



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA

BRUNO MASCARENHAS DA SILVA

ANÁLISE PREDITIVA DO CRESCIMENTO DA GERAÇÃO DISTRIBUÍDA
FOTOVOLTAICA NO BRASIL

FORTALEZA

2022

BRUNO MASCARENHAS DA SILVA

ANÁLISE PREDITIVA DO CRESCIMENTO DA GERAÇÃO DISTRIBUÍDA
FOTOVOLTAICA NO BRASIL

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia Elétrica.

Orientador: Prof. Dr. Raphael Amaral da Câmara.

FORTALEZA

2022

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

- S578a Silva, Bruno Mascarenhas da.
Análise preditiva do crescimento da geração distribuída fotovoltaica no Brasil / Bruno Mascarenhas da Silva. – 2022.
49 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia, Curso de Engenharia Elétrica, Fortaleza, 2022.
Orientação: Prof. Dr. Raphael Amaral da Câmara.
1. Solar fotovoltaica. 2. Análise preditiva. 3. Regressão linear. 4. Regressão polinomial. I. Título.
CDD 621.3
-

BRUNO MASCARENHAS DA SILVA

ANÁLISE PREDITIVA DO CRESCIMENTO DA GERAÇÃO DISTRIBUÍDA
FOTOVOLTAICA NO BRASIL

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Engenharia Elétrica do
Centro de Tecnologia da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Engenharia Elétrica.

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Raphael Amaral da Câmara (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Paulo Peixoto Praça
Universidade Federal do Ceará (UFC)

Prof. Dr. Ernande Eugênio Campelo Morais
Universidade Federal do Ceará (UFC)

Aos meus pais, Antônia e Gilvan.

AGRADECIMENTOS

Aos meus pais, Antônia e Gilvan, por apoiarem e acreditarem no caminho que escolhi seguir.

À minha amiga de longa data, Mônica, por estar sempre presente e por todo o suporte que sempre me deu.

Aos amigos que fiz na UFC: Alice, Carolina, Maria Annalyanne, Nathanal e Raul, pela companhia e por terem tornado minha jornada na universidade muito melhor.

Ao Prof. Dr. Raphael Amaral, pela orientação e por toda a paciência ao longo do desenvolvimento deste trabalho.

Ao Prof. Dr. Paulo Peixoto Praça e ao Prof. Dr. Ernande Eugênio Campelo Morais, por terem aceitado o convite para compor a banca de avaliação.

À Tecsys Jr. e às pessoas maravilhosas que pude conhecer dentro desse projeto, por todos os momentos vividos e por todo o crescimento pessoal e profissional que me proporcionaram.

Ao Movimento Empresa Jr., por todo o investimento em capacitar e preparar melhor os estudantes para o mercado de trabalho.

“Você não pode esperar construir um mundo melhor sem melhorar as pessoas. Cada um de nós deve trabalhar para o nosso próprio aprimoramento.”

(Marie Curie)

RESUMO

Uma vez que a sociedade percebeu os aspectos negativos da emissão de gases poluentes causadores do efeito estufa, ganharam muito destaque as fontes renováveis de energia, como as usinas hidrelétricas, nucleares, eólicas e solar fotovoltaica. Apesar da primeira célula fotovoltaica ter sido criada em 1883, foi apenas na última década que os avanços tecnológicos tornaram a fabricação de módulos fotovoltaicos acessível e vantajosa para os consumidores, ocasionando um grande crescimento dessa modalidade no Brasil e no mundo. Tendo em vista esse aumento recente, o trabalho busca estimar a tendência de crescimento da potência instalada de usinas de micro e minigeração fotovoltaica no país, utilizando a linguagem de programação *python* para desenvolver modelos de regressão linear e polinomial.

Palavras-chave: solar fotovoltaica; análise preditiva; regressão linear; regressão polinomial.

ABSTRACT

Once society realized the negative aspects of the emission of polluting gases that cause the greenhouse effect, renewable energy sources, such as hydroelectric, nuclear, wind and solar photovoltaic plants have gained a lot of prominence. Although the first photovoltaic cell was created in 1883, it was only in the last decade that technological advances allowed manufacturing of photovoltaic modules to be accessible and economically viable for consumers, causing a huge growth of this technology in Brazil and the world. In view of this recent increase, this study seeks to estimate the growth trend of the installed capacity of micro and mini photovoltaic power plants in the country, using python programming language to develop linear and polynomial regression models.

Keywords: photovoltaic solar; predictive analytics; linear regression; polynomial regression.

LISTA DE FIGURAS

Figura 1 – Código exemplo para regressão linear.....	24
Figura 2 – Sistema linear obtido.....	26
Figura 3 – Código de exemplo para regressão polinomial.....	27

LISTA DE GRÁFICOS

Gráfico 1	– Reta de regressão linear obtida.....	25
Gráfico 2	– Curva de regressão polinomial obtida.....	28
Gráfico 3	– Previsão de Potência Instalada no Brasil.....	32
Gráfico 4	– Curvas de Previsão de Potência Instalada no Nordeste.....	33
Gráfico 5	– Retas de Previsão de Potência Instalada no Nordeste.....	33
Gráfico 6	– Curvas de Previsão de Potência Instalada no Sudeste.....	34
Gráfico 7	– Retas de Previsão de Potência Instalada no Sudeste.....	34
Gráfico 8	– Curvas de Previsão de Potência Instalada no Norte.....	35
Gráfico 9	– Retas Previsão de Potência Instalada no Norte.....	35
Gráfico 10	– Curvas de Previsão de Potência Instalada no Sul.....	36
Gráfico 11	– Retas de Previsão de Potência Instalada no Sul.....	36
Gráfico 12	– Curvas de Previsão de Potência Instalada no Centro-Oeste.....	37
Gráfico 13	– Retas de Previsão de Potência Instalada no Centro-Oeste.....	37

LISTA DE TABELAS

Tabela 1 – Valores de R^2 Ajustado - Brasil.....	38
Tabela 2 – Valores da Raiz do Erro Médio Quadrático – Brasil.....	38
Tabela 3 – Valores de R^2 Ajustado – Estados.....	39
Tabela 4 – Valores da Raiz do Erro Médio Quadrático - Estados.....	40

LISTA DE ABREVIATURAS E SIGLAS

ABSOLAR	Associação Brasileira de Energia Solar Fotovoltaica
ANEEL	Agência Nacional de Energia Elétrica
CCST	Centro de Ciência do Sistema Terrestre
CEPEL	Centro de Pesquisa Elétrica
CSV	Valores Separados por Vírgula
ENAP	Escola Nacional de Administração Pública
EPE	Empresa de Pesquisa Energética
FV	Fotovoltaico
GC	Geração Centralizada
GD	Geração Distribuída
INEE	Instituto Nacional de Eficiência Energética
INPE	Instituto Nacional de Pesquisas Espaciais
LABREN	Laboratório de Modelagem e Estudos de Recursos Renováveis de Energia
MMQ	Método dos Mínimos Quadrados
ONS	Operador Nacional do Sistema
PDGD	Painel de Dados de Micro e Minigeração Distribuída
RMSE	Raiz do Erro Quadrático Médio
SIN	Sistema Interligado Nacional
TCC	Trabalho de Conclusão de Curso
UF	Unidade Federativa

LISTA DE SÍMBOLOS

°C	Graus Celsius
%	Porcentagem
kW	Quilowatt
MW	Megawatt
Σ	Somatório
$\sqrt{\quad}$	Raiz Quadrada
V	Volts
A	Amperes
R ²	Coeficiente de Determinação
α	Alfa
β	Beta
μ	Mu
ε	Epsilon
θ	Teta

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Contextualização	16
1.2	Motivação	18
1.3	Objetivos	18
1.3.1	<i>Objetivo Geral</i>	18
1.3.2	<i>Objetivos Específicos</i>	18
1.4	Estrutura do Trabalho	19
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	Geração Centralizada	20
2.2	Geração Distribuída	20
2.3	Atlas Brasileiro de Energia Solar.....	22
2.4	Linguagem de Programação Python.....	22
2.5	Modelos de Regressão	23
2.4.1	<i>Regressão Linear Simples</i>	23
2.4.2	<i>Regressão Linear Simples no Python</i>	24
2.4.3	<i>Método dos Mínimos Quadrados</i>	25
2.4.4	<i>Regressão Polinomial no Python</i>	27
2.5	Métricas de Performance	28
2.5.1	<i>Raiz do Erro Quadrático Médio</i>	28
2.5.2	<i>Coefficiente de Determinação Ajustado</i>	29
3	METODOLOGIA E RESULTADOS	30
3.1	Dados Utilizados	30
3.2	Processamento de Dados	31
3.3	Gráficos de Regressão	32
3.3.1	<i>Região Nordeste</i>	33
3.3.2	<i>Região Sudeste</i>	34
3.3.3	<i>Região Norte</i>	35
3.3.4	<i>Região Sul</i>	36
3.3.5	<i>Região Centro-Oeste</i>	37
3.4	Métricas e Discussões	38
4	CONCLUSÃO	42

REFERÊNCIAS	43
APÊNDICE A – EQUAÇÕES DA REGRESSÃO LINEAR	47
APÊNDICE B – EQUAÇÕES DA REGRESSÃO POLINOMIAL	48
ANEXO A – DATASHEET PAINEL SOLAR CANADIAN	49

1 INTRODUÇÃO

Nesta seção, tem-se, inicialmente, uma breve revisão sobre a importância da eletricidade, as principais fontes de geração renováveis e o surgimento das células fotovoltaicas. Em seguida, são apresentados os principais fatores que motivaram a escolha deste tema e, por fim, são apresentados os objetivos e a estrutura deste trabalho.

1.1 Contextualização

Para uma boa parcela da população mundial, a eletricidade é um elemento indispensável. Uma vez que foi desenvolvida a tecnologia para gerar, transmitir e distribuir energia elétrica através de grandes distâncias, abriu-se o caminho para uma enorme evolução social, econômica e técnico-científica. Tal feito teve um impacto tão significativo na forma como a sociedade vive e relaciona-se, que não seria viável abrir mão desse serviço atualmente, uma vez que dependem dela tanto as tarefas mais simples, como conservar alimentos em uma geladeira, como também em aplicações essenciais, tais quais as encontradas em centros hospitalares, redes de comunicação e toda a automação industrial existente. Por conta dessa dependência, fez-se necessário aperfeiçoar e diversificar a produção de energia.

O carvão e o gás natural, por serem abundantes na natureza, foram e continuam sendo as principais matérias primas para a geração de energia elétrica do mundo, tendo sido responsáveis por aproximadamente 59% da geração total do ano de 2021 (RITCHIE; ROSER; ROSADO, 2022). Contudo, o uso excessivo dessas fontes é a causa de muitos problemas ambientais e humanitários, fatores que têm feito com que os países busquem fontes alternativas de energia para o futuro.

As fontes renováveis de energia mais relevantes no mundo hoje são a hidrelétrica, nuclear, eólica e solar fotovoltaica (RITCHIE; ROSER; ROSADO, 2022). O Brasil é um dos países que se destaca no uso das hidrelétricas, uma vez que, até setembro de 2022, essa matriz energética compôs cerca de 61% de toda a geração nacional (ONS, 2022). Ainda que seja uma fonte com baixa emissão de gases poluentes quando está em operação, as hidrelétricas dependem da geografia local, da incidência de chuva e de grandes áreas que sirvam como reservatório de água, sendo difíceis de implementar em países sem estes requisitos. Por conta disso, as hidrelétricas tiveram sua parcela da geração mundial caindo de 20% em 1985 para 15% em 2021 (RITCHIE; ROSER; ROSADO, 2022).

As usinas nucleares, por sua vez, utilizam o urânio, que é um recurso abundante, não emitem gases poluentes, produzem energia de forma barata, não dependem de condições climáticas e não precisam de grandes áreas para operar. Todavia, além dos dejetos radioativos, acidentes em usinas nucleares se mostraram catastróficos, com o principal exemplo a explosão de um dos reatores da usina nuclear em Chernobil, em 1986, então controlada pela União Soviética. O acidente causou a morte de milhares pessoas e contaminou uma grande área com radiação, de modo que ainda hoje existem áreas contaminadas que não podem ser populadas (KORTOV; USTYANTSEV, 2012). Por conta deste e outros acidentes, a energia nuclear se tornou um tópico polêmico, e sua participação na geração de energia global caiu de aproximadamente 15% em 1985 para 9,84% em 2021 (RITCHIE; ROSER; ROSADO, 2022).

Diferente das outras fontes renováveis abordadas, que tiveram queda na contribuição da geração global de energia, a fonte eólica tem mostrado um expressivo aumento na última década, com sua participação global na geração de energia subindo de 1,61% em 2010 para 6,54% em 2021 (RITCHIE; ROSER; ROSADO, 2022). Apesar de dependerem das condições climáticas, os aerogeradores têm se mostrado eficazes na produção de energia e, além de causarem pouco impacto ambiental, podem ser instalados em diversas geografias, e até mesmo no mar.

Por último, tem-se a energia solar fotovoltaica, principal tópico de estudo deste trabalho, cuja participação no cenário global cresceu de 0,16% em 2010 para 3,63% em 2021 (RITCHIE; ROSER; ROSADO, 2022). A primeira célula fotovoltaica foi criada em 1883, por Charles Fritts, que conseguiu uma eficiência de aproximadamente 1% utilizando selênio como seu principal elemento. Entretanto, foi apenas em 1954 que as células fotovoltaicas de silício (com cerca de 4% de eficiência) conhecidas e utilizadas atualmente foram desenvolvidas. Daryl Chapin, Calvin Fuller e Gerald Pearson foram os percussores dessa tecnologia que tem se popularizado e revolucionado a geração de energia elétrica no mundo (RICHARDSON, 2022).

Com o aumento da demanda dessa tecnologia, os módulos fotovoltaicos atuais já ultrapassam os 20% de eficiência, com diversos fabricantes focados em elevar cada vez mais esse valor. A significativa melhora alcançada nos últimos anos fez com que essa fonte energética se tornasse muito mais viável financeiramente e, uma vez que os sistemas são modulares, com baixa complexidade e pouca necessidade de manutenção, foi possível popularizar o seu acesso, possibilitando famílias comuns terem o seu próprio sistema de geração solar, de forma que essa tecnologia não ficou restrita apenas às grandes concessionárias de energia.

1.2 Motivação

De acordo com Greener (2022), a potência instalada em novas instalações de Geração Distribuída (GD), no primeiro semestre do ano de 2022, foi 51% maior em relação ao mesmo período do ano anterior. Essa tendência de aumento não é exclusiva desse período, uma vez que é observada desde os últimos quatro anos.

Com este acentuado aumento, faz-se necessário desenvolver métodos que possam prever o avanço desse tipo de geração no país, além de movimentar o mercado, fomentar órgãos de pesquisa e embasar decisões técnicas governamentais de países quanto aos seus investimento e incentivos em plantas energéticas. Nos trabalhos de Cunha e Sobel (2021) e de Urzagasti *et al.* (2021), recursos de programação são utilizados a fim de prever a geração de energia das usinas já existentes a partir de parâmetros ambientais, como o histórico meteorológico e valores de irradiação diário.

Desse modo, esse trabalho tem como motivação analisar e extrapolar os dados de potência instalada, utilizando recursos computacionais, com ênfase na geração distribuída fotovoltaica (FV), uma vez que a estimativa desses valores pode gerar insumo para que se antecipe os recursos humanos e materiais necessários para o crescimento esperado nos próximos anos.

1.3 Objetivos

Neste tópico, serão elencados os resultados esperados ao fim deste trabalho.

1.3.1 Objetivo Geral

O principal objetivo deste Trabalho de Conclusão de Curso (TCC) é analisar a tendência do aumento da potência instalada referente às usinas fotovoltaicas nos estados do Brasil, considerando tanto a micro como a minigeração dentro do contexto da GD.

1.3.2 Objetivos Específicos

Objetiva-se, utilizando a linguagem de programação Python, elaborar algoritmos preditivos utilizando Regressão Linear e Regressão Polinomial que permitam:

- a) Comparar o aumento da potência instalada em cada estado;

- b) Avaliar a tendência de aumento entre os estados e a média nacional;
- c) Obter insumos sobre a precisão dos modelos utilizados.

1.4 Estrutura do Trabalho

Este trabalho encontra-se dividido em quatro capítulos. No Capítulo 1, tem-se a introdução ao tema abordado, com a contextualização das fontes renováveis de energia no mundo, e também a motivação e os objetivos do trabalho.

No Capítulo 2, desenvolve-se a fundamentação teórica quem embasou o estudo desenvolvido neste trabalho. Nele, são explicados o princípio de funcionamento dos módulos fotovoltaicos, os modelos de geração adotados no país, as ferramentas matemáticas de regressão utilizadas e, por fim, as métricas estatísticas escolhidas para avaliar a performance dos modelos de regressão.

No Capítulo 3, tem-se a apresentação dos dados utilizados e o processamento feito em cima destes. A partir disso, apresenta-se os resultados obtidos, em forma de gráficos e tabelas com as métricas para cada modelo, as quais são analisadas e discutidas ao fim da seção.

No Capítulo 4, encerra-se o trabalho com a conclusão sobre os resultados obtidos e as sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção serão apresentados conceitos e definições referentes à geração fotovoltaica no Brasil e aos tópicos matemáticos utilizados. Inicialmente, será abordado o funcionamento dos módulos fotovoltaicos, seguido pelos modelos de geração distribuído e centralizado, adotados no país. Por fim, serão demonstradas as fórmulas para regressão linear e regressão polinomial, as quais são utilizadas nos modelos preditivos deste trabalho.

2.1 Geração Centralizada

A geração centralizada (GC) é composta por usinas de grande porte, tanto no âmbito da energia solar como nas outras modalidades de geração, cuja potência instalada costuma ser na ordem dos MW (*megawatt*) de potência. Na GC, as usinas exigem um grande aporte financeiro para serem construídas e começarem a operar, por conta disso, pertencem majoritariamente ao setor privado, que comercializa a energia gerada nos leilões realizados pelos governos estaduais (BARROS; SAUAIA; KOLOSZUK, 2019).

No Brasil, a GC é o modelo de geração predominante, com destaque para as grandes usinas hidrelétricas, as quais compõem 60,2% da capacidade de geração instalada no país (EPE, 2022). Essa característica, aliada ao fato de que o sistema elétrico nacional é interligado, faz com que se tenha muita robustez e segurança nos sistemas de geração, transmissão e distribuição de energia nacionais.

Contudo, devido a sua composição em grandes unidades geradoras, sendo a maioria responsável por manter cargas de regiões completas do país, o sistema interligado nacional (SIN) torna-se sensível aos problemas enfrentados no processo de geração (podendo ser de natureza técnica ou climática), tendo como consequência direta o desbalanceamento do seu fluxo de potência. A solução amplamente adotada para esse cenário parte do aumento da geração através de combustíveis fósseis, que impactam na elevação da tarifa de energia para o consumidor e causam impactos negativos ao meio ambiente pela emissão excessiva de gases poluentes.

2.2 Geração Distribuída

Quando a energia elétrica é gerada pelos próprios consumidores, em usinas de menor potência, próximas ou juntas à unidade consumidora, tem-se a chamada Geração

Distribuída (INEE, s.d.). Dessa forma, origina-se o conceito de prossumidores (clientes que consomem e produzem energia), os quais tanto consomem energia da rede de distribuição como também injetam o excedente de energia elétrica gerada (TAVEIRA, 2020).

De acordo com a Agência Nacional de Energia Elétrica (ANEEL, 2022), qualquer fonte renovável pode ser utilizada nessa modalidade. Contudo, existem limitações nas dimensões que a unidade geradora pode possuir, de modo que a potência não pode ultrapassar 5 MW, sendo os sistemas com até 75 kW classificados como microgeração distribuída, enquanto aqueles que possuem potência maior que 75 kW (e menor ou igual ao limite máximo de 5 MW) são classificados como minigeração distribuída.

Na GD, quando a geração acontece na unidade consumidora, a energia é utilizada diretamente pelo cliente. Todavia, caso a energia que esteja sendo produzida ultrapasse a que está sendo consumida, tem-se um excedente que passa por um medidor bidirecional da concessionária e é injetado na rede. Essa energia extra é convertida em créditos, os quais possuem validade de 60 meses e que são descontados na conta de energia da unidade prossumidora ou de outras unidades consumidoras pertencentes à mesma pessoa, se estas estiverem dentro da área de fornecimento da mesma concessionária (ANEEL, 2022).

Em relação à geração distribuída através de painéis fotovoltaicos, tem-se, além do baixo impacto ambiental e da diversificação da matriz energética, diversas outras vantagens que justificam a adoção desse modelo de geração de energia: uma vez que a produção acontece próxima ao ponto de injeção na rede, que, por sua vez, está próxima a outras unidades consumidoras, tem-se poucas perdas no sistema; a geração por parte dos próprios consumidores também reduz a carga do sistema e pode adiar a necessidade de novos investimentos para a expansão das redes de transmissão e distribuição (ANEEL, 2022); os painéis fotovoltaicos possuem uma vida útil entre 25 e 35 anos, tendo a maior parte dos materiais utilizados em sua construção sendo recicláveis (PORTALSOLAR, s.d.).

No entanto, apesar de todos os pontos positivos, existem também pontos que impedem ou dificultam a implementação dessa tecnologia, como o custo que ainda é consideravelmente alto e, para a maioria dos consumidores, só é possível adquirir seu próprio sistema através de financiamentos, o que retarda o retorno do capital investido. Outro fator de grande relevância atualmente são as mudanças na legislação ou nos impostos cobrados sobre esse tipo de geração, que podem mudar drasticamente o rendimento financeiro das instalações, tornando-as menos vantajosas e desincentivando a adesão de novos interessados nesse sistema de geração. Além disso, não há como armazenar energia de forma barata e não há produção à noite, fazendo com que essa fonte não ajude a desafogar as redes de distribuição no horário de

pico (entre 18 e 21 horas), e finalmente, a geração fotovoltaica também é fortemente dependente das condições climáticas, uma vez que sombreamento e altas temperaturas ocasionam a diminuição da energia gerada, fazendo com que algumas regiões sejam menos propícias para essas instalações (DIAS, 2005).

2.3 Atlas Brasileiro de Energia Solar

O Atlas Brasileiro de Energia Solar consiste em um livro e em um banco de dados desenvolvidos pelo Laboratório de Modelagem e Estudos de Recursos Renováveis de Energia (LABREN), pelo Centro de Ciência do Sistema Terrestre (CCST) e pelo Instituto Nacional de Pesquisas Espaciais (INPE).

Este projeto é um dos mais importantes no contexto da energia solar fotovoltaica, uma vez que sua base de dados, disponibilizada gratuitamente, serve de insumo para os mais variados projetos, seja ele acadêmico ou de implementação no mercado.

Um dos destaques do Atlas é o mapa de irradiação nacional, o qual utiliza diversos sensores espalhados no território para gerar médias diárias, mensais e anuais de irradiação em todo o território nacional. Essa informação é extremamente significativa, uma vez que as usinas fotovoltaicas são mais vantajosas em locais com maior irradiação (PEREIRA et al., 2017).

2.4 Linguagem de Programação *Python*

O *Python* se tornou, nos últimos anos, uma das linguagens de programação mais populares do mundo no mercado. Por se tratar de uma linguagem de alto nível, ou seja, mais próxima da linguagem falada e escrita, é mais fácil de aprender, o que agiliza seu uso e implementação por novos usuários. Além disso, é uma linguagem muito versátil, uma vez que apresenta diversas bibliotecas complementares para os mais diversos tipos de aplicações, como ciência de dados, aprendizado de máquina, redes neurais, desenvolvimento de software, entre outros (BERKELEY, s.d.).

Apesar de sua popularidade, o *Python* não é a linguagem mais otimizada para a maioria dos processos. Por ser uma linguagem de alto nível, muitas das configurações que são feitas automaticamente acabam aumentando o tempo de execução dos programas. Contudo, com a disponibilidade de processamento da tecnologia atual, esse aumento no tempo de processamento é quase desprezível na maioria dos processos (BOBBY, 2019).

2.5 Modelos de Regressão

Nos modelos de regressão deste trabalho, tem-se como objetivo encontrar uma equação que demonstre a relação entre um preditor x e uma variável dependente y , a partir dos pontos obtidos com os dados. Nesta seção, tem-se a fundamentação teórica a respeito dos métodos de regressão utilizados, bem como uma breve demonstração de sua aplicação através de ferramentas computacionais.

2.5.1 Regressão Linear Simples

Os modelos de regressão são amplamente utilizados para classificar qualitativamente e quantitativamente a relação entre duas ou mais variáveis, com o intuito de entender melhor o comportamento de um determinado espaço amostral e, possivelmente, extrapolar os resultados obtidos a fim de prever situações não observadas.

O modelo de regressão linear simples é um dos modelos preditivos mais básicos utilizados, e tem como objetivo relacionar, através da equação de uma reta, uma variável dependente, a qual será chamada de Y , com uma única variável independente, a qual será chamada de X (CHEIN, 2019). Vale ressaltar que este modelo, apesar de poder indicar relações de causa-efeito, não é capaz de provar que sejam verdades.

Com isso, em um experimento com um número $i = (1, \dots, n)$ de dados medidos, tem-se i equações como a Equação 1, em que E_i são valores aleatórios e independentes com média zero:

$$Y_i = \beta_0 + \beta_1 X_i + E_i \quad (1)$$

Na Equação 1, β_0 e β_1 são os parâmetros que se deseja encontrar, de modo que $\beta_0 + E_i$ é o coeficiente linear, o ponto de interseção da reta com o eixo das ordenadas, enquanto que β_1 é o coeficiente angular, igual à tangente do ângulo formado entre a reta e o eixo das abscissas (BASTOS; GUIMARÃES; SEVERO, 2015).

Para cada valor de X_i , Y_i é uma variável aleatória cuja média μ_{Y_i} é igual a $\beta_0 + \beta_1 X_i$, podendo-se reescrever a Equação 1 como a Equação 2:

$$Y_i = \mu_{Y_i} + E_i \quad (2)$$

Por fim, uma vez que sejam calculadas as médias e os parâmetros sejam estimados, tem-se a Equação 3, que é a equação de uma reta:

$$Y = \alpha + \beta X \quad (3)$$

2.5.2 Regressão Linear Simples no Python

Apesar de ser um dos modelos de regressão mais simples, quando o volume de dados é muito alto, torna-se inviável realizar todos os cálculos à mão. Por conta disso, os modelos de regressão linear obtidos neste trabalho foram obtidos utilizando-se a linguagem de programação *Python*, com o *software JupyterLab*. Esta seção não tem como objetivo detalhar todos os passos necessários para a elaboração de um modelo regressão linear complexo, mas visa apenas mostrar, de forma sucinta, uma rápida aplicação das ferramentas presentes na linguagem de programação utilizada.

Figura 1 - Código exemplo para regressão linear

```

1. import numpy as np
2. from sklearn.linear_model import LinearRegression
3. import matplotlib.pyplot as plt
4. valores_X = np.array([0,1,2,3,4,5])
5. valores_Y = np.array([5, 11, 9, 15, 17, 16])
6. x = valores_X.reshape(-1,1)
7. y = valores_Y
8. reg_lin = LinearRegression()
9. reg_lin.fit(x, y)
10. plt.figure(1, figsize = [16,9])
11. plt.scatter(valores_X,valores_Y, color='black', label =
'Valores atribuídos')
12. plt.plot(valores_X, reg_lin.predict(x), 'k', label =
'Regressão Linear')
13. plt.xlabel('X')
14. plt.ylabel('Y')
15. plt.legend(fontsize='xx-large')

```

Fonte: elaborado pelo autor.

No exemplo elaborado, disponível na Figura 1, foram utilizadas três bibliotecas: *numpy*, que dá acesso à estrutura de dados de vetores; *sklearn*, que possui a função *LinearRegression*, utilizada para obter os parâmetros da Equação 3; e a *matplotlib*, a qual

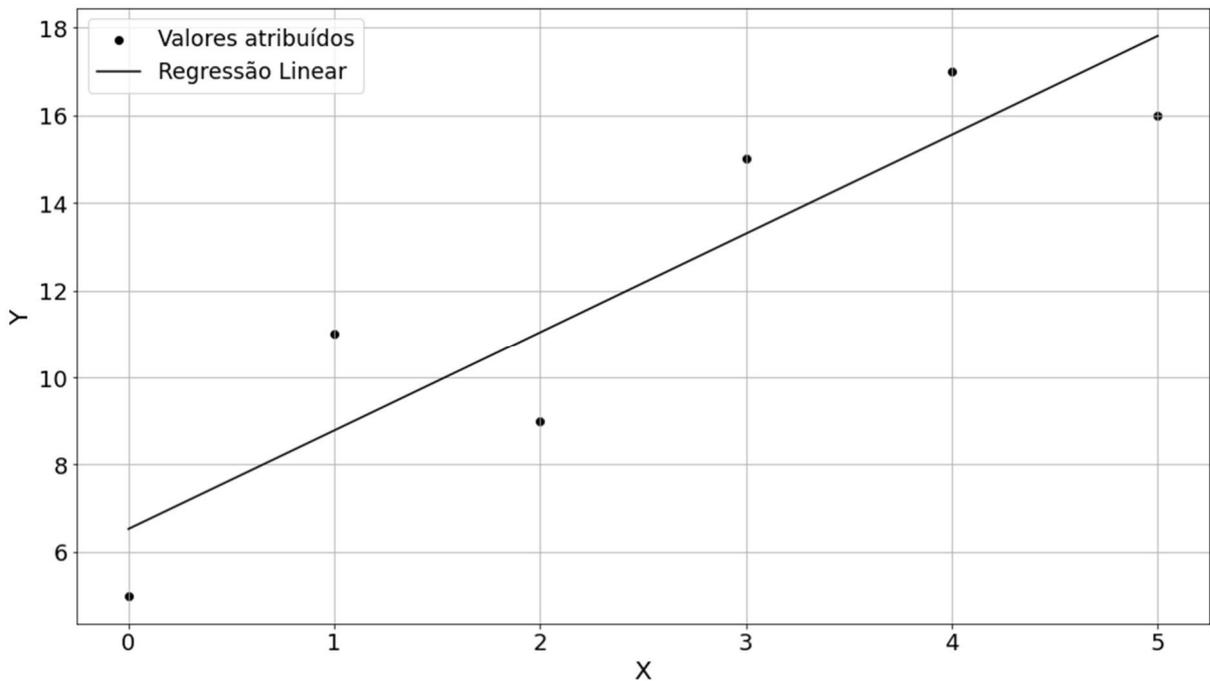
permite visualizar os dados através de um gráfico. As linhas de código 1 a 3 foram usadas para importar essas bibliotecas.

Em seguida, nas linhas 4 a 7, são criadas as variáveis `valores_X` e `valores_Y`, com valores atribuídos aleatoriamente para o exemplo. Nelas, utiliza-se a estrutura de vetor, chamada de *array* da biblioteca *numpy*. Por fim, transpõe-se o vetor `valores_X` e o passa para a variável `x`, o mesmo procedimento é realizado para `valores_Y`, agora chamado de `y`. O intuito desta etapa é separar os valores iniciais dos valores utilizados no modelo de regressão.

Nas linhas 8 e 9, tem-se a criação do modelo e sua execução com as variáveis `x` e `y`, criadas anteriormente.

Com isso, tem-se os parâmetros da reta obtida através do código `reg_lin.predict(x)`, e utiliza-se isto para criar um gráfico (ver Gráfico 1) que permita a visualização do resultado obtido a partir da sequência explicada. As linhas 10 a 15 mostram o código elaborado para exibir os dados iniciais e a reta obtida com a regressão.

Gráfico 1 – Reta de regressão linear obtida



Fonte: elaborado pelo autor.

2.5.3 Método dos Mínimos Quadrados

O método dos mínimos quadrados (MMQ) consiste em uma forma de estimar os coeficientes de uma função de regressão que se deseja obter, podendo ser utilizado tanto para regressões lineares quanto para modelos mais complexos, como a regressão polinomial.

Segundo Morettin e Bussab (2010), para se obter os parâmetros através do MMQ, deve-se fazer algumas suposições sobre as variáveis envolvidas: a variável x é controlada e não varia aleatoriamente, a média dos erros é zero, os erros não são correlacionados e possuem a mesma variabilidade em torno de x .

Com o objetivo de aproximar n pontos amostrais (x_i, y_i) , $i = 1, 2, \dots, j$, através de um polinômio de grau n , tem-se, inicialmente, uma função como a Equação 4, em que \hat{y}_i é o valor esperado para a variável y na observação i , ou seja, é o valor de y da curva a ser obtida, α_n são os coeficientes que se desejam encontrar e x_i é o valor observado (SILVA; MATTOS, 2017).

$$\hat{y}_i = \alpha_0 + \alpha_1 x_i + \alpha_2 x_i^2 + \dots + \alpha_n x_i^n \quad (4)$$

Todavia, para cada ponto observado, y_i está a uma distância ε_i da curva de regressão, que se pode considerar o erro que se deseja diminuir. Com isso, tem-se agora a Equação 5 e, a partir dela, obtém-se a Equação 6, na qual deseja-se encontrar o valor mínimo de S para se obter os coeficientes (MATUCHESKI, 2021).

$$y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 x_i^2 + \dots + \alpha_n x_i^n + \varepsilon_i \quad (5)$$

$$S = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

Por fim, ao substituir as expressões das Equações 4 e 5 na Equação 6, derivar S em relação à $\alpha_0, \dots, \alpha_n$, e igualar as derivadas a zero, obtém-se o sistema linear $AX = B$ (ver Figura 2) que, quando resolvido, encontram-se os coeficientes de regressão.

Figura 2 – Sistema linear obtido

$$\begin{bmatrix} \sum_{i=1}^n 1 & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \dots & \sum_{i=1}^n x_i^m \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 & \dots & \sum_{i=1}^n x_i^{m+1} \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 & \sum_{i=1}^n x_i^5 & \dots & \sum_{i=1}^n x_i^{m+2} \\ \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 & \sum_{i=1}^n x_i^5 & \sum_{i=1}^n x_i^6 & \dots & \sum_{i=1}^n x_i^{m+3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_i^m & \sum_{i=1}^n x_i^{m+1} & \sum_{i=1}^n x_i^{m+2} & \sum_{i=1}^n x_i^{m+3} & \dots & \sum_{i=1}^n x_i^{2m} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n y_i x_i \\ \sum_{i=1}^n y_i x_i^2 \\ \sum_{i=1}^n y_i x_i^3 \\ \vdots \\ \sum_{i=1}^n y_i x_i^m \end{bmatrix}$$

Fonte: adaptado de (MATUCHESKI, 2021).

2.5.4 Regressão Polinomial no Python

Com o sistema obtido na Figura 2 da Seção 2.5.3, percebe-se que o auxílio de ferramentas computacionais é essencial para se desenvolver curvas polinomiais a partir de um grande volume de dados. Com isso, esta seção visa mostrar, de forma sucinta, como obter uma curva de regressão polinomial utilizando a linguagem de programação *Python*.

Figura 3 – Código de exemplo para regressão polinomial

```

1. import numpy as np
2. import matplotlib.pyplot as plt
3. x = np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10])
4. y = np.array([4, 5, 10, 22, 28, 40, 50, 76, 92, 105])
5. reg_pol = np.polyfit(x, y, 2)
6. eq = np.poly1d(reg_pol)
7. plt.figure(1, figsize = [16,9])
8. plt.scatter(x,y, color='k', label = 'Valores atribuídos')
9. plt.plot(x, eq(x), color='k', label = 'Regressão polinomial')
10. plt.legend(fontsize='xx-large')
11. plt.xlabel('X')
12. plt.ylabel('Y')

```

Fonte: elaborado pelo autor.

No exemplo elaborado, disponível na Figura 3, utilizou-se as bibliotecas *numpy* e *matplotlib* do *Python*. Com a primeira, é possível declarar as listas com os valores para as variáveis *x* e *y*, e também é através desta biblioteca que se pode gerar o modelo de regressão polinomial. A segunda biblioteca, por sua vez, é utilizada para gerar o gráfico que permite a visualização dos dados e da curva de regressão obtida.

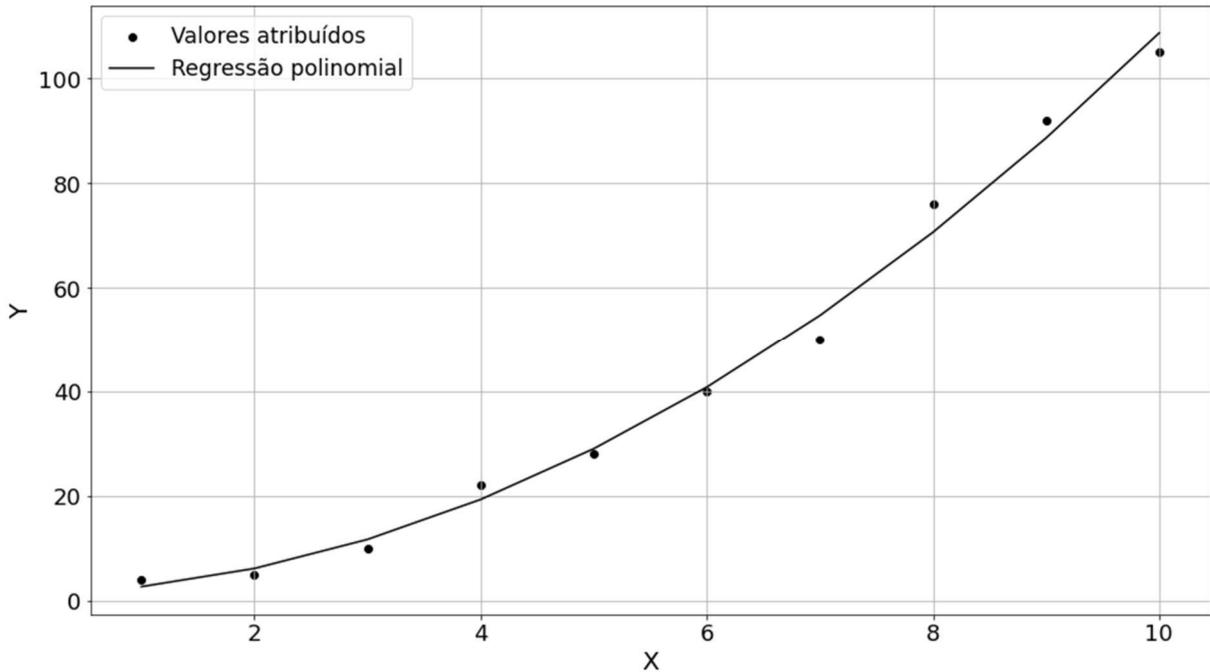
Nas linhas 1 e 2, tem-se o código para importar as bibliotecas. Já nas linhas 3 e 4, vê-se os valores que foram atribuídos às variáveis do exemplo.

Nas linhas 5 e 6, tem-se, a criação do modelo de regressão polinomial, a qual recebe como parâmetros as variáveis *x*, *y* e o grau do polinômio que se deseja obter, bem como a equação criada a partir das constantes obtidas no modelo de regressão, sendo necessária para a criação do gráfico.

Para finalizar o exemplo, tem-se, nas linhas 7 a 12, o código utilizado para gerar o gráfico de visualização do modelo, que pode ser visto no Gráfico 2. Por motivos didáticos,

escolheu-se os valores atribuídos às variáveis de modo que estes fiquem próximos à uma parábola, de modo que o grau 2 escolhido para o polinômio de regressão seja o mais adequado para este caso.

Gráfico 2 – Curva de regressão polinomial obtida



Fonte: elaborado pelo autor.

2.6 Métricas de Performance

Para que os resultados dos modelos obtidos tenham significado, faz-se necessário o uso de alguma ferramenta matemática que possa quantificar o quanto cada gráfico de regressão representa os dados analisados. Com isso, foram escolhidos dois modelos amplamente utilizados na avaliação de modelos de regressão, a Raiz do Erro Quadrático Médio e o Coeficiente de Determinação Ajustado, os quais são abordados nas seções 2.6.1 e 2.6.2.

2.6.1 Raiz do Erro Quadrático Médio

A raiz do erro quadrático médio (RMSE, do inglês *root mean squared error*) é utilizada para medir a precisão de modelos matemáticos. O cálculo consiste na raiz da média do quadrado da diferença entre um valor estimado (Θ_i) e um valor medido (θ_i). A Equação 7 mostra o cálculo do RMSE (MORETTIN; BUSSAB, 2010).

Este método de avaliação é mais sensível à valores muito fora do padrão, também conhecidos como *outliers*, uma vez que a diferença do erro é elevada ao quadrado. Contudo, por ser uma métrica que assume a mesma unidade de medida dos dados analisados, pode facilitar a interpretação dos resultados e do comportamento dos valores estudados.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\theta_i - \hat{\theta}_i)^2}{n}} \quad (7)$$

2.6.2 Coeficiente de Determinação Ajustado

O Coeficiente de Determinação Ajustado (ou R^2 Ajustado) é uma outra métrica frequentemente utilizada para avaliar como modelos de regressão se adequam aos dados, indicando, de forma percentual, quanta variação dos dados é levada em consideração no modelo (MONTGOMERY; RUNGER, 2009).

O Coeficiente de Determinação, sem ajuste, é encontrado a partir da soma dos quadrados total (SQ_t), dos resíduos (SQ_{res}) e explicada (SQ_{ex}), de modo que, para um número amostral n , sendo y_i um valor observado, \bar{y}_i a média dos valores observados, e \hat{y}_i o valor estimado de y_i , tem-se as Equações 8 a 11 (MONTGOMERY; RUNGER, 2009).

$$SQ_t = \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad (8)$$

$$SQ_{res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

$$SQ_{ex} = \sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2 \quad (10)$$

$$R^2 = \frac{SQ_{ex}}{SQ_t} = 1 - \frac{SQ_{res}}{SQ_t} \quad (11)$$

Por fim, com o valor de R^2 encontrado na Equação 11, pode-se fazer um ajuste que permite o uso desta ferramenta matemática para casos com múltiplas variáveis explicativas, as quais podem distorcer o resultado obtido para R^2 , uma vez que este pode aumentar mesmo que as variáveis adicionais não sejam significativas para explicar o comportamento dos dados. A Equação 12 mostra o cálculo de R^2 Ajustado, em que k é o número de variáveis independentes.

$$\bar{R}^2 = 1 - \frac{(1-R^2)(n-1)}{n-k+1} \quad (12)$$

3 METODOLOGIA E RESULTADOS

Com a tecnologia atual, tem-se que a maioria das pessoas e empresas usufruem de alguma forma da internet. Essa informatização vem gerando um volume de dados cada vez maior, que precisam ser armazenados de forma que possam também ser analisados. Por conta disso, a popularização dos bancos de dados deu origem ao termo “*big data*”, o qual resume uma complexa gama de assuntos e metodologias para o tratamento e processamento de grandes volumes de dados. (CALDAS; SILVA, 2016).

Nesta seção, serão apresentados os dados utilizados, disponibilizados pela Empresa de Pesquisa Energética (EPE), o processamento feito em cima desses dados e também os resultados obtidos, com os gráficos dos modelos de regressão e a *performance* de cada um deles.

3.1 Dados Utilizados

Como já mencionado, os dados foram obtidos através do *site* da EPE, a qual disponibiliza um Painel de Dados de Micro e Minigeração Distribuída (PDGD) com as principais métricas a respeito da GD no Brasil, de 2013 a 2021.

O arquivo é disponibilizado no formato CSV (*comma-separated values*, em tradução livre, valores separados por vírgulas), apresentando as seguintes colunas de dados:

- “ano”: ano de instalação da unidade geradora;
- “nome_4md”: nome da concessionária de energia responsável;
- “classe”: classe de consumo do consumidor, podendo ser Poder Público, Residencial, Comercial, Industrial, Rural, Serviço Público ou Iluminação Pública;
- “subgrupo”: subgrupo do consumidor;
- “modalidade”: identifica se a instalação é geração na própria UC, autoconsumo remoto, geração compartilhada ou condomínio;
- “uf”: sigla do estado;
- “município”: nome do município;
- “fonte_resumo”: fonte de geração energética, podendo ser fotovoltaica, eólica, termelétrica ou hidroelétrica;
- “mini_micro”: identifica se é micro ou minigeração;
- “subsistema”: subsistema no qual a usina foi conectada;

- “regiao”: sigla com a região do país;
- “segmento”: identifica se o consumidor é comercial ou residencial, alta tensão ou baixa tensão, remoto ou não;
- “pot_grupo”: faixa de potência da usina;
- “atbt”: identifica se é alta tensão ou baixa tensão;
- “potencia_mw”: potência instalada da usina, em MW;
- “qtde_u_csrecebem_os_creditos”: quantidade de unidades consumidoras que recebem créditos da usina;
- “n_sistemas”: número de sistemas interligados;
- “investimento_milhoes”: valor estimado da usina, em milhões de reais.

3.2 Processamento de Dados

Para realizar a manipulação dos dados e a criação dos modelos de regressão, utilizou-se a linguagem de programação *Python*, através do *software JupyterLab*, os quais já foram brevemente apresentados nas Seções 2.5.2 e 2.5.4. Dentro do software escolhido, o trabalho foi desenvolvido utilizando o recurso de *Notebooks*, os quais permitem a execução de blocos de código dentro do mesmo arquivo.

Com os dados carregados no *script*, é necessário garantir a consistência dos dados e, se necessário, realizar correções, a fim de evitar utilizar intervalos com muitos dados faltantes ou com erros de grafia.

Neste caso, nenhuma correção foi necessária, pois não houveram campos com dados nulos e nem problemas de grafia. Com isso, na primeira etapa do tratamento dos dados, foram separadas as colunas relevantes para a criação dos modelos: “ano”, “uf”, “fonte_resumo” e “potencia_mw”. Em seguida, foram selecionadas apenas as linhas que apresentaram “Fotovoltaica” como parâmetro da coluna “fonte_resumo”, para que as outras fontes geradoras não influenciassem nos gráficos obtidos.

Por fim, foi criado um *script* que será executado para cada um dos estados, que segue a estrutura dos *scripts* apresentados nas Seções 2.5.2 e 2.5.4, mais o cálculo das métricas escolhidas. Para cada estado específico, o código realiza a soma cumulativa das potências instaladas em cada ano, chama as funções de regressão escolhidas para este trabalho (linear e polinomial com grau 2), calcula as métricas de performance R^2 Ajustado e Raiz do Erro Quadrático Médio e gera os gráficos apresentados na Seção 3.3.

3.3 Gráficos de Regressão

Com o *script* desenvolvido, foram gerados onze gráficos, sendo um para o país (ver Gráfico 3) e os restantes para cada uma das regiões, separados entre modelo polinomial e modelo linear. Todos estes serão apresentados nesta seção, organizados em tópicos por região.

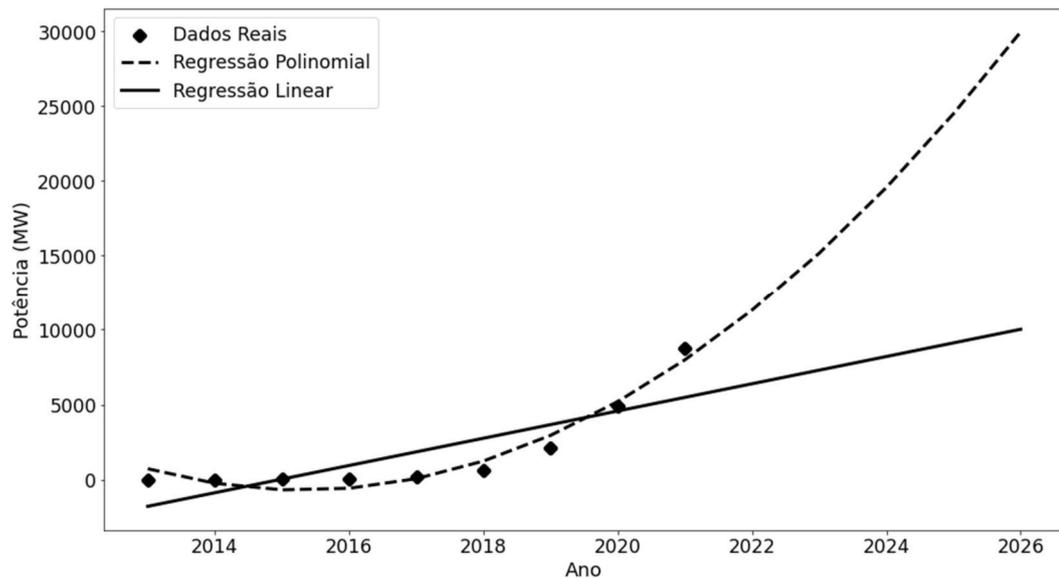
Nos gráficos desta sessão, pode-se ver os valores reais de potência instalada acumulada (pontos em formato de losango) de micro e miniusinas fotovoltaicas. A linha reta, no Gráfico 3, exibe a previsão linear nacional, enquanto que, nos Gráficos 4 a 13, as linhas pontilhadas exibem tanto as curvas como as retas de regressão obtidas, a depender do que for especificado na legenda de cada gráfico.

No Gráfico 3, tem-se as informações no contexto nacional. Neste gráfico, percebe-se como as previsões entre os modelos são discrepantes, uma vez que a regressão polinomial indica, em 2026 uma potência aproximadamente três vezes maior que a de 2021, enquanto a regressão linear prevê para o futuro um valor que já foi atingido.

Na Seção 3.4 são mostradas as medidas de R^2 Ajustado e Raiz do Erro Quadrático Médio, a fim de determinar a confiabilidade de cada modelo.

As equações das retas e curvas obtidas no *script* estão disponíveis nos Apêndices A e B.

Gráfico 3 – Previsão de Potência Instalada no Brasil



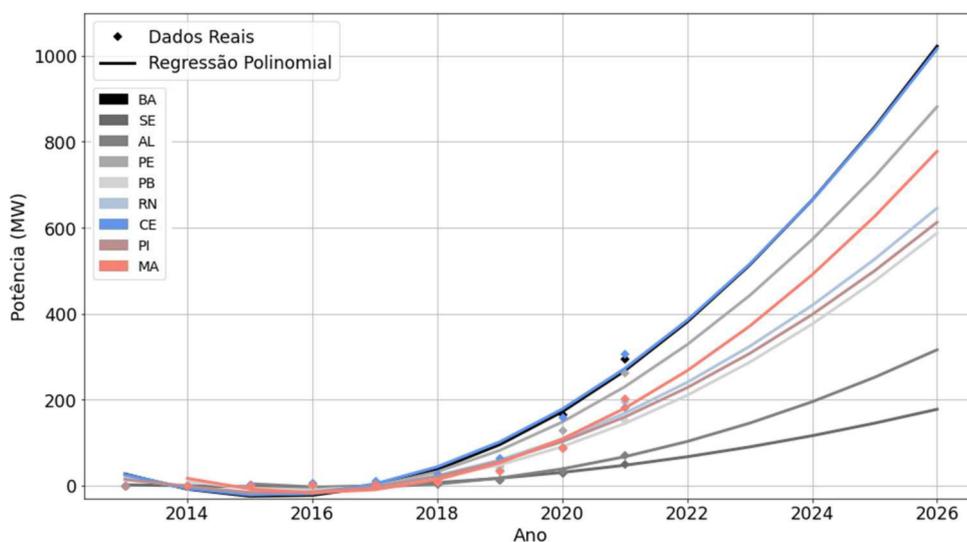
Fonte: elaborado pelo autor.

3.3.1 Região Nordeste

No Nordeste, terceira maior região tratando-se de GD fotovoltaica, se destaca o estado do Ceará, sendo o que possui a maior potência instalada no ramo, seguido pela Bahia e por Pernambuco, como o segundo e o terceiro lugar, respectivamente.

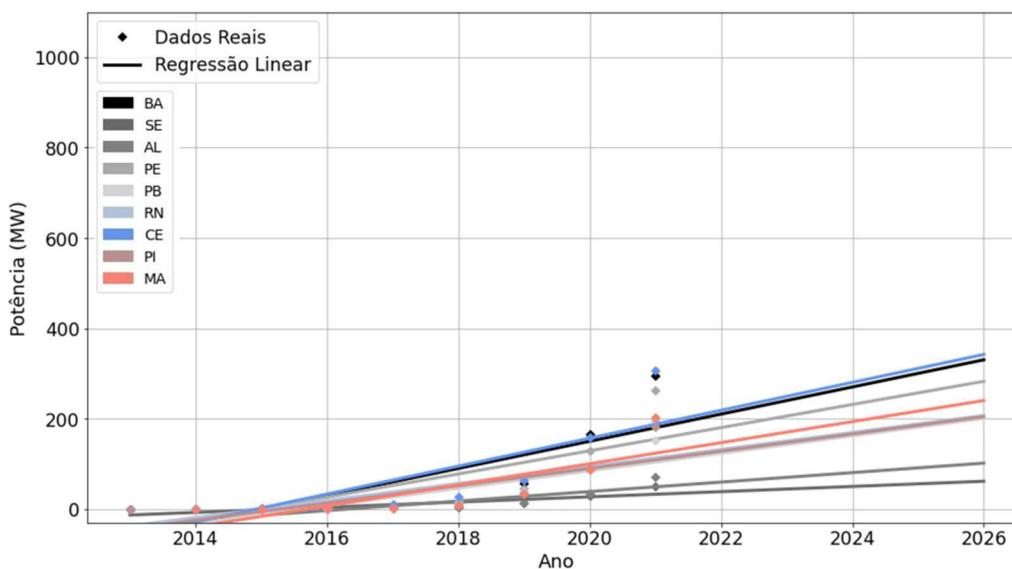
Nos Gráficos 4 e 5 tem-se os resultados detalhados para os estados desta região.

Gráfico 4 – Curvas de Previsão de Potência Instalada no Nordeste



Fonte: elaborado pelo autor.

Gráfico 5 – Retas de Previsão de Potência Instalada no Nordeste



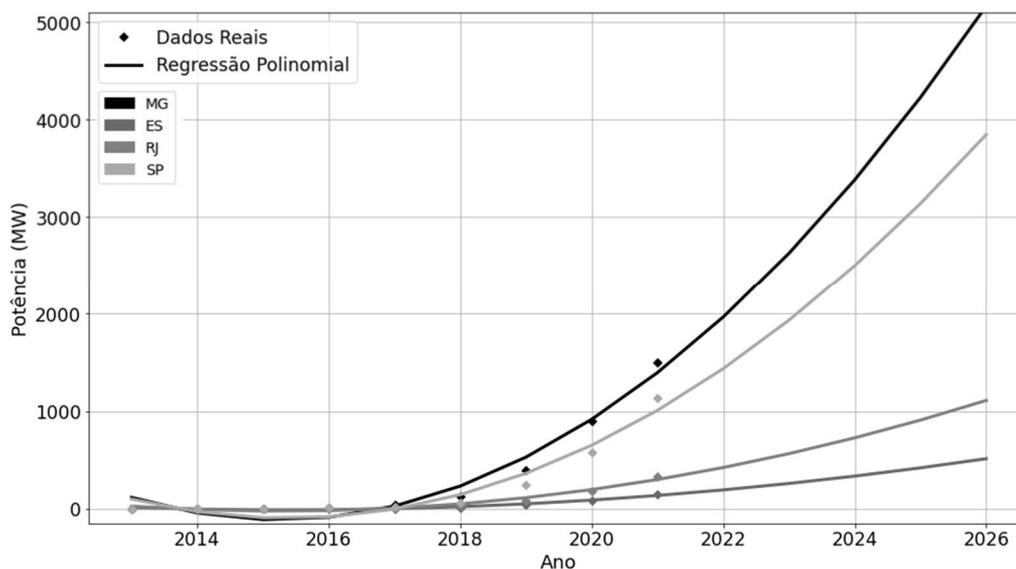
Fonte: elaborado pelo autor.

3.3.2 Região Sudeste

A região Sudeste se destaca por ser a maior produtora de energia através desta modalidade, com Minas Gerais e São Paulo sendo os maiores tanto regionalmente como nacionalmente.

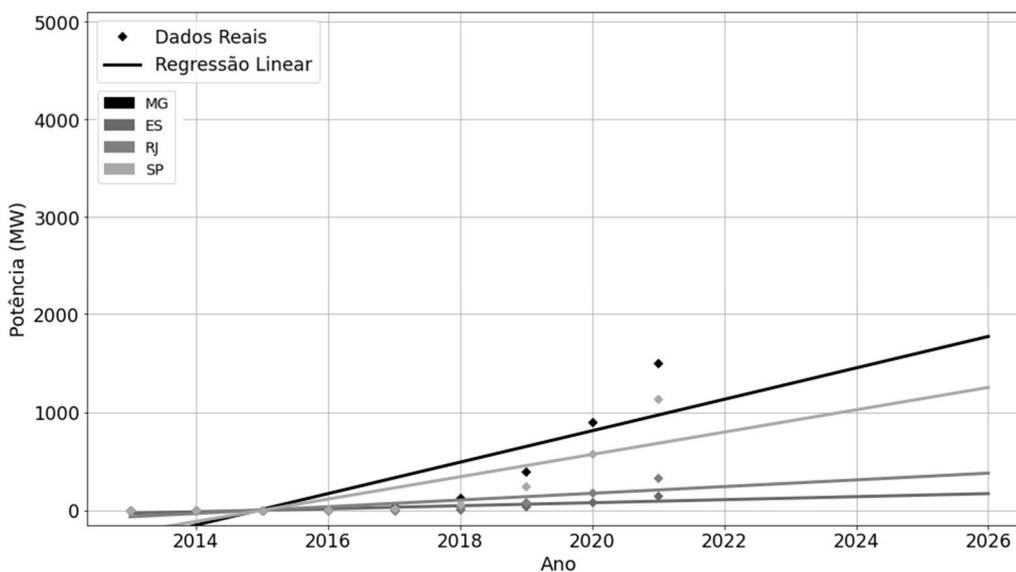
Os Gráficos 6 e 7 mostram os detalhes para estes e os outros estados desta região.

Gráfico 6 – Curvas de Previsão de Potência Instalada no Sudeste



Fonte: elaborado pelo autor.

Gráfico 7 – Retas de Previsão de Potência Instalada no Sudeste



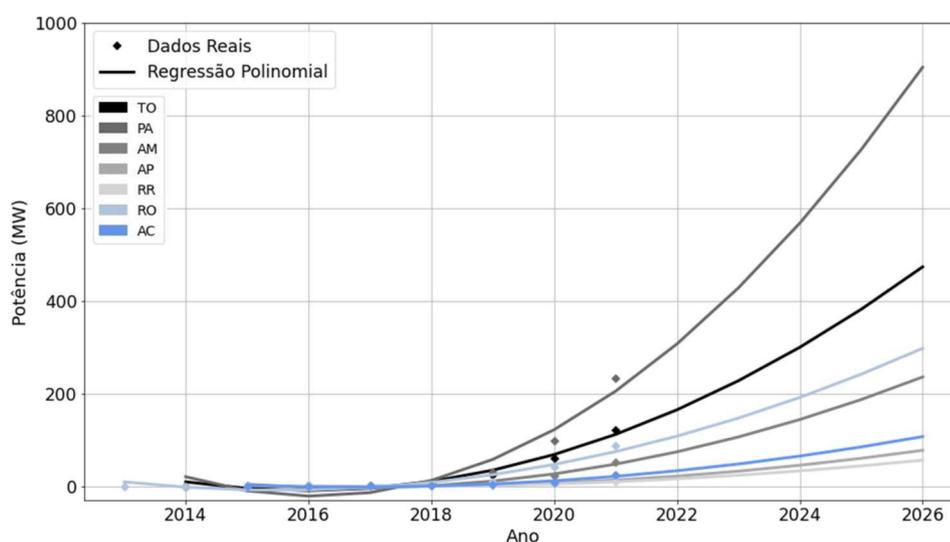
Fonte: elaborado pelo autor.

3.3.3 Região Norte

A região Norte, por sua vez, é a que possui menor participação da produção de energia através de micro e minigeração distribuída com painéis solares, sendo de lá os três estados com a menor potência instalada: Roraima, Amapá e Acre. Destaca-se o Pará, que fica em 13º no ranking nacional.

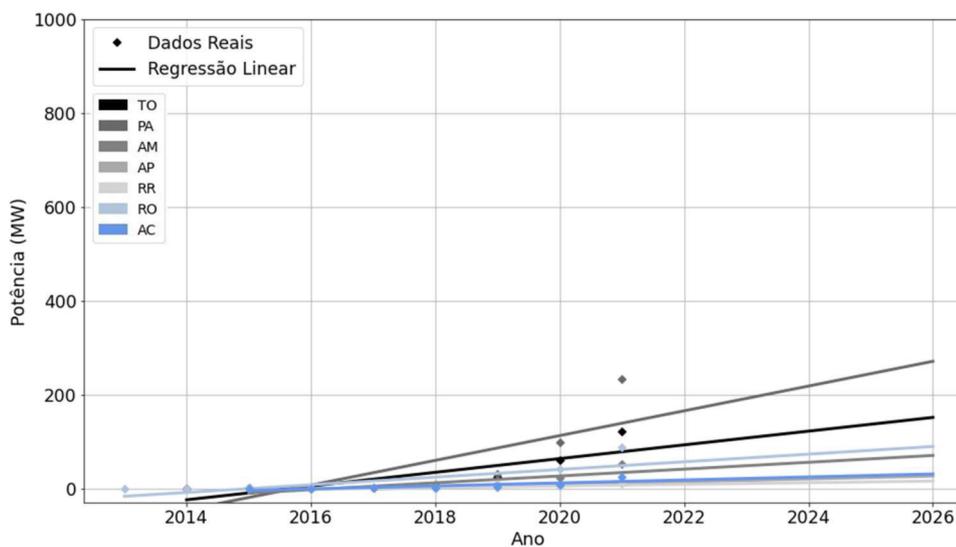
Nos Gráficos 8 e 9, tem-se os resultados para todos os estados

Gráfico 8 – Curvas de Previsão de Potência Instalada no Norte



Fonte: elaborado pelo autor.

Gráfico 9 – Retas Previsão de Potência Instalada no Norte



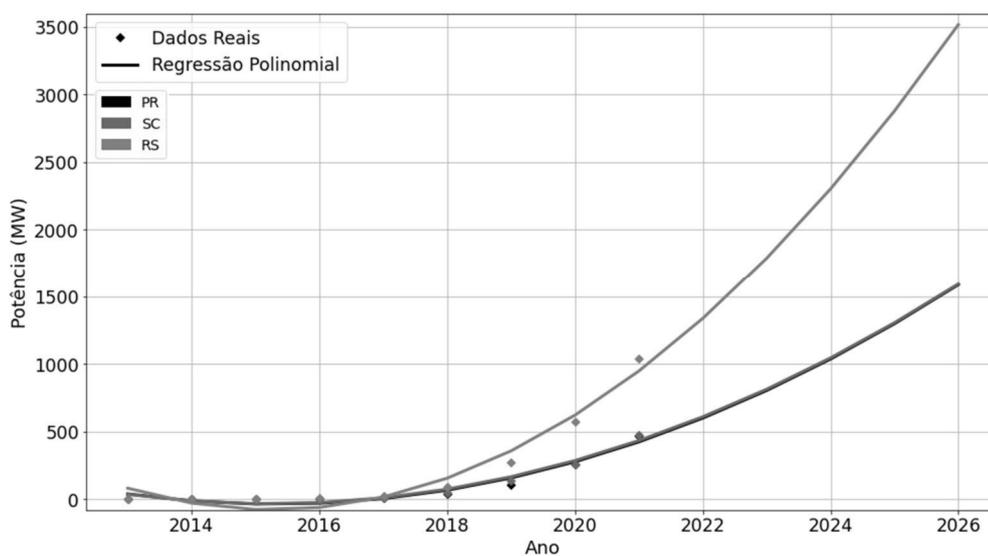
Fonte: elaborado pelo autor.

3.3.4 Região Sul

A região Sul fica atrás apenas da região Sudeste no ranking, com os estados do Rio Grande do Sul, Santa Catarina e Paraná sendo o 3º, 5º e 6º colocados, respectivamente, no ranking nacional.

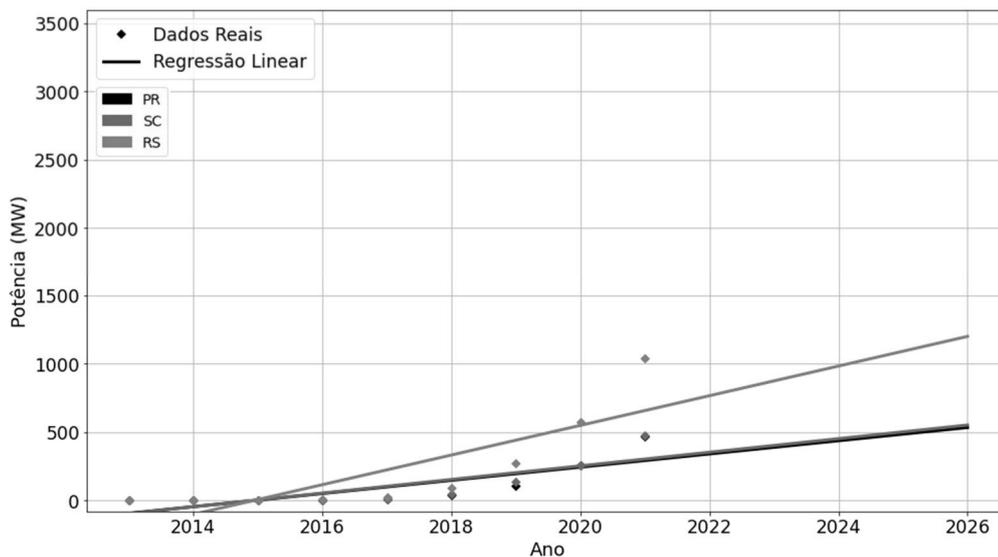
Os Gráficos 10 e 11 mostram os modelos obtidos para esses estados.

Gráfico 10 – Curvas de Previsão de Potência Instalada no Sul



Fonte: elaborado pelo autor.

Gráfico 11 – Retas de Previsão de Potência Instalada no Sul



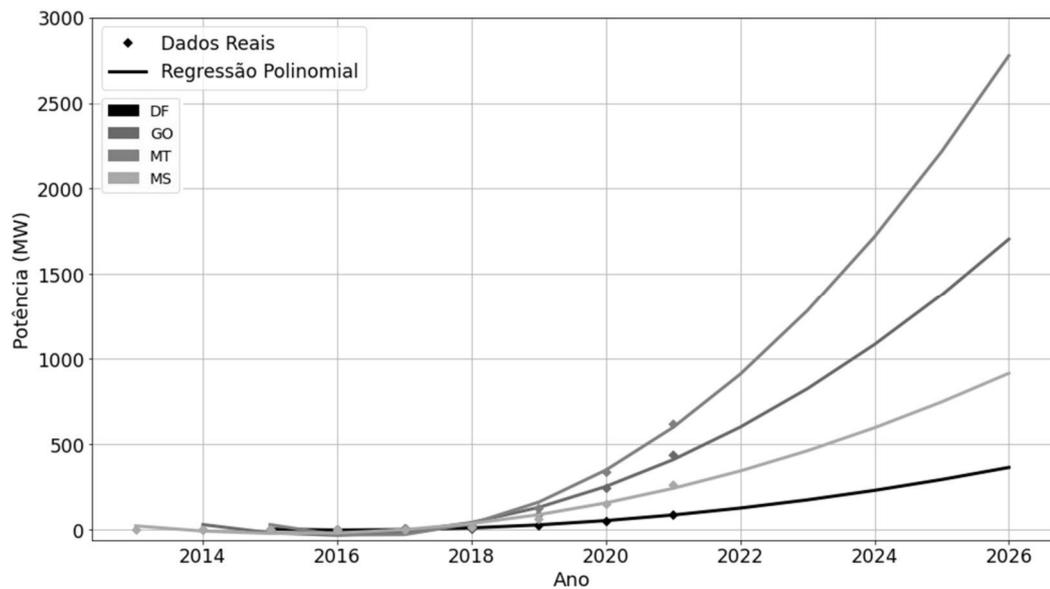
Fonte: elaborado pelo autor.

3.3.5 Região Centro-Oeste

O Centro-Oeste é a penúltima região no ranking de GD fotovoltaica, contudo, o estado do Mato Grosso se destaca, tendo a quarta maior potência instalada do país.

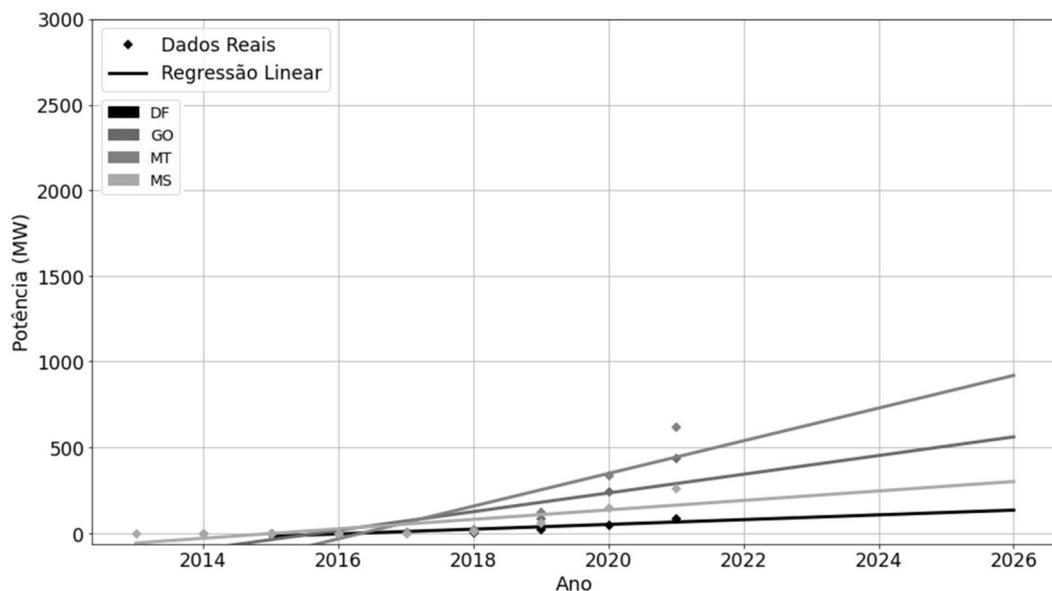
Os resultados obtidos desta região estão nos Gráficos 12 e 13.

Gráfico 12 – Curvas de Previsão de Potência Instalada no Centro-Oeste



Fonte: elaborado pelo autor.

Gráfico 13 – Retas de Previsão de Potência Instalada no Centro-Oeste



Fonte: elaborado pelo autor.

3.4 Métricas e Discussões

Com os gráficos da Seção 3.3, tem-se a confirmação de que todos os estados nacionais tiveram crescimento na potência instalada nos últimos anos. Esses valores tendem a continuar crescendo, mas existe a dificuldade em determinar o quanto eles crescerão.

Visualmente, pode-se chegar à conclusão de que a reta de regressão linear obtida não é apropriada para estimar o comportamento dos dados no intervalo escolhido. Entretanto, faz-se necessário utilizar alguma métrica que confirme ou conteste essa afirmação, por isso, para avaliar o quanto cada modelo pode explicar os dados reais, usou-se as medidas de R^2 Ajustado e a Raiz do Erro Quadrático Médio, explicados nas Seções 2.6.1 e 2.6.2.

As Tabelas 1 e 2 mostram os valores obtidos em relação ao Gráfico 3, referente aos valores totais do Brasil, enquanto as tabelas 3 e 4 mostram as métricas para os resultados dos estados.

Tabela 1 – Valores de R^2 Ajustado - Brasil

País	Linear	Polinomial	Dif. Percentual
Brasil	0,611	0,949	36%

Fonte: elaborado pelo autor.

Tabela 2 – Valores da Raiz do Erro Médio Quadrático - Brasil

País	Linear	Polinomial	Dif. Percentual
Brasil	1681,869	606,849	64%

Fonte: elaborado pelo autor.

Com os valores das Tabelas 1 e 2, tem-se que, em ambas as métricas utilizadas, a curva de regressão polinomial representa melhor os dados estudados, tendo um valor de R^2 Ajustado 36% maior que a regressão linear e ficando bem próximo do valor máximo 1. Se tratando do RMSE, o erro para a curva de regressão polinomial foi de pouco mais de um terço do erro calculado para a reta.

Os valores de diferença percentual das Tabelas 1 e 2 representam a média dos estados, todavia, não há correlação entre os valores de cada uma das diferentes métricas, de modo que estes não devem ser comparados individualmente. Porém, ambos permitem concluir que a regressão polinomial representou os dados de forma mais precisa.

Tabela 3 – Valores de R² Ajustado - Estados

Estado	Linear	Polinomial	Dif. Percentual
Ceará	0,604	0,934	35%
Bahia	0,579	0,935	38%
Sergipe	0,573	0,964	41%
Alagoas	0,669	0,971	31%
Pernambuco	0,560	0,916	39%
Paraíba	0,675	0,976	31%
Rio Grande do Norte	0,557	0,910	39%
Piauí	0,551	0,920	40%
Maranhão	0,576	0,928	38%
Minas Gerais	0,635	0,959	34%
Espírito Santo	0,604	0,946	36%
Rio de Janeiro	0,624	0,951	34%
São Paulo	0,582	0,933	38%
Tocantins	0,616	0,954	35%
Pará	0,553	0,922	40%
Amazonas	0,605	0,950	36%
Amapá	0,790	0,989	20%
Roraima	0,553	0,890	38%
Rondônia	0,500	0,887	44%
Acre	0,554	0,914	39%
Paraná	0,614	0,949	35%
Santa Catarina	0,641	0,960	33%
Rio Grande do Sul	0,628	0,956	34%
Distrito Federal	0,769	0,989	22%
Goiás	0,643	0,964	33%
Mato Grosso	0,697	0,983	29%
Mato Grosso do Sul	0,603	0,949	36%
Média	0,613	0,944	36%

Fonte: elaborado pelo autor.

Tabela 4 – Valores da Raiz do Erro Médio Quadrático - Estados

Estado	Linear	Polinomial	Dif. Percentual
Ceará	58,056	23,645	59%
Bahia	59,224	23,298	61%
Sergipe	10,625	3,085	71%
Alagoas	12,959	3,865	70%
Pernambuco	52,376	22,864	56%
Paraíba	27,490	7,452	73%
Rio Grande do Norte	38,648	17,411	55%
Piauí	37,672	15,905	58%
Maranhão	40,340	16,579	59%
Minas Gerais	282,957	94,384	67%
Espírito Santo	29,034	10,729	63%
Rio de Janeiro	61,741	22,243	64%
São Paulo	222,818	89,394	60%
Tocantins	23,460	8,120	65%
Pará	47,639	19,903	58%
Amazonas	10,210	3,626	64%
Amapá	1,869	0,432	77%
Roraima	2,184	1,081	51%
Rondônia	18,596	8,822	53%
Acre	4,896	2,151	56%
Paraná	89,040	32,328	64%
Santa Catarina	86,943	28,988	67%
Rio Grande do Sul	194,373	67,226	65%
Distrito Federal	13,498	2,930	78%
Goiás	82,515	26,240	68%
Mato Grosso	110,310	25,864	77%
Mato Grosso do Sul	51,695	18,481	64%
Média	-	-	64%

Fonte: elaborado pelo autor.

Nas Tabelas 3 e 4, é possível observar o resultado individual por estado para cada uma das equações obtidas. Pode-se destacar o caso do Amapá, que possui a menor diferença

percentual de R^2 Ajustado, o que pode ter sido causado pelo número menor de anos considerados, por conta da falta de dados ou pelos dados serem nulos. Já em Rondônia tem-se o oposto, pois este foi o estado com maior variação, de modo que a regressão linear foi a menos precisa.

Na tabela com os valores de RMSE, não é possível comparar os valores absolutos entre os estados, uma vez que estes variam com a amplitude dos dados de cada estado. Entretanto pode-se ver que Amapá e Mato Grosso foram os estados com maior variação entre os modelos, enquanto Roraima foi o que teve menos variação.

Nos estados com maior variação há, portanto, um comportamento que se afasta mais do comportamento linear e aproxima-se de uma curva parabólica, enquanto que, nos estados com menor variação entre os modelos de regressão, os dados tendem a se comportar um pouco mais próximo de uma reta.

4 CONCLUSÃO

Com os resultados obtidos, fica claro o expressivo crescimento da Geração Distribuída Fotovoltaica no Brasil. As inovações neste setor têm permitido que mais pessoas usufruam de sua própria geração de energia, economizando e contribuindo de forma positiva para o meio ambiente. Contudo, essas novidades ainda são muito recentes, de modo que o mercado deste setor tende a crescer muito mais, gerando novos empregos e criando demanda por fontes de qualificação profissional e infraestrutura em áreas remotas que sejam vantajosas para essa tecnologia.

Também pode-se notar a importância da análise de dados e a quantidade de informações que podem ser obtidas através de um banco de dados relativamente simples, uma vez que modelos preditivos contribuem para um crescimento mais organizado e eficiente do mercado e da indústria, evitando surpresas e gargalos de produção pela falta de insumos.

Os modelos utilizados neste trabalho, por sua vez, são apenas uma introdução prática ao tema e não podem ser considerados uma verdade absoluta como forma de orientação à investimentos. As equações mostram a tendência a partir dos valores já existentes, não considerando parâmetros externos que podem alterar bruscamente o comportamento dos dados.

Ademais, a linguagem de programação *python* e as bibliotecas disponíveis para a análise de dados também se demonstraram muito eficazes, uma vez que o script pôde ser feito em poucas linhas, utilizando na maioria das vezes funções já existentes.

Em trabalhos futuros, é possível revisar os modelos de regressão escolhidos utilizando configurações diferentes, como com um intervalo de tempo menor, selecionando apenas os dados de anos mais recentes, ou utilizando polinômios com mais graus, tendo o cuidado para evitar o super ajuste das curvas. É possível também aumentar a complexidade e a confiabilidade do sistema com a criação de modelos preditivos utilizando redes neurais, as quais permitiriam adicionar mais parâmetros que possam influenciar no comportamento dos dados.

Além disso, é possível também utilizar a metodologia de estudo deste trabalho com novos parâmetros que se deseja prever, como, por exemplo, a relação entre a curva de vida útil de um painel fotovoltaico correlacionada à um fator de sombreamento, ou a correlação entre a capacidade nominal de micro inversores em relação aos inversores convencionais.

REFERÊNCIAS

- BALANÇO de Energia. **ONS**. 2022. Disponível em: <http://www.ons.org.br/paginas/energia-agora/balanco-de-energia>. Acesso em: 18 set. 2022.
- BALANÇO Energético Nacional 2022. **EPE**. 2022. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2022>. Acesso em: 12 set. 2022.
- BARROS, R.; SAUAIA, R.; KOLOSZUK, R. Perspectivas para a Geração Centralizada Solar Fotovoltaica no Brasil. **ABSOLAR**. 2019. Disponível em: <https://www.absolar.org.br/artigos/artigo-perspectivas-para-a-geracao-centralizada-solar-fotovoltaica-no-brasil/>. Acesso em: 21 set. 2022.
- BASTOS, E. V. P.; GUIMARÃES, J. C. F. de; SEVERO, E. A. Linear regression model for investment analysis of an oil company. **Revista Produção e Desenvolvimento**, [S. l.], v. 1, n. 1, p. 77–88, 2015. DOI: 10.32358/rpd.2015.v1.62. Disponível em: <https://revistas.cefet-rj.br/index.php/producaoedesenvolvimento/article/view/e62>. Acesso em: 1 out. 2022.
- BOBBY. **Why Python is Popular Despite Being (Super) Slow**. 2019. Disponível em: <https://medium.com/@trungluongquang/why-python-is-popular-despite-being-super-slow-83a8320412a9>. Acesso em: 18 dez. 2022.
- CALDAS, M. S.; CLAUDINO SILVA, E. C. Fundamentos e aplicação do Big Data: como tratar informações em uma sociedade de *yottabytes*. **Bibliotecas Universitárias: pesquisas, experiências e perspectivas**, [S. l.], v. 3, n. 1, 2016. Disponível em: <https://periodicos.ufmg.br/index.php/revistarbu/article/view/3086>. Acesso em: 19 nov. 2022.
- CHEIN, F. **Introdução aos modelos de regressão linear**: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas. Brasília: ENAP, 2019.

CUNHA, H.; SOBEL, L. Aplicação de Redes Neurais Artificiais na Predição de Geração de Energia Fotovoltaica no Nordeste do Brasil. **Revista de Engenharia e Pesquisa Aplicada**, v. 6, n. 5, p. 73-80, 19 nov. 2021.

DATASHEETS. **CanadianSolar**. 2021. Disponível em: <https://support.csisolar.com/hc/pt-br/articles/4413715629453-HiKu-CS3W-P-MS>. Acesso em: 13 nov. 2022.

DIAS, M. V. X. **Geração distribuída no Brasil: oportunidades e barreiras**. 2005. 143 f. Dissertação (Mestrado em Engenharia de Energia) - Instituto de Engenharia Mecânica, Universidade Federal de Itajubá, Itajubá, 2005.

ESTUDO Estratégico: Geração Distribuída 2022 | Mercado Fotovoltaico 1º Semestre. **Greener**. 2022. Disponível em: <https://www.greener.com.br/estudo/estudo-estrategico-geracao-distribuida-2022-mercado-fotovoltaico-1-semester/>. Acesso em: 29 nov. 2022.

GERAÇÃO Distribuída. **ANEEL**. 2022. Disponível em: <https://www.gov.br/aneel/pt-br/assuntos/geracao-distribuida>. Acesso em: 22 set. 2022.

HALLIDAY, D.; RESNICK, R.; WALKER, J. **Fundamentos de Física**. 10. ed. Rio de Janeiro: LTC, 2012. v.3.

KORTOV, V.; USTYANTSEV, Y. Chernobyl accident: Causes, consequences and problems of radiation measurements. **Radiation Measurements**, v. 55, p. 12-16, 2013.

MATUCHESKI, S. **Método dos Mínimos Quadrados**. 2021. 139 f. Dissertação (Mestrado em Matemática - Nível de Mestrado Profissional) - Universidade Estadual do Oeste do Paraná, Cascavel, 2021.

MONTGOMERY, D. C.; RUNGER, G. C. **Estatística Aplicada e Probabilidade para Engenheiros**. 4. ed. Rio de Janeiro: LTC, 2009.

MORETTIN, P. A.; BUSSAB, W. O. **Estatística Básica**. 6. ed. São Paulo: Saraiva, 2010.

MOST In-Demand Programming Languages. **Berkeley**. s.d. Disponível em: <https://bootcamp.berkeley.edu/blog/most-in-demand-programming->

languages/#:~:text=According%20to%20Stack%20Overflow's%202020,(PDF%2C%202.4%20MB). Acesso em: 18 dez. 2022.

O que é Geração Distribuída. **INEE**. s.d. Disponível em:

[http://www.inee.org.br/forum_ger_distrib.asp#:~:text=Gera%C3%A7%C3%A3o%20Distribu%C3%ADda%20\(GD\)%20%C3%A9%20uma,incluir%20pot%C3%Aancias%20cada%20vez%20menores](http://www.inee.org.br/forum_ger_distrib.asp#:~:text=Gera%C3%A7%C3%A3o%20Distribu%C3%ADda%20(GD)%20%C3%A9%20uma,incluir%20pot%C3%Aancias%20cada%20vez%20menores). Acesso em: 22 set. 2022.

O Sistema em Números. **ONS**. 2022. Disponível em: <http://www.ons.org.br/paginas/sobre-o-sin/o-sistema-em-numeros>. Acesso em: 18 set. 2022.

PAINEL de Dados de Micro e Minigeração Distribuída (PDGD). **EPE**. 2022. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/painel-de-dados-de-micro-e-minigeracao-distribuida-pdgd->. Acesso em: 12 set. 2022.

PEREIRA, E. B.; MARTINS, F. R.; GONÇALVES, A. R.; COSTA, R. S.; LIMA, F. L.; RÜTHER, R.; ABREU, S. L.; TIEPOLO, G. M.; PEREIRA, S. V.; SOUZA, J. G. **Atlas brasileiro de energia solar**. 2.ed. São José dos Campos: INPE, 2017. 80p. Disponível em: <http://doi.org/10.34024/978851700089>. Acesso em: 17 dez. 2022.

PINHO, J. T.; GALDINO, M. A. **Manual de Engenharia Para Sistemas Fotovoltaicos**. Centro de Pesquisas de Energia Elétrica. Rio de Janeiro, p. 530. 2014. Disponível em: <http://www.cresesb.cepel.br/index.php?section=publicacoes&task=livro&cid=481>. Acesso em: 15 nov. 2022.

RECICLAGEM de Painéis Solares. **Portal Solar**. s.d. Disponível em:

<https://www.portalsolar.com.br/reciclagem-de-paineis-solares>. Acesso em: 25 set. 2022.

RICHARDSON, L. The History of Solar Energy. **EnergySage**, 2022. Disponível em: <https://news.energysage.com/the-history-and-invention-of-solar-panel-technology/>. Acesso em: 12 set. 2022.

RITCHIE, H.; ROSER, M.; ROSADO, P. Energy. **Our World in Data**. 2022. Disponível em: <https://ourworldindata.org/energy>. Acesso em: 20 set. 2022.

SILVA, J. R. **Uma sequência de ensino investigativa para a abordagem do efeito fotovoltaico**. 2020. 130 f. Dissertação (Mestrado Nacional Profissional em Ensino de Física) - Instituto de Ciências Exatas, Universidade Federal Fluminense, Volta Redonda, 2020.

SILVA, V. M.; DE MATTOS, V. L. D. O método dos mínimos quadrados no ajuste de um modelo polinomial. **Scientia Plena**, [S. l.], v. 13, n. 4, 2017. DOI:

10.14808/sci.plena.2017.049910. Disponível em:

<https://scientiaplena.org.br/sp/article/view/3689>. Acesso em: 24 out. 2022.

TAVEIRA, I. M. **A influência da geração fotovoltaica no fator de potência de unidades prosumidoras: uma análise técnica e financeira**. 2019. 81 f. Monografia (Graduação em Engenharia Elétrica) – Universidade Federal do Ceará, Fortaleza, 2019.

URZAGASTI, C. A.; MACIEL, J. N.; WENTZ, V. H.; LEDESMA, J. J. G.; ANDO JUNIOR, O. H. Comparação da Acurácia de Modelos de Redes Neurais Artificiais na Predição da Irradiância Solar e Geração de Energia Fotovoltaica. *In*: CONGRESSO LATINO-AMERICANO DE SOFTWARE LIVRE E TECNOLOGIAS ABERTAS (LATINOWARE), 18. , 2021, Online. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 53-58. DOI: <https://doi.org/10.5753/latinoware.2021.19905>.

APÊNDICE A – EQUAÇÕES DA REGRESSÃO LINEAR

UF	Equações Lineares
Brasil	$(9,073 \cdot 10^2)x - 1,828 \cdot 10^6$
Ceará	$(3,089 \cdot 10)x - 6,223 \cdot 10^4$
Bahia	$(3,003 \cdot 10)x - 6,051 \cdot 10^4$
Sergipe	$(5,743)x - 1,157 \cdot 10^4$
Alagoas	$(1,049 \cdot 10)x - 2,115 \cdot 10^4$
Pernambuco	$(2,562 \cdot 10)x - 5,162 \cdot 10^4$
Paraíba	$(1,932 \cdot 10)x - 3,894 \cdot 10^4$
Rio Grande do Norte	$(1,881 \cdot 10)x - 3,790 \cdot 10^4$
Piauí	$(1,909 \cdot 10)x - 3,848 \cdot 10^4$
Maranhão	$(2,330 \cdot 10)x - 4,697 \cdot 10^4$
Minas Gerais	$(1,601 \cdot 10^2)x - 3,226 \cdot 10^5$
Espírito Santo	$(1,545 \cdot 10)x - 3,112 \cdot 10^4$
Rio de Janeiro	$(3,417 \cdot 10)x - 6,886 \cdot 10^4$
São Paulo	$(1,137 \cdot 10^2)x - 2,291 \cdot 10^5$
Tocantins	$(1,462 \cdot 10)x - 2,948 \cdot 10^4$
Pará	$(2,638 \cdot 10)x - 5,318 \cdot 10^4$
Amazonas	$(7,287)x - 1,469 \cdot 10^4$
Amapá	$(3,054)x - 6,161 \cdot 10^3$
Roraima	$(1,713)x - 3,456 \cdot 10^3$
Rondônia	$(8,160)x - 1,644 \cdot 10^4$
Acre	$(3,184)x - 6,421 \cdot 10^3$
Paraná	$(4,826 \cdot 10)x - 9,723 \cdot 10^4$
Santa Catarina	$(4,973 \cdot 10)x - 1,002 \cdot 10^5$
Rio Grande do Sul	$(1,084 \cdot 10^2)x - 2,185 \cdot 10^5$
Distrito Federal	$(1,381 \cdot 10)x - 2,783 \cdot 10^4$
Goiás	$(5,426 \cdot 10)x - 1,094 \cdot 10^5$
Mato Grosso	$(9,486 \cdot 10)x - 1,913 \cdot 10^5$
Mato Grosso do Sul	$(2,745 \cdot 10)x - 5,532 \cdot 10^4$

Fonte: elaborado pelo autor.

APÊNDICE B – EQUAÇÕES DA REGRESSÃO POLINOMIAL

UF	Equações Polinomiais
Brasil	$(2,681 \cdot 10^2)x^2 - (1,081 \cdot 10^6)x + 1,089 \cdot 10^9$
Ceará	$(9,064)x^2 - (3,653 \cdot 10^4)x + 3,681 \cdot 10^7$
Bahia	$(9,307)x^2 - (3,752 \cdot 10^4)x + 3,780 \cdot 10^7$
Sergipe	$(1,565)x^2 - (6,308 \cdot 10^3)x + 6,356 \cdot 10^6$
Alagoas	$(3,571)x^2 - (1,440 \cdot 10^4)x + 1,452 \cdot 10^7$
Pernambuco	$(8,055)x^2 - (3,247 \cdot 10^4)x + 3,272 \cdot 10^7$
Paraíba	$(5,774)x^2 - (2,328 \cdot 10^4)x + 2,346 \cdot 10^7$
Rio Grande do Norte	$(5,898)x^2 - (2,377 \cdot 10^4)x + 2,396 \cdot 10^7$
Piauí	$(5,571)x^2 - (2,245 \cdot 10^4)x + 2,263 \cdot 10^7$
Maranhão	$(8,025)x^2 - (3,236 \cdot 10^4)x + 3,262 \cdot 10^7$
Minas Gerais	$(4,560 \cdot 10)x^2 - (1,838 \cdot 10^5)x + 1,852 \cdot 10^8$
Espírito Santo	$(4,612)x^2 - (1,859 \cdot 10^4)x + 1,873 \cdot 10^7$
Rio de Janeiro	$(9,845)x^2 - (3,968 \cdot 10^4)x + 3,998 \cdot 10^7$
São Paulo	$(3,489 \cdot 10)x^2 - (1,406 \cdot 10^5)x + 1,417 \cdot 10^8$
Tocantins	$(4,803)x^2 - (1,937 \cdot 10^4)x + 1,952 \cdot 10^7$
Pará	$(9,445)x^2 - (3,808 \cdot 10^4)x + 3,839 \cdot 10^7$
Amazonas	$(2,755)x^2 - (1,111 \cdot 10^4)x + 1,121 \cdot 10^7$
Amapá	$(1,087)x^2 - (4,386 \cdot 10^3)x + 4,424 \cdot 10^6$
Roraima	$(7,608 \cdot 10^{-1})x^2 - (3,070 \cdot 10^3)x + 3,096 \cdot 10^6$
Rondônia	$(2,798)x^2 - (1,128 \cdot 10^4)x + 1,137 \cdot 10^7$
Acre	$(1,269)x^2 - (5,120 \cdot 10^3)x + 5,163 \cdot 10^6$
Paraná	$(1,418 \cdot 10)x^2 - (5,716 \cdot 10^4)x + 5,760 \cdot 10^7$
Santa Catarina	$(1,401 \cdot 10)x^2 - (5,647 \cdot 10^4)x + 5,690 \cdot 10^7$
Rio Grande do Sul	$(3,118 \cdot 10)x^2 - (1,257 \cdot 10^5)x + 1,266 \cdot 10^8$
Distrito Federal	$(3,804)x^2 - (1,534 \cdot 10^4)x + 1,546 \cdot 10^7$
Goiás	$(1,707 \cdot 10)x^2 - (6,883 \cdot 10^4)x + 6,938 \cdot 10^7$
Mato Grosso	$(3,096 \cdot 10)x^2 - (1,248 \cdot 10^5)x + 1,259 \cdot 10^8$
Mato Grosso do Sul	$(8,253)x^2 - (3,326 \cdot 10^4)x + 3,352 \cdot 10^7$

Fonte: elaborado pelo autor.

ANEXO A – DATASHEET PAINEL SOLAR CANADIAN



HiKu HIGH POWER MONO PERC MODULE 435 W ~ 460 W CS3W-435 | 440 | 445 | 450 | 455 | 460MS

MORE POWER

-  26 % more power than conventional modules
-  Up to 4.5 % lower LCOE
Up to 2.7 % lower system cost
-  42°C Low NMOT: 42 ± 3 °C
Low temperature coefficient (Pmax): -0.35 % / °C
-  Better shading tolerance

MORE RELIABLE

-  Lower internal current, lower hot spot temperature
-  Minimizes micro-crack impacts
-  Heavy snow load up to 5400 Pa, wind load up to 3600 Pa*

* For detailed information, please refer to Installation Manual.



12 Years Enhanced Product Warranty on Materials and Workmanship*

25 Years Linear Power Performance Warranty*

1st year power degradation no more than 2%
Subsequent annual power degradation no more than 0.55%

*According to the applicable Canadian Solar Limited Warranty Statement.

MANAGEMENT SYSTEM CERTIFICATES*

ISO 9001:2015 / Quality management system
ISO 14001:2015 / Standards for environmental management system
ISO 45001: 2018 / International standards for occupational health & safety

PRODUCT CERTIFICATES*

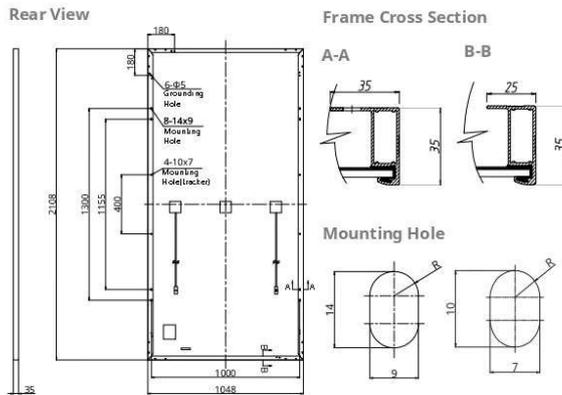
IEC 61215 / IEC 61730 / CE / MCS / INMETRO
FSEC (US Florida) / UL 61730 / IEC 61701 / IEC 62716 / IEC 60068-2-68
UNI 9177 Reaction to Fire: Class 1 / Take-e-way



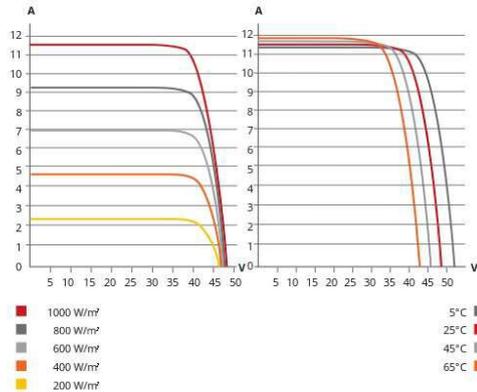
* The specific certificates applicable to different module types and markets will vary, and therefore not all of the certifications listed herein will simultaneously apply to the products you order or use. Please contact your local Canadian Solar sales representative to confirm the specific certificates available for your product and applicable in the regions in which the products will be used.

CSI Solar Co., Ltd. is committed to providing high quality solar products, solar system solutions and services to customers around the world. Canadian Solar was recognized as the No. 1 module supplier for quality and performance/price ratio in the IHS Module Customer Insight Survey, and is a leading PV project developer and manufacturer of solar modules, with over 50 GW deployed around the world since 2001.

ENGINEERING DRAWING (mm)



CS3W-435MS / I-V CURVES



ELECTRICAL DATA | STC*

CS3W	435MS	440MS	445MS	450MS	455MS	460MS
Nominal Max. Power (Pmax)	435 W	440 W	445 W	450 W	455 W	460 W
Opt. Operating Voltage (Vmp)	40.5 V	40.7 V	40.9 V	41.1 V	41.3 V	41.5 V
Opt. Operating Current (Imp)	10.75 A	10.82 A	10.89 A	10.96 A	11.02 A	11.09 A
Open Circuit Voltage (Voc)	48.5 V	48.7 V	48.9 V	49.1 V	49.3 V	49.5 V
Short Circuit Current (Isc)	11.42 A	11.48 A	11.54 A	11.60 A	11.66 A	11.72 A
Module Efficiency	19.7%	19.9%	20.1%	20.4%	20.6%	20.8%
Operating Temperature	-40°C ~ +85°C					
Max. System Voltage	1500V (IEC/UL) or 1000V (IEC/UL)					
Module Fire Performance	TYPE 1 (UL 61730 1500V) or TYPE 2 (UL 61730 1000V) or CLASS C (IEC 61730)					
Max. Series Fuse Rating	20 A					
Application Classification	Class A					
Power Tolerance	0 ~ + 10 W					

* Under Standard Test Conditions (STC) of irradiance of 1000 W/m², spectrum AM 1.5 and cell temperature of 25°C.

ELECTRICAL DATA | NMOT*

CS3W	435MS	440MS	445MS	450MS	455MS	460MS
Nominal Max. Power (Pmax)	325 W	328 W	332 W	336 W	339 W	343 W
Opt. Operating Voltage (Vmp)	37.8 V	37.9 V	38.1 V	38.3 V	38.5 V	38.7 V
Opt. Operating Current (Imp)	8.59 A	8.65 A	8.71 A	8.76 A	8.82 A	8.87 A
Open Circuit Voltage (Voc)	45.6 V	45.8 V	46.0 V	46.2 V	46.4 V	46.6 V
Short Circuit Current (Isc)	9.21 A	9.26 A	9.31 A	9.36 A	9.41 A	9.45 A

* Under Nominal Module Operating Temperature (NMOT), irradiance of 800 W/m² spectrum AM 1.5, ambient temperature 20°C, wind speed 1 m/s.

MECHANICAL DATA

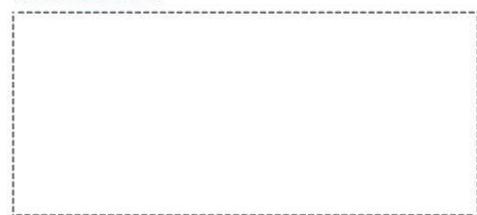
Specification	Data
Cell Type	Mono-crystalline
Cell Arrangement	144 [2 X (12 X 6)]
Dimensions	2108 X 1048 X 35 mm (83.0 X 41.3 X 1.38 in)
Weight	24.3 kg (53.6 lbs)
Front Cover	3.2 mm tempered glass
Frame	Anodized aluminium alloy, crossbar enhanced
J-Box	IP68, 3 bypass diodes
Cable	4 mm ² (IEC), 12 AWG (UL)
Cable Length (Including Connector)	500 mm (19.7 in) (+) / 350 mm (13.8 in) (-) or customized length*
Connector	T4 series or H4 UTX or MC4-EVO2
Per Pallet	30 pieces
Per Container (40' HQ)	660 pieces

* For detailed information, please contact your local Canadian Solar sales and technical representatives.

TEMPERATURE CHARACTERISTICS

Specification	Data
Temperature Coefficient (Pmax)	-0.35 % / °C
Temperature Coefficient (Voc)	-0.27 % / °C
Temperature Coefficient (Isc)	0.05 % / °C
Nominal Module Operating Temperature	42 ± 3°C

PARTNER SECTION



* The specifications and key features contained in this datasheet may deviate slightly from our actual products due to the on-going innovation and product enhancement. CSI Solar Co., Ltd. reserves the right to make necessary adjustment to the information described herein at any time without further notice.
Please be kindly advised that PV modules should be handled and installed by qualified people who have professional skills and please carefully read the safety and installation instructions before using our PV modules.

CSI Solar Co., Ltd.
199 Lushan Road, SND, Suzhou, Jiangsu, China, 215129, www.csisolar.com, support@csisolar.com