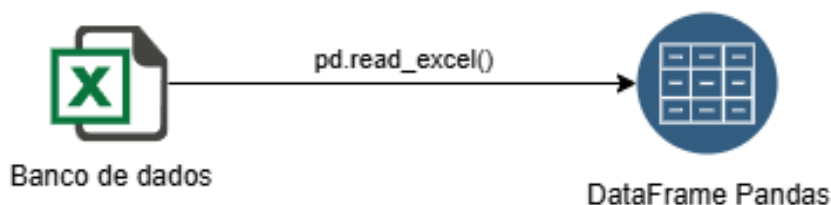
3.3.1 Extração (*Extract*)

A fase de Extração é o ponto de partida fundamental do processo ETL e seu principal objetivo é estabelecer a conexão com as fontes de dados originais para transferir as informações brutas para um ambiente de preparação, ou *staging area*. Neste trabalho, esse ambiente foi a memória do ambiente de análise em *Python*. O sucesso desta etapa é crucial, pois ela garante que a matéria-prima para todas as fases subsequentes seja uma cópia fiel e completa dos dados de origem, sem que nenhuma informação seja perdida ou corrompida no processo.

Um desafio comum em projetos de análise de dados, e presente neste estudo, é a necessidade de integrar informações de fontes heterogêneas, ou seja, sistemas que armazenam dados de natureza e formato distintos. Conforme detalhado na seção 3.1, foram utilizadas duas fontes principais: o banco de dados central da empresa, com informações operacionais quantitativas e em formato de série temporal (energia gerada, irradiação, entre outros), e o banco de ordens de serviço, com dados de manutenção de natureza mais descritiva e baseada em eventos (datas de limpeza). A integração de sistemas tão díspares é um dos principais motivadores para a adoção de um processo ETL estruturado.

Para a implementação técnica da Extração foi utilizada a biblioteca Pandas, que oferece estruturas de dados e funções de alto desempenho para manipulação de dados tabulares. Especificamente, a função `pd.read_excel()` foi empregada para realizar a leitura programática dos arquivos de planilha, abstraindo a complexidade do formato `.xlsx` e convertendo seu conteúdo em *DataFrames*, como demonstra a Figura 7.

Figura 7 - Representação da conversão de arquivos Excel para DataFrames em memória.



Fonte: Autoria própria.

Um *DataFrame* do Pandas é uma estrutura de dados bidimensional, análoga a uma tabela, que permite a manipulação eficiente de grandes volumes de dados em memória. Ao final desta fase, os dados brutos encontravam-se carregados nos *DataFrames*, preservando sua estrutura original e prontos para serem submetidos à rigorosa etapa de Transformação.

3.3.2 Transformação (Transform)

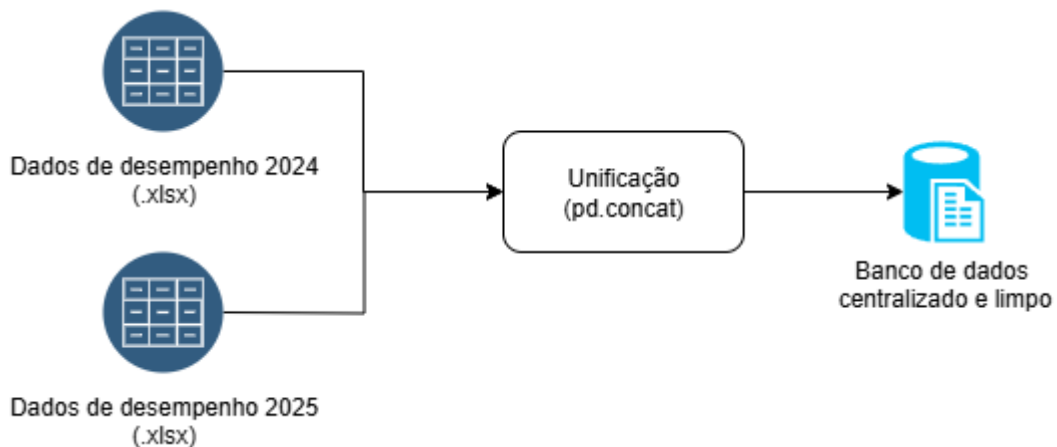
A etapa de Transformação é o núcleo do processo ETL, onde os dados brutos, extraídos na fase anterior, são submetidos a uma série de tratamentos para garantir sua qualidade, consistência e relevância para a análise. Esta fase é a mais complexa e envolve procedimentos de limpeza, padronização, enriquecimento de dados e filtragem, convertendo a matéria-prima informacional em um conjunto de dados coeso e pronto para a investigação estatística.

Nesse sentido, o primeiro passo da transformação foi a padronização e limpeza dos dados

para assegurar a uniformidade entre as diferentes fontes. Um requisito fundamental para análises temporais é a consistência no formato das datas; para isso, todas as colunas de data foram convertidas para o tipo `datetime` padrão do *Python*, utilizando a função `pd.to_datetime` da biblioteca *Pandas*. Similarmente, as colunas que deveriam conter valores numéricos, como 'Energia Gerada Diária', foram convertidas para o formato `float` através do comando `pd.to_numeric`. Notavelmente, foi utilizado o parâmetro `errors='coerce'`, uma estratégia robusta que transforma quaisquer entradas de texto ou erros de digitação em valores nulos (`NaN` - *Not a Number*), prevenindo interrupções no script.

Após o tratamento individual, os *DataFrames* de desempenho de 2024 e 2025 foram unificados em uma única base de dados centralizada através da função `pd.concat()`. Para garantir a integridade e a confiabilidade deste conjunto de dados, foram então removidas todas as linhas que continham valores ausentes ou fisicamente implausíveis na coluna de eficiência. O resultado desse processo, um banco de dados unificado e tratado, é ilustrado na Figura 8.

Figura 8 - Fluxo de unificação e preparação da base de dados de desempenho.



Fonte: Autoria própria.

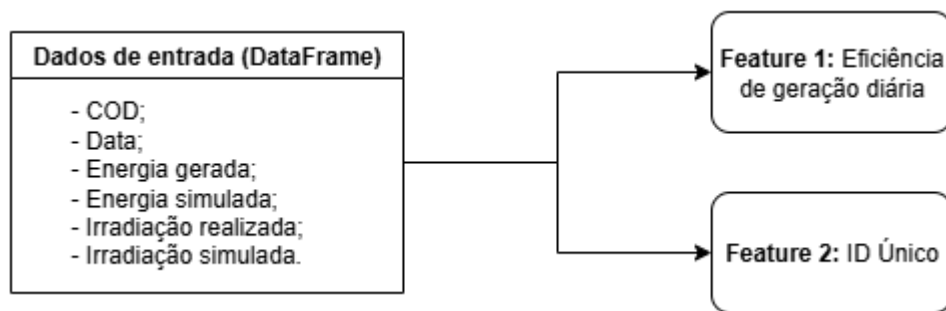
Por fim, todas as linhas que continham valores ausentes ou fisicamente implausíveis na coluna de eficiência foram removidas, garantindo a integridade e a confiabilidade do conjunto de dados final.

Superada a limpeza inicial, o processo seguiu para o enriquecimento dos dados através da *Feature Engineering*, uma etapa que consiste em criar novas variáveis de maior valor analítico a partir das existentes. A principal *feature* resultante foi a 'Eficiência de Geração diária', calculada a partir das colunas de energia e irradiação (conforme Equação 2.2). A criação desta variável foi o passo mais importante da transformação, pois ela normaliza o desempenho

da usina em relação às condições climáticas, permitindo uma comparação justa da performance ao longo do tempo e isolando o impacto de fatores operacionais, como a sujeidade.

Para permitir a integração entre os dados de desempenho e os registros de manutenção, foi criado um identificador único (ID). A base para esta chave foi o campo 'COD', um código alfanumérico de três caracteres já existente nas planilhas originais, onde cada usina do portfólio possui um 'COD' exclusivo que funciona como seu principal identificador em todas as fontes de dados. Este código foi então concatenado com a data da ocorrência, formatada textualmente no padrão AAAA-MM-DD, resultando em um ID único e legível para cada registro diário de cada usina. A importância do 'COD' como componente deste identificador é central para o projeto, pois ele é utilizado posteriormente para agregar todos os dados pertencentes a uma mesma planta, viabilizando a geração de gráficos individualizados e a consulta específica de suas datas de limpeza, como apresentado nos capítulos a seguir.

Figura 9 - Processo de *Feature Engineering* para enriquecimento dos dados.



Fonte: Autoria própria.

Por fim, para focar a análise exclusivamente no escopo do problema de pesquisa, foi realizada a filtragem e organização final do conjunto de dados de manutenção. Essa filtragem foi implementada através da função `.str.contains()` da biblioteca Pandas, que realizou uma busca textual na coluna 'Plano de Manutenção' por registros que contivessem a expressão "LIMPEZA MÓDULOS FOTOVOLTAICOS". Essa abordagem permitiu isolar de forma precisa e automatizada apenas os eventos de interesse para o estudo. Após a filtragem, o *DataFrame* resultante foi organizado com a função `.sort_values()`, que ordenou os dados primeiramente pelo 'COD' da usina e, em seguida, cronologicamente pela 'Data' do evento. Essa dupla ordenação foi essencial para estruturar os dados de forma lógica e sequencial, preparando o terreno para a fase final de Carga.

3.3.3 Carga (Load)

A fase de Carga é a etapa final do processo ETL, na qual os dados já transformados são movidos para seu destino final, onde serão consumidos para análise e geração de relatórios. Neste trabalho, a carga foi realizada em duas frentes principais: o carregamento das estruturas de dados finais em memória e a persistência dos resultados visuais em arquivos de imagem. Optou-se por apresentar esta seção de forma sucinta, uma vez que os artefatos de dados gerados nesta etapa são os insumos diretos para a fase de análise, e seu detalhamento técnico será mais bem contextualizado na seção subsequente, "Análise Estatística e Validação de Dados".

O principal produto carregado ao final do processo foi o *DataFrame* `banco_de_dados`, que consolidou o histórico de desempenho de todas as usinas em uma única base de dados unificada, limpa e enriquecida. Paralelamente, foi carregado o *DataFrame* `df_limpezas`, contendo a lista curada e organizada de todos os eventos de limpeza que serviram como gatilho para a análise comparativa. Essas duas estruturas de dados representam a entrega final do pipeline de preparação, servindo como a fonte de dados primária e confiável para a subsequente fase de análise estatística.

A segunda frente da etapa de Carga consistiu na persistência dos artefatos visuais gerados a partir dos dados tratados. No contexto deste trabalho, a carga também se refere ao ato de salvar os gráficos, que são uma forma de carregar as informações relevantes de um ambiente de memória volátil para um formato permanente. Para cada usina analisada, foram gerados gráficos de dispersão e *boxplots* comparativos para ilustrar o comportamento da eficiência antes e depois da limpeza. A persistência desses gráficos foi realizada através da função `plt.savefig()` da biblioteca *Matplotlib*, que exportou cada visualização para um arquivo de imagem no formato .png. Esse processo garantiu que os resultados visuais fossem salvos de forma sistemática e pudessem ser posteriormente incorporados e discutidos neste documento.

3.4 Análise Estatística e Validação de Dados

Com os dados devidamente tratados, organizados e carregados pelo processo ETL, a fase de análise estatística foi conduzida para quantificar objetivamente o impacto da limpeza dos módulos sobre a eficiência de geração das usinas. A abordagem foi dividida em cinco pilares metodológicos: a definição de uma janela de análise consistente para a comparação; o cálculo de métricas de desempenho para mensurar o impacto; a criação de visualizações de dados para a interpretação dos resultados; a validação estatística dos resultados por meio de teste de

hipótese; e, por fim, a tentativa de uma modelagem preditiva para o ciclo de limpeza. Antes de iniciar, foi determinado o escopo da amostra: a análise compreende um total de 42 usinas, totalizando 42 eventos de limpeza registrados e avaliados, ou seja, um evento para cada usina.

3.4.1 Definição da Janela de Análise

O pilar da análise consistiu na comparação do desempenho de cada usina em períodos anterior e posterior a um evento de limpeza. Para cada evento identificado no *DataFrame* `df_limpezas`, a data da intervenção foi definida como uma data-marco (D). A partir dela, foram estabelecidas duas janelas temporais simétricas de 30 dias: o período antes (D-30 a D-1) e o período depois (D+1 a D+30), conforme ilustra a Figura 10.

Figura 10 - Divisão dos períodos de análise.



Fonte: Autoria própria.

A escolha da janela de análise de 30 dias fundamenta-se na necessidade de reduzir os efeitos da variabilidade inerente às séries temporais de dados climáticos, especialmente aquelas relacionadas à irradiância solar. Tais séries apresentam oscilações significativas decorrentes de fatores como cobertura de nuvens, precipitação, variações sazonais e demais condições meteorológicas transitórias, que, quando consideradas em períodos curtos, podem distorcer a representatividade das médias (LOHMANN; MONAHAN, 2018).

Ao adotar um intervalo desta dimensão, minimiza-se a influência de eventos isolados e assegura-se uma base amostral mais robusta para a comparação entre os períodos. Essa abordagem também contribui para atenuar o impacto de ocorrências operacionais pontuais, como indisponibilidades temporárias de equipamentos, que poderiam comprometer a avaliação do desempenho real da usina. Ademais, a utilização de janelas temporais mais extensas tende a reduzir substancialmente a variabilidade estatística e aumentar a confiabilidade das estimativas obtidas, promovendo resultados mais consistentes e representativos da realidade operacional (LOHMANN; MONAHAN, 2018).

3.4.2 Cálculo das Métricas de Desempenho e Análise Agregada

Para cada evento de limpeza, a melhoria no desempenho foi quantificada individualmente para cada usina, antes de ser agregada. Utilizando um laço de repetição sobre os eventos em `df_limpezas`, o código filtrou o `banco_de_dados` para a janela de 60 dias de cada usina e calculou a eficiência média para o período antes ($\bar{E}^{Antes}$) e depois ($\bar{E}^{Depois}$), utilizando a função `.mean()` da biblioteca Pandas.

A partir dessas médias, foram calculadas duas métricas de impacto: a Variação Absoluta (Δ), que representa o ganho percentual direto, e a Variação Relativa ($\Delta\%$), que normaliza o ganho pela eficiência inicial da usina, conforme as equações 3.1 e 3.2, respectivamente. Os resultados de cada evento foram armazenados em um *DataFrame* final, o `df_resultados`, permitindo a análise agregada.

$$\Delta = \bar{E}^{Depois} - \bar{E}^{Antes} \quad (3.1)$$

$$\Delta\% = \frac{\bar{E}^{Depois} - \bar{E}^{Antes}}{\bar{E}^{Antes}} \quad (3.2)$$

A análise agregada dos resultados, considerando as métricas de todas as usinas no banco de dados, não se limitou ao cálculo de um único valor central, mas empregou um conjunto de estatísticas descritivas para fornecer uma compreensão mais completa do fenômeno. Nesse sentido, a média foi calculada para estabelecer o ganho de eficiência esperado, representando fisicamente o valor médio de melhoria que se pode antecipar ao realizar uma limpeza. Por ser a medida de tendência central mais comum, ela oferece um valor único que resume o efeito geral, embora seja sensível a valores extremos (*outliers*), que podem inflacionar ou subestimar o resultado.

Para contornar essa sensibilidade, a mediana foi utilizada como uma medida de tendência central mais robusta, representando o ganho de eficiência típico, ou seja, o valor que divide a amostra exatamente ao meio, fornecendo um retrato mais fiel do resultado mais frequente na prática. Complementarmente, para quantificar a dispersão dos resultados, calculou-se o desvio padrão. Um valor elevado para esta métrica, indica fisicamente que o impacto da limpeza não é uniforme entre as usinas, sugerindo que fatores locais, como tipo de sujeira e clima, influenciam significativamente o ganho de eficiência.

Em conjunto, essas três métricas oferecem um panorama completo: a média aponta o

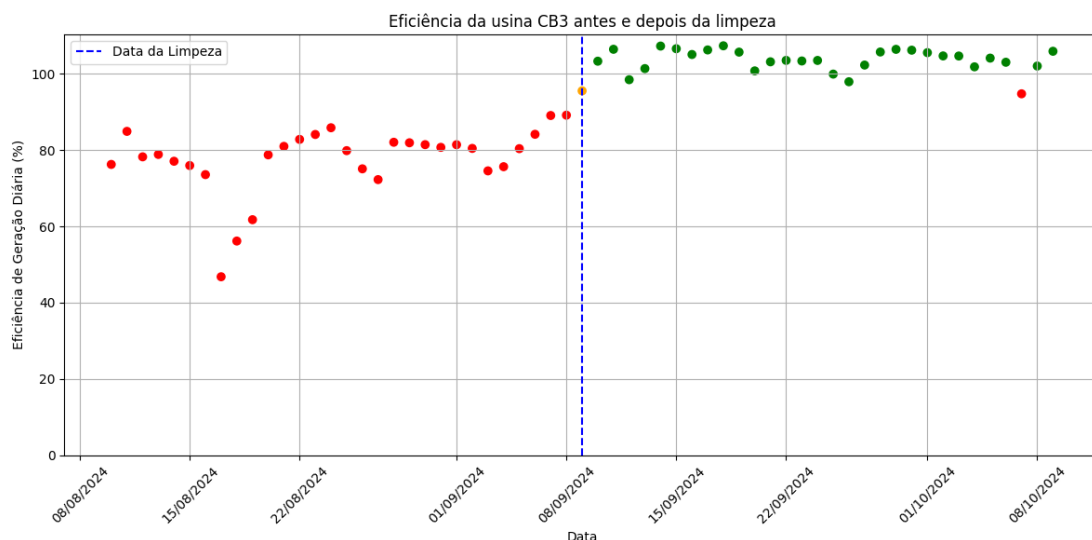
ganho esperado, a mediana confirma o ganho típico, e o desvio padrão revela a consistência (ou a falta dela) desse ganho entre os diferentes casos analisados.

3.4.3 Visualização de Dados

Para complementar a análise numérica e permitir uma interpretação mais intuitiva dos dados, foram desenvolvidas duas abordagens de visualização, implementadas com as bibliotecas *Matplotlib* e *Seaborn*. A escolha destes dois tipos de gráficos foi estratégica para oferecer visões complementares do fenômeno: uma cronológica e outra estatística.

A primeira abordagem consistiu na criação de gráficos de dispersão temporal para cada evento de limpeza. O objetivo deste gráfico é fornecer uma representação visual clara do comportamento da eficiência ao longo do tempo. Para sua construção, utilizou-se a função `plt.scatter()` para plotar a eficiência diária (eixo Y) em função da data (eixo X). Um elemento central para a análise é a inclusão de uma linha vertical, implementada com a função `plt.axvline()`, que marca a data exata da limpeza, permitindo uma inspeção visual imediata da mudança de comportamento entre os períodos antes e depois. Adicionalmente, os pontos foram codificados por cores com base em faixas de eficiência de geração pré-definidas (verde para $\geq 97\%$, laranja entre 95% e 97% , e vermelho para $< 95\%$), oferecendo um diagnóstico rápido da condição operacional da planta. A escolha deste tipo de gráfico se justifica por sua capacidade de mostrar a sequência temporal dos dados, permitindo a observação de tendências e variações diárias de forma intuitiva. A Figura 11 é um exemplo de um dos gráficos de dispersão gerados.

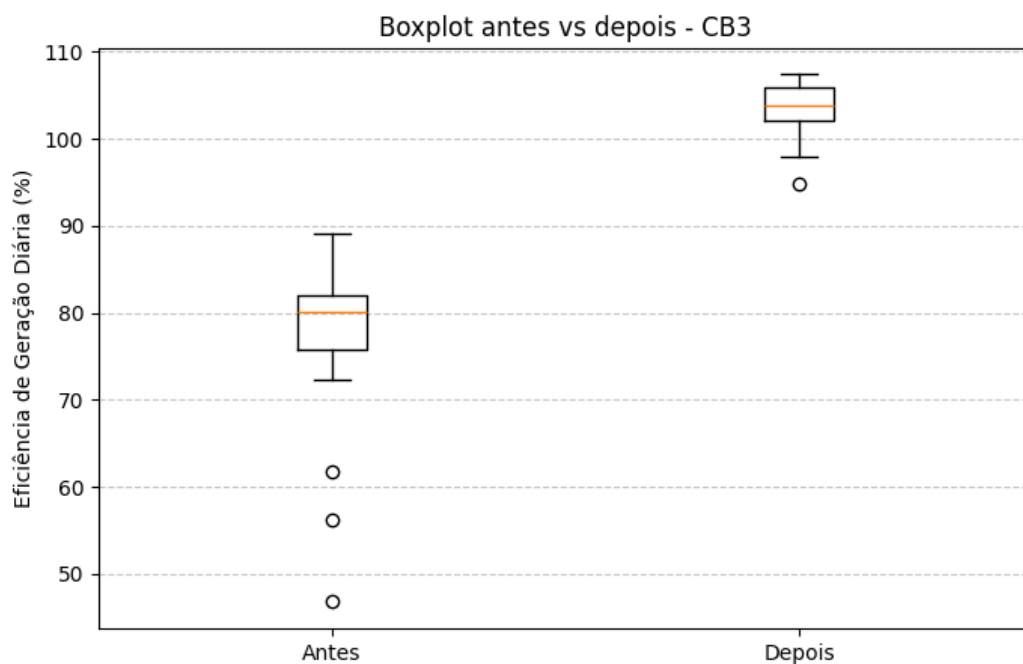
Figura 11 - Gráfico de dispersão da usina Capim Branco III.



Fonte: Autoria própria.

A segunda abordagem, de natureza estatística, utilizou *boxplots* comparativos. O propósito deste gráfico é contrastar de forma robusta a distribuição dos valores de eficiência dos períodos anterior e posterior à limpeza. Para sua geração com a função `plt.boxplot()`, os dados de eficiência de cada janela de 30 dias foram primeiro segregados em dois conjuntos de dados distintos. A escolha do *boxplot* é estratégica pois, diferentemente do gráfico de dispersão que mostra cada ponto individualmente, ele resume a distribuição estatística de cada período. Isso permite uma comparação rigorosa de indicadores-chave como a mediana (tendência central), a dispersão dos dados (representada pela altura da caixa, ou intervalo interquartil) e a identificação de valores atípicos (*outliers*). Dessa forma, enquanto o gráfico de dispersão oferece uma visão intuitiva da cronologia, o *boxplot* fornece a síntese estatística, oferecendo em conjunto uma análise visual completa e aprofundada do impacto da limpeza. A Figura 12 é um exemplo de um dos gráficos de dispersão gerados.

Figura 12 - Gráfico *boxplot* da usina Capim Branco III.



Fonte: Autoria própria.

3.4.4 Validação Estatística com Teste de Hipótese

Embora a análise descritiva apresente uma tendência clara de melhoria na eficiência, é fundamental, do ponto de vista científico, verificar se essa diferença é estatisticamente significativa ou se poderia ter ocorrido por mera variação aleatória dos dados. Para conferir tal validade estatística, foi realizado um teste de hipótese, uma ferramenta de inferência que

permite tomar decisões baseadas em probabilidades.

O teste foi estruturado em torno de duas hipóteses concorrentes. A primeira é a hipótese nula (H_0), que representa a premissa de que não há efeito ou mudança. No contexto deste trabalho, ela postula que a limpeza dos módulos não possui um efeito real, e que a diferença média de eficiência observada entre os períodos antes e depois é estatisticamente igual a zero. Em contrapartida, a hipótese alternativa (H_1) postula que o efeito observado é real, ou seja, que a diferença média de eficiência entre os dois períodos é significativamente diferente de zero.

O método específico escolhido foi o Teste t Pareado, e essa escolha é crucial para a natureza do estudo. Diferentemente de testes para amostras independentes, o teste t pareado é utilizado quando as medições são feitas sobre os mesmos sujeitos em condições distintas. Neste caso, os sujeitos são as mesmas usinas, avaliadas em dois momentos: antes e depois da limpeza. Essa abordagem pareada é mais poderosa, pois controla a variabilidade inerente a cada usina e foca exclusivamente no impacto da intervenção (a limpeza) para cada caso.

O mecanismo do teste opera sobre as diferenças individuais calculadas para cada par de observações ($\text{Eficiência_Depois} - \text{Eficiência_Antes}$). A partir do conjunto dessas diferenças, o teste calcula uma estatística (a estatística t) que mede o quão distante da hipótese nula (diferença média de zero) os dados da amostra se encontram. O resultado principal do teste é o p-valor, que representa a probabilidade de se obter uma diferença média amostral tão ou mais extrema que a observada, assumindo que a hipótese nula seja verdadeira. Em termos práticos, ele responde à pergunta: "Se a limpeza não fizesse nada, qual seria a chance de, por puro acaso, eu encontrar o resultado que encontrei?".

Para a tomada de decisão, foi adotado um nível de significância (α) padrão de 0,05. A regra de decisão é a seguinte: se o p-valor calculado for inferior a 0,05, a hipótese nula é rejeitada, concluindo-se que o efeito da limpeza é estatisticamente significativo. Caso contrário, não haveria evidências suficientes para rejeitar a hipótese nula. A implementação do teste foi realizada com a função `ttest_rel` da biblioteca `scipy.stats` em *Python*.

3.4.5 Análise Preditiva com Regressão Linear

Como etapa final e exploratória da análise, buscou-se desenvolver um modelo preditivo para estimar o ciclo de manutenção ótimo para uma usina. O objetivo era modelar a taxa de perda de eficiência devido à sujidade e, a partir dela, prever em quanto tempo a usina retornaria ao seu patamar de desempenho pré-limpeza, indicando a necessidade de uma nova intervenção. Para isso, adotou-se uma abordagem de estudo de caso, focando em usinas específicas para uma

análise mais aprofundada.

A preparação dos dados para o modelo iniciou-se com o isolamento dos registros de eficiência de uma usina para um período estendido de 90 dias subsequentes à data da limpeza. Para transformar a série temporal em uma variável quantitativa adequada para a modelagem, foi realizada uma etapa de *Feature Engineering*, na qual foi criada a variável `dias_apos_limpeza`. Essa *feature*, que serviu como a variável independente (eixo X) do modelo, representa a contagem de dias desde a data da manutenção para cada registro. A Eficiência de Geração diária foi definida como a variável dependente (eixo Y). O desempenho de referência, ou seja, a eficiência que o modelo buscava prever o retorno, foi estabelecido como a eficiência média da usina nos 30 dias anteriores à limpeza, valor já calculado na análise descritiva.

A modelagem da relação entre as variáveis foi realizada por meio de uma Regressão Linear Simples, implementada com a classe `LinearRegression` da biblioteca *Scikit-learn* em *Python*. Este modelo estatístico busca encontrar a reta que melhor se ajusta à dispersão dos dados, descrevendo a tendência linear entre o tempo transcorrido e a eficiência. O treinamento do modelo foi executado pela função `.fit()`, que calcula os parâmetros da reta a partir dos dados de entrada (X e Y).

Após o treinamento, o modelo forneceu os dois parâmetros fundamentais da equação da reta (equação 3.3). O coeficiente angular (m), interpretado neste contexto como a taxa de sujidade, que representa a perda média de eficiência por dia. O intercepto (b) representa a eficiência inicial teórica, imediatamente após a limpeza. Com estes parâmetros, o tempo (em dias) para a eficiência retornar ao patamar pré-limpeza foi então calculado algebricamente, isolando a variável independente na equação da reta.

$$Y = m.X + b \quad (3.3)$$

4 RESULTADOS

Neste capítulo, são apresentados os resultados obtidos a partir da aplicação da metodologia de análise de dados descrita anteriormente. A exposição dos achados seguirá uma ordem lógica, partindo da análise agregada do impacto da limpeza para o conjunto total de usinas, passando pela validação estatística desses resultados e, por fim, aprofundando em estudos de caso individuais para ilustrar a variabilidade do fenômeno.

4.1 Análise Agregada do Impacto da Limpeza

A análise estatística realizada para todas as 42 usinas com registros de limpeza permitiu quantificar o ganho de eficiência associado a esse procedimento, conforme apresentado na Tabela 1.

Tabela 1 - Resumo das estatísticas de ganho de eficiência após a limpeza.

Métrica estatística	Variação absoluta	Variação relativa
Média	11,51%	12,78%
Mediana	10,18%	11,24%
Desvio padrão	7,39%	9,34%

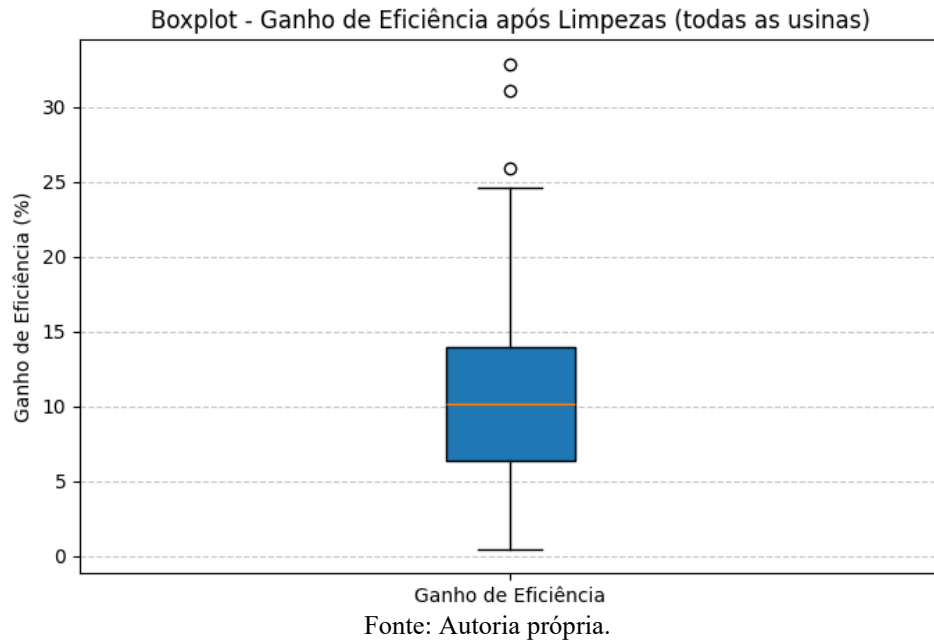
Fonte: Autoria própria.

A diferença absoluta média de eficiência após a limpeza foi de 11,51%, enquanto a mediana ficou em 10,18%. O desvio padrão das diferenças absolutas foi de 7,39%, indicando uma variação significativa nos resultados entre as diferentes usinas analisadas. No caso da diferença relativa, a média foi de 12,78%, com mediana de 11,24% e desvio padrão de 9,34%, sugerindo que, proporcionalmente, algumas usinas apresentaram ganhos bastante expressivos, enquanto outras tiveram aumentos mais modestos.

Para ilustrar a distribuição desses ganhos de eficiência, a Figura 13 apresenta o *boxplot* das variações relativas para todos os eventos de limpeza analisados. Observa-se que a maior parte dos resultados, representada pela caixa do gráfico (o intervalo interquartil), se concentra no intervalo de aproximadamente 5% a 15%, o que indica um padrão consistente de melhoria após a remoção da sujeira. A mediana, situada próxima a 11%, reforça que, para a maioria das usinas, os ganhos com a limpeza foram expressivos e dentro de uma faixa relativamente

previsível.

Figura 13 - *Boxplot* das variações relativas de eficiência após limpeza.



No entanto, o gráfico também evidencia a presença de valores atípicos (*outliers*), com ganhos superiores a 25% e outros próximos de zero. Esses casos extremos indicam situações operacionais distintas do padrão. Os ganhos muito elevados, por exemplo, provavelmente estão associados a usinas que apresentavam níveis de sujeira significativamente mais altos, seja por um intervalo maior entre manutenções ou por condições ambientais mais severas, como a escassez de chuvas. Por outro lado, as usinas com ganhos próximos de zero provavelmente já se encontravam em um estado relativamente limpo antes da intervenção, tornando o impacto da manutenção menos acentuado.

4.2 Validação Estatística do Resultado

A análise descritiva, que apontou um ganho médio absoluto de 11,51% na eficiência, foi corroborada por um Teste t Pareado para verificar se o efeito observado era real e não fruto de variações aleatórias. O resultado do teste indicou uma estatística t de 10.0959 e um p-valor de 0.0000000000.

O p-valor representa a probabilidade de se obter um ganho de eficiência tão ou mais expressivo que o encontrado caso a limpeza não tivesse nenhum efeito real (a hipótese nula). O valor obtido, embora exibido como zero, representa na verdade uma probabilidade

extremamente pequena, tão próxima de zero que, para fins práticos de formatação, é arredondada. Isso significa que a chance de o ganho de eficiência observado ser uma mera coincidência é praticamente nula.

Sendo o p-valor, portanto, significativamente inferior ao nível de significância de 0,05 estabelecido na metodologia, a hipótese nula foi rejeitada. Este resultado permite afirmar, com um nível de confiança de 95%, que o aumento na eficiência de geração não é uma ocorrência casual, mas sim um efeito estatisticamente significativo diretamente atribuível ao procedimento de limpeza dos módulos fotovoltaicos. A validação estatística confere, portanto, um alto grau de confiabilidade aos resultados apresentados, confirmando que a manutenção é um fator causal na melhoria do desempenho das usinas.

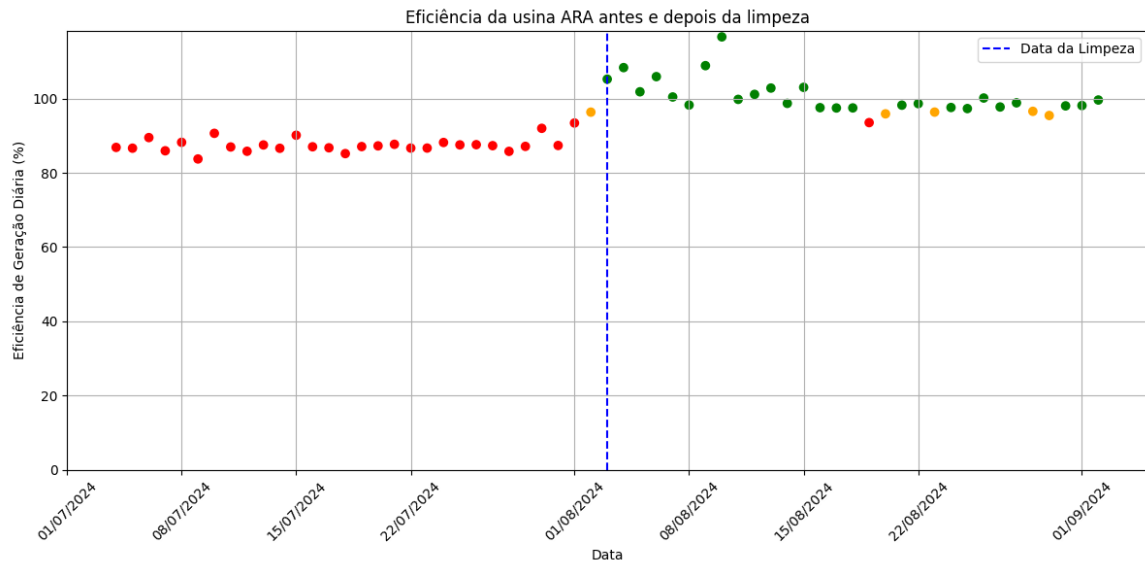
4.3 Análise de Desempenho por Usina: Estudos de Caso

Após a análise agregada, foram investigados casos individuais para ilustrar os diferentes perfis de comportamento observados. A geração de gráficos de dispersão e *boxplots* para cada usina permitiu uma análise aprofundada, revelando desde casos que exemplificam perfeitamente a média de ganhos até *outliers* que ajudam a compreender a variabilidade dos resultados. A seguir, são apresentados três estudos de caso representativos.

4.3.1 Estudo de Caso I: Araguari – O Exemplo Típico

A usina Araguari (ARA) serve como um exemplo representativo do impacto positivo e consistente da limpeza. O gráfico de dispersão temporal (Figura 14) demonstra visualmente a transição de desempenho. No período anterior à manutenção, a eficiência da usina operava de forma constante, porém em um patamar inferior, com a grande maioria dos pontos indicados em vermelho (eficiência $< 95\%$). Após a data da limpeza, demarcada pela linha vertical, observa-se um salto imediato e sustentado no desempenho, com os pontos passando para a cor verde (eficiência $\geq 97\%$), indicando uma recuperação clara da capacidade de geração.

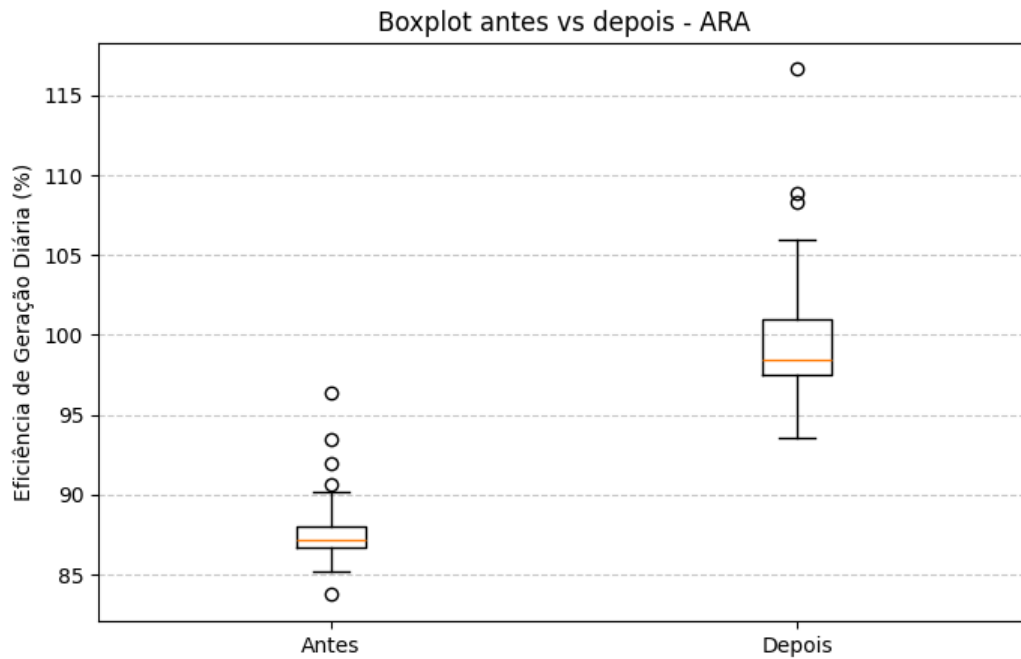
Figura 14 - Dispersão da eficiência antes e depois da limpeza para a usina Araguari.



Fonte: Autoria própria.

A análise estatística dessa transição, apresentada no *boxplot* comparativo (Figura 15), corrobora a observação visual. A distribuição antes possui uma mediana próxima a 87%, enquanto a distribuição depois exibe um deslocamento acentuado para cima, com uma nova mediana de aproximadamente 98%. É notável que a totalidade da caixa do período depois (o intervalo interquartil) se encontra acima da mediana do período antes, reforçando a magnitude da melhoria. O comportamento da usina Araguari, portanto, alinha-se perfeitamente com os resultados médios encontrados para o conjunto de dados.

Figura 15 - *Boxplot* comparativo antes vs. depois da limpeza para a usina Araguari.

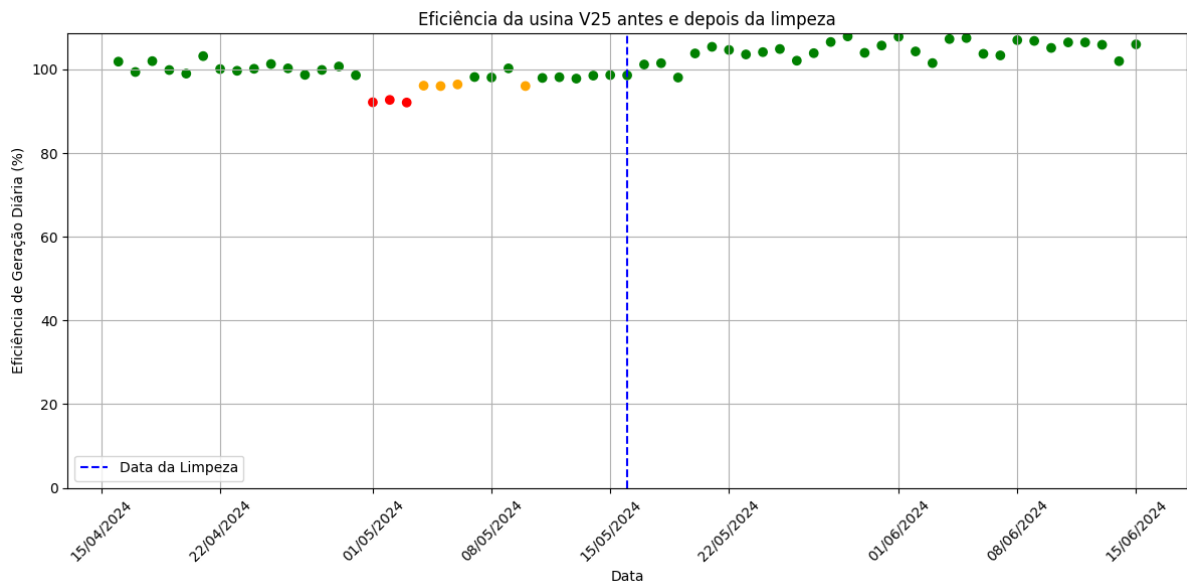


Fonte: Autoria própria.

4.3.2 Estudo de Caso II: Lagoa Formosa 2 – O Outlier de Baixo Impacto

A usina Lagoa Formosa II (V25) exemplifica um dos casos *outliers* identificados na análise agregada, onde o ganho de eficiência foi mínimo. O gráfico de dispersão (Figura 16) revela o porquê: a usina já operava com altíssima eficiência no período anterior à limpeza, com quase todos os pontos na faixa verde. Embora haja uma pequena queda de desempenho nos dias que antecedem a manutenção, não havia um déficit de geração significativo a ser recuperado.

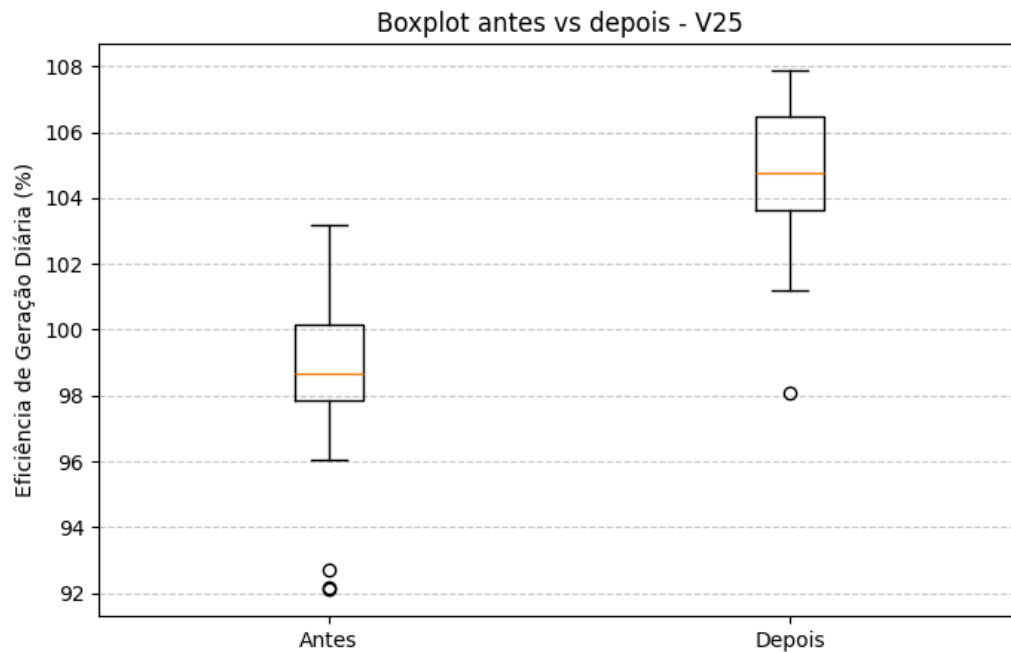
Figura 16 - Dispersão da eficiência antes e depois da limpeza para a usina Lagoa Formosa II.



Fonte: Autoria própria.

O *boxplot* comparativo (Figura 17) quantifica essa observação. Embora exista um leve aumento na mediana após a limpeza, as distribuições antes e depois se sobrepõem consideravelmente. Este cenário sugere que a limpeza foi realizada de forma preventiva ou em uma planta com baixo acúmulo de sujeira, possivelmente devido a chuvas recentes. Este caso é importante por ilustrar que o benefício da limpeza é condicional ao estado inicial do sistema, explicando os ganhos próximos de zero vistos na análise agregada.

Figura 17 - *Boxplot* comparativo antes vs. depois da limpeza para a usina Lagoa Formosa II.

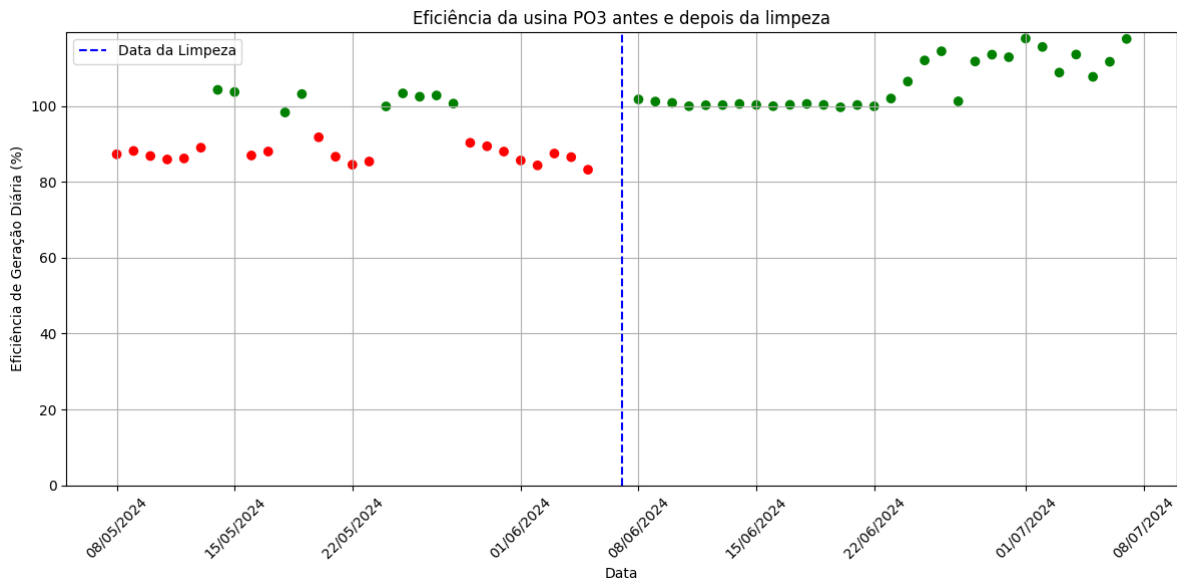


Fonte: Autoria própria.

4.3.3 Estudo de Caso III: Poconé III – O Exemplo de Alta Variabilidade

A usina Poconé III (PO3) foi selecionada para demonstrar a eficácia da análise mesmo em cenários de alta variabilidade operacional. O gráfico de dispersão (Figura 18) mostra uma oscilação diária de eficiência muito maior que a da usina Araguari, tanto antes quanto depois da limpeza.

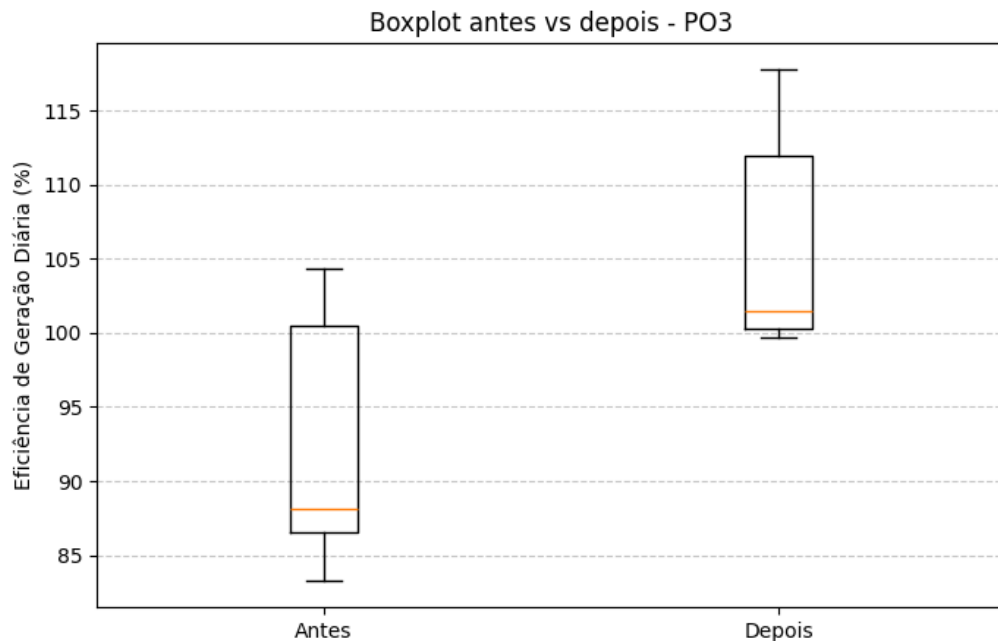
Figura 18 - Dispersão da eficiência antes e depois da limpeza para a usina Poconé III.



Fonte: Autoria própria.

Apesar dessa instabilidade, o *boxplot* comparativo (Figura 19) consegue extrair um sinal claro de melhoria. A caixa antes é visualmente "alta", confirmando a grande dispersão nos dados, com uma mediana em torno de 88%. A caixa depois, embora também demonstre variabilidade, está posicionada em um patamar de eficiência muito superior, com uma mediana acima de 100%, indicando um superdesempenho em relação à simulação de referência, conforme justificado na seção 2.3. O fato de o ponto mais baixo da distribuição depois ser superior à mediana da distribuição antes prova que, estatisticamente, houve um ganho inequívoco. Este caso evidencia a robustez do método de análise por *boxplot*, que é capaz de confirmar o impacto positivo da manutenção mesmo quando os dados diários são ruidosos.

Figura 19 - *Boxplot* comparativo antes vs. depois da limpeza para a usina Poconé III.



Fonte: Autoria própria.

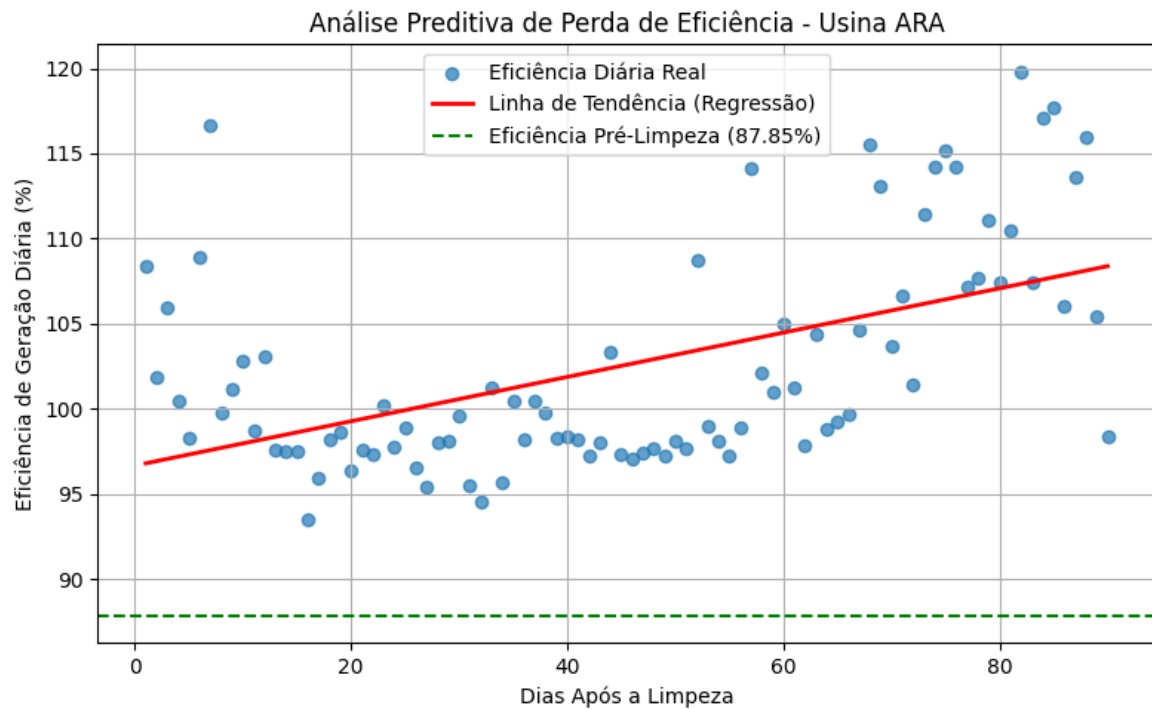
4.4 Análise Preditiva: Desafios e Limitações do Modelo

Como desdobramento da análise de ganhos, buscou-se desenvolver um modelo preditivo para estimar o ciclo de manutenção ótimo, conforme a metodologia descrita na seção 3.4.5. A abordagem consistiu em aplicar uma regressão linear simples para modelar a queda de eficiência em função dos dias transcorridos após a limpeza. Assim como na seção 4.3, foram selecionadas as usinas Araguari, Lagoa Formosa II e Poconé III, três usinas com diferentes perfis, para um estudo de caso aprofundado. Os resultados, no entanto, revelaram a alta complexidade do fenômeno, e a ineficácia do modelo simples para esta tarefa constituiu, por si só, um resultado relevante.

O primeiro estudo de caso foi a usina Araguari. Conforme apresentado na Figura 20, a aplicação do modelo de regressão linear resultou em uma inesperada tendência positiva de eficiência, subindo ao invés de cair após a limpeza. Este resultado, fisicamente inconsistente com o efeito isolado da sujeira, expõe uma limitação fundamental do modelo de regressão linear simples para este problema: sua natureza univariada. O modelo assume que a passagem do tempo é a única variável a influenciar o desempenho, quando na realidade, a eficiência é um fenômeno influenciado por múltiplas variáveis. Fatores como variações diárias de irradiação, temperatura ambiente (que afeta diretamente a performance dos módulos) e a sazonalidade

(como a ocorrência de chuvas) impactam o resultado simultaneamente. Portanto, a tendência positiva observada sugere que o efeito combinado dessas outras variáveis foi mais forte do que o impacto negativo do acúmulo de sujeira no período analisado, tornando a previsão do ciclo de limpeza inviável com esta abordagem.

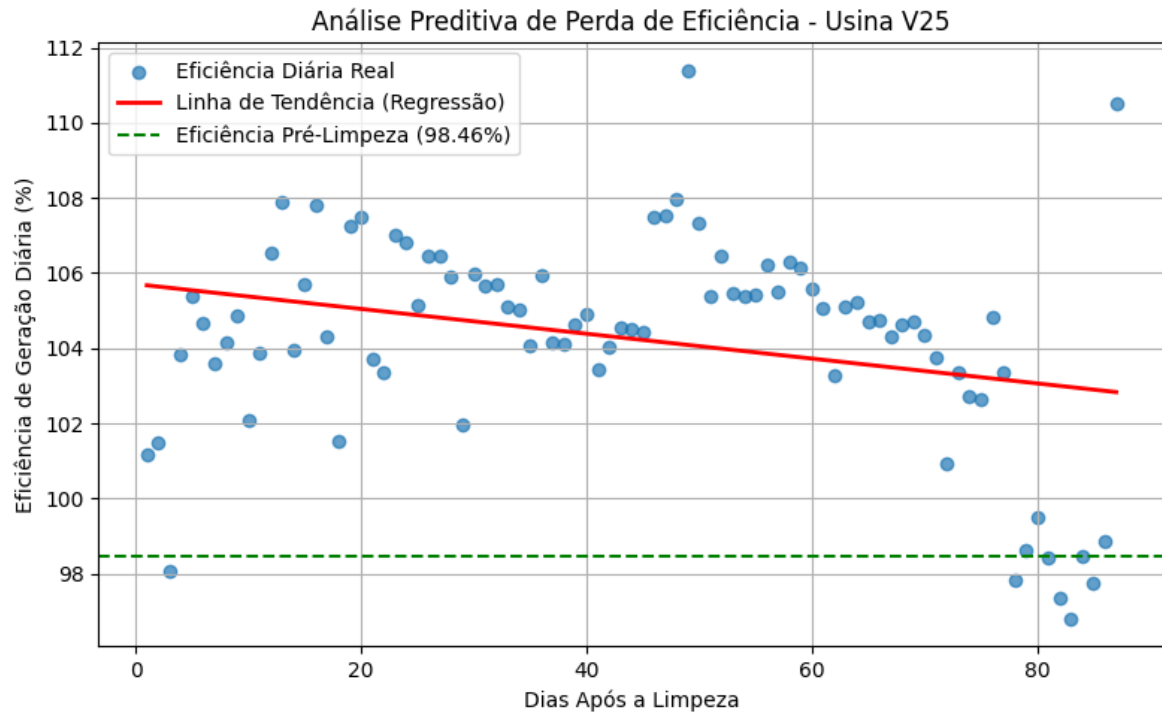
Figura 20 - Resultado da análise preditiva para a usina Araguari.



Fonte: Autoria própria.

A análise da usina Lagoa Formosa II, ilustrada na Figura 21, identificou uma tendência de queda, como esperado. O modelo estimou uma taxa de sujeira de $-0,0330\%$ de perda de eficiência por dia. Com base nessa taxa, a previsão indicou que a usina atingiria seu patamar pré-limpeza de $98,46\%$ em aproximadamente 219 dias. Embora plausível, um ciclo de limpeza superior a sete meses pode ser considerado longo, sugerindo que o efeito da sujeira foi lento ou mitigado por outros fatores.

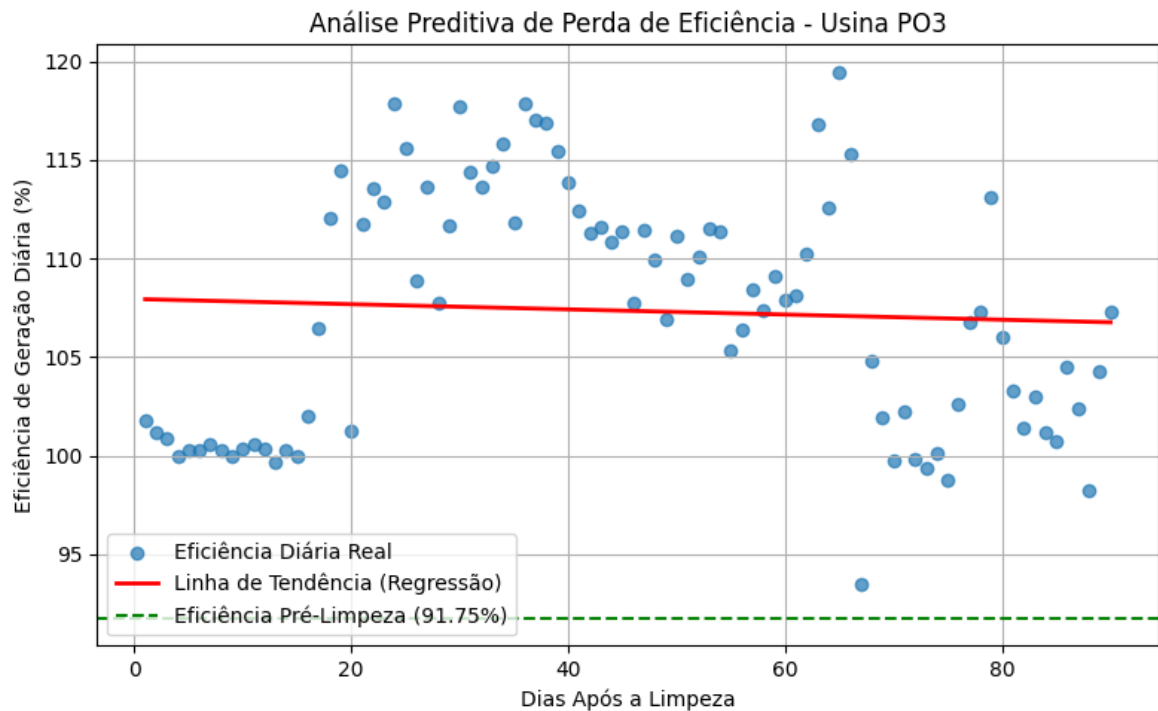
Figura 21 - Resultado da análise preditiva para a usina Lagoa Formosa II.



Fonte: Autoria própria.

Por fim, o estudo de caso da usina Poconé III (Figura 22) também revelou uma tendência negativa, porém ainda mais lenta, com uma taxa de sujidade estimada em apenas -0,0131% ao dia. Isso resultou em uma previsão logisticamente impraticável de 1233 dias, ou mais de três anos, para que a usina retornasse ao seu nível de eficiência pré-limpeza de 91,75%.

Figura 22 - Resultado da análise preditiva para a usina Poconé III.



Fonte: Autoria própria.

A inconsistência dos resultados obtidos nos três casos evidencia que a perda de eficiência não é uma função exclusiva do tempo, mas sim um fenômeno multivariado. A tendência positiva em Araguari sugere a interferência de outras ocorrências operacionais, enquanto as previsões extensas para V25 e PO3 indicam a influência de fatores ambientais, como chuvas. Conclui-se, portanto, que a previsão do ciclo de limpeza por meio de uma regressão linear simples é inviável, e que a determinação da real complexidade do fenômeno é o principal resultado desta análise exploratória.

5 CONCLUSÃO

O presente trabalho teve como objetivo central analisar o impacto da sujeidade nos módulos fotovoltaicos e quantificar a eficácia da limpeza como estratégia de mitigação. Por meio de uma robusta metodologia de tratamento e análise de dados de 42 usinas, pode-se afirmar que este objetivo foi plenamente alcançado. A pesquisa não apenas mediu as perdas de desempenho, mas também validou estatisticamente a relação causal entre a manutenção e a recuperação da eficiência de geração.

A análise quantitativa revelou que a limpeza dos módulos gera um impacto positivo e significativo, resultando em um ganho médio absoluto de 11,51% e um ganho médio relativo de 12,78% na eficiência. A validação destes resultados por meio de um Teste t Pareado (p-valor ≈ 0) conferiu um alto grau de confiabilidade a esta conclusão, provando que a melhoria observada não é fruto do acaso, mas sim um efeito direto do procedimento de limpeza.

Além da quantificação do ganho, um dos aprendizados mais relevantes deste trabalho foi a compreensão da complexidade e variabilidade do fenômeno no mundo real. A análise de casos individuais demonstrou que o benefício da limpeza é condicional, dependendo diretamente do nível de sujeidade prévio de cada usina. A reflexão mais profunda, contudo, surgiu da tentativa de criar um modelo preditivo. A ineficácia de um modelo de regressão linear simples para prever o ciclo de limpeza ideal foi um resultado fundamental, pois revelou a forte influência de variáveis de confusão, como a ocorrência de chuvas e outras manutenções operacionais, que mascaram e alteram a tendência de degradação por sujeidade.

Essa compreensão da complexidade do problema aponta um caminho claro para trabalhos futuros. Conclui-se que um modelo preditivo robusto para este fim exigiria uma abordagem multivariada, que incorpore um conjunto de dados mais rico, incluindo séries temporais de dados pluviométricos e um histórico detalhado de todas as intervenções operacionais na planta. A demonstração da necessidade e das variáveis para tal modelo é uma contribuição importante deste trabalho. Em síntese, este estudo não apenas cumpre seu objetivo de quantificar o valor da limpeza, mas também elucida as complexidades envolvidas em sua otimização, oferecendo subsídios técnicos para a tomada de decisão e para o desenvolvimento de práticas de manutenção mais inteligentes e eficientes no crescente setor fotovoltaico brasileiro.

REFERÊNCIAS BIBLIOGRÁFICAS

ALBA ENERGIA. *ALBA Cinturão Solar e o potencial para geração de energia - ALBA*. Disponível em: <<https://albaenergia.com.br/cinturao-solar-e-o-potencial-para-geracao-de-energia/>>. Acesso em: 21 jan. 2025.

ARAÚJO, Danielly Norberto; MARQUES DE CARVALHO, Paulo César; DUPONT, Ivonne Montero. Efeitos da acumulação de sujeira sobre o desempenho de módulos fotovoltaicos. *Revista Tecnologia*, v. 40, n. 2, 5 nov. 2019.

BRASIL. *Brasil bate recorde de expansão da energia solar em 2023* — Ministério de Minas e Energia. Disponível em: <<https://www.gov.br/mme/pt-br/assuntos/noticias/brasil-bate-recorde-de-expansao-da-energia-solar-em-2023>>. Acesso em: 21 jan. 2025.

CARNEIRO, Joaquim. *ELECTROMAGNETISMO B MÓDULOS FOTOVOLTAICOS CARACTERÍSTICAS E ASSOCIAÇÕES 2º Ano do Mestrado Integrado em Engenharia Civil 1º Semestre (2010)*. . [S.l: s.n.], 2010.

CORREA, Rafael Santos; NASCIMENTO, José Fernando Silva Do; NEVES, Fernando Frachone. PRÁTICAS DE INSPEÇÃO E MANUTENÇÃO PREVENTIVA EM SISTEMAS FOTOVOLTAICOS. *Ciência & Tecnologia*, v. 15, n. 1, p. e1514, 14 jun. 2023.

DE CAMPOS, João. *DETERMINAÇÃO DO TEMPO ÓTIMO PARA LIMPEZA DE PAINÉIS FOTOVOLTAICOS PARA OBTENÇÃO DA MELHOR PRODUTIVIDADE-ESTUDO DE CASO DOS SFVCR's IMPLANTADOS NA UTFPR*. . [S.l: s.n.], 2017.

DURISCH, Wilhelm *et al*. Efficiency model for photovoltaic modules and demonstration of its application to energy yield estimation. *Solar Energy Materials and Solar Cells*, v. 91, n. 1, p. 79–84, 5 jan. 2007.

FARIA, Roberto. Efeito fotovoltaico e fotocondutividade de dispositivos poliméricos. 2000.

FERNANDES COELHO, Thays; VALDÉS SERRA, Juan Carlos. *Tecnologias para Reciclagem de Sistemas Fotovoltaicos: Impactos Ambientais*. . [S.l: s.n.], 2018.

GÓES, Patrícia Fagundes; TANIMOTO, Armando Hirohumi. *TECNOLOGIAS E PARÂMETROS AMBIENTAIS PARA A ESCOLHA DE UMA PLACA GERADORA DE ENERGIA SOLAR FOTOVOLTAICA*. *Revista Scientia*. [S.l: s.n.], 2021.

JUNIOR, Afrânio. *IMPACTO DA LIMPEZA DE PAINÉIS FOTOVOLTAICOS PARA MELHOR PRODUTIVIDADE*. . [S.l: s.n.], 2021.

LIMA, Ariane A. *et al*. Uma revisão dos princípios da conversão fotovoltaica de energia. *Revista Brasileira de Ensino de Física*, v. 42, 2019.

LOHMANN, Gerald M.; MONAHAN, Adam H. Effects of temporal averaging on short-term irradiance variability under mixed sky conditions. *Atmospheric Measurement Techniques*, v. 11, n. 5, p. 3131–3144, 31 maio 2018.

MANI, Monto; PILLAI, Rohit. *Impact of dust on solar photovoltaic (PV) performance: Research status, challenges and recommendations. Renewable and Sustainable Energy Reviews*. [S.l.]: Elsevier Ltd. , 2010

OLIVEIRA, Adriana Tenir Egéa De *et al.* A energia solar fotovoltaica: transformação, evolução, aspectos ambientais e abordagens na sala de aula. *Research, Society and Development*, v. 11, n. 9, p. e25811932533, 20 jul. 2022.

PORTAL SOLAR. *Energia Solar no Brasil - Portal Solar - Tudo sobre Energia Solar Fotovoltaica | Portal Solar*. Disponível em: <<https://www.portalsolar.com.br/energia-solar-no-brasil.html>>. Acesso em: 21 jan. 2025a.

PORTAL SOLAR. *História da Energia Solar: Quem Criou a Energia Solar? | Portal Solar*. Disponível em: <https://www.portalsolar.com.br/historia-e-origem-da-energia-solar?utm_source=chatgpt.com>. Acesso em: 21 jan. 2025b.

RABELO, Pamela. *ANÁLISE DA INFLUÊNCIA DA TEMPERATURA NA EFICIÊNCIA DE GERAÇÃO DE SISTEMAS DE ENERGIA SOLAR FOTOVOLTAICA COM OTIMIZADORES DE POTÊNCIA*. . Fortaleza: [s.n.], 2022.

RAMPINELLI, Giuliano Arns; KRENZINGER, Arno; ROMERO, Faustino Chenlo. Descrição e Análise de Inversores Utilizados em Sistemas Fotovoltaicos. *Revista Ciencias Exatas e Naturais*, v. 15, n. 1, 2013.

RASHID, Mahnoor *et al.* Effect of dust accumulation on the performance of photovoltaic modules for different climate regions. *Heliyon*, v. 9, n. 12, 1 dez. 2023.

ROSSI, Catherine; LEITE, Gelon; RAMPINELLI, Giuliano. *SIMULAÇÃO E ANÁLISE DE DESEMPENHO DE TECNOLOGIAS DE MÓDULOS FOTOVOLTAICOS EM DIFERENTES MUNICÍPIOS BRASILEIROS*. . [S.l: s.n.], 2014.

SARVER, Travis; AL-QARAGHULI, Ali; KAZMERSKI, Lawrence L. *A comprehensive review of the impact of dust on the use of solar energy: History, investigations, results, literature, and mitigation approaches. Renewable and Sustainable Energy Reviews*. [S.l: s.n.]. , abr. 2013

SEENIVASAN, Dhamotharan. ETL in a World of Unstructured Data: Advanced Techniques for Data Integration. *International Journal of Management*, v. 11, p. 1, 2021. Disponível em: <<http://www.ijmra.us>,<http://www.ijmra.us>,>.