

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

ANDRÉ FELIPE DE MOURA VIEIRA

ERIC GUSTAVO MARTINS

JOÃO VICTOR MOINHO GIROTO

RODRIGO QUEIROZ DE LOYOLA

**SISTEMAS FOTOVOLTAICOS DE GERAÇÃO DE ENERGIA:** Modelo de previsão  
de curto prazo a partir de dados meteorológicos.

CURITIBA

2022

ANDRÉ FELIPE DE MOURA VIEIRA  
ERIC GUSTAVO MARTINS  
JOÃO VICTOR MOINHO GIROTO  
RODRIGO QUEIROZ DE LOYOLA

**SISTEMAS FOTOVOLTAICOS DE GERAÇÃO DE ENERGIA:** Modelo de previsão de curto prazo a partir de dados meteorológicos.

**PHOTOVOLTAIC ENERGY GENERATION SYSTEMS:** Short-term forecast model based on meteorological data.

Trabalho de conclusão de curso de graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Controle e Automação do curso de Engenharia de Controle e Automação da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. Jorge Assade Leludak

Curitiba  
2022



[4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

ANDRÉ FELIPE DE MOURA VIEIRA  
ERIC GUSTAVO MARTINS  
JOÃO VICTOR MOINHO GIROTO  
RODRIGO QUEIROZ DE LOYOLA

**SISTEMAS FOTOVOLTAICOS DE GERAÇÃO DE ENERGIA:** Modelo de previsão  
de curto prazo a partir de dados meteorológicos.

Trabalho de conclusão de curso de  
graduação apresentado como requisito  
para obtenção do título de Bacharel em  
Engenharia de Controle e Automação  
do curso de Engenharia de Controle e  
Automação da Universidade  
Tecnológica Federal do Paraná  
(UTFPR).

Data da aprovação: 01/dezembro/2022

---

Daniel Balieiro Silva  
Professor (Mestrado)  
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ (UTFPR)

---

José da Silva Maia  
Professor (Mestrado)  
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ (UTFPR)

---

Rafael Solheid Jacobs  
Professor (Graduação em Sistemas de Telecomunicações)  
FACULDADE DE TECNOLOGIA DE CURITIBA (FATEC)

Curitiba  
2022

## RESUMO

Este trabalho de conclusão de curso, realizado por discentes do curso de Engenharia de Controle e Automação, possui como premissa verificar a possibilidade de se realizar um sistema de previsão de eficiência de um painel fotovoltaico utilizando dados de uma estação solarimétrica local. Para tal foram utilizados os dados dos anos de 2021 e 2022 de uma estação em Curitiba, fornecida pelo Laboratório de Energia Solar da Universidade Tecnológica Federal do Paraná (LABENS). Realizou-se modelos lineares utilizando Excel, assim como a previsão de eficiência futura das placas solares. Foi realizado, também, um modelo linear regressivo (ARX) como prova de conceito, utilizando os dados de um mês. Foi verificado que os modelos possuem um grande grau de complexidade matemática por conta do caos matemático, assim como o grande peso computacional dos modelos, tornando modelos precisos onerosos computacionalmente. Observou-se que os modelos gerados com dados referentes a um espaço de tempo menor, com uma previsão de eficiência das placas a curto prazo, possuem um grau de confiabilidade maior do que os modelos gerados com dados de grandes espaços de tempo. Observou-se, também, que a escolha das variáveis de entrada é extremamente importante, uma vez que modelos gerados com entradas lineares possuem um resultado melhor do que com entradas lineares e não lineares. Isso faz com que a modelagem de sistemas baseados em dados meteorológicos se torne extremamente difícil, necessitando de computadores de alta performance para gerar modelos precisos.

Palavras-chave: predições; potência; modelo; previsão; meteorológicos.

## **ABSTRACT**

This thesis, carried out by students of the Control and Automation Engineering course, has the premise to verify the possibility of carrying out a system for predicting the efficiency of a photovoltaic panel using data from a local solarimetric station. For this, data from the years 2021 and 2022 from a station in Curitiba, provided by the Solar Energy Laboratory of the Federal Technological University of Paraná (LABENS) were used. Linear models were carried out using Excel, as well as the prediction of future efficiency of solar panels. A linear regressive model (ARX) was also performed as a proof of concept, using data from one month. It was found that the models have a high degree of mathematical complexity due to the mathematical chaos, as well as the large computational weight of the models, making accurate models computationally costly. It was observed that the models generated with data referring to a shorter period of time, with a forecast of efficiency of the panels in the short term, have a higher degree of reliability than the models generated with data from large periods of time. It was also observed that the choice of input variables is extremely important, since models generated with linear inputs have a better result than with linear and non-linear inputs. This makes the modeling of systems based on meteorological data extremely difficult, requiring high-performance computers to generate accurate models.

Keywords: predictions; power; model; preview; meteorological.

## ÍNDICE DE FIGURAS

|                                                                                  |    |
|----------------------------------------------------------------------------------|----|
| Figura 1: Tipos diferentes de modelos. Isermann, adaptado pelos autores. ....    | 27 |
| Figura 2: Excel – Janela de Opções.....                                          | 36 |
| Figura 3: Excel - Gerenciador de Suplementos.....                                | 36 |
| Figura 4: Excel - Janela Análise de Dados.....                                   | 37 |
| Figura 5: Excel - Janela da Função Regressão. ....                               | 37 |
| Figura 6: Gráfico Comparativo - Regressão da variável pac-cdteCT. ....           | 43 |
| Figura 7: Gráfico Comparativo - Regressão da variável pac-cigsCT. ....           | 44 |
| Figura 8: Gráfico Comparativo - Regressão da variável pac-mon1CT.....            | 44 |
| Figura 9: Gráfico Comparativo - Regressão da variável pac-mon2CT.....            | 45 |
| Figura 10: Gráfico Comparativo - Regressão da variável pac-pol1CT. ....          | 45 |
| Figura 11: Gráfico Comparativo - Regressão da variável pac-pol2CT. ....          | 46 |
| Figura 12: Excel - Planilha de previsão. ....                                    | 48 |
| Figura 13: Gráfico Planilha de Previsão - press-so10CT. ....                     | 49 |
| Figura 14: Gráfico Planilha de Previsão - rain-so10CT. ....                      | 49 |
| Figura 15: Gráfico Planilha de Previsão - tilt_avg-so01CT.....                   | 50 |
| Figura 16: Gráfico Planilha de Previsão - tp_sfc-so10CT.....                     | 50 |
| Figura 17: Gráfico Planilha de Previsão - ws_max-so10CT. ....                    | 51 |
| Figura 18: Gráfico Planilha de Previsão - humid-so10CT. ....                     | 52 |
| Figura 19: Segundo gráfico comparativo - Regressão Linear - pac-cdte. ....       | 53 |
| Figura 20: Segundo gráfico comparativo - Regressão Linear - pac-mon1. ....       | 53 |
| Figura 21: Segundo gráfico comparativo - Regressão Linear - pac-pol1. ....       | 54 |
| Figura 22: Segundo gráfico comparativo - Regressão Linear - pac-cigs.....        | 54 |
| Figura 23: Segundo gráfico comparativo - Regressão Linear - pac-mon2. ....       | 55 |
| Figura 24: Segundo gráfico comparativo - Regressão Linear - pac-pol2. ....       | 55 |
| Figura 25: Gráfico comparativo - Primeiro Dia - pac-cdte. ....                   | 56 |
| Figura 26: Gráfico comparativo - Primeiro Dia - pac-mon1. ....                   | 57 |
| Figura 27: Gráfico comparativo - Primeiro Dia – pac-pol1.....                    | 57 |
| Figura 28: Gráfico comparativo - Primeiro Dia – pac-cigs.....                    | 58 |
| Figura 29: Gráfico comparativo - Primeiro Dia – pac-mon2. ....                   | 58 |
| Figura 30: Gráfico comparativo - Primeiro Dia – pac-pol2.....                    | 59 |
| Figura 31: Primeira Ampliação - Gráfico comparativo da variável pac-cdteCT. .... | 65 |
| Figura 32: Primeira Ampliação - Gráfico comparativo da variável pac-cigsCT.....  | 66 |
| Figura 33: Primeira Ampliação - Gráfico comparativo da variável pac-mon1CT. .... | 66 |

|                                                                                                      |    |
|------------------------------------------------------------------------------------------------------|----|
| Figura 34: Primeira Ampliação - Gráfico comparativo da variável pac-mon2CT.....                      | 67 |
| Figura 35: Primeira Ampliação - Gráfico comparativo da variável pac-pol1CT.....                      | 68 |
| Figura 36: Primeira Ampliação - Gráfico comparativo da variável pac-pol2CT.....                      | 68 |
| Figura 37: Gráfico comparativo sem dados nulos - variável pac-cdteCT.....                            | 70 |
| Figura 38: Gráfico comparativo sem dados nulos - variável pac-cigsCT.....                            | 70 |
| Figura 39: Gráfico comparativo sem dados nulos - variável pac-mon1CT.....                            | 71 |
| Figura 40: Gráfico comparativo sem dados nulos - variável pac-mon2CT.....                            | 72 |
| Figura 41: Gráfico comparativo sem dados nulos - variável pac-pol1CT.....                            | 72 |
| Figura 42: Gráfico comparativo sem dados nulos - variável pac-pol2CT.....                            | 73 |
| Figura 43: gráfico previsão tilt_avg sem dados nulos e eixo x sequencial.....                        | 74 |
| Figura 44: gráfico previsão tilt_avg - primeiras 24 horas com eixo x sequencial.....                 | 74 |
| Figura 45: gráfico previsão tilt_avg com todos os dados, no modo de eixo x como data/hora.....       | 75 |
| Figura 46: gráfico previsão tilt_avg - primeiras 24 horas, com eixo x como data/hora.....            | 75 |
| Figura 47: Tecnologia cdte - exposição da previsão de todos os dados.....                            | 76 |
| Figura 48: Tecnologia cdte - exposição da previsão das primeiras 24 horas.....                       | 76 |
| Figura 49: Tecnologia cdte - previsão de todos os dados com ajuste no eixo X.....                    | 77 |
| Figura 50: Tecnologia cdte - primeiras 24h com ajuste no eixo X.....                                 | 78 |
| Figura 51: Tecnologia cdte - previsão de todos os dados com eixo X sequencial... ..                  | 79 |
| Figura 52: Tecnologia cdte - primeiras 24 horas com eixo X sequencial.....                           | 79 |
| Figura 53: Tecnologia cdte - exposição da previsão de todos os dados após ajuste no eixo X.....      | 80 |
| Figura 54: Tecnologia cdte - exposição da previsão das primeiras 24 horas após ajuste no eixo X..... | 80 |
| Figura 55: Segunda Ampliação - Gráfico comparativo da variável pac-cdteCT.....                       | 81 |
| Figura 56: Segunda Ampliação - Gráfico comparativo da variável pac-cigsCT.....                       | 82 |
| Figura 57: Segunda Ampliação - Gráfico comparativo da variável pac-mon1CT.....                       | 82 |
| Figura 58: Segunda Ampliação - Gráfico comparativo da variável pac-pol1CT.....                       | 83 |
| Figura 59: Segunda Ampliação - Gráfico comparativo da variável pac-mon2.....                         | 83 |
| Figura 60: Segunda Ampliação - Gráfico comparativo da variável pac-pol2.....                         | 84 |
| Figura 61: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-cdteCT..... | 85 |

|                                                                                                           |     |
|-----------------------------------------------------------------------------------------------------------|-----|
| Figura 62: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-cigsCT.....      | 85  |
| Figura 63: Segunda Ampliação sem valores negativos - Gráfico comparativo da variávelpac-mon1CT.....       | 86  |
| Figura 64: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-pol1CT.....      | 86  |
| Figura 65: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-mon2CT.....      | 87  |
| Figura 66: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-pol2CT.....      | 87  |
| Figura 67: Tecnologia cdte - exposição da previsão de todos os dados após 2ª ampliação.....               | 88  |
| Figura 68: Tecnologia cdte - exposição da previsão das primeiras 24h após 2ª ampliação.....               | 89  |
| Figura 69: Tecnologia cdte - exposição da previsão de todos os dados após 2ª ampliação e ajuste em X..... | 90  |
| Figura 70: tecnologia cdte - exposição da previsão das primeiras 24h após 2ª ampliação e ajuste em X..... | 90  |
| Figura 71: FIT sobre dados dos testes.....                                                                | 103 |
| Figura 72: Ordem do Sistema SISO Regressivo.....                                                          | 104 |
| Figura 73: Ordem do Sistema MISO Regressivo.....                                                          | 104 |
| Figura 74: Ordem do Sistema SISO TOOLBOX.....                                                             | 104 |
| Figura 75: Ordem do Sistema MISO TOOLBOX.....                                                             | 105 |
| Figura 76: gráfico com dados climatológicos para Curitiba 1991-2021.....                                  | 108 |
| Figura 77: ANEXO 1 Resultados - Forecast SISO Regressivo.....                                             | 115 |
| Figura 78: ANEXO 1 Resultados – Modelo Toolbox MISO x Dados Reais.....                                    | 116 |
| Figura 79: ANEXO 1 Resultados – Método Regressivo SISO x Dados Reais.....                                 | 117 |
| Figura 80: ANEXO 1 Resultados - Forecast SISO Toolbox.....                                                | 118 |
| Figura 81: ANEXO 1 Resultados - Forecast MISO Toolbox.....                                                | 119 |
| Figura 82: ANEXO 1 Resultados – Método Regressivo MISO x Dados Reais.....                                 | 120 |
| Figura 83: ANEXO 1 Resultados – Modelo Toolbox MISO x Dados Reais.....                                    | 121 |
| Figura 84: ANEXO 1 Resultados - Forecast MISO Regressivo.....                                             | 122 |

## ÍNDICE DE TABELAS

|                                                                                                          |     |
|----------------------------------------------------------------------------------------------------------|-----|
| Tabela 1: Variável pac-cdteCT.....                                                                       | 39  |
| Tabela 2: Variável pac-cigsCT. ....                                                                      | 39  |
| Tabela 3: variável pac-mon1CT.....                                                                       | 40  |
| Tabela 4: Variável pac-mon2CT.....                                                                       | 40  |
| Tabela 5: Variável pac-pol1CT.....                                                                       | 41  |
| Tabela 6: Variável pac-pol2CT.....                                                                       | 41  |
| Tabela 7: Primeira Ampliação - Regressão da variável pac-cdteCT.....                                     | 60  |
| Tabela 8: Primeira Ampliação - Regressão da variável pac-cigsCT. ....                                    | 61  |
| Tabela 9: Primeira Ampliação - Regressão da variável pac-mon1CT.....                                     | 61  |
| Tabela 10: Primeira Ampliação - Regressão da variável pac-mon2CT.....                                    | 62  |
| Tabela 11: Primeira Ampliação - Regressão da variável pac-pol1CT.....                                    | 62  |
| Tabela 12: Primeira Ampliação - Regressão da variável pac-pol2CT.....                                    | 63  |
| Tabela 13: Erro médio de previsões - JUN/2021. ....                                                      | 99  |
| Tabela 14: Tabela de erros médios de previsões – Período de 1 ano e eixo sequencial.<br>.....            | 100 |
| Tabela 15: Tabela de erros médios de previsões – Período de 1 ano e eixo ajustado.<br>.....              | 100 |
| Tabela 16: Tabela de erros médios – Período de 1 ano, eixo sequencial e valores<br>nulos filtrados. .... | 101 |
| Tabela 17: Tabela de erros médios – Período de 1 ano, eixo ajustado e valores nulos<br>filtrados.....    | 101 |
| Tabela 18: Tabela de erros médios - Regressão JUN/2021 e eixo sequencial.....                            | 102 |
| Tabela 19: Tabela de erros médios - Regressão JUN/2021 e eixo ajustado. ....                             | 102 |
| Tabela 20: Estudo de linearidade entre variáveis. ....                                                   | 103 |

## LISTA DE SIGLAS E ABREVIATURAS

|             |                                                                           |
|-------------|---------------------------------------------------------------------------|
| GHI         | Global Horizontal Irradiance                                              |
| GW          | Gigawatt                                                                  |
| SPV         | Solar Photovoltaic Energy                                                 |
| SVM         | Support Vector Machine                                                    |
| AR          | Auto Regressive Model                                                     |
| ARX         | Auto Regressive Model with eXogenous inputs                               |
| ARMA        | Auto Regressive Model with Moving Average                                 |
| ARMAX       | Auto Regressive Model with Moving Average and eXogenous inputs            |
| NARMAX      | Non linear Auto Regressive Model with Moving Average and eXogenous inputs |
| INMET       | Instituto Nacional de Meteorologia                                        |
| ETS         | Error, trend and Seasonal Exponential Smoothing Method                    |
| humid-so    | Umidade atmosférica, em porcentagem                                       |
| press-so    | Pressão atmosférica, em milibares                                         |
| rain-so     | Precipitação, em milímetros por hora                                      |
| tp_sfc-so   | Temperatura do ar superficial, em graus Celcius                           |
| ws_max-so   | Velocidade máxima do vento, em metros por segundo                         |
| tilt_avg-so | Média da radiação no plano inclinado, em Watts por metro quadrado         |
| pac-cdte    | Potência da rede AC - inversor de 1,5kW                                   |
| pac-cigs    | Potência da rede AC - inversor de 1,5kW                                   |
| pac-mon1    | Potência da rede AC - inversor de 3kW (monocristalino 1)                  |
| pac-mon2    | Potência da rede AC - inversor de 3kW (monocristalino 2)                  |
| pac-pol1    | Potência da rede AC - inversor de 3kW (policristalino 1)                  |
| pac-pol2    | Potência da rede AC - inversor de 3kW (policristalino 2)                  |
| na          | Ordem do polinômio A(q) - modelagem ARX                                   |
| nb          | Vetor de ordem do polinômio B(q) - modelagem ARX                          |
| nk          | Vetor de atraso no tempo - modelagem ARX                                  |

# SUMÁRIO

|          |                                                        |           |
|----------|--------------------------------------------------------|-----------|
| <b>1</b> | <b>INTRODUÇÃO</b>                                      | <b>13</b> |
| 1.1      | TEMA                                                   | 13        |
| 1.1.1    | Delimitação do tema                                    | 15        |
| 1.2      | PROBLEMAS E PREMISSAS                                  | 17        |
| 1.3      | OBJETIVOS                                              | 18        |
| 1.3.1    | Objetivo geral                                         | 18        |
| 1.3.2    | Objetivos específicos                                  | 18        |
| 1.4      | JUSTIFICATIVA                                          | 19        |
| 1.5      | PROCEDIMENTOS METODOLÓGICOS                            | 20        |
| 1.5.1    | Metodologia Para Análise De Dados                      | 20        |
| 1.5.2    | Metodologia Para Obtenção De Dados                     | 21        |
| <b>2</b> | <b>FUNDAMENTAÇÃO TEÓRICA</b>                           | <b>22</b> |
| 2.1      | PRESSÃO ATMOSFÉRICA                                    | 23        |
| 2.2      | PRECIPITAÇÃO                                           | 23        |
| 2.3      | UMIDADE                                                | 24        |
| 2.4      | TEMPERATURA DO AR SUPERFICIAL                          | 24        |
| 2.5      | IRRADIAÇÃO                                             | 24        |
| 2.5.1    | Emissão Solar                                          | 24        |
| 2.5.2    | Fatores Que Influenciam Na Irradiação Sobre A Terra    | 25        |
| 2.5.2.1  | Translação Terrestre                                   | 25        |
| 2.5.2.2  | Radiação Direta x Difusa                               | 25        |
| 2.6      | MODELAGEM DE SISTEMAS                                  | 26        |
| 2.6.1    | Modelos lineares                                       | 28        |
| 2.6.1.1  | Modelo Autorregressivo exógeno com média móvel (ARMAX) | 28        |
| 2.6.1.2  | Modelo Autorregressivo exógeno (ARX)                   | 29        |
| 2.6.2    | Modelos não lineares (NARMAX)                          | 30        |
| 2.6.3    | CRITÉRIO DE INFORMAÇÃO DE AKAIKE                       | 31        |
| <b>3</b> | <b>DESENVOLVIMENTO</b>                                 | <b>32</b> |
| 3.1      | ANÁLISE INICIAL                                        | 32        |
| 3.2      | OBSTÁCULOS ENCONTRADOS                                 | 33        |
| 3.3      | SEGUNDA ANÁLISE                                        | 33        |
| 3.4      | IMPLEMENTAÇÃO POR SOFTWARES                            | 35        |
| 3.4.1    | Excel                                                  | 35        |
| 3.4.1.1  | Regressão linear                                       | 35        |
| 3.4.1.2  | Método Dos Mínimos Quadrados                           | 38        |
| 3.4.1.3  | Aplicação da Regressão Linear                          | 39        |

|           |                                                                |            |
|-----------|----------------------------------------------------------------|------------|
| 3.4.1.4   | Cálculo de Previsão de Energia Gerada                          | 47         |
| 3.4.1.4.1 | Suavização Exponencial                                         | 47         |
| 3.4.1.5   | Primeira Ampliação do Conjunto de Dados                        | 60         |
| 3.4.1.5.1 | Resultados da Regressão Linear após Primeira Ampliação         | 60         |
| 3.4.1.5.2 | Resultados de Previsão Após Primeira Ampliação de Dados        | 73         |
| 3.4.1.6   | Segunda Ampliação do Conjunto de Dados                         | 81         |
| 3.4.1.6.1 | Filtragem de Valores Negativos                                 | 85         |
| 3.4.1.7   | Resultados da Previsão após Segunda Ampliação de Dados         | 88         |
| 3.4.2     | MatLab                                                         | 91         |
| 3.4.2.1   | Modelagem Caixa-Preta                                          | 91         |
| 3.4.2.1.1 | Escolha de Variáveis de Entrada e Saída                        | 92         |
| 3.4.2.1.2 | Modelagem ARX                                                  | 93         |
| <b>4</b>  | <b>RESULTADOS E DISCUSÕES</b>                                  | <b>95</b>  |
| 4.1       | SELEÇÃO DO MÉTODO A SER APLICADO                               | 95         |
| 4.1.1     | Primeira Aplicação                                             | 96         |
| 4.1.2     | Segunda Aplicação                                              | 96         |
| 4.1.3     | Definição do Método Para Excel                                 | 96         |
| 4.1.4     | Ampliação do Conjunto de Dados e Análise de Resultados Obtidos | 97         |
| 4.2       | RESULTADOS POR REGRESSÃO LINEAR                                | 98         |
| 4.3       | RESULTADOS DA PREVISÃO POR EXCEL                               | 99         |
| 4.3.1     | Considerando Apenas Dados de JUN/2021                          | 99         |
| 4.3.2     | Considerando Dados de 1 Ano e Previsão Sequencial              | 100        |
| 4.3.2.1   | Valores Nulos Presentes                                        | 100        |
| 4.3.2.1.1 | Eixo X Sequencial                                              | 100        |
| 4.3.2.1.2 | Eixo X Ajustado                                                | 100        |
| 4.3.2.2   | Valores Nulos Filtrados                                        | 101        |
| 4.3.2.2.1 | Eixo Sequencial                                                | 101        |
| 4.3.2.2.2 | Eixo x Ajustado                                                | 101        |
| 4.3.2.3   | Regressão de Junho/2021                                        | 102        |
| 4.3.2.3.1 | Eixo X Sequencial                                              | 102        |
| 4.3.2.3.2 | Eixo X Ajustado                                                | 102        |
| 4.4       | RESULTADOS DA PREVISÃO POR MATLAB                              | 103        |
| 4.4.1     | ESTUDO DE LINEARIDADE                                          | 103        |
| 4.4.2     | MODELO SISO REGRESSIVO                                         | 103        |
| 4.4.3     | MODELO MISO REGRESSIVO                                         | 105        |
| 4.4.4     | Método TOOLBOX                                                 | 106        |
| <b>5</b>  | <b>CONCLUSÃO</b>                                               | <b>107</b> |

|                |                    |            |
|----------------|--------------------|------------|
| 5.1            | APLICAÇÃO EM EXCEL | 107        |
| 5.2            | MODELAGEM ARX      | 108        |
| <b>ANEXO 1</b> |                    | <b>115</b> |
| <b>ANEXO 2</b> |                    | <b>123</b> |

# 1 INTRODUÇÃO

## 1.1 TEMA

Atualmente, em escala global, as formas de geração de energia predominantes são dependentes de combustíveis fósseis ou a nuclear, que além de serem formas prejudiciais ao meio ambiente, são dependentes de recursos finitos e assim, formas de energia renovável vêm ganhando espaço, como a energia eólica, biomassa e solar, visto que vento, matéria orgânica e luz solar são recursos abundantes e fornecidos naturalmente (GLOBAL RESOURCES OUTLOOK 2019, United Nations Environment).

Países como Alemanha, Japão, Itália, Espanha e Estados Unidos foram os primeiros a incentivar a pesquisa em grande escala e, juntos, representavam 88% da produção mundial de energia fotovoltaica no ano de 2011 (Rev. Virtual Quim., 2015). Contudo, por se tratar de um campo relativamente novo, a geração fotovoltaica apresenta um rendimento inferior, e sua aplicação precisa ser mais cautelosa, tanto pelo seu custo de instalação e manutenção, quanto por seu retorno não ser linear e de difícil previsão. Porém, é indiscutível que possui um potencial enorme, visto que sua fonte é abundante, limpa, e “inesgotável” (Revista Brasileira de Energia Solar Ano 9 Volume IX Número 1 Julho de 2018).

De um modo geral, a geração de energia fotovoltaica é possível graças ao efeito fotovoltaico, onde a radiação solar, após incidir nos painéis solares, causa a excitação dos elétrons, fazendo com que estes se movimentam, gerando assim energia elétrica. No material semicondutor (de silício, por exemplo), os cristais se organizam fazendo ligações covalentes com átomos vizinhos: cada átomo compartilha um elétron de sua camada de valência com um elétron de valência do átomo vizinho. Dessa forma, a camada de valência dos átomos no material semicondutor fica cheia, pois os elétrons ficam presos nela, e o átomo fica em estado estável (CASTRO, 2002).

Para que um elétron se desloque, é preciso que ganhe energia suficiente para passar da camada de valência para a camada de condução: isso pode ocorrer quando um fóton contendo energia suficiente atinge um elétron que esteja na camada de valência, fazendo com que o mesmo ganhe a energia necessária para se mover para

a camada de condução. Ao se mover para a camada de condução, o elétron deixa uma “lacuna” em seu lugar, criando um par elétron-lacuna (CASTRO, 2002).

Para que os elétrons se desloquem, é preciso que ganhem energia suficiente para passarem da camada de valência para a camada de condução. Quando um fóton, que contém energia suficiente, atinge um elétron na camada de valência, este elétron ganha a energia necessária para mover-se para a camada de condução, e fazendo isso, deixa uma “lacuna” em seu lugar. Quando tudo isso acontece, diz-se que o fóton criou um par elétron-lacuna (CASTRO, 2002 e CRESESB, 2006).

Entretanto, uma célula feita somente por cristais de silício não é capaz de produzir energia, pois, embora os elétrons saltem para a camada de condução, eles se recombinam com as lacunas. Por este motivo, é feita a dopagem do silício (processo que introduz outros materiais de modo a alterar as propriedades elétricas), que permite a criação de duas camadas na célula fotovoltaica: uma tipo  $p$  (com excesso de cargas positivas), e uma tipo  $n$  (com excesso de cargas negativas). Deste modo, na região onde esses dois materiais se encontram (a chamada junção  $p-n$ ) é formado um campo elétrico, que separa os elétrons (energizados pelos fótons, que saltaram para a camada de condução) para um terminal negativo, e as lacunas (deixadas pelos elétrons) para um terminal positivo, criando assim uma diferença de potencial, que como sabemos, é capaz de induzir uma corrente elétrica (CASTRO, 2002 e CRESESB, 2006).

Como esse efeito é dependente da radiação solar, o processo de geração de energia fotovoltaica pode ocorrer em dias nublados, tendo um menor rendimento. No campo da geração através de sistemas fotovoltaicos, já podem ser encontrados diversos métodos de cálculo e estimativa de potencial de geração, utilizando dados de localização geográfica, ambientais e meteorológicas, os quais são de grande importância para projetar os locais ideais de captação de radiação solar para um melhor desempenho.

Como esse efeito é dependente da radiação solar, o processo de geração de energia fotovoltaica pode ocorrer em dias nublados, tendo um menor rendimento. No campo da geração através de sistemas fotovoltaicos, já podem ser encontrados diversos métodos de cálculo e estimativa de potencial de geração, utilizando dados de localização geográfica, ambientais e meteorológicas, os quais são de grande importância para projetar os locais ideais de captação de radiação solar para um melhor desempenho (CASTRO, 2002).

Com um banco de dados de previsões meteorológicas em mãos é possível desenvolver métodos para se estimar a quantidade de radiação solar possível de ser adquirida em um momento futuro próximo, porém tal processo lida com dados não-lineares, pois os fenômenos que interferem na geração de energia fotovoltaica são naturais e totalmente aleatórios.

As dificuldades encontradas na manipulação dos dados adquiridos através dos sistemas de previsões meteorológicas serão discutidas, juntamente com o sistema de correções até o momento, no decorrer da **FUNDAMENTAÇÃO TEÓRICA**.

### **1.1.1 Delimitação do tema**

A geração de energia solar fotovoltaica depende fundamentalmente da irradiação solar ao nível do solo, incidente em um plano horizontal. Esta irradiação varia devido a múltiplos fatores meteorológicos, tais como: nebulosidade, precipitação pluviométrica, temperatura ambiente, pressão atmosférica, direção do vento, umidade, entre outras. Graças a tantas variáveis, existe um grande espectro de estudos e pesquisas para desenvolvimento de métodos de previsão que podem ser subdivididos entre: físicos, estatísticos e híbridos.

Nas abordagens físicas, é feito o uso de equações analíticas para modelar um sistema de obtenção de energia solar fotovoltaica. Já na abordagem estatística, é utilizado o histórico da região em modelos estocásticos para induzir a aprendizagem da máquina para efetuar a previsão.

Dentro das abordagens de previsão, ainda existem outras maneiras de subdividi-las entre:

- Muito Curto Prazo - Consiste em uma previsão para menos de 60 minutos;
- Curto Prazo - Consiste em uma previsão de até uma semana.

Como segue a proposta deste trabalho, serão avaliadas as características da previsão de curto prazo.

Usufruir de todos os métodos de previsão podem trazer diversos benefícios, pois permite que a energia solar fotovoltaica possa participar dos planejamentos de geração de energia, participando mais efetivamente no mercado de eletricidade, podendo atender diversas casas com um bom nível de confiança.

Os benefícios dessas previsões são:

- Garantir o abastecimento e fornecimento energético;
- Facilitar planejamento energético em pequena e grande escala;
- Fornecer cálculos precisos para ajudar na aceitação pública do método de obtenção de energia.

## 1.2 PROBLEMAS E PREMISSAS

Apesar da constante evolução no ramo de geração e distribuição de energia solar, do aumento do uso e implantação dessa tecnologia, e dos incentivos, investimentos e iniciativas, a proporção da geração fotovoltaica ainda está longe do cenário ideal.

Diversos são os motivos que justificam a atual situação, como custos de implementação, carência de incentivos governamentais, a própria eficiência dos painéis solares mais amplamente comercializados, e até mesmo a falta de conhecimento e informações sobre o tema por parte da maioria da sociedade. Na intenção de incentivar o crescimento desse setor e aumentar o interesse por projetos com painéis solares, são estudadas estratégias alternativas que possam amenizar as desvantagens, equilibrando a balança e tornando os projetos mais viáveis (TIEPOLO, 2013).

Uma estratégia que pode aumentar significativamente a viabilidade de projetos seria a implementação de uma análise mais detalhada acerca dos diversos fatores que influenciam na geração de energia, identificando variáveis-chaves e comparando seu impacto no resultado final.

### **Como a coleta e análise de dados meteorológicos de períodos passados pode auxiliar na previsão da geração futura?**

Há muito se sabe que estudando os fatores locais e ambientais intrínsecos em um determinado projeto, pode-se assegurar uma melhor eficiência e garantir melhores resultados, e diversos estudos sobre o assunto já foram e ainda vêm sendo desenvolvidos, de maneira que o entendimento sobre o impacto de cada uma das possíveis variáveis na geração fotovoltaica não é mais um mistério.

O que pode ser feito então é procurar por meios de utilizar esse conhecimento da melhor forma possível, não deixando de implementar o que já é compreendido e indo ainda além, desenvolvendo métodos de análise e aplicações que permitam usar essas informações de maneira sistemática.

## **1.3 OBJETIVOS**

Os objetivos presentes neste trabalho visam o estudo e desenvolvimento de um método de análise de dados meteorológicos para uma previsão na geração de energia fotovoltaica.

### **1.3.1 Objetivo geral**

Determinar as variáveis que afetarão a geração, de maneira que o nível de influência de cada variável no resultado possa ser quantificado e comparável, para então desenvolver um método capaz de prever, com certo grau de precisão, a quantidade de energia gerada num futuro próximo.

### **1.3.2 Objetivos específicos**

Para o desenvolvimento de um método eficiente é necessário o cumprimento dos seguintes objetivos:

- a) planejar um método para coleta e armazenamento de informações meteorológicas;
- b) desenvolver um algoritmo que receba todas as informações e calcule, com certa precisão, a quantidade de energia prevista para ser gerada no dia desejado;
- c) verificar a confiabilidade do método criado para garantir uma margem de erro máxima que seja aceitável;
- d) analisar a viabilidade do método.

## 1.4 JUSTIFICATIVA

Neste trabalho foram criados dois sistemas de obtenção de dados, sendo o primeiro através do EXCEL e o segundo através do MATLAB, visando uma melhor aproximação e maior confiabilidade nos resultados obtidos após um comparativo entre métodos.

A possibilidade de poder prever a quantidade de energia elétrica que será gerada em determinada planta, usina, ou até mesmo para uso residencial, acrescentaria em muito para os sistemas de geração fotovoltaica, pois além de aumentar a eficiência geral, otimizaria e facilitaria o planejamento do mesmo, melhorando significativamente o resultado. As vantagens e benefícios em termos econômicos seriam imensas. Olhando para uma escala maior, serviria para despertar interesses, esclarecer dúvidas e incertezas, e assim solucionar possíveis problemas e fatores que normalmente seriam suficientes para descartar a possibilidade de um projeto, tornando-se mais um incentivo que motivaria a implementação de plantas de geração de energia fotovoltaica. (MONDOC B.; POP F, 2010)

## 1.5 PROCEDIMENTOS METODOLÓGICOS

Nesta seção serão apresentados os métodos escolhidos para a pesquisa e organização de informações, dados, programas e técnicas para futuras análises.

### 1.5.1 Metodologia Para Análise De Dados

A presente pesquisa tem como intenção utilizar de uma análise eficiente e difundida, bem como uma abordagem que efetue a previsão com o máximo de tempo possível com a utilização de estatísticas de curto prazo.

Ao mencionar curto prazo, significa que é a abordagem que providencia a previsão com a maior quantidade de tempo até o evento ocorrer, neste caso, de 5 a 7 dias, sendo essa a estimativa que necessita da obtenção de dados de maior qualidade. Idealmente, os sensores de medição precisam estar em perfeito funcionamento, livre de erros ou perturbações provocadas por falhas nos mesmos. Outro fator importante é o bom funcionamento dos painéis fotovoltaicos ou componentes, sejam eles os inversores ou transformadores presentes no circuito.

Erros ou discrepâncias nos dados obtidos irão afetar os ajustes finos do modelo e a qualidade das previsões.

Segundo Caldas, Melo, Pessanha e Falcão (2019, p. 4 e 5), a geração de energia solar fotovoltaica (Solar Photovoltaic Energy ou SPV, em inglês) na estratégia SVM (Support Vector Machine), que é um modelo de aprendizagem de máquina, tem forte dependência na incidência de irradiação solar no plano horizontal (Global Horizontal Irradiance ou GHI, em inglês), a estratégia adotada consiste na filtragem dos dados das medições de GHI e na estimação do seu efeito na geração SPV, por meio de equação de regressão linear e o tratamento de dados compreende cinco etapas:

1. Tratamento de valores nulos de GHI e geração, no horário de 4h00 a 21h00;
2. Filtragem dos registros superestimados em cada hora;
3. Identificação e preenchimento dos perfis diários com lacunas de dados nos registros horários de GHI;
4. Análise de agrupamentos de dados;
5. Ajuste de um novo modelo de regressão linear simples.

Devemos notar que o método SVM é um modelo difundido e popular para análise de dados. A escolha de um método específico ainda está pendente e sujeito

a alterações, pois, o software preferível de análise de dados será o Microsoft Excel, presente no pacote Microsoft Office.

### **1.5.2 Metodologia Para Obtenção De Dados**

Um sistema de energia fotovoltaica precisa, para a previsão da potência gerada, de dados meteorológicos precisos, com resultados que exemplificam todos os horários do dia, sendo esses dados:

- Previsão da temperatura ambiente;
- Previsão da precipitação pluviométrica;
- Previsão da cobertura de nuvens;
- Previsão do tempo (de forma geral).

A previsão do tempo é um dado comum, que pode ser obtido por qualquer um, com bom grau de confiabilidade e precisão. A própria ferramenta da Google que, ao pesquisar por “clima”, já fornece dados para até 7 dias adiante e, se o usuário estiver com a localização habilitada, fornecerá dados específicos para aquela localidade (precisão de bairro, 10km<sup>2</sup> em média para bairros maiores), com a opção de se informar de forma horária, com gráfico de temperatura, chance de chuva, umidade, velocidade e direção do vento. (CALDAS, MELO, PESSANHA E FALCÃO, 2019)

Para efeitos de consistência, o sistema desconsidera os períodos da madrugada e da noite, pois a obtenção de energia é nula. Segundo o método SVM, a cada hora H, é gerado um conjunto de dados para “treinar” a máquina e gerar uma possível previsão. (CALDAS, MELO, PESSANHA E FALCÃO, 2019)

## 2 FUNDAMENTAÇÃO TEÓRICA

A energia solar no Brasil representa apenas 1,7% da matriz energética, porém é uma área que vem crescendo consideravelmente nos últimos anos (especialmente nas regiões Sul e Sudeste) com o incentivo de benefícios econômicos e ambientais, principalmente em residências, onde é instalada com a intenção de reduzir a conta de energia. Já no Paraná, segundo a COPEL (principal companhia de distribuição de energia do estado), a energia solar tem sido cada vez mais adotada pela população, com o número de unidades consumidoras chegando a 7,7 mil em instalações somente em 2019, ocupando o 4º lugar dentre os estados em geração distribuída segundo a Agência Nacional de Energia Elétrica (Aneel) e a Associação Brasileira de Energia Solar Fotovoltaica (ABSOLAR), somando 268,9 MW de potência instalada. (PORTAL SOLAR, 2014-2022)

Olhando tanto para o Brasil quanto para o Paraná, temos condições muito favoráveis para o incentivo à implantação de geração fotovoltaica, justificando a elaboração de pesquisas sobre esta forma de geração de energia elétrica, em especial relacionadas à distribuição espacial da radiação solar no território paranaense. (TIEPOLO, 2018)

O custo fotovoltaico já é menor do que o valor cobrado pelas distribuidoras de energia na maior parte do país (segundo dados da ANEEL), além da oferta de equipamentos e serviços de instalação estarem cada vez mais acessíveis. Por outro lado, outro fator a ser considerado é a variabilidade do fornecimento de energia desses sistemas (fotovoltaicos) ao longo do dia: a produção de energia solar começa por volta das 6h, atinge o pico às 12h (em média) e vai reduzindo até zerar quando o sol se põe. Portanto, no caso de haver um grande número de sistemas fotovoltaicos instalados na rede, essa variação pode gerar instabilidade. (DANTAS E POMPERMAYER, IPEA, 2018)

Isso evidencia mais uma vez a importância de estudos como o presente trabalho, que além de contribuir com o aumento do incentivo à implantação de sistemas fotovoltaicos, também auxiliam na busca por soluções para as possíveis dificuldades a serem encontradas.

Neste tópico serão apresentadas informações teóricas, definições e dados de suma importância para o desenvolvimento deste trabalho, do mesmo modo, necessários para seu entendimento.

## 2.1 PRESSÃO ATMOSFÉRICA

No geral, se analisarmos de maneira direta, os valores de pressão atmosférica não afetam tanto os valores da média da radiação. Os valores de pressão atmosférica são os que menos variam dentre as variáveis analisadas (mantendo-se sempre levemente abaixo de 1 bar, como é esperado), assim podemos dizer que neste estudo esta é a variável mais estável. Entretanto, sua influência na radiação e conseqüentemente na geração ocorre de maneira indireta, já que a pressão atmosférica tem maior relação/influência nas outras variáveis (precipitação, umidade, temperatura e velocidade do vento), assim como estas outras variáveis também tem influência na variação da pressão.

A nível de superfície, a pressão atmosférica varia também devido a mudanças na densidade do ar. A densidade do ar pode variar conforme a temperatura (maiores temperaturas significam mais energia cinética no ar, e conseqüentemente menor densidade, e menor pressão), e/ou conforme a umidade relativa do ar (já que quanto maior a umidade maior a quantidade de água no ar, e conseqüentemente menor o peso da coluna de ar, já que o ar úmido é menos denso que o ar seco, então menor será pressão). (UFPEL, MARIANO, 2014)

Conforme mencionado acima, o ar úmido por ser menos denso tende a subir. Ao alcançar altas altitudes, encontrará menores pressões (fazendo-o se expandir) e menores temperaturas (fazendo-o resfriar e condensar), levando a formação de nuvens e, posteriormente, precipitação. Por esta razão a precipitação geralmente está ligada à situações de baixa pressão atmosférica. (CREF UFRGS, 2019)

## 2.2 PRECIPITAÇÃO

Dentre as variáveis analisadas neste estudo, a precipitação é, claramente, a variável cujos valores (quando positivos) mais afetam a média da radiação, de maneira negativa obviamente. Isso ocorre pois quando está chovendo a nebulosidade é alta (a nebulosidade impede que parte da radiação vinda do sol chegue até os painéis no nível do solo), e geralmente quanto maior a precipitação, maior a nebulosidade, e conseqüentemente menor a radiação nos painéis. Como neste estudo não se dispõe de dados da nebulosidade, faz-se uso destes dados de precipitação.

## **2.3 UMIDADE**

Embora os valores de precipitação (quando positivos) sejam os que mais impactam nos valores da média da radiação, pode-se perceber que, dentre as variáveis analisadas, a umidade se mostra sendo a mais diretamente ligada à média da radiação (pensando-se em linearidade). Assim, conclui-se que isso se deve ao fato da umidade estar diretamente ligada com a precipitação, por motivos triviais. Em geral, quanto maior a precipitação, maior a umidade, e menor a radiação. É claro que existem também períodos em que a precipitação é nula e mesmo assim a umidade é alta, e mesmo nestes momentos pode-se notar que a radiação tende a ser afetada negativamente. De toda forma, pode-se observar que quanto maior a umidade relativa do ar, menor a média de radiação nos painéis.

## **2.4 TEMPERATURA DO AR SUPERFICIAL**

A temperatura do ambiente influencia diretamente a temperatura dos painéis fotovoltaicos e, portanto, a sua eficiência na geração de energia. Sabe-se que os módulos solares funcionam melhor em certas condições climáticas estáveis e com baixas temperaturas, porém as condições climáticas estão em constante mudança, por isso o desempenho de sistemas fotovoltaicos varia conforme as condições de temperatura que está sujeito (InterEspaço, 2018).

## **2.5 IRRADIAÇÃO**

### **2.5.1 Emissão Solar**

Dada a distância entre o Sol e a Terra, apenas uma mínima parte da radiação solar emitida atinge a superfície da Terra, entretanto, essa quantidade de energia corresponde aproximadamente a dez mil vezes à demanda global de energia. Sendo assim, utilizando apenas 0,01 % desta energia, pode-se satisfazer a demanda energética total da humanidade. (PortalEnergia, 2004)

O Sol fornece anualmente, para a atmosfera terrestre, aproximadamente  $5,445 \times 10^{24}$  joules ou  $1,5 \times 10^{18}$  kWh de energia. (GEPEA, 2004)

## 2.5.2 Fatores Que Influenciam Na Irradiação Sobre A Terra

A intensidade da radiação solar (fora da atmosfera terrestre) depende da distância entre o Sol e a Terra, que varia um pouco durante o ano. Portanto a irradiância ( $E_0$ ) também varia, sendo seu valor médio em torno de  $E_0 = 1.367 \text{ W/m}^2$ . No entanto, obviamente nem toda essa quantidade atinge a superfície terrestre. A atmosfera reduz a radiação solar que chega até a superfície através da reflexão, absorção e dispersão. Assim, o nível de irradiância na Terra atinge um total aproximado de  $1.000 \text{ W/m}^2$  ao meio-dia, em boas condições climáticas, independentemente da localização. Agora, quando se adiciona a quantidade total da radiação solar que incide na superfície terrestre durante o período de um ano, obtém-se a irradiação global anual (medida em  $\text{kWh/m}^2$ ), que varia significativamente conforme a localização. (PortalEnergia, 2004)

### 2.5.2.1 Translação Terrestre

Em seu movimento em torno do Sol (anual), a Terra descreve em trajetória elíptica um plano que é inclinado de aproximadamente  $23,5^\circ$  com relação ao plano equatorial. Esta inclinação é responsável pela variação da elevação do Sol no horizonte em relação à mesma hora, ao longo dos dias, o que dá origem às estações do ano, e dificulta os cálculos da posição do Sol (e consequentemente do nível de radiação) para uma determinada data. A posição angular do Sol, ao meio dia solar, em relação ao plano do Equador (Norte positivo) é chamada de Declinação Solar ( $\delta$ ), ângulo este que varia (de acordo com o dia do ano) entre  $-23,45^\circ$  e  $23,45^\circ$ . (CRESESB, 2006)

Sabe-se que radiação solar que atinge o topo da atmosfera terrestre provém da região da fotosfera solar (camada com aproximadamente 300 km de espessura e temperatura superficial da ordem de 5800 K), porém, esta radiação não é constante/regular, pois existe a influência das camadas mais externas do Sol (cromosfera e coroa), além dos fenômenos que ocorrem na estrela. (CRESESB, 2006)

### 2.5.2.2 Radiação Direta x Difusa

Quando a radiação solar incide no limite superior da atmosfera terrestre, sofre uma série de reflexões, dispersões e absorções durante o seu percurso até o solo. A

incidência total da radiação solar sobre um corpo localizado no solo é a soma das componentes direta, difusa e refletida: radiação direta é a radiação proveniente diretamente do disco solar e que não sofreu nenhuma mudança de direção (além da provocada pela refração atmosférica); radiação difusa é aquela recebida por um corpo após a direção dos raios solares ter sido modificada por reflexão ou espalhamento na atmosfera; a radiação refletida depende das características do solo e da inclinação do equipamento captador. (GEPEA, 2004)

## **2.6 MODELAGEM DE SISTEMAS**

Quando se trata de sistemas de alta complexidade, em que o equacionamento matemático se torna problemático, uma possível solução é modelar o sistema utilizando apenas dados medidos empiricamente. Tal modelo é chamado de sistema caixa-preta, como é mostrado na figura 2. Um sistema caixa-preta tenta prever um modelo de forma recursiva através da leitura dos dados de entrada e saída do sistema, fazendo com que o modelo seja preciso dentro de uma faixa de erro para o conjunto de dados iniciais que geraram o modelo. Como consequência, se há correlação entre os dados de entrada e saída, o modelo gerado com os dados iniciais deve ser capaz de prever a saída para um segundo conjunto de dados, ainda dentro de uma faixa de erro. Tal processo está dentro da área de identificação de sistemas. Serão discutidos nas sessões 2.1.1.1 e 2.1.1.2 os modelos ARMAX, ARX (modelagem linear usando regressores recursivos) e NARMAX (modelagem não linear usando regressores recursivos). (ISERMANN, 2011; DUARTE, 2006)

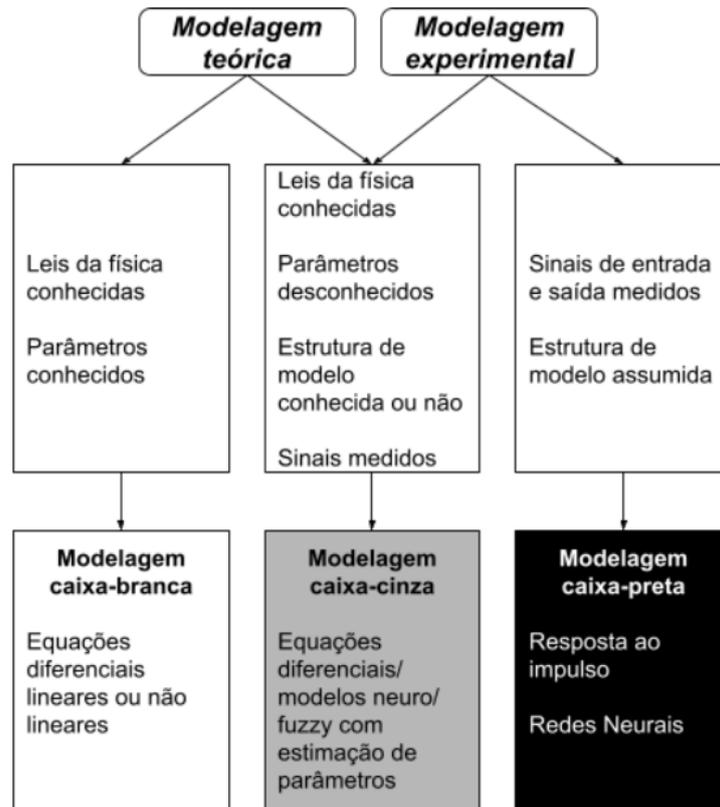


Figura 1: Tipos diferentes de modelos. Isermann, adaptado pelos autores.

Para entender o que são os modelos que serão apresentados a seguir, é importante inicialmente definir o que são modelos lineares autorregressivos (modelos AR), e o que são séries temporais.

**Séries temporais** é o nome dado a um conjunto de observações de uma variável  $y(k)$ , tal qual todas as medidas e observações são feitas em um instante  $k$  (BROCKWELL & DAVIS, 2002). A importância de conhecer séries temporais, no contexto deste trabalho, é entender que através do estudo das observações de uma série temporal, é possível realizar a análise da mesma encontrando, por exemplo, o comportamento dinâmico do modelo (DUARTE, 2006). Uma possível aplicação de séries temporais é a modelagem de um sistema Auto Regressivo (AR), como mostrado na equação 1. (DUARTE, 2006)

$$y(k) = \sum_{j=1}^{n_y} a_j y(k-j) + e(k) \quad \text{Equação 1: Modelo AR}$$

Observa-se que o modelo da equação 1 prevê o próximo passo da série temporal  $y(k)$  através do passo anterior da série temporal  $y(k-j)$  e dos parâmetros da

parte auto regressiva do modelo  $a_j$ . O resíduo  $e(k)$  é dado pelo erro um passo a frente entre a série observada e a série obtida pelo modelo. (DUARTE, 2006)

Vale lembrar que os modelos descritos neste capítulo levam em consideração que o sistema é invariante no tempo e, com a exceção do modelo NARMAX, lineares. (AGUIRRE, 2004, BILLINGS, 2013)

## 2.6.1 Modelos lineares

Quando se trata de modelagem linear, há vários tipos de modelos diferentes que podem ser empregados para alcançar um resultado capaz de obter o resultado esperado. Nesta seção serão descritos os modelos Autorregressivo Exógeno de média móvel, também chamado de ARMAX (*Auto-Regressive Moving Average eXogenous*) (Isermann, Munchhof) e um modelo linear mais simples, o modelo ARX (*Auto-Regressive eXogenous*), que se mostrou extremamente eficaz na aplicação de monitoramento de geração de energia fotovoltaica, como foi mostrado por Lazzaretti (2020, tradução do autor): “Modelos lineares ARX são um dos modelos dinâmicos mais simples. Portanto, sua implementação computacional consome menos tempo, o que o torna ideal para a implementação de sistemas de monitoramento na área fotovoltaica”.

### 2.6.1.1 Modelo Autorregressivo exógeno com média móvel (ARMAX)

Este modelo é o mais complexo entre todos os modelos lineares que serão vistos neste trabalho. A modelagem ARMAX leva em conta, além das séries temporais do modelo estimado e o resíduo, como no modelo AR, uma filtragem de média móvel e as entradas do sistema, que se trata de entradas exógenas (ISERMANN, 2011).

O modelo ARMAX é obtido através do modelo geral para modelagem discreta, dada pelas equações a seguir:

$$A(q)y(k) = \frac{B(q)}{F(q)}u(k) + \frac{C(q)}{D(q)}v(k) \quad \text{Equação 2: Modelo geral para modelagem de sistemas discretos (Duarte, 2006; Aguirre, 2004)}$$

$$A(q) = 1 + a_1q^{-1} + \dots + a_{n_y}q^{-n_y};$$

$$B(q) = b_1q^{-1} + \dots + b_{n_u}q^{-n_u};$$

$$C(q) = 1 + c_1q^{-1} + \dots + c_{n_c}q^{-n_c};$$

$$D(q) = 1 + d_1q^{-1} + \dots + d_{n_d}q^{-n_d};$$

$$F(q) = 1 + f_1q^{-1} + \dots + f_{n_f}q^{-n_f}; \quad \text{Equação 3: Polinômios do modelo geral da equação 2} \\ \text{(Duarte, 2006)}$$

Na equação 2, o operador  $(k)$  é a representação de um ruído branco, e “ $q$ ” é o operador de atraso. Logo, um modelo armax é dado pelas relações 4, mostradas a seguir, em que o resíduo  $e(k)$  é dado pela última parcela do modelo geral. Logo, o resíduo é dado por  $e(k) = \frac{C(q)}{A(q)}u(k)$ , tal qual a parcela  $\frac{C(q)}{A(q)}$  é dada pela filtragem do sistema através de um filtro ARMAX desconsiderando as entradas exógenas, ou seja, por uma filtragem ARMA (*Auto-Regressive Moving Average*). (AGUIRRE, 2004; DUARTE, 2006)

$$y(k) = \frac{B(q)}{A(q)}u(k) + \frac{C(q)}{A(q)}v(k)$$

$$y(k) = a_0 + \sum_{j=1}^{n_y} a_j y(k-j) + \sum_{i=1}^{n_u} b_i u(k-i) + e(k) \quad \text{Equação 4: representação} \\ \text{de um modelo ARMAX, Duarte 2006, Billings 2013}$$

Observa-se que na equação 4 (representação de um modelo ARMAX) há uma complexidade muito maior que em um modelo AR, pois a quantidade de regressores é muito maior, uma vez que há regressores na série temporal  $y(k)$  assim como nas entradas exógenas  $u(k)$  (além da adição de um filtro ARMA). Isso faz com que o sistema seja, no geral, muito mais preciso e muito mais lento computacionalmente quando comparado a modelos lineares mais simples. Logo, para sistemas em que a previsão deve ser rápida, sistemas como o ARX podem ser uma opção válida (LAZZARETTI, 2020).

### 2.6.1.2 Modelo Autorregressivo exógeno (ARX)

Como já comentado, o modelo ARX é um dos modelos mais simples para obter um modelo linear. O modelo ARX, como o ARMAX, é um modelo não autônomo que leva em consideração as entradas exógenas do sistema. O modelo ARX é o equivalente não autônomo do sistema autônomo AR. (DUARTE, 2006)

O modelo ARX, como o modelo ARMAX, utiliza a o modelo geral para modelagem de sistemas discretos (equação 4), todavia sem utilizar uma filtragem

ARMA complexa. Logo, a filtragem realizada pelo modelo ARX é uma filtragem auto-regressiva com os pólos iguais aos pólos do processo, ou seja as raízes do polinômio  $A(q)$ . Logo, o resíduo  $e(k)$  é dado pelo seguinte processo branco filtrado  $e(k) = \frac{v(k)}{A(q)}$ .

O modelo ARX é dado pelas equações representadas a seguir.

$$y(k) = \frac{B(q)}{A(q)}u(k) + \frac{1}{A(q)}v(k),$$

$$y(k) = a_0 + \sum_{j=1}^{n_y} a_j y(k-j) + \sum_{i=1}^{n_u} b_i u(k-i) + e(k) \quad \text{Equação 5: representação de um modelo ARX, Duarte 2006, Billings 2013}$$

Observa-se que a substituição do filtro ARMA pelo filtro auto regressivo faz com que o sistema se torne mais simples que o ARMAX, o tornando ideal para aplicações em que é necessária velocidade computacional (LAZZARETTI, 2020).

## 2.6.2 Modelos não lineares (NARMAX)

Para aplicações em que o conjunto de dados não é linear há a necessidade de se utilizar um modelo que não assuma a linearidade do sistema, como é feito em todos os casos anteriores (AR, ARMAX e ARX). No entanto, tentar modelar um sistema não linear vem com o custo computacional atrelado ao fato de se estar trabalhando com um conjunto de dados de alto grau de complexidade.

Este trabalho não possui a pretensão de aplicar modelos não lineares, porém, como modelos de previsão meteorológica são altamente não lineares, entende-se a importância de se descrever uma possível solução não linear para o problema de geração fotovoltaica.

Modelos NARMAX são extremamente complexos, no sentido de relacionar as entradas e as saídas de tal forma que se o mapeamento não linear  $y(k)$ , para um resíduo não branco  $\xi(k)$  como mostrado na equação 6.

$$y(k) = F(y^{|k-1|}, u^{|k-1|}, \varepsilon^{|k-1|}) + \varepsilon(k)$$

$$y^{[k]} = [y(k-1), y(k-2), \dots, y(k-n_y)]$$

$$u^{[k]} = [u(k-1), u(k-2), \dots, u(k-n_u)]$$

$$\varepsilon^{[k]} = [\varepsilon(k-1), \varepsilon(k-2), \dots, \varepsilon(k-n_\varepsilon)]$$

*Equação 6: forma simplificada do mapeamento não linear NARMAX*

À primeira vista a modelagem NARMAX não aparenta ser tão complicada, todavia entende-se que as entradas e saídas são não lineares e o resíduo é uma série independente que pode ser linear ou não, tornando a resolução do sistema longa e complexa, havendo a necessidade de se realizar uma série de aproximações durante o processo. Isso faz com que a solução de um sistema NARMAX seja inviável para a proposta do trabalho. (Billings, 2013)

### **2.6.3 CRITÉRIO DE INFORMAÇÃO DE AKAIKE**

O critério de informação de Akaike é uma forma tradicionalmente usada para encontrar modelos que tenham, objetivamente, o melhor ajuste sobre um certo conjunto de dados. A equação é mostrada na Equação 7, onde  $N$  é o número de observações,  $E$  é o erro médio quadrático do modelo e  $K$  é o número de parâmetros. O resultado é dado por AIC, tal qual o modelo que possui o menor AIC, é o mais adequado para tal modelo e conjunto de dados (Emiliano, 2010).

$$AIC = N \ln(E/N) + 2K \quad \text{Equation 7: Critério de informação de Akaike.}$$

### 3 DESENVOLVIMENTO

#### 3.1 ANÁLISE INICIAL

Analisando a planilha e as variáveis meteorológicas que nela constam, optou-se por utilizar como base 6 variáveis/grandezas:

- **humid-so10CT** (%) - umidade;
- **press-so10CT** (mb) - pressão atmosférica;
- **rain-so10CT** (mm/h) - precipitação;
- **tp\_sfc-so10CT** (°C) - temperatura do ar superficial;
- **ws\_max-so10CT** (m/s) - velocidade máxima do vento;
- **tilt\_avg-so01CT** (W/m<sup>2</sup>) - média da radiação;

Como o objetivo é estimar e prever a quantidade de energia gerada por um sistema fotovoltaico fazendo uso de informações e variáveis meteorológicas, e partindo dos dados disponíveis, a melhor escolha a princípio foi basear-se nestas 5 variáveis citadas acima encontradas na planilha de dados do mês de junho de 2021, planilha esta que possui dados que foram coletados de minuto a minuto.

Dentre as variáveis presentes nos dados disponíveis, a média da radiação é a que está mais fortemente ligada à geração fotovoltaica. Portanto, buscou-se encontrar uma relação entre as demais variáveis para com a média da radiação, pois a maneira com que estas 4 variáveis influenciam e alteram a média da radiação é consequentemente a mesma maneira com que alteram a geração de energia, já que a quantidade de radiação recebida/incidente é diretamente proporcional à quantidade de energia gerada em um sistema fotovoltaico. (BEIGELMAN, 2013)

Então, com o intuito de analisar as informações disponíveis, foi realizada uma filtragem na planilha de dados, de modo que só restassem valores de medições “confiáveis” para uma tentativa de aproximação e simulação matemática: como os tempos de amostragem de alguns dos sensores divergem, em alguns períodos carecem informações, com medições nulas; outro ponto foi a remoção das medições de períodos noturnos, onde a radiação é nula;

Após a filtragem dos dados, obteve-se uma planilha reduzida com valores “válidos” e utilizáveis de umidade, pressão atmosférica, precipitação, temperatura do ar superficial e média da radiação, coincidentes no tempo. Assim, obteve-se uma

planilha com valores de todas as 5 variáveis escolhidas amostradas pelo período das 6:50 às 17:50 horas, de todos os dias do mês de junho de 2021.

### **3.2 OBSTÁCULOS ENCONTRADOS**

No presente projeto, a obtenção de dados ocorreu através das tabelas do INMET (Instituto Nacional de Meteorologia), que disponibiliza dados de Data, Hora (UTC), Umidade (%), Pressão (hPa), Vento (Velocidade em m/s e direção em graus), Nebulosidade (Décimos), Insolação (horas), Temperatura Máxima (Graus), Temperatura Mínima (Graus) e Chuva (milímetros). Para cumprir com o propósito do projeto o dado necessário é de irradiância solar, expressa em Watts por metro quadrado (SANDIA NATIONAL LABORATORIES, 2021).

O banco de dados extraído tem início na data de primeiro de Janeiro de 2000 e foi atualizado até o dia 15 de Maio de 2021.

Como a irradiância ausente na base de dados do INMET, foi utilizada a insolação, em horas, que, após a obtenção do modelo, seria multiplicada por 1 kiloWatt, que é considerada a média da irradiância no solo (SOLARVIEW, 2020). Como os dados são diários e têm-se essa informação da média da irradiância, esta multiplicação resulta na quantia de energia que pode ser gerada naquele determinado dia.

### **3.3 SEGUNDA ANÁLISE**

Conforme descrito anteriormente, inicialmente, utilizando as 6 variáveis mencionadas no tópico 3.1 (umidade, pressão atmosférica, precipitação, temperatura do ar superficial, velocidade média do vento e média da radiação), tentou-se encontrar a relação entre as 5 primeiras com a média da radiação. O objetivo foi encontrar uma equação onde a umidade, pressão, precipitação, temperatura e velocidade do vento fossem as “entradas” e a radiação fosse a “saída”.

Após muita pesquisa e tentativas falhas, o método encontrado que mais se aproximou do que seria uma possível solução foi a regressão linear, onde, têm-se as variáveis umidade, pressão, precipitação, temperatura e velocidade do vento como X (entrada) e a radiação como Y (saída).

Entretanto, após novos ensaios, esta abordagem não se mostrou viável devida a aleatoriedade observada dos resultados. Isso se deve ao fato das variáveis “de

entrada” possuem muita influência entre si, e não somente na radiação: sabemos por exemplo, que a precipitação influencia na umidade, na temperatura; que a velocidade do vento tem ligação com a pressão atmosférica, com a temperatura e precipitação também. Ou seja, trata-se de um sistema muito complexo e dinâmico, cuja tentativa de linearização, nestas condições, se mostra ineficaz.

Por esta razão, optou-se por outra abordagem: considerar todas estas variáveis tratadas até o momento como sendo variáveis de entrada (incluindo a média da radiação), e como saída, utilizar os valores de potência gerada:

- **pac-cdteCT** (W)- Potência da rede AC - inversor de 1,5kW;
- **pac-cigsCT** (W)- Potência da rede AC - inversor de 1,5kW;
- **pac-mon1CT** (W)- Potência da rede AC - inversor de 3kW (monocristalino 1);
- **pac-mon2CT** (W)- Potência da rede AC - inversor de 3kW (monocristalino 2);
- **pac-pol1CT** (W)- Potência da rede AC - inversor de 3kW (policristalino 1);
- **pac-pol2CT** (W)- Potência da rede AC - inversor de 3kW (policristalino 2);

Utilizando agora o método da regressão linear com estas novas variáveis “X” e “Y”, obteve-se um resultado muito mais satisfatório.

Inicialmente pensou-se em considerar o período do dia (hora) no cálculo (neste caso, na regressão linear), mas isso só faria sentido na tentativa anterior, onde a regressão tentava estimar a equação para o valor da média da radiação (tilt\_avg, como saída), já que esta varia conforme o horário do dia (assim cada hora deveria ter um peso diferente). Porém, agora que o alvo (saída) da regressão são os valores de potência gerados, a média da radiação agora está entre as variáveis de entrada na regressão. Dessa forma, a informação de horário não entra no cálculo, pois ela de certa forma já está contida nos valores de radiação.

No modelo de regressão utilizado, pode-se ter inúmeras variáveis de entrada “X”, porém apenas uma variável de saída “Y”. Neste caso, como o presente trabalho dispõe de 6 valores de potência (6 variáveis de saída), foram realizadas 6 regressões lineares, uma para cada variável de potência.

## 3.4 IMPLEMENTAÇÃO POR SOFTWARES

Com tais dados a disposição, iniciou-se a montagem do primeiro sistema de previsão através do software **Excel** e, posteriormente, houve a implementação através do software **MatLab**.

### 3.4.1 Excel

Para a implementação através do Excel fez-se uso da Regressão Linear, a qual será devidamente justificada (tópico 4.1.3) como o método de maior confiabilidade para tal análise.

#### 3.4.1.1 Regressão linear

A regressão linear é uma expressão que efetua a tentativa de estipular o relacionamento entre duas variáveis, adicionando uma equação linear aos dados observados. Uma variável é a explanatória, e a outra é a dependente, sendo, neste caso, a data a explanatória e a insolação a dependente (Yale, 1997).

A equação linear tem a forma de:

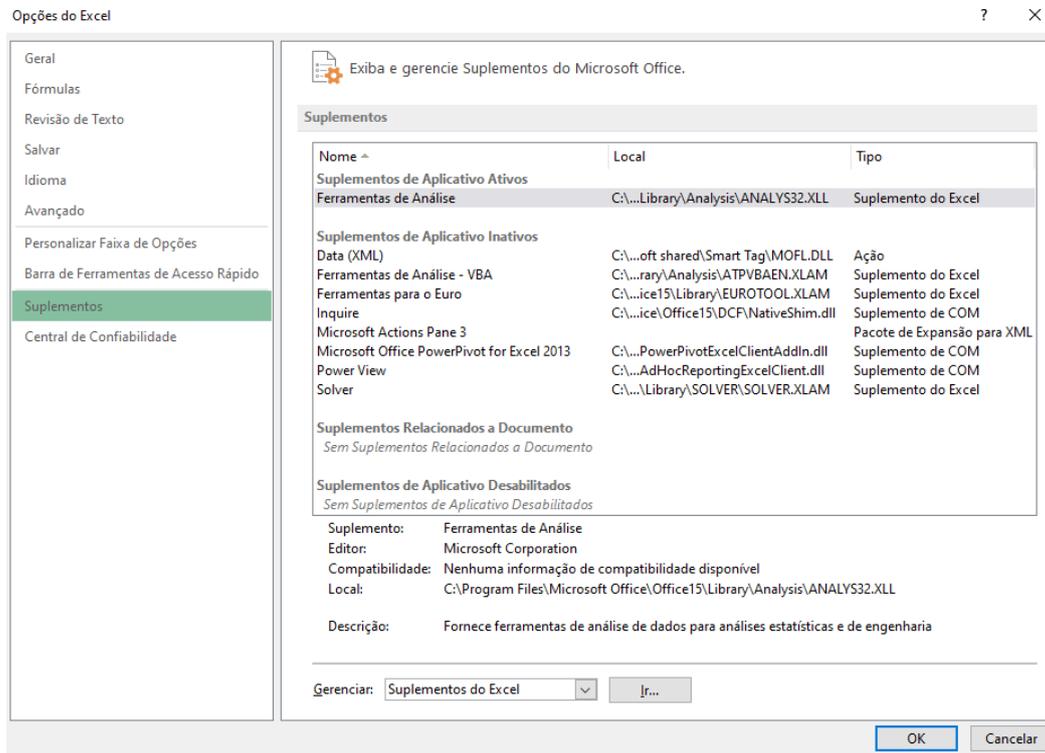
$$Y = a + b \cdot x \quad \text{Equation 8: equação de linearização}$$

Para a realização do cálculo da potência gerada, foi necessário encontrar uma fórmula que relacionasse todas as demais variáveis medidas com a potência gerada. O método escolhido foi a regressão linear.

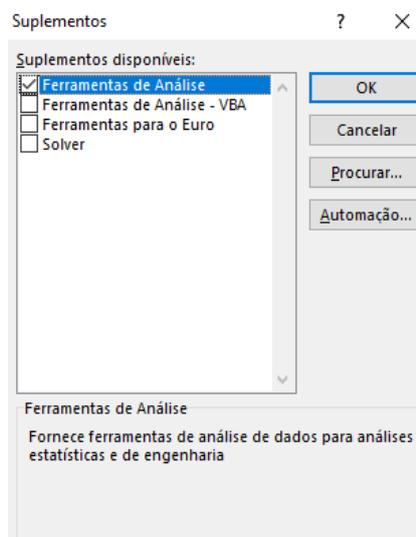
A análise de regressão linear é usada para prever o valor de uma variável com base no valor de outra ou outras, estimando o valor da variável de saída através da realização de uma soma ponderada dos valores das variáveis de entrada. Sendo assim, a variável que se deseja prever é chamada de variável dependente, já a variável que é usada para prever o valor da anterior é chamada de variável independente. Este tipo de análise estima os coeficientes da equação linear, envolvendo uma ou mais variáveis independentes que melhor preveem o valor da variável dependente. Dessa forma, a regressão linear se ajusta a uma linha reta que minimiza as discrepâncias entre os valores de saída previstos e reais.

No caso tratado neste trabalho, como já explicando anteriormente, as variáveis independentes ou "de entrada" foram as medições humid, press, rain, tp-sfc, ws\_max e tilt\_avg, já as variáveis dependentes ou "de saída" seriam as medições de potência pac-cdte, pac-cigs, pac-mon1, pac-mon2, pac-pol1 e pac-pol2.

A ferramenta utilizada foi a própria ferramenta de análise de dados do excel. Esta funcionalidade pode ser habilitada em **Arquivo > Opções > Suplementos > Gerenciar: Ir > Ferramentas de Análise > Ok**, como exposto pelas figuras 2 e 3 a seguir:

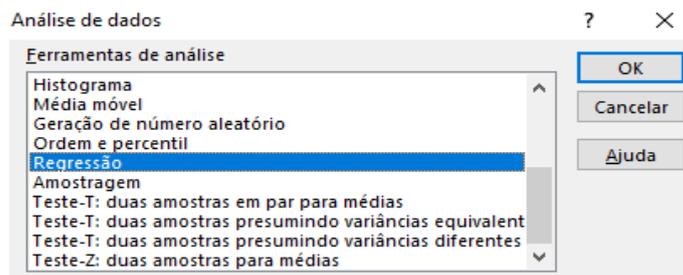


**Figura 2: Excel – Janela de Opções**

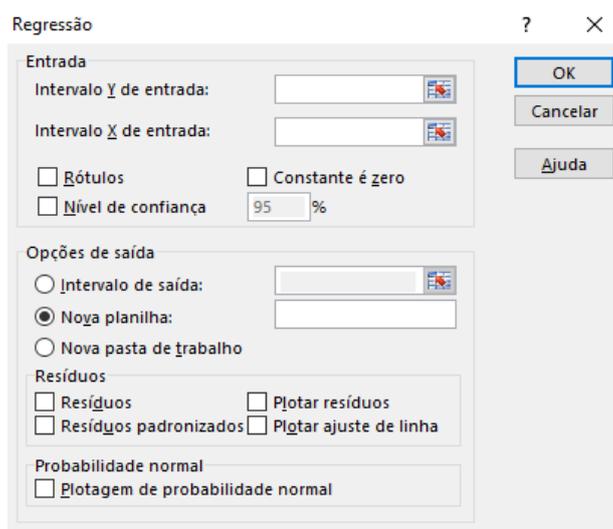


**Figura 3: Excel - Gerenciador de Suplementos.**

Após habilitar esta funcionalidade, pode-se acessar na aba **DADOS > Análise de Dados > Regressão**, como ilustrado pelas figuras 4 e 5 a seguir:



**Figura 4: Excel - Janela Análise de Dados.**



**Figura 5: Excel - Janela da Função Regressão.**

Em “**Intervalo Y de entrada**” são selecionados os dados de potência (variável dependente), e em “**Intervalo X de entrada**” são selecionados os dados de umidade, pressão, precipitação, temperatura, velocidade do vento e média da radiação. Para facilitar a plotagem dos resultados recomenda-se habilitar “**Rótulos**” e, ao selecionar as colunas dos intervalos de entrada, incluir a primeira linha que contém os nomes das variáveis. Em seguida, em “**Opções de saída**” marcar a opção “**Nova planilha**”

A ferramenta de análise Regressão do excel executa uma análise de regressão linear usando o método de "quadrados mínimos" para ajustar uma linha em um conjunto de observações. (Fonte: Fórum Microsoft)

### 3.4.1.2 Método Dos Mínimos Quadrados

É uma técnica de otimização matemática que permite, de forma aproximada, retirar informação de sistemas “impossíveis”, também conhecidos como sistemas que simplesmente não possuem soluções. Esta técnica procura encontrar o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados dos erros obtidos na aproximação (as diferenças entre o valor estimado e os dados observados). (UFRGS 2020)

O resultado da regressão linear retorna valor da intersecção (valor constante, que serve como ponto de partida/inicial) e um coeficiente para cada uma das variáveis independentes (de entrada). A equação que calcula a potência gerada apresenta a seguinte forma:

$$Y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + [\dots] + \alpha_n x_n \quad \text{Equação 9: Calculo de potência gerada.}$$

Onde  $Y$  é a variável dependente (saída, neste caso, a potência gerada),  $\alpha_0$  é a intersecção,  $\alpha_1$  é o coeficiente referente à respectiva variável independente de mesmo índice, e  $x_1$  a variável independente (neste caso, humid, press, rain, tp-sfc, ws\_max e/ou tilt\_avg).

### 3.4.1.3 Aplicação da Regressão Linear

Realizadas as regressões lineares para os 6 valores de potência, obteve-se os seguintes resultados para cada variável de entrada:

- **pac-cdteCT**

|                 | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-----------------|---------------------|--------------------|
| Interseção      | 365,111853          | 1201,89178         |
| humid-so10CT    | -0,397390314        | 0,189402976        |
| press-so10CT    | -0,406260589        | 1,312814958        |
| rain-so10CT     | 1,417869059         | 17,4596835         |
| tp_sfc-so10CT   | -0,476937892        | 0,687658642        |
| ws_max-so10CT   | 0,111001634         | 0,091517051        |
| tilt_avg-so01CT | 1,302128114         | 0,009433479        |

*Tabela 1: Variável pac-cdteCT.*

Os coeficientes indicados pela tabela acima (Tabela 1) são os pesos que cada variável terá no cálculo da potência **pac-cdteCT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-cigsCT**

|                 | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-----------------|---------------------|--------------------|
| Interseção      | 2348,974895         | 1344,364313        |
| humid-so10CT    | -0,644463769        | 0,211854849        |
| press-so10CT    | -2,570201286        | 1,468436351        |
| rain-so10CT     | 1,604400199         | 19,52935847        |
| tp_sfc-so10CT   | -3,008727343        | 0,769173859        |
| ws_max-so10CT   | 0,184573048         | 0,102365504        |
| tilt_avg-so01CT | 1,501475013         | 0,010551726        |

*Tabela 2: Variável pac-cigsCT.*

Da mesma forma, os coeficientes indicados pela tabela acima (Tabela 2) são os pesos que cada variável terá no cálculo da potência **pac-cigs**, já o erro padrão significa o quanto este peso é incerto.

- **pac-mon1CT**

|                 | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-----------------|---------------------|--------------------|
| Interseção      | -1720,316169        | 2176,678128        |
| humid-so10CT    | -0,940687507        | 0,343017002        |
| press-so10CT    | 1,713083358         | 2,377564814        |
| rain-so10CT     | 8,82390099          | 31,62024388        |
| tp_sfc-so10CT   | -0,878019357        | 1,245379618        |
| ws_max-so10CT   | 0,782920822         | 0,165741348        |
| tilt_avg-so01CT | 2,057383511         | 0,01708444         |

*Tabela 3: variável pac-mon1CT.*

De forma semelhante, os coeficientes citados na tabela acima (Tabela 3) são os pesos que cada variável terá no cálculo da potência **pac-mon1CT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-mon2CT**

|                 | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-----------------|---------------------|--------------------|
| Interseção      | -2260,149307        | 2056,517516        |
| humid-so10CT    | -0,591918917        | 0,324081206        |
| press-so10CT    | 2,270381957         | 2,246314521        |
| rain-so10CT     | 19,03693716         | 29,87469051        |
| tp_sfc-so10CT   | -0,536785895        | 1,176630098        |
| ws_max-so10CT   | 0,76773997          | 0,156591818        |
| tilt_avg-so01CT | 1,902956658         | 0,016141316        |

*Tabela 4: Variável pac-mon2CT.*

Os coeficientes citados na tabela acima (Tabela 4) são os pesos que cada variável terá no cálculo da potência **pac-mon2CT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-pol1CT**

|                 | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-----------------|---------------------|--------------------|
| Interseção      | 1008,456462         | 1930,222702        |
| humid-so10CT    | -0,814927648        | 0,304178737        |
| press-so10CT    | -1,145245644        | 2,108363897        |
| rain-so10CT     | 7,337426207         | 28,04002658        |
| tp_sfc-so10CT   | -3,41990441         | 1,104370913        |
| ws_max-so10CT   | 0,490311012         | 0,146975204        |
| tilt_avg-so01CT | 1,808792403         | 0,015150046        |

*Tabela 5: Variável pac-pol1CT.*

Da mesma forma, os coeficientes citados na tabela acima (Tabela 5) são os pesos que cada variável terá no cálculo da potência **pac-pol1CT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-pol2CT**

|                 | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-----------------|---------------------|--------------------|
| Interseção      | 776,9527525         | 1411,52985         |
| humid-so10CT    | -0,875140404        | 0,22243929         |
| press-so10CT    | -0,892893076        | 1,541800629        |
| rain-so10CT     | 10,9626536          | 20,50506114        |
| tp_sfc-so10CT   | -3,195384417        | 0,807602412        |
| ws_max-so10CT   | 0,500374802         | 0,107479768        |
| tilt_avg-so01CT | 1,96102085          | 0,011078899        |

*Tabela 6: Variável pac-pol2CT.*

Por último, os coeficientes citados na tabela acima (Tabela 6) são os pesos que cada variável terá no cálculo da potência **pac-pol2CT**, já o erro padrão significa o quanto este peso é incerto.

Observando os valores dos coeficientes e o erro padrão obtidos, pode-se verificar, a grosso modo, qual o grau de influência que cada uma das variáveis analisadas

causa na potência gerada. Valores positivos nos coeficientes indicam uma influência “positiva” na potência gerada, enquanto que valores negativos indicam uma influência “negativa” na potência gerada. Sendo assim, constataram-se como “verdadeiras” a maior parte das especulações realizadas previamente à esta abordagem, puramente empíricas e presunçosas, mas baseadas na observação dos dados estudados:

- percebe-se que, como esperado, para todos os casos, o coeficiente relacionado à média da radiação (tilt\_avg-so01CT) apresenta um valor positivo, pois quanto maior a radiação incidente nos módulos, maior a potência gerada;
- o coeficiente relacionado à temperatura (tp\_sfc-so10CT), também como esperado, apresenta valores negativos para todos os casos, pois sabe-se que a temperatura elevada (apesar de muitas vezes ser relacionada com altos níveis de radiação solar) reduz o desempenho do sistema fotovoltaico;
- o coeficiente relacionado à umidade (humid-so10CT) novamente, como esperado, apresenta valores negativos para todos os casos. Fica evidente quando observado, que períodos com valores de umidade mais altos apresentam valores de radiação (e conseqüentemente potência gerada) mais baixos;
- o coeficiente relacionado à velocidade máxima do vento (ws\_max-so10CT) apresenta valores positivos para todos os casos;
- o coeficiente relacionado à pressão atmosférica (press-so10CT) apresenta valores negativos para a maioria dos casos, sendo positivo apenas para as potências onde é utilizada tecnologia monocristalina (pac-mon1CT e pac-mon2);

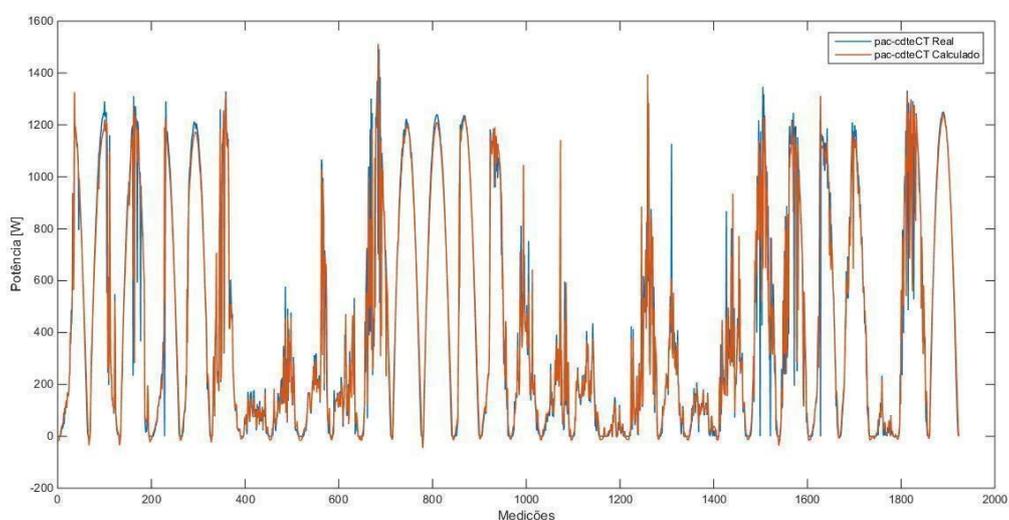
já o coeficiente relacionado à precipitação (rain-so10CT) apresenta valores positivos para todos os casos, o que é totalmente inesperado e contra intuitivo. Entretanto, analisando novamente os dados, constatou-se que a regressão linear definiu coeficientes positivos para a precipitação devido ao fato da grande maioria dos dados desta variável serem iguais à zero (apenas 5,72% apresentam valores positivos diferentes de zero), e ainda as poucas medições positivas apresentam valores múltiplos de 0.2 (0.2, 0.4, 0.6, 0.8, 1.0, 1.2, [...]). Isso também fica claro quando observados os valores de Erro padrão dos coeficientes relacionados à precipitação, que são consideravelmente maiores que o Erro padrão referente às outras variáveis.

Ou seja, os valores dos coeficientes e erros referentes à precipitação são consequências da forma com que foi feita a amostragem desta variável

Tendo em mãos os coeficientes referentes a cada uma das variáveis, foi possível montar a equação para determinar a potência (no caso, foram 6, uma para cada variável de potência), e assim calcular e obter valores teóricos, com o objetivo de comparar estes valores calculados com os valores de potência reais (os valores das potências medidos).

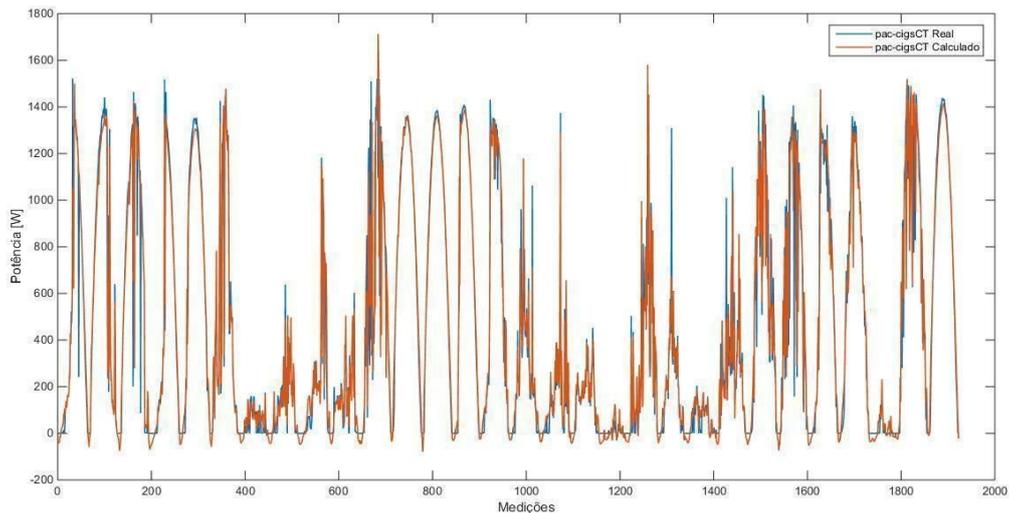
Comparando os valores de potência calculados e reais para cada uma das medições (conforme exposto anteriormente, estas medições são de Junho de 2021, realizadas com intervalos de 10 minutos), considerando somente os valores válidos e diurnos (considerando somente as medições de períodos com valor da média da radiação positiva), obteve-se um resultado significativamente satisfatório.

Podemos comparar os resultados nos gráficos abaixo:



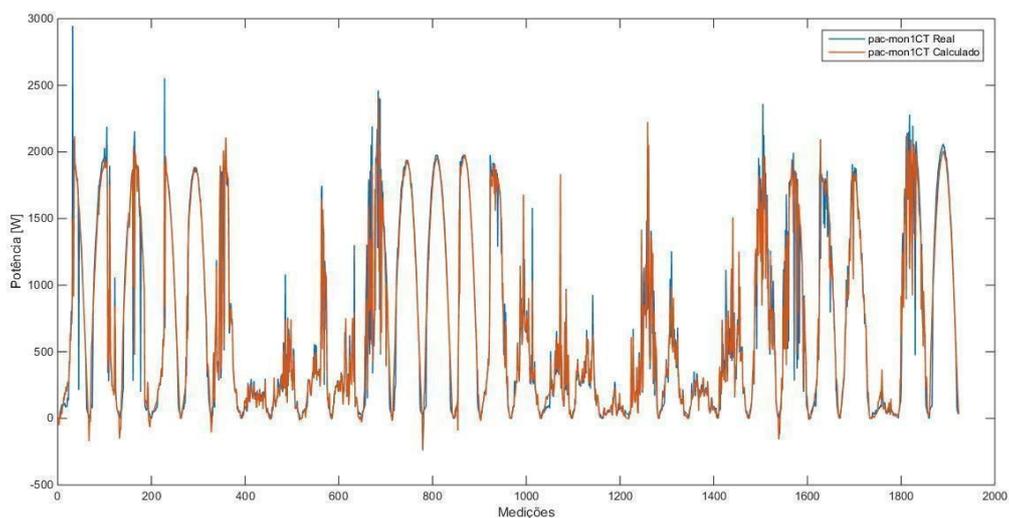
**Figura 6: Gráfico Comparativo - Regressão da variável pac-cdteCT.**

A linha azul do gráfico acima (Figura 6) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante similaridade entre o valor das duas, conforme sua amplitude (eixo Y) varia ao longo do tempo/medições (eixo X).



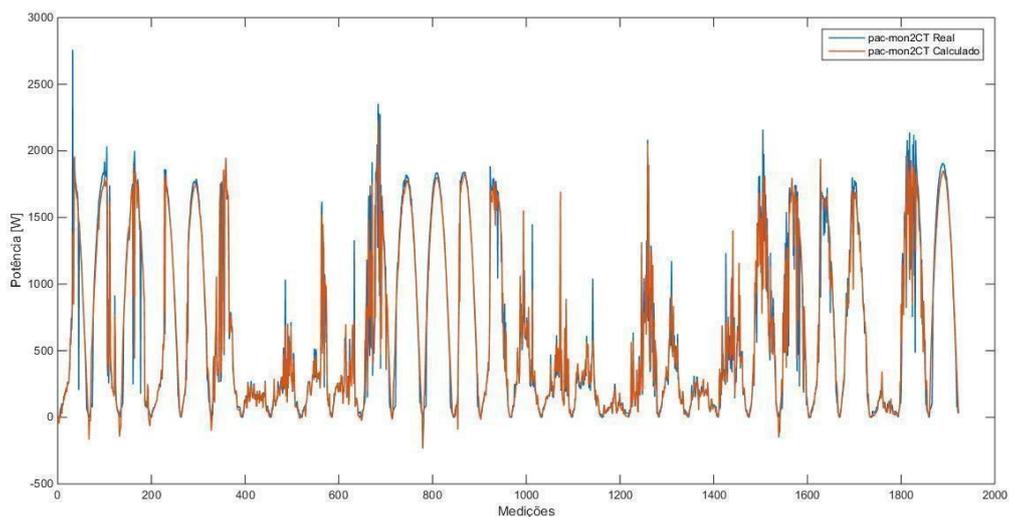
**Figura 7: Gráfico Comparativo - Regressão da variável *pac-cigsCT*.**

Da mesma forma, a linha azul do gráfico acima (Figura 7) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante similaridade entre o valor das duas, conforme sua amplitude (eixo Y) varia ao longo do tempo/medições (eixo X).



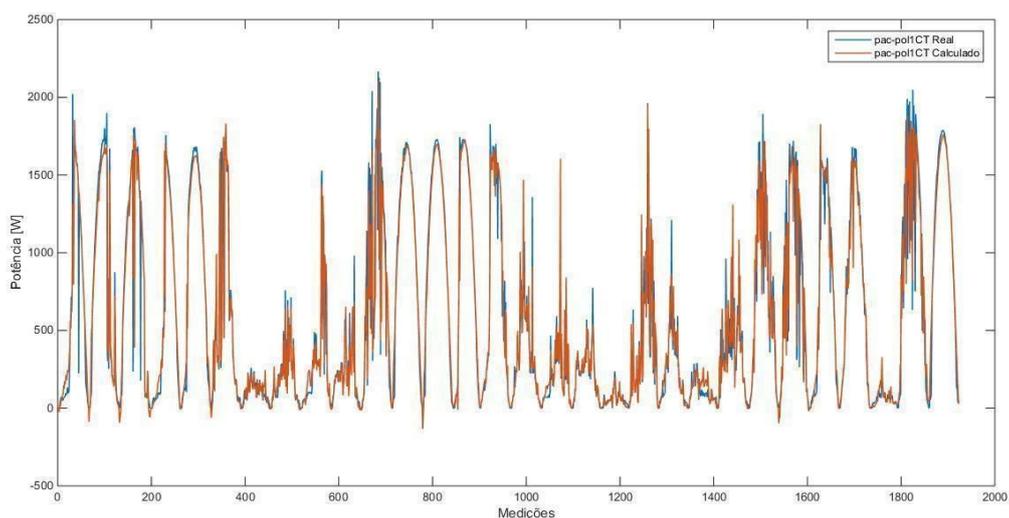
**Figura 8: Gráfico Comparativo - Regressão da variável *pac-mon1CT*.**

De modo semelhante, a linha azul do gráfico acima (Figura 8) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante similaridade entre o valor das duas, conforme sua amplitude (eixo Y) varia ao longo do tempo/medições (eixo X).



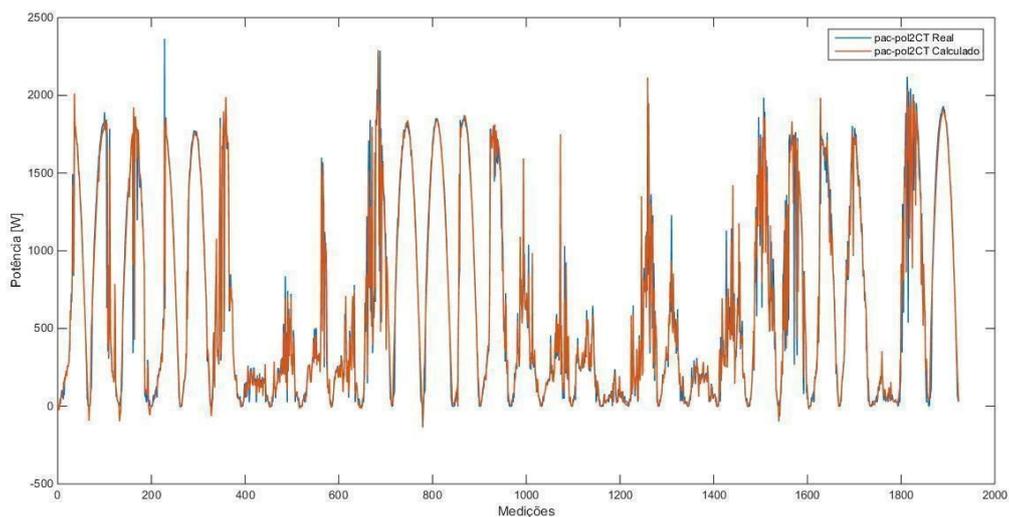
**Figura 9: Gráfico Comparativo - Regressão da variável pac-mon2CT.**

A linha azul do gráfico acima (Figura 9) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante similaridade entre o valor das duas, conforme sua amplitude (eixo Y) varia ao longo do tempo/medições (eixo X).



**Figura 10: Gráfico Comparativo - Regressão da variável pac-pol1CT.**

Da mesma forma, a linha azul do gráfico acima (Figura 10) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante similaridade entre o valor das duas, conforme sua amplitude (eixo Y) varia ao longo do tempo/medições (eixo X).



**Figura 11: Gráfico Comparativo - Regressão da variável pac-pol2CT.**

Por fim, a linha azul do gráfico acima (Figura 11) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante similaridade entre o valor das duas, conforme sua amplitude (eixo Y) varia ao longo do tempo/medições (eixo X).

Nota-se uma grande semelhança entre os valores calculados (teóricos) e os valores medidos (reais), para todas as 6 potências analisadas, evidenciando a precisão do cálculo de regressão linear do excel e reforçando a escolha da abordagem empregada nesta análise.

### 3.4.1.4 Cálculo de Previsão de Energia Gerada

Após realizado o cálculo/estimativa da potência gerada baseada nas variáveis climáticas, o próximo passo é a aplicação de técnicas de previsão de dados, com o intuito de prever de maneira satisfatória a potência gerada num futuro próximo.

Dado o grau de assertividade obtido com a abordagem da regressão linear para a estimativa de potência, e tendo em mãos a equação da potência teórica (para as 6 potências analisadas), a alternativa mais plausível consiste na aplicação de modelos e técnicas de previsão de dados nas variáveis independentes (umidade, pressão, precipitação, temperatura, velocidade do vento e média da radiação), pois com os valores “futuros” destas variáveis, pode-se estimar estes valores de potência utilizando a equação obtida com a regressão linear.

Através da funcionalidade “**Planilha de previsão**”, presente na aba “**Dados**” e na subseção “**Previsão**”, a ferramenta utiliza uma interface personalizável que faz uso da função “**PREVISÃO.ETS**” para prever dados futuros baseando-se nos dados passados.

Essa função utiliza o método de suavização exponencial (tópico 3.1.4.3), que é um modelo de série temporal, contendo um componente de erro (E), um componente de tendência (T) e um componente de sazonalidade (S). Este modelo, especificamente, utiliza a versão AAA, ou seja, todos os componentes são aditivos.

#### 3.4.1.4.1 Suavização Exponencial

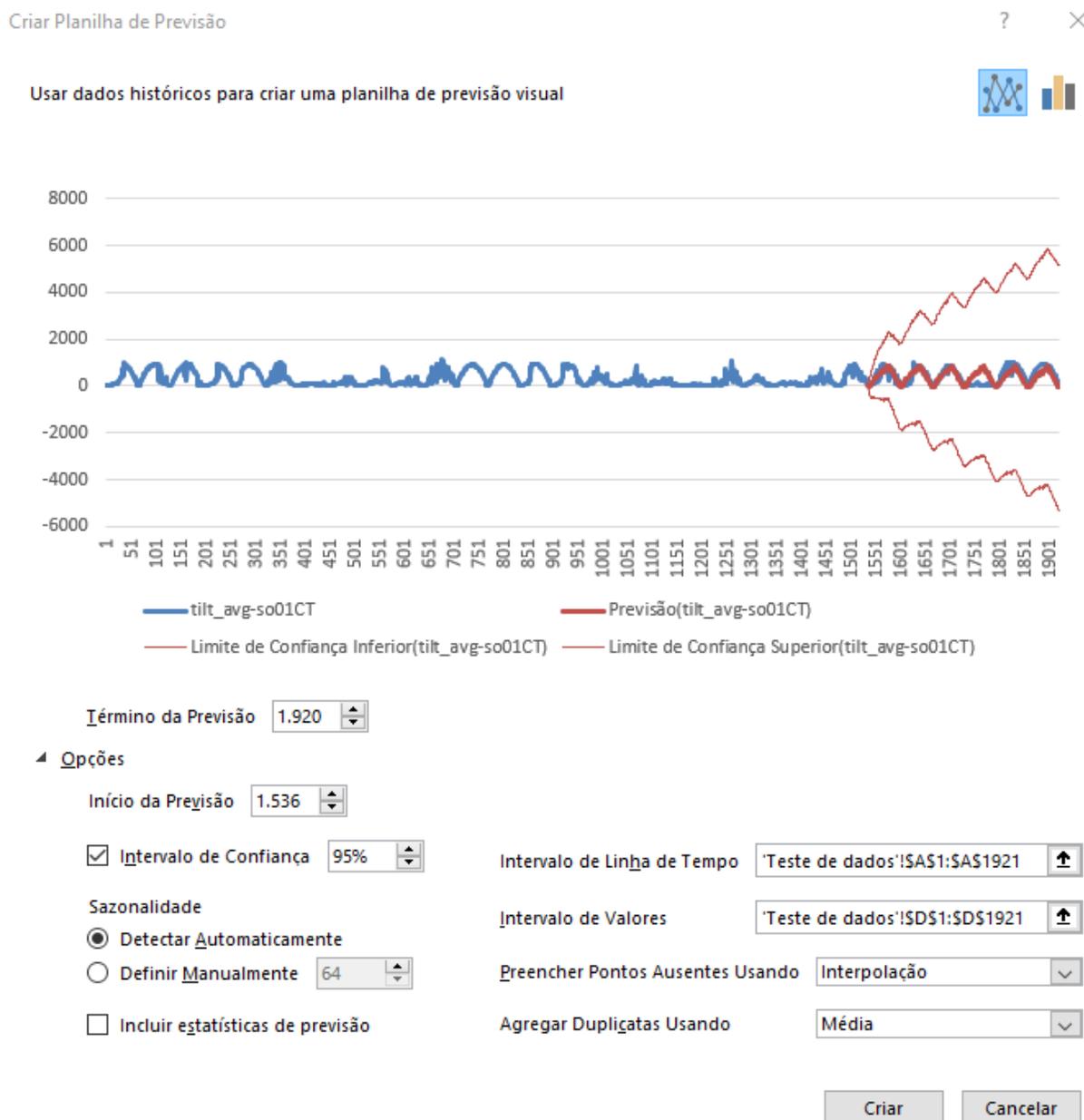
A suavização exponencial é utilizada em sistemas de previsão onde não há uma sazonalidade clara e definida. As previsões que a utilizam possuem médias com peso de observações passadas, com os pesos gradativamente diminuindo conforme as observações envelhecem. Portanto, é possível concluir que quanto mais nova a observação, maior o peso da mesma no algoritmo (HYNDMAN, 2018).

A equação de suavização exponencial tem a forma de:

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots, \quad \text{Equação 10: Suavização exponencial.}$$

Onde  $\alpha$ , situado entre 0 e 1, é o parâmetro de suavização. A previsão *one-step-ahead* (um passo à frente) para o tempo  $T+1$  é a média com peso das observações das séries  $y_1, \dots, y_t$ . A taxa pela qual os pesos reduzem é controlada pelo parâmetro  $\alpha$ . (HYNDMAN, 2018)

Com os dados separados, ao selecionar a ferramenta, a interface apresentada é a seguinte:



**Figura 12: Excel - Planilha de previsão.**

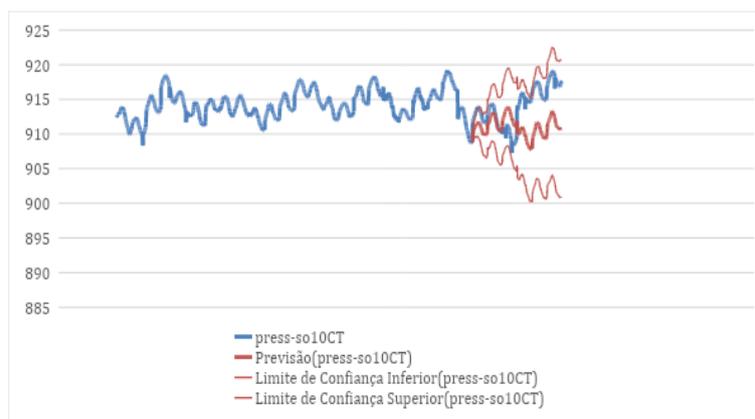
Através da interface da ferramenta de planilha de previsão do Microsoft Excel, ilustrada acima, é possível controlar a previsão em termos de sazonalidade, preenchimento de dados ausentes ou agregação de duplicatas.

Considerando o uso de 80% dos dados para previsão, e subsequente comparação, dos 20% restantes, as 6 variáveis consideradas tiveram suas planilhas de previsão criadas. Existe a possibilidade de definir a sazonalidade manualmente, que foi o primeiro método utilizado, onde o gráfico que melhor se alinhasse com os

dados reais dos 20%, seria a sazonalidade definida para aquela variável em específico. Entretanto, este método apresentou resultados arbitrários e com possível *overfitting* em relação ao sistema em questão.

Mantendo-se com a detecção automática de sazonalidade e utilizando 80% dos dados para prever os próximos 20%, os gráficos obtidos foram:

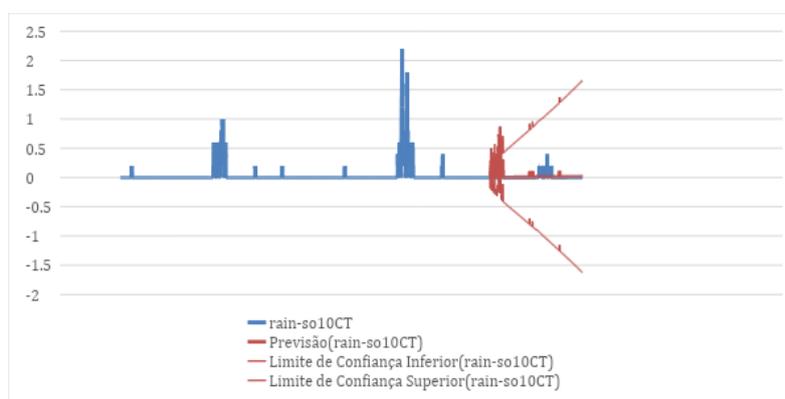
- **press-so10CT:**



**Figura 13: Gráfico Planilha de Previsão - press-so10CT.**

A figura acima (Figura 13) representa o resultado da previsão da entrada de pressão atmosférica, onde a linha azul corresponde aos dados medidos, a linha vermelha espessa corresponde à previsão e as linhas vermelhas finas correspondem ao intervalo de confiança de 95%. Para os dados das primeiras medições, que se dão no início da linha vermelha espessa, é possível observar uma boa correspondência.

- **rain-so10CT:**

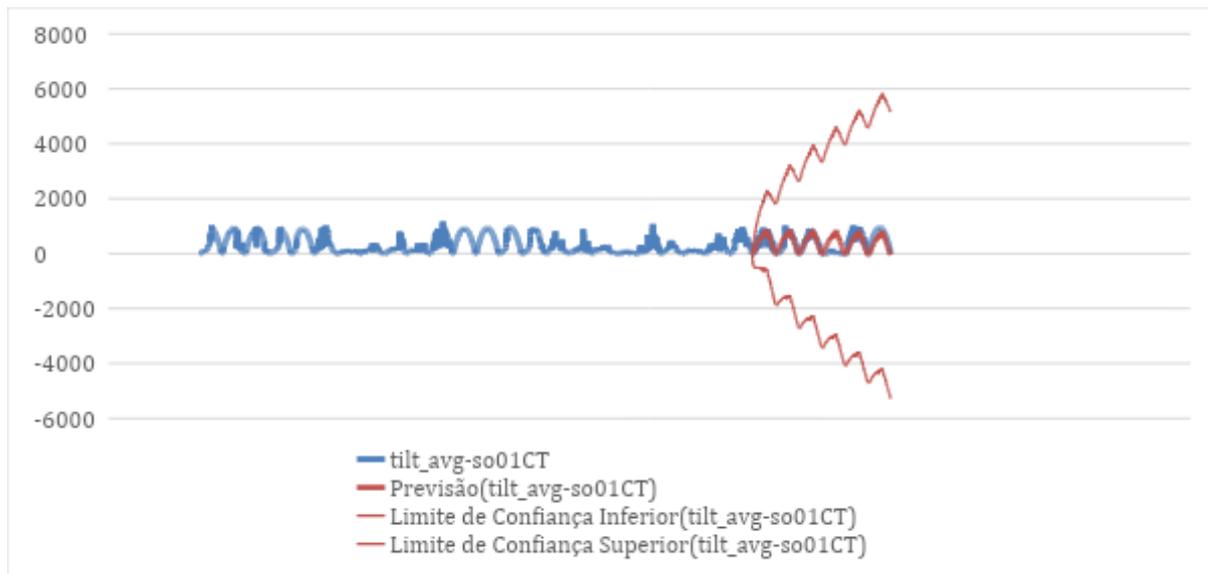


**Figura 14: Gráfico Planilha de Previsão - rain-so10CT.**

A figura acima (Figura 14) representa o resultado da previsão da entrada de precipitação, onde a linha azul corresponde aos dados medidos, a linha vermelha espessa corresponde à previsão e as linhas vermelhas finas correspondem ao

intervalo de confiança de 95%. Para esta entrada foi observado o grande risco de prever precipitação quando não houver, pois causa grande impacto no resultado final.

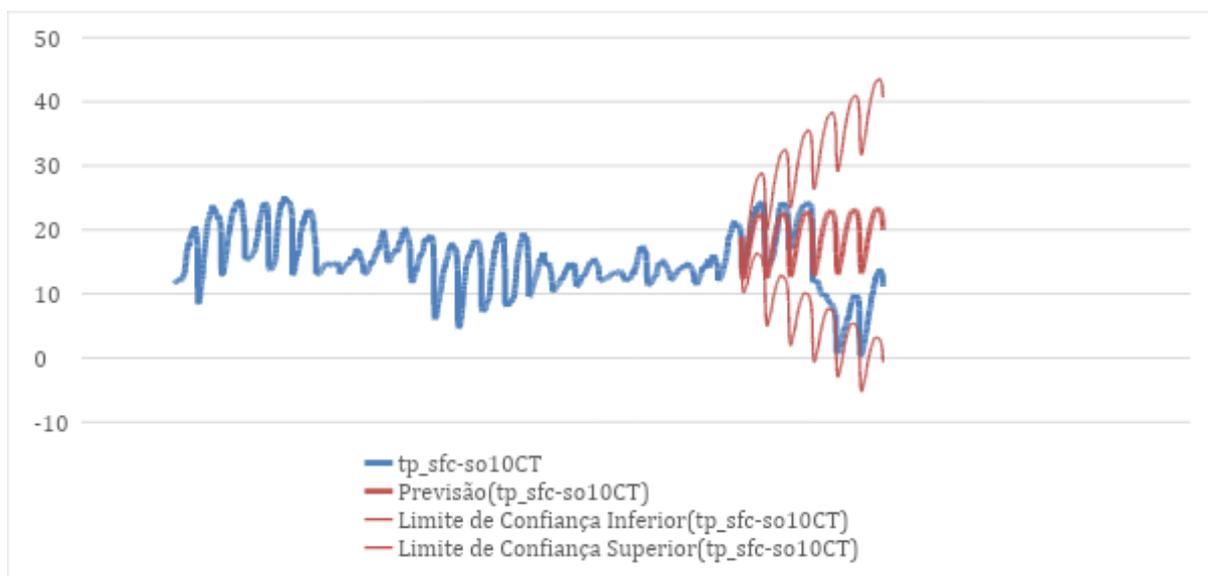
- tilt\_avg-so01CT:



**Figura 15: Gráfico Planilha de Previsão - tilt\_avg-so01CT.**

A figura acima (Figura 15) representa o resultado da previsão da entrada de média da radiação incidente, onde a linha azul corresponde aos dados medidos, a linha vermelha espessa corresponde à previsão e as linhas vermelhas finas correspondem ao intervalo de confiança de 95%. Para esta entrada sendo a que oferece dados mais lineares, a previsão é a mais satisfatória.

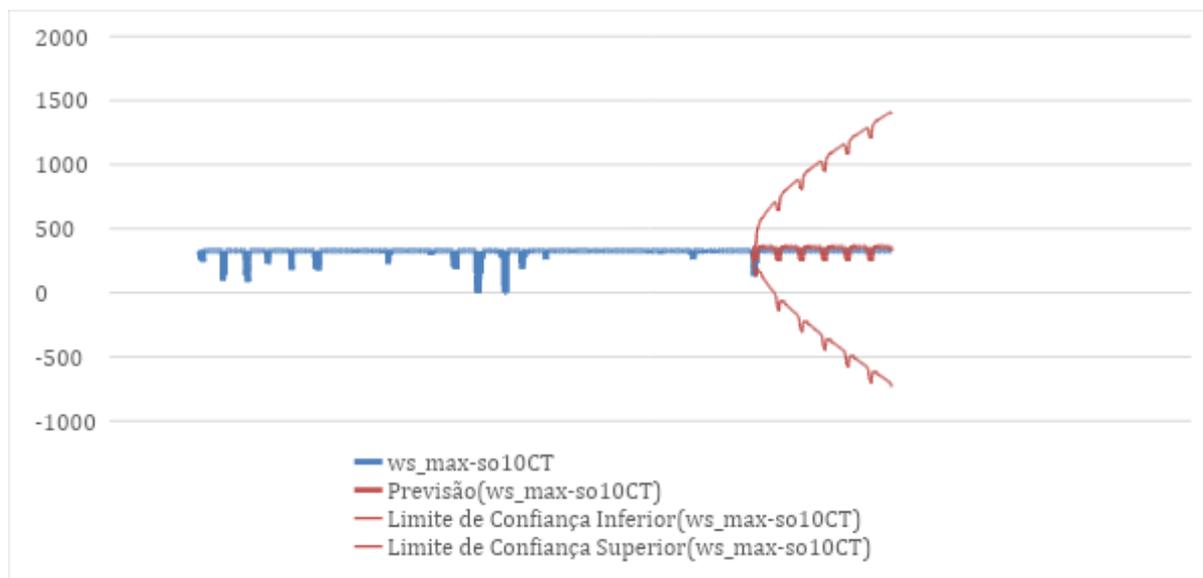
- tp\_sfc-so10CT:



**Figura 16: Gráfico Planilha de Previsão - tp\_sfc-so10CT.**

A figura exposta anteriormente (Figura 16) representa o resultado da previsão da entrada de temperatura do ar superficial, onde a linha azul corresponde aos dados medidos, a linha vermelha espessa corresponde à previsão e as linhas vermelhas finas correspondem ao intervalo de confiança de 95%. Para os dados das primeiras medições, que se dão no início da linha vermelha espessa, é possível observar uma boa correspondência.

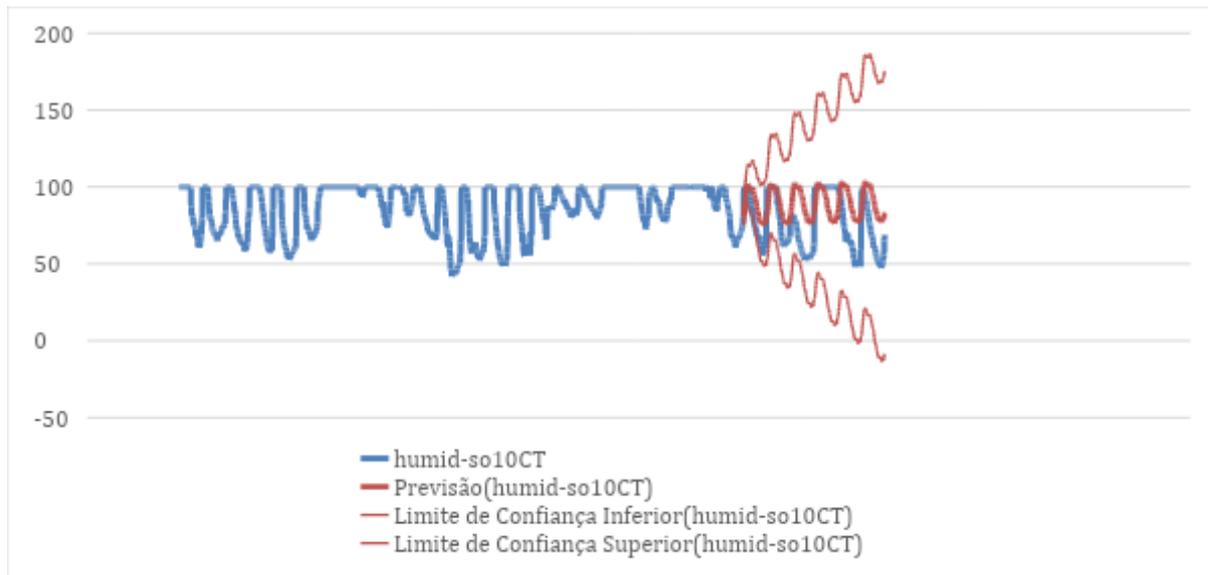
- `ws_max-so10CT`:



**Figura 17: Gráfico Planilha de Previsão - `ws_max-so10CT`.**

A figura exposta anteriormente (Figura 17) representa o resultado da previsão da entrada de velocidade máxima do vento, onde a linha azul corresponde aos dados medidos, a linha vermelha espessa corresponde à previsão e as linhas vermelhas finas correspondem ao intervalo de confiança de 95%. Apesar do menor impacto nos resultados finais, a previsão tende a errar a incidência de ventos.

- humid-so10CT:

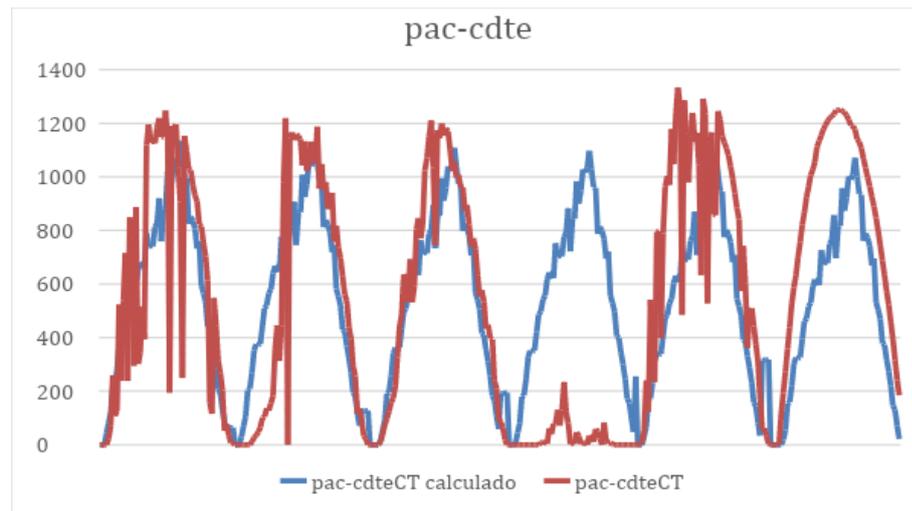


**Figura 18: Gráfico Planilha de Previsão - humid-so10CT.**

A figura exposta anteriormente (Figura 18) representa o resultado da previsão da entrada de umidade, onde a linha azul corresponde aos dados medidos, a linha vermelha espessa corresponde à previsão e as linhas vermelhas finas correspondem ao intervalo de confiança de 95%. Para os dados das primeiras medições, que se dão no início da linha vermelha espessa, é possível observar uma boa correspondência.

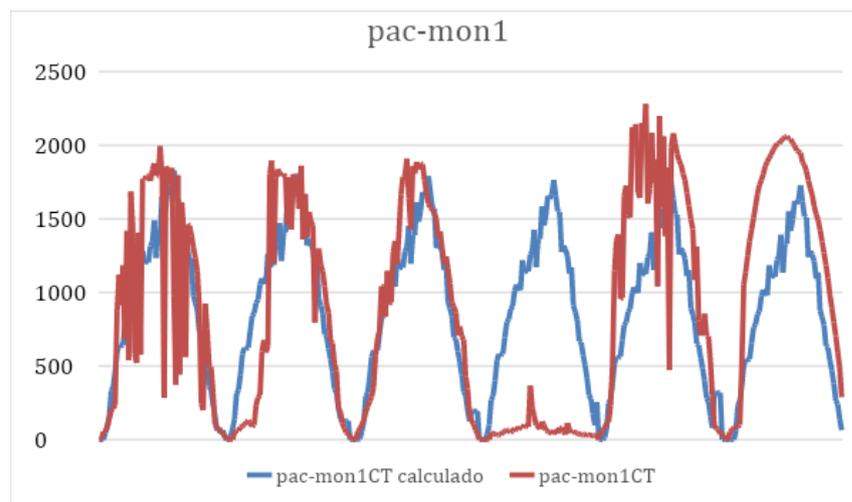
Fica evidente que o gráfico que apresenta melhor precisão é o da variável `tilt_avg-so01CT`, o qual possui o maior índice de linearidade. Os dados gerados são então inseridos na planilha de regressão, com o fim de verificar a precisão das previsões individuais, agora combinadas onde geram como saída a potência das tecnologias da usina.

As comparações estão presentes nas imagens a seguir:



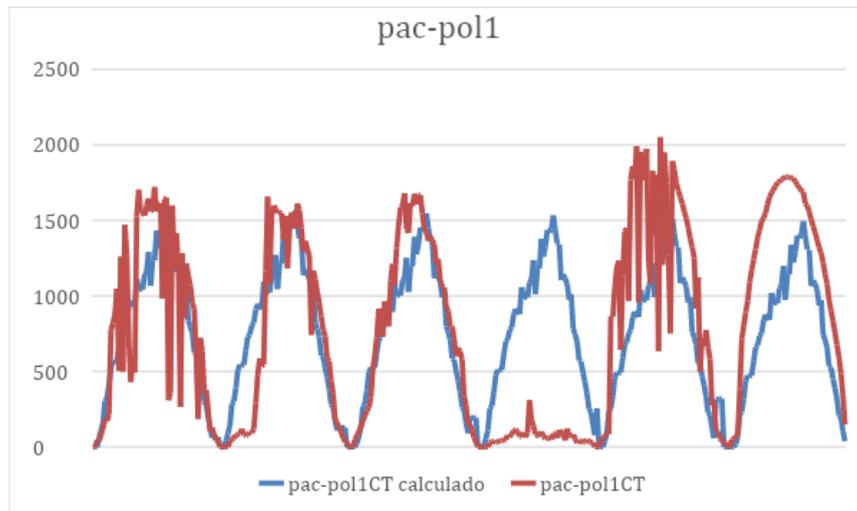
**Figura 19: Segundo gráfico comparativo - Regressão Linear - pac-cdte.**

A figura acima (Figura 19) diz respeito ao resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. Excluindo o antepenúltimo dia das medições, os resultados das previsões se mostram relativamente satisfatórios.



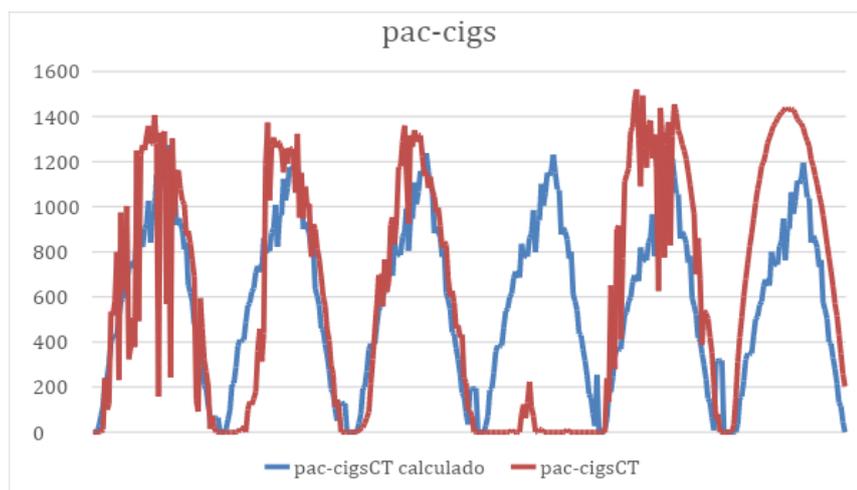
**Figura 20: Segundo gráfico comparativo - Regressão Linear - pac-mon1.**

Ja a figura acima (Figura 20) diz respeito ao resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. Excluindo o antepenúltimo dia das medições, os resultados das previsões se mostram relativamente satisfatórios.



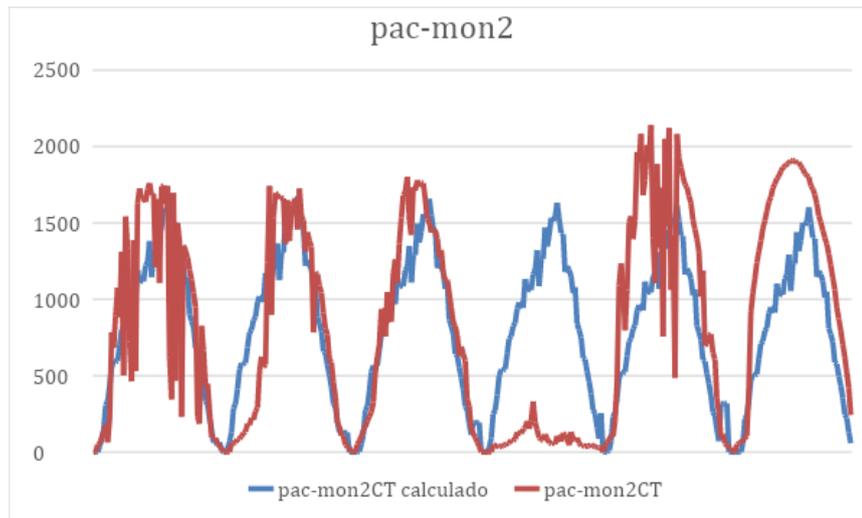
**Figura 21: Segundo gráfico comparativo - Regressão Linear - pac-pol1.**

A figura acima (Figura 21) representa o resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. Excluindo o antepenúltimo dia das medições, os resultados das previsões se mostram relativamente satisfatórios.



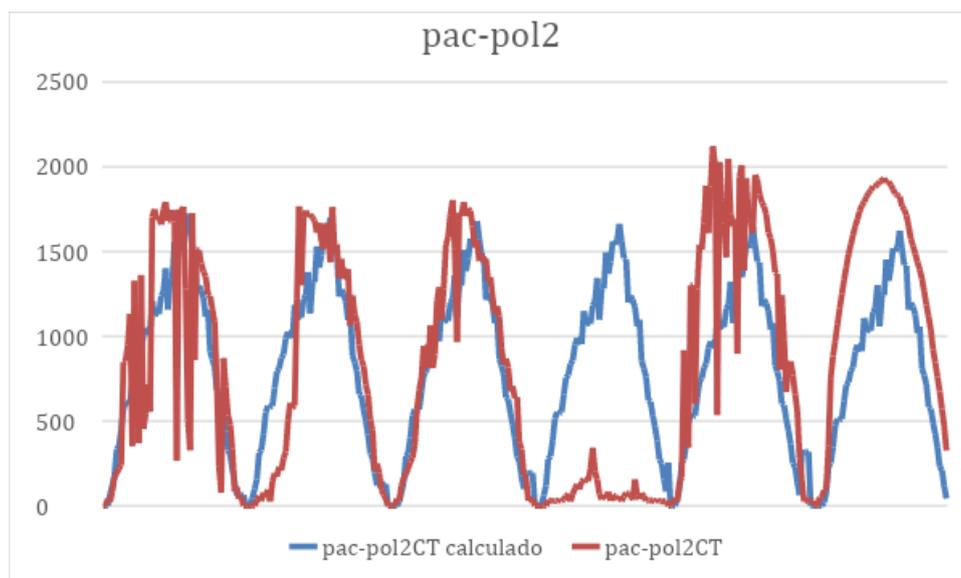
**Figura 22: Segundo gráfico comparativo - Regressão Linear - pac-cigs.**

A figura acima (Figura 22) diz respeito ao resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. Excluindo o antepenúltimo dia das medições, os resultados das previsões se mostram relativamente satisfatórios.



**Figura 23: Segundo gráfico comparativo - Regressão Linear - pac-mon2.**

A figura acima (Figura 23) representa o resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. Excluindo o antepenúltimo dia das medições, os resultados das previsões se mostram relativamente satisfatórios.



**Figura 24: Segundo gráfico comparativo - Regressão Linear - pac-pol2.**

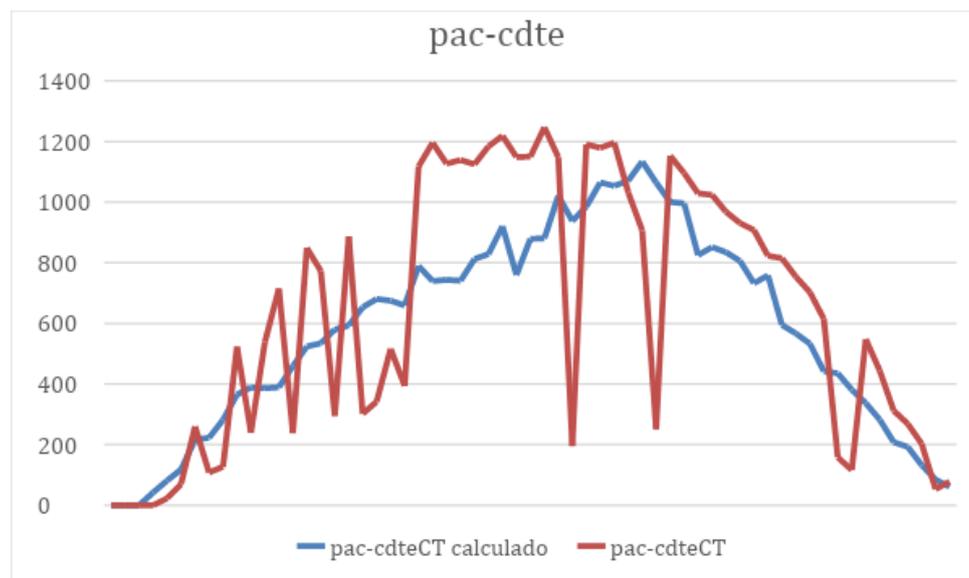
Enquanto a figura acima (Figura 24) representa o resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. Excluindo o antepenúltimo dia das medições, os resultados das previsões se mostram relativamente satisfatórios.

Com isso, percebe-se imediatamente duas coisas:

- Um pico de dados que difere demasiadamente do restante do conjunto de dados;
- Como cada segmento do gráfico corresponde a um dia, dados dos primeiros dias aparentam ter mais intersecções com dados reais.

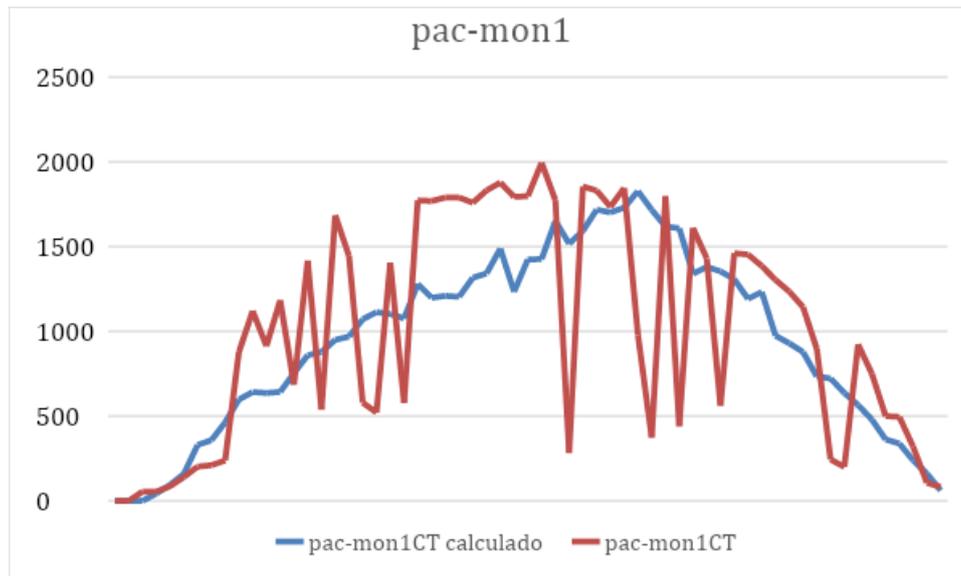
O segmento com maior divergência corresponde a um dia onde todas as medições foram praticamente nulas, apontando um possível mal funcionamento da usina, ou um dia específico onde praticamente não houve geração de energia.

Considerando apenas os dados do primeiro dia de previsão, os dados são:



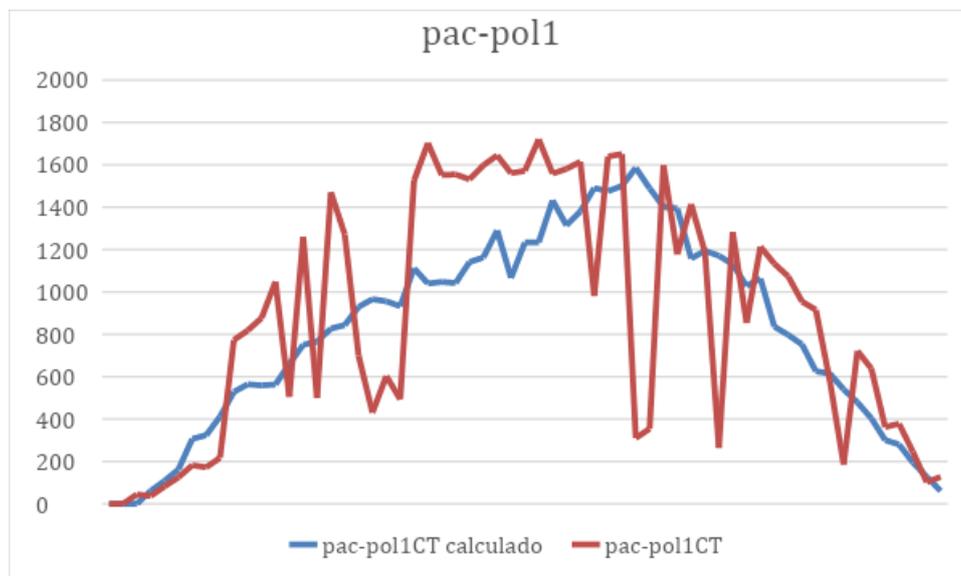
**Figura 25: Gráfico comparativo - Primeiro Dia - pac-cdte.**

A figura acima (Figura 25) ilustra a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. reflete um resultado relativamente satisfatório.



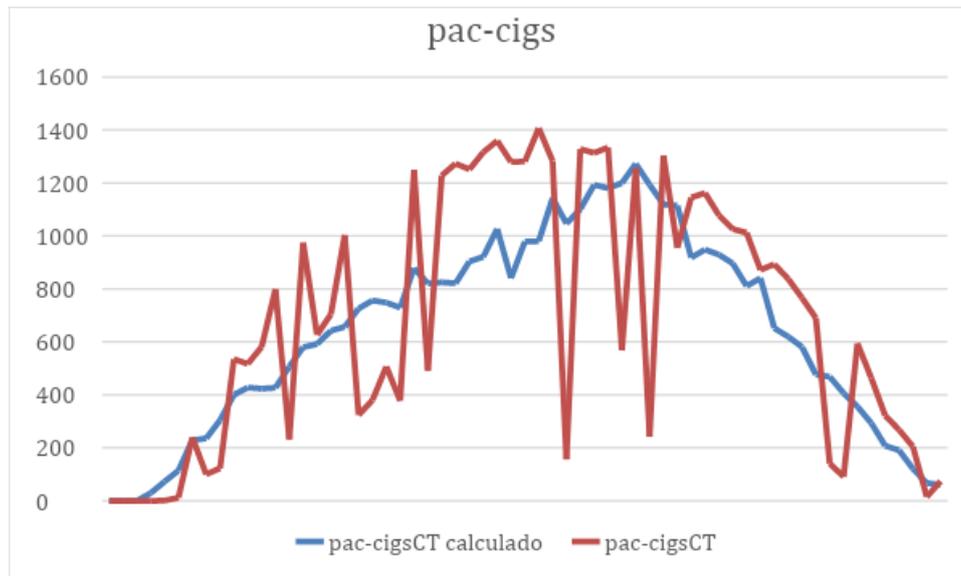
**Figura 26: Gráfico comparativo - Primeiro Dia - pac-mon1.**

Já a figura acima (Figura 26) ilustra a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. reflete um resultado relativamente satisfatório.



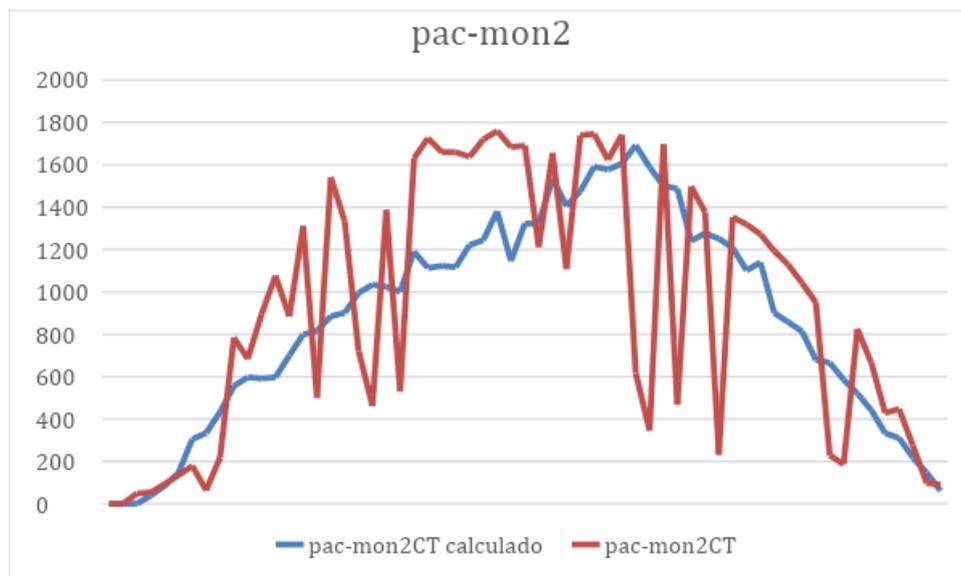
**Figura 27: Gráfico comparativo - Primeiro Dia – pac-pol1.**

A figura acima (Figura 27) representa a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. reflete um resultado relativamente satisfatório.



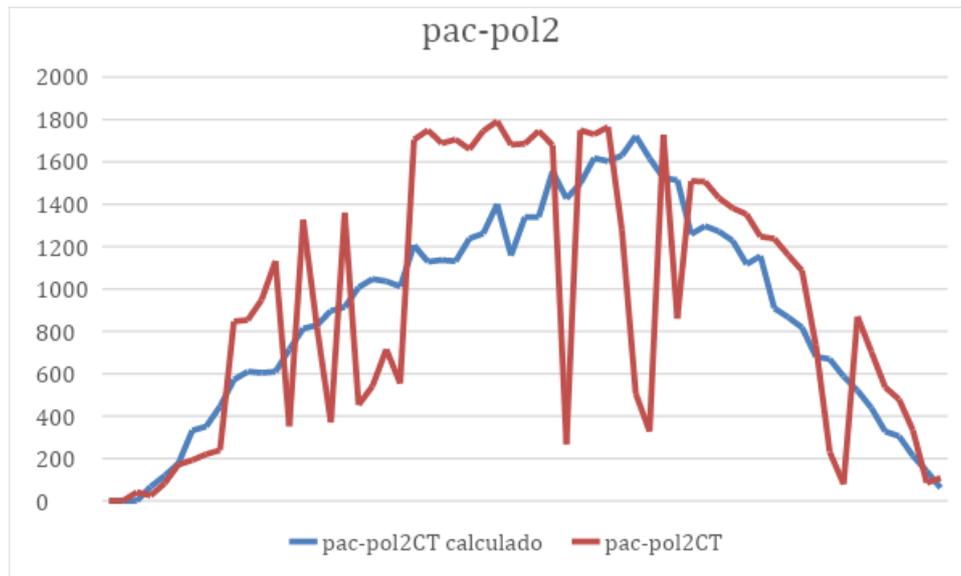
**Figura 28: Gráfico comparativo - Primeiro Dia – pac-cigs.**

Enquanto a figura acima (Figura 28) ilustra a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. reflete um resultado relativamente satisfatório.



**Figura 29: Gráfico comparativo - Primeiro Dia – pac-mon2.**

Já a figura acima (Figura 29) representa a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. reflete um resultado relativamente satisfatório.



**Figura 30: Gráfico comparativo - Primeiro Dia – pac-pol2.**

Já a figura acima (Figura 30) ilustra a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha vermelha corresponde aos dados medidos e a linha azul corresponde ao resultado calculado da previsão com a equação da regressão linear. Reflete um resultado relativamente satisfatório.

Como a linha azul corresponde aos valores previstos, nota-se que o comportamento dos dois gráficos é diferente, onde nos valores reais há variação durante todo o dia, e nos valores calculados, o que prevalece são valores com comportamento menos variante, dada a natureza linear do algoritmo de previsão e a não linearidade da maioria das variáveis utilizadas.

### 3.4.1.5 Primeira Ampliação do Conjunto de Dados

Após a realização dos diversos testes e análises, que acabaram por culminar na abordagem descrita acima (regressão linear através do microsoft excel, método dos mínimos quadrados), seguida da apreciação e conferência dos resultados, estes tidos como válidos e consideravelmente satisfatórios, decidiu-se assim definir esta abordagem como válida.

Desta forma, dando seguimento ao estudo proposto por este trabalho, o conjunto de dados utilizados até o momento foi substituído por um conjunto maior (mantendo as variáveis previamente escolhidas), desta vez do período de novembro do ano de 2021 até outubro do ano de 2022, utilizando assim aproximadamente 12 vezes mais medições para o método de regressão linear.

#### 3.4.1.5.1 Resultados da Regressão Linear após Primeira Ampliação

Realizando agora as regressões lineares com os dados do período citado acima, obteve-se os seguintes resultados:

- **pac-cdteCT**

|                   | <i>Coefficientes</i> | <i>Erro padrão</i> |
|-------------------|----------------------|--------------------|
| <i>Interseção</i> | -1003,708428         | 315,6614696        |
| <i>tp_sfc</i>     | 11,28887237          | 0,595538591        |
| <i>humid</i>      | -6,580876968         | 0,175352068        |
| <i>press</i>      | 1,679955125          | 0,342982196        |
| <i>rain</i>       | 2,299970164          | 12,6285932         |
| <i>ws_max</i>     | -0,047206306         | 0,017316932        |
| <i>tilt_avg</i>   | 0,095608486          | 0,008795006        |

*Tabela 7: Primeira Ampliação - Regressão da variável pac-cdteCT.*

Os coeficientes citados através da tabela acima (Tabela 7) são os pesos de cada variável dentro do cálculo da potência **pac-cdteCT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-cigsCT**

|                   | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-------------------|---------------------|--------------------|
| <i>Interseção</i> | -1039,585767        | 351,5544008        |
| <i>tp_sfc</i>     | 14,06154336         | 0,663255521        |
| <i>humid</i>      | -7,723755445        | 0,195290833        |
| <i>press</i>      | 1,76639679          | 0,381981686        |
| <i>rain</i>       | 13,50674127         | 14,06455315        |
| <i>ws_max</i>     | -0,036016322        | 0,019285989        |
| <i>tilt_avg</i>   | 0,154622413         | 0,00979506         |

**Tabela 8: Primeira Ampliação - Regressão da variável pac-cigsCT.**

Os coeficientes citados através da tabela acima (Tabela 8) são os pesos de cada variável dentro do cálculo da potência **pac-cigsCT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-mon1CT**

|                   | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-------------------|---------------------|--------------------|
| <i>Interseção</i> | -1292,931081        | 506,2565297        |
| <i>tp_sfc</i>     | 21,67867757         | 0,955122274        |
| <i>humid</i>      | -10,465845          | 0,281228906        |
| <i>press</i>      | 2,24598037          | 0,550073395        |
| <i>rain</i>       | 26,38966655         | 20,25368436        |
| <i>ws_max</i>     | 0,031134265         | 0,027772823        |
| <i>tilt_avg</i>   | 0,247341378         | 0,014105393        |

**Tabela 9: Primeira Ampliação - Regressão da variável pac-mon1CT.**

Os coeficientes citados através da tabela acima (Tabela 9) são os pesos de cada variável dentro do cálculo da potência **pac-mon1CT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-mon2CT**

|                   | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-------------------|---------------------|--------------------|
| <i>Interseção</i> | -1191,180576        | 478,2870579        |
| <i>tp_sfc</i>     | 20,13339327         | 0,902354035        |
| <i>humid</i>      | -9,886557072        | 0,265691676        |
| <i>press</i>      | 2,096169115         | 0,519683145        |
| <i>rain</i>       | 27,73243714         | 19,13471637        |
| <i>ws_max</i>     | 0,001929176         | 0,02623844         |
| <i>tilt_avg</i>   | 0,235203949         | 0,013326103        |

*Tabela 10: Primeira Ampliação - Regressão da variável pac-mon2CT.*

Os coeficientes citados através da tabela acima (Tabela 10) são os pesos de cada variável dentro do cálculo da potência **pac-mon2CT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-pol1CT**

|                   | <i>Coeficientes</i> | <i>Erro padrão</i> |
|-------------------|---------------------|--------------------|
| <i>Interseção</i> | -1188,354968        | 445,5029329        |
| <i>tp_sfc</i>     | 18,25687095         | 0,840502293        |
| <i>humid</i>      | -9,362387205        | 0,247479874        |
| <i>press</i>      | 2,061194111         | 0,484061531        |
| <i>rain</i>       | 22,63968713         | 17,82312969        |
| <i>ws_max</i>     | -0,012455939        | 0,02443993         |
| <i>tilt_avg</i>   | 0,225329772         | 0,012412667        |

*Tabela 11: Primeira Ampliação - Regressão da variável pac-pol1CT.*

Os coeficientes citados através da tabela acima (Tabela 11) são os pesos de cada variável dentro do cálculo da potência **pac-pol1CT**, já o erro padrão significa o quanto este peso é incerto.

- **pac-pol2CT**

|                   | <i>Coefficientes</i> | <i>Erro padrão</i> |
|-------------------|----------------------|--------------------|
| <i>Interseção</i> | -1149,213038         | 468,408593         |
| <i>tp_sfc</i>     | 19,65938767          | 0,883716958        |
| <i>humid</i>      | -9,677625699         | 0,260204122        |
| <i>press</i>      | 2,033561359          | 0,508949692        |
| <i>rain</i>       | 20,31567337          | 18,73951098        |
| <i>ws_max</i>     | 0,009208233          | 0,025696516        |
| <i>tilt_avg</i>   | 0,233870094          | 0,013050868        |

*Tabela 12: Primeira Ampliação - Regressão da variável pac-pol2CT.*

Os coeficientes citados através da tabela acima (Tabela 12) são os pesos de cada variável dentro do cálculo da potência **pac-pol2CT**, já o erro padrão significa o quanto este peso é incerto.

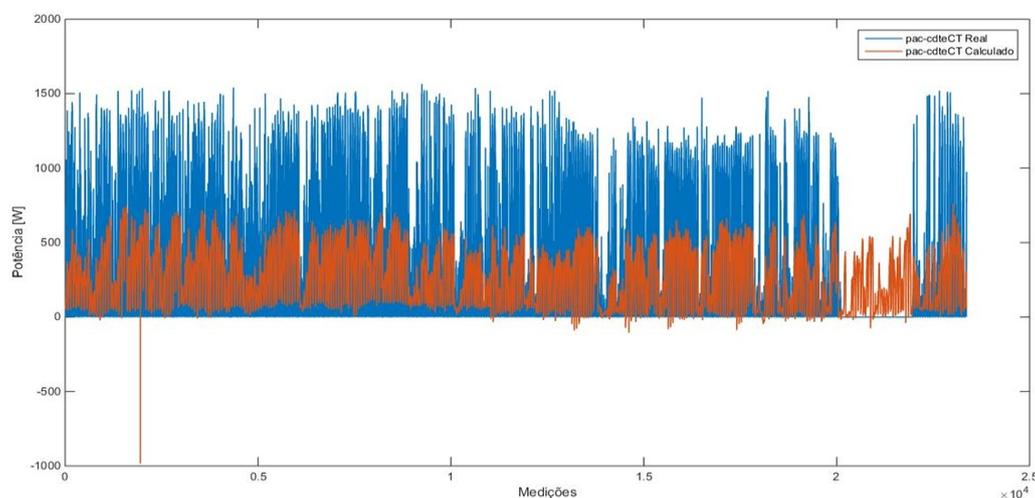
Observando desta vez os valores dos coeficientes e erro padrão obtidos, verificou-se algumas diferenças comparando com a primeira análise realizada (com menor quantidade de dados). Assim como observado anteriormente, valores positivos nos coeficientes indicam uma influência “positiva” na potência gerada, enquanto que valores negativos indicam uma influência “negativa” na potência gerada:

- Nota-se novamente, como já esperado, que em todos os casos o coeficiente relacionado à média da radiação (*tilt\_avg-so01CT*) apresenta valor positivo, pois quanto maior a radiação, maior a potência gerada;
- o coeficiente relacionado à umidade (*humid-so10CT*) novamente, como na análise anterior, apresenta valores negativos para todos os casos. Desta vez, entretanto, o módulo dos coeficientes aumentaram significativamente, evidenciando então que quando considerando um maior número de dados (analisando um período mais longo), esta variável torna-se mais relevante na potência gerada;
- o coeficiente relacionado à temperatura (*tp\_sfc-so10CT*), assim como o coeficiente da umidade, apresenta valores significativamente maiores para todos os casos nesta análise com maior quantidade de dados, se mostrando que a temperatura se torna mais influente quando analisada por períodos mais longos.

Entretanto, nota-se que agora apresenta valores positivos. Apesar de sabermos que a temperatura elevada reduz o desempenho do sistema fotovoltaico, altas temperaturas quase sempre significam altos níveis de radiação solar. Assim sendo, nota-se que as desvantagens de altas temperaturas (desempenho, perdas, etc.) para sistemas fotovoltaicos evidenciam-se a curto prazo, enquanto que suas vantagens (maiores níveis de radiação solar atingindo os módulos) evidenciam-se a longo prazo;

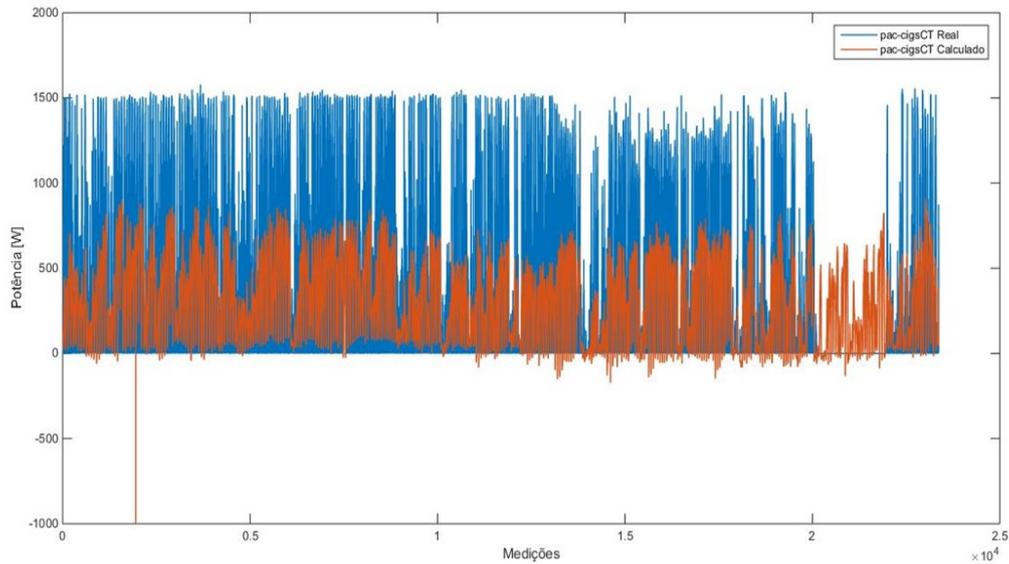
- o coeficiente relacionado à velocidade máxima do vento ( $ws\_max-so10CT$ ) que anteriormente apresentou valores positivos para todos os casos, agora apresenta valores positivos em alguns e negativos em outros, entretanto, o principal a destacar seria o módulo (que diminuiu) e não o sinal: notou-se que ao aplicar o método num conjunto de dados maior, esta variável tornou-se menos relevante na potência gerada;
- o coeficiente relacionado à pressão atmosférica ( $press-so10CT$ ) apresenta desta vez valores positivos, para a todos os casos, ao contrário da primeira análise com menos dados. Quanto ao módulo, este se manteve próximo;
- o coeficiente relacionado à precipitação ( $rain-so10CT$ ) apresenta valores positivos para todos os casos, de maneira totalmente inesperada, conforme observado na análise anterior ( o motivo, conforme já discutido, seria devido a grande maioria dos dados desta variável serem iguais à zero, e ainda as poucas medições positivas apresentam valores múltiplos de 0.2. A questão dos valores de Erro padrão dos coeficientes relacionados à precipitação serem consideravelmente maiores que o Erro padrão referente às outras variáveis também se mantém (consequências da natureza da amostragem desta variável);

Novamente, com posse dos coeficientes (pesos) referentes a cada uma das variáveis, montou-se as equações para determinar os valores das seis potências e calcular os valores teóricos, para comparar com os valores de potência reais (os valores das potências medidos). Abaixo têm-se os resultados apresentados pelos gráficos:



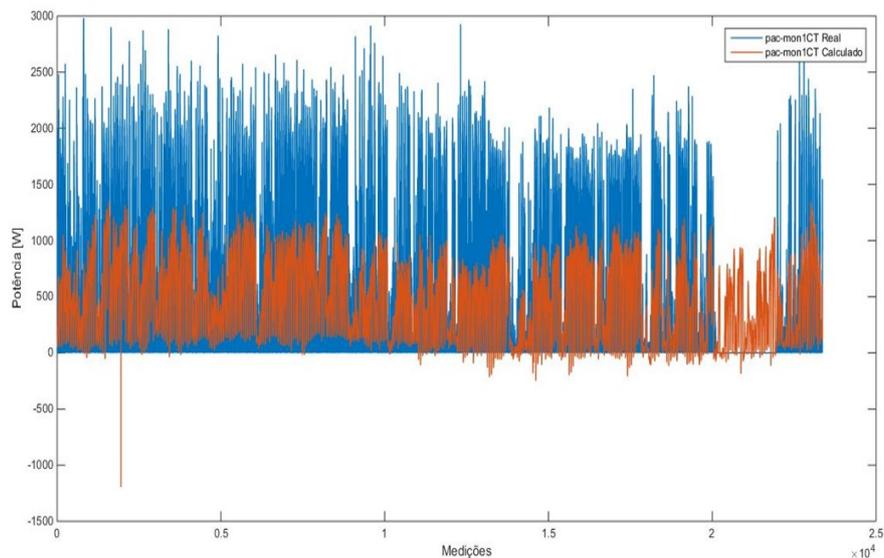
**Figura 31: Primeira Ampliação - Gráfico comparativo da variável *pac-cdteCT*.**

Na figura representada acima (Figura 31), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante discrepância entre seus valores, principalmente em sua amplitude (eixo Y), devido à grande quantidade de medições nulas neste conjunto de dados.



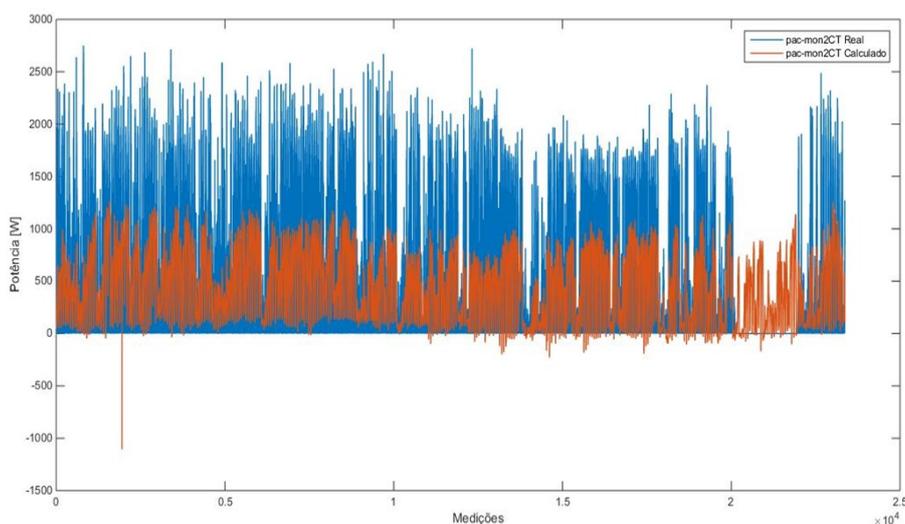
**Figura 32: Primeira Ampliação - Gráfico comparativo da variável pac-cigsCT.**

Já na figura representada acima (Figura 32), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante discrepância entre seus valores, principalmente em sua amplitude (eixo Y), devido à grande quantidade de medições nulas neste conjunto de dados.



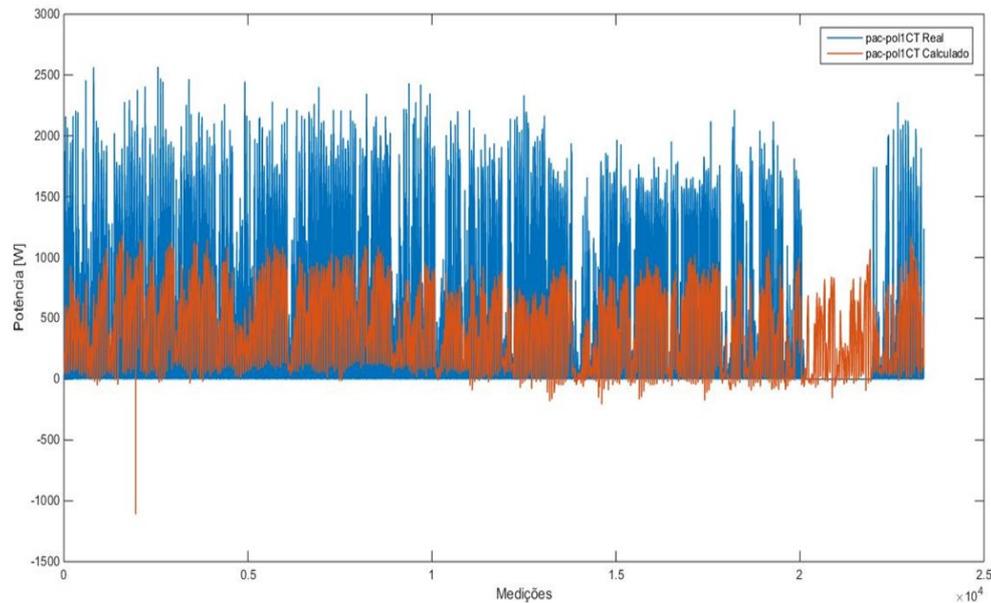
**Figura 33: Primeira Ampliação - Gráfico comparativo da variável pac-mon1CT.**

Já na figura representada anteriormente (Figura 33), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante discrepância entre seus valores, principalmente em sua amplitude (eixo Y), devido à grande quantidade de medições nulas neste conjunto de dados.



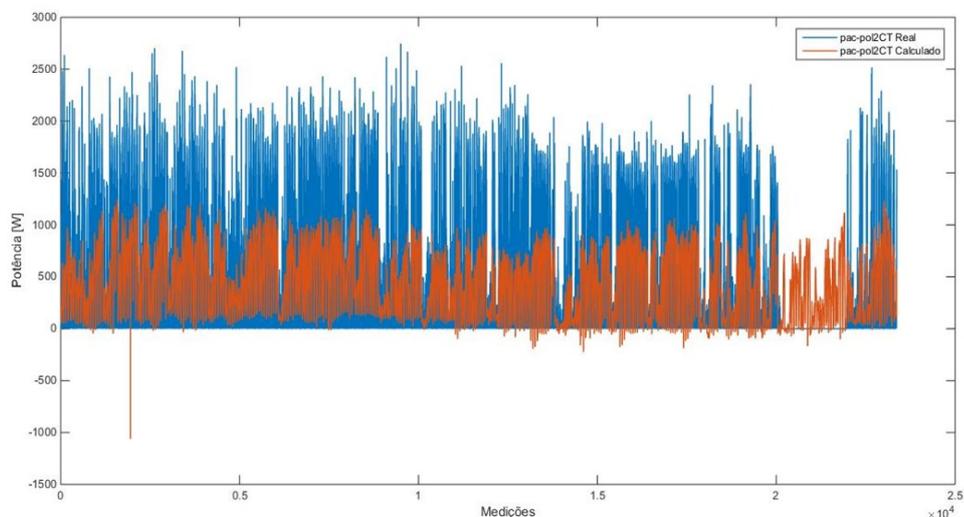
**Figura 34: Primeira Ampliação - Gráfico comparativo da variável pac-mon2CT.**

A figura representada acima (Figura 34), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante discrepância entre seus valores, principalmente em sua amplitude (eixo Y), devido à grande quantidade de medições nulas neste conjunto de dados.



**Figura 35: Primeira Ampliação - Gráfico comparativo da variável *pac-pol1CT*.**

A figura representada acima (Figura 35), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se bastante discrepância entre seus valores, principalmente em sua amplitude (eixo Y), devido à grande quantidade de medições nulas neste conjunto de dados.



**Figura 36: Primeira Ampliação - Gráfico comparativo da variável *pac-po2CT*.**

Enquanto figura representada acima (Figura 36), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de

potência calculado. Nota-se bastante discrepância entre seus valores, principalmente em sua amplitude (eixo Y), devido à grande quantidade de medições nulas neste conjunto de dados.

Nota-se desta vez que a semelhança entre os valores calculados (teóricos) e os valores medidos (reais) para todas as seis potências analisadas, mostram-se consideravelmente menor com relação à análise anterior (com o conjunto de dados do mês de junho de 2021), portanto, neste caso, a precisão do cálculo de regressão linear diminuiu.

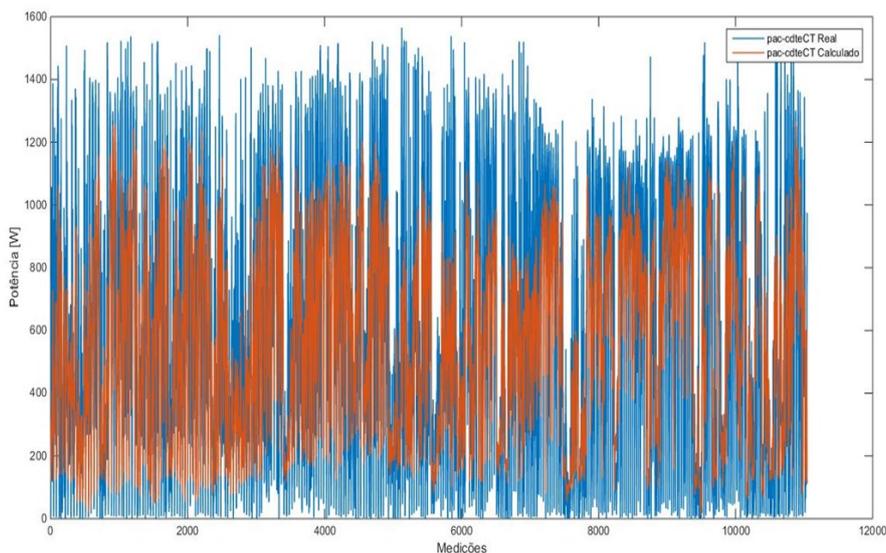
Entretanto, percebeu-se que o principal motivo foi o próprio conjunto de dados analisados: foram encontrados, proporcionalmente, muito mais “medições nulas” dos valores das potências durante o período de novembro de 2021 e outubro de 2022 do que nas medições do mês de junho de 2021 (por exemplo, em determinado dia às 14:10 registrada uma potência igual a 679; após, às 14:20, registrada uma potência igual à 0; às 14:30 e às 14:40, igual à 0 novamente; só então às 14:50, registrada uma potência igual à 942;). Estes valores nulos de potências medidos se devem, provavelmente, à problemas nos sensores e/ou na coleta e armazenamento dos dados.

Porém, se analisados os gráficos, percebe-se que a maior discrepância observada entre os valores medidos e os valores calculados está em seu módulo (intensidade), enquanto que a sincronia entre as variações de potência gerada ao longo do tempo se manteve (com exceção de um período por volta do mês de setembro de 2022 onde não houveram dados das potências geradas, possivelmente por problemas na aquisição dos dados).

Numa tentativa de melhorar os resultados obtidos, realizou-se uma limpeza nos dados deste período considerado (novembro/2021~outubro/2022): foram removidas e desconsideradas as medições nos momentos em que o valor das potências era igual a zero, com o objetivo de solucionar o problema da falta de medições e melhorar a qualidade deste conjunto de dados.

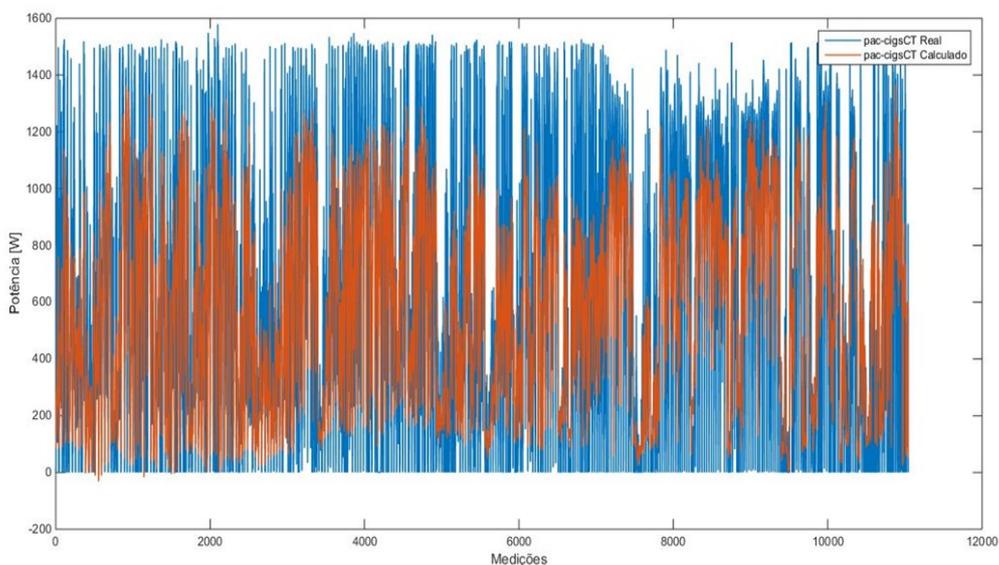
Foram então removidos os dados “nulos” do conjunto e reaplicado o método da regressão linear. Apesar de percebida uma melhora nos valores retornados, os resultados da análise neste conjunto de dados ainda se mostram insatisfatórios.

Abaixo seguem os resultados:



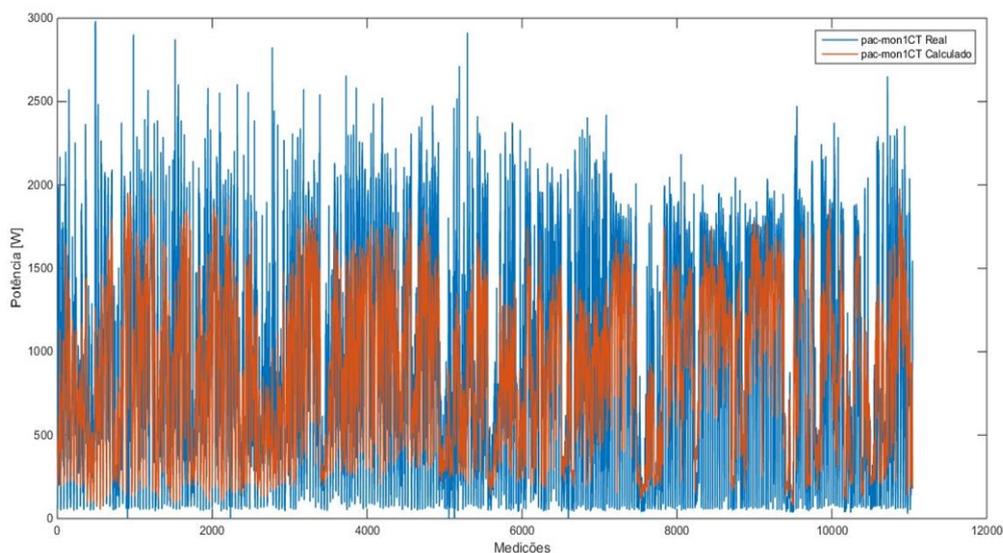
**Figura 37: Gráfico comparativo sem dados nulos - variável *pac-cdteCT*.**

A linha azul do gráfico representado pela figura acima (Figura 37) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se que eliminando os dados nulos do conjunto, a diferença entre seus valores de amplitude (eixo Y) diminuiu. Ainda assim, os resultados deste conjunto continuam insatisfatórios.



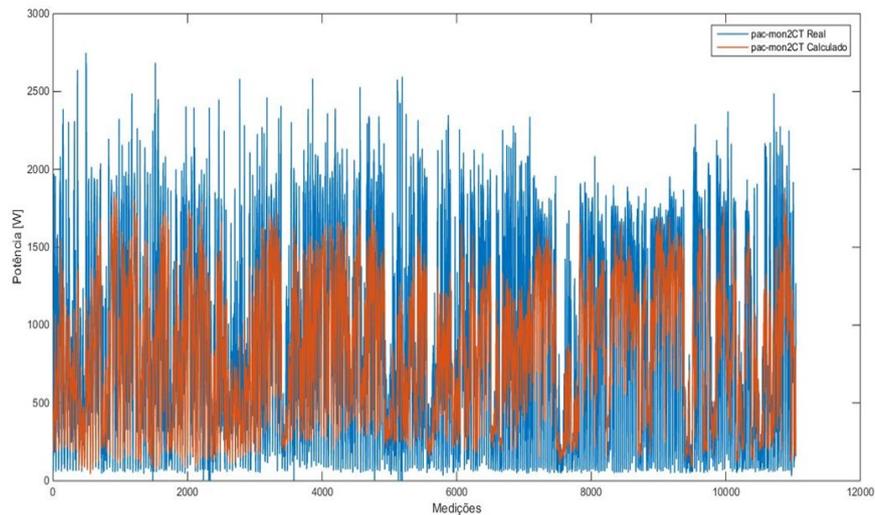
**Figura 38: Gráfico comparativo sem dados nulos - variável *pac-cigsCT*.**

A linha azul do gráfico representado pela figura anteriormente (Figura 38) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se que eliminando os dados nulos do conjunto, a diferença entre seus valores de amplitude (eixo Y) diminuiu. Ainda assim, os resultados deste conjunto continuam insatisfatórios



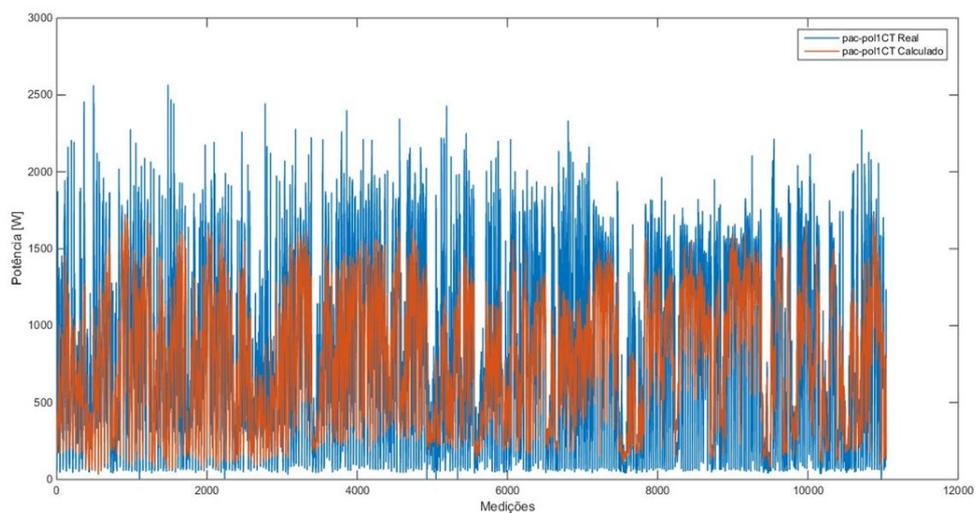
**Figura 39: Gráfico comparativo sem dados nulos - variável pac-mon1CT.**

A linha azul do gráfico representado pela figura acima (Figura 39) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se que eliminando os dados nulos do conjunto, a diferença entre seus valores de amplitude (eixo Y) diminuiu. Ainda assim, os resultados deste conjunto continuam insatisfatório



**Figura 40: Gráfico comparativo sem dados nulos - variável pac-mon2CT.**

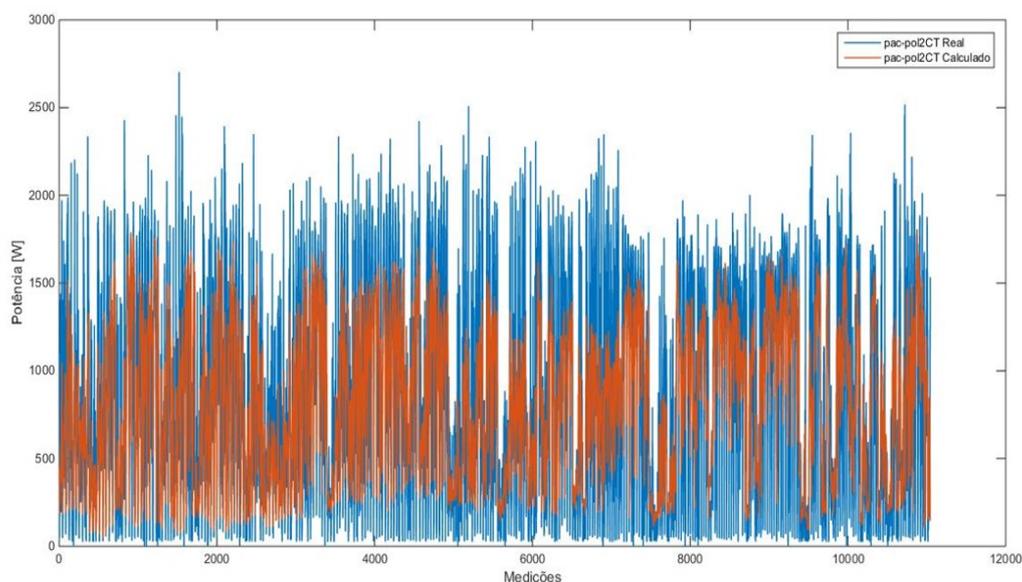
A linha azul do gráfico representado pela figura acima (Figura 40) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se que eliminando os dados nulos do conjunto, a diferença entre seus valores de amplitude (eixo Y) diminuiu. Ainda assim, os resultados deste conjunto continuam insatisfatório



**Figura 41: Gráfico comparativo sem dados nulos - variável pac-pol1CT.**

A linha azul do gráfico representado pela figura acima (Figura 41) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se que eliminando os dados nulos do conjunto, a diferença

entre seus valores de amplitude (eixo Y) diminuiu. Ainda assim, os resultados deste conjunto continuam insatisfatório



**Figura 42: Gráfico comparativo sem dados nulos - variável pac-pol2CT.**

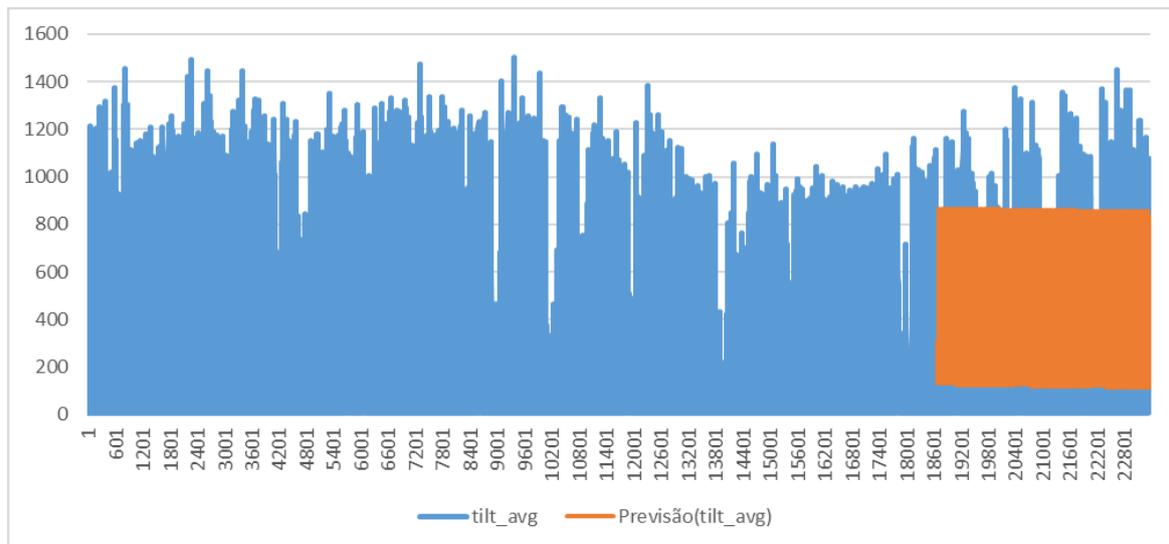
A linha azul do gráfico representado pela figura acima (Figura 42) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se que eliminando os dados nulos do conjunto, a diferença entre seus valores de amplitude (eixo Y) diminuiu. Ainda assim, os resultados deste conjunto continuam insatisfatório.

### **3.4.1.5.2 Resultados de Previsão Após Primeira Ampliação de Dados**

As regressões lineares são efetuadas com as saídas e, com as entradas são feitas as planilhas de previsão do excel. Considerando o mesmo intervalo de dados demonstrado acima, as entradas passaram pelo sistema de previsão do excel como demonstrado anteriormente, seguindo a mesma regra de 80% de análise e 20% de teste.

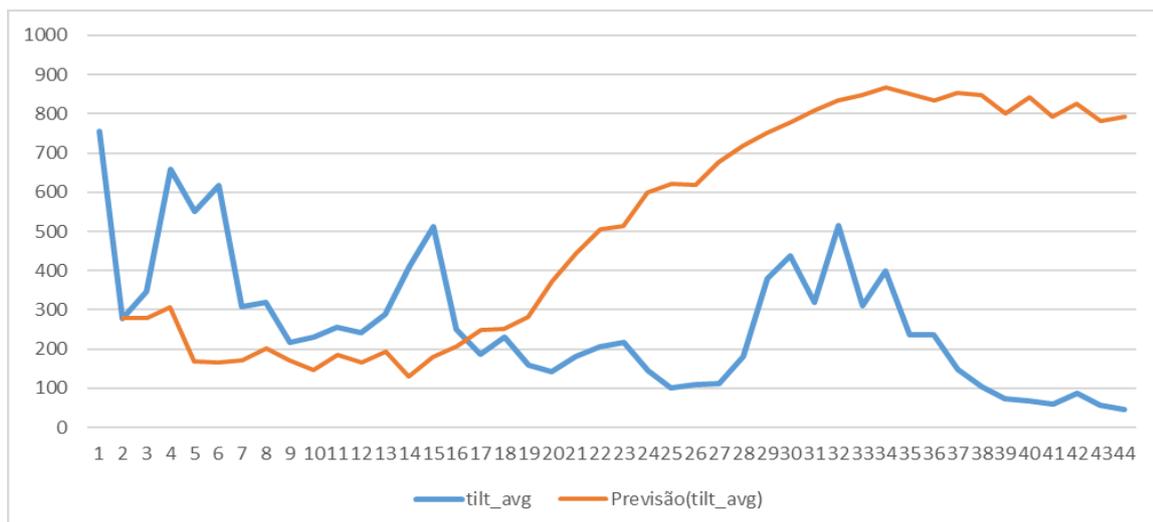
Para cada entrada, foram feitos dois testes: um onde o eixo x é considerado como apenas uma sequência de dados de valor 1 até a linha final, cerca de 23 mil linhas, método utilizado pois, para trabalhar com o sistema data/hora (modo nativo do banco de dados), o espaçamento entre intervalos precisa ser perfeito, e também seguindo o mesmo número de linhas e dados da planilha “limpa” de dados, onde os

valores nulos foram eliminados. O segundo teste é feito com o “ajuste à perfeição” destes intervalos de dados, e considerando 100% dos dados disponíveis, cerca de 53 mil linhas, fornecendo 2 modelos diferentes, com os seguintes gráficos:



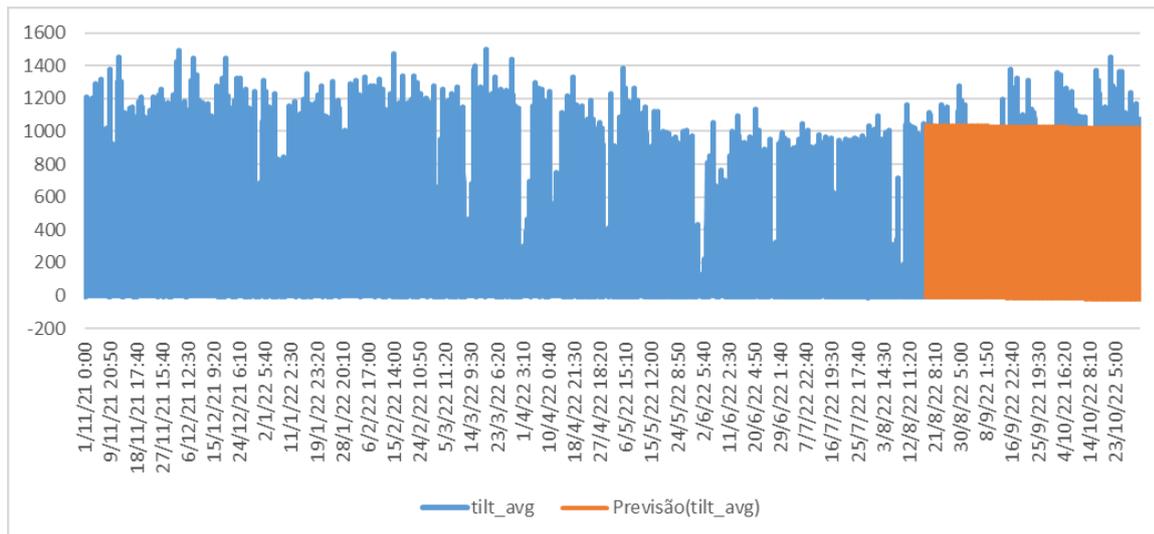
**Figura 43: gráfico previsão tilt\_avg sem dados nulos e eixo x sequencial.**

A figura acima (Figura 43) é resultado da previsão da média da radiação incidente, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado das previsões. Percebe-se um resultado insatisfatório.



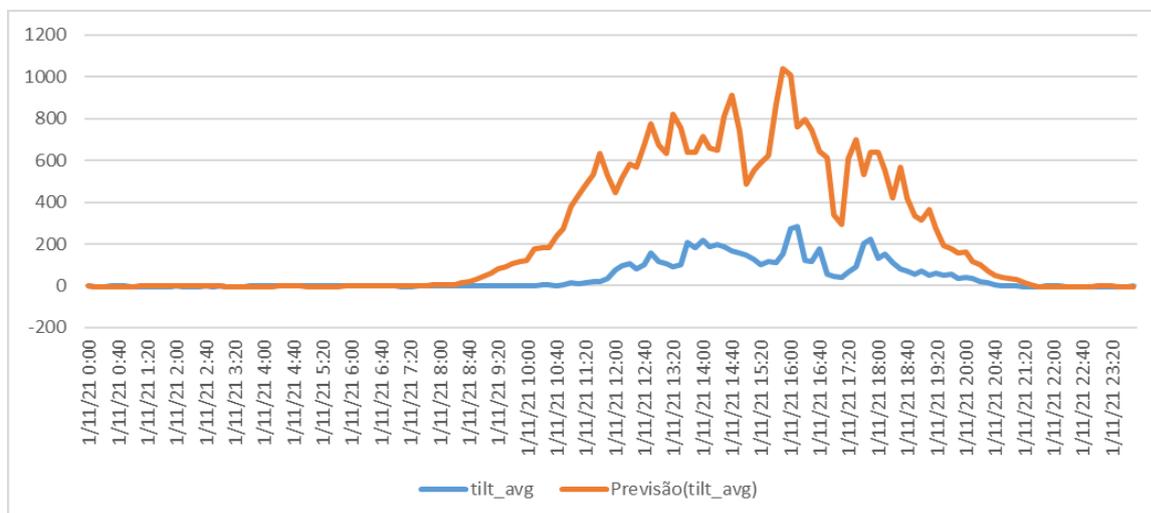
**Figura 44: gráfico previsão tilt\_avg - primeiras 24 horas com eixo x sequencial.**

Já a figura acima (Figura 44) representa a visualização das primeiras 24 horas do resultado da previsão da média da radiação incidente, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado das previsões. Nota-se a diferença entre o padrão das linhas, indicando o resultado insatisfatório mencionado anteriormente.



**Figura 45: gráfico previsão tilt\_avg com todos os dados, no modo de eixo x como data/hora.**

De modo semelhante, a figura acima (Figura 45) ilustra o resultado da previsão da média da radiação incidente, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado das previsões. Percebe-se um resultado insatisfatório.



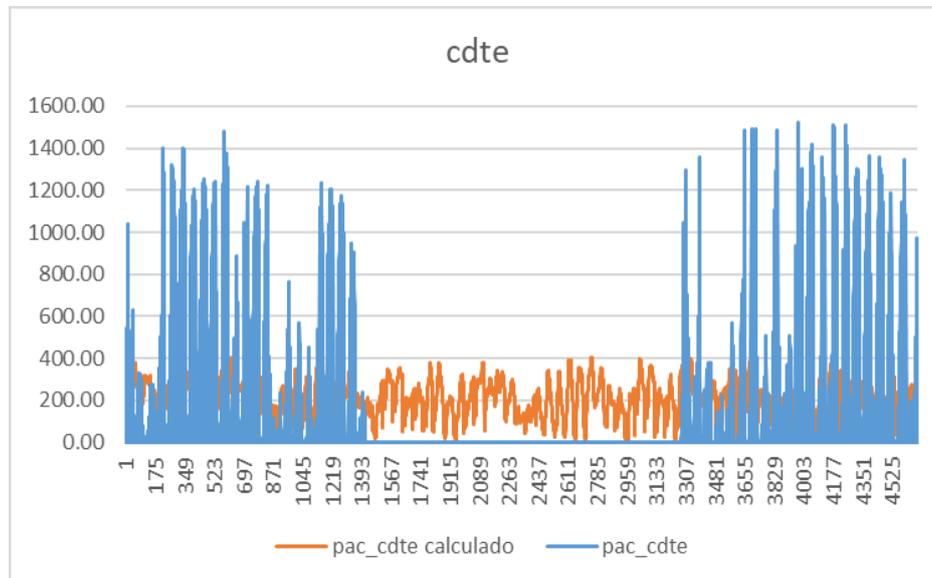
**Figura 46: gráfico previsão tilt\_avg - primeiras 24 horas, com eixo x como data/hora.**

Por fim, a figura acima (Figura 46) ilustra a visualização das primeiras 24 horas do resultado da previsão da média da radiação incidente, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado das previsões. Nota-se uma similaridade no padrão, porém com amplitude diferente.

O modelo ajustado para o intervalo de dados correto mostrou-se mais robusto, capaz de gerar dados comparativos, principalmente para o início das previsões, pois leva em consideração a sazonalidade dia/hora/minuto.

Ao implementar no modelo de menor precisão, que foi a primeira e mais rudimentar tentativa após a aquisição de mais dados, os resultados foram expostos a seguir.

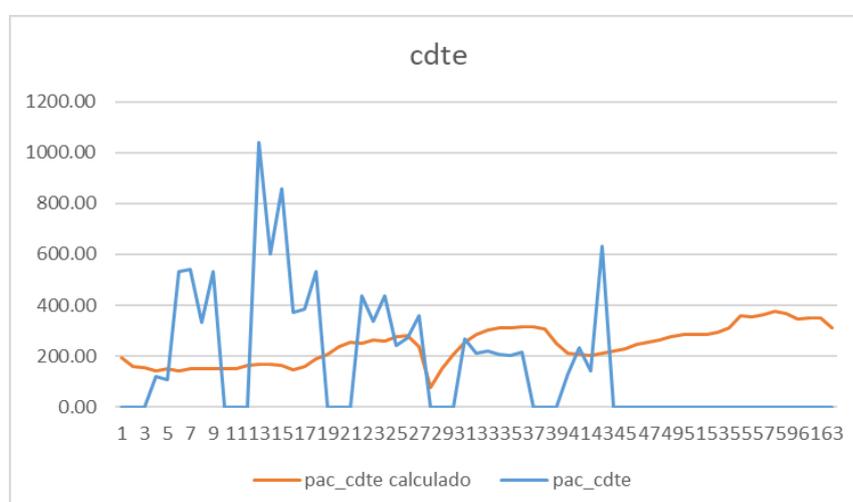
Considerando o eixo x sequencial:



**Figura 47: Tecnologia cdte - exposição da previsão de todos os dados.**

A figura representada acima (Figura 47) é resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Reflete resultados completamente insatisfatórios.

No mesmo modelo, porém apenas com dados das primeiras 24 horas, tem-se o seguinte resultado:

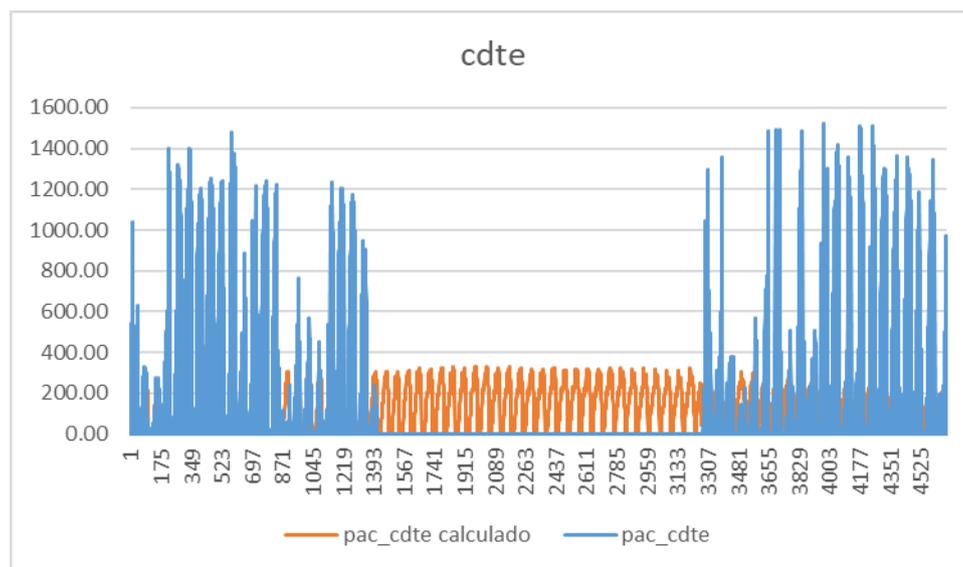


**Figura 48: Tecnologia cdte - exposição da previsão das primeiras 24 horas.**

A figura representada anteriormente (Figura 48) ilustra a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Reflete resultados completamente insatisfatórios.

Por ser a primeira tentativa, observa-se através dos gráficos uma correspondência equivalente ao esperado, um sistema com valores divergentes e incapazes de prever as potências diárias.

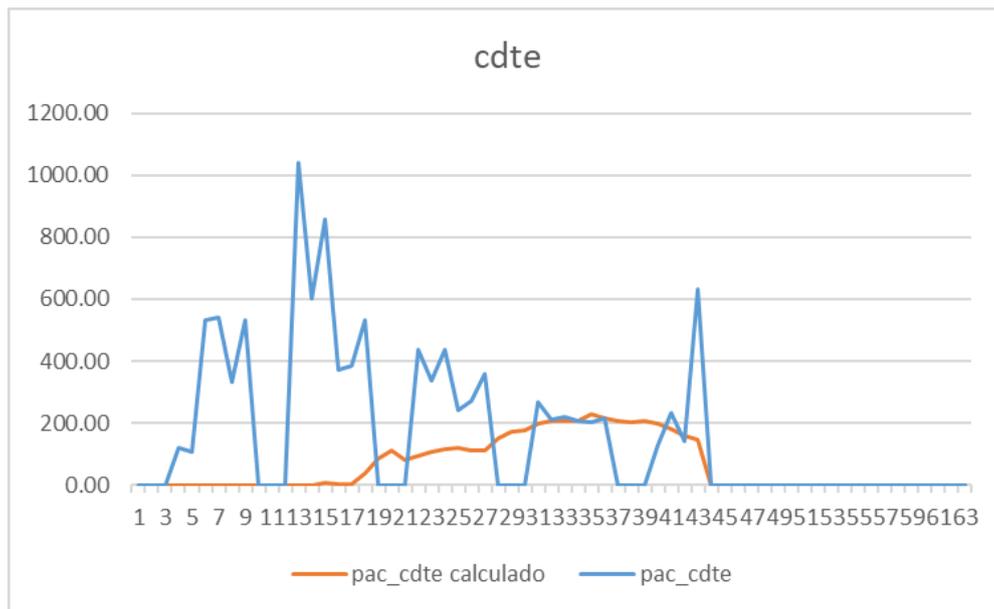
Considerando o eixo x ajustado:



**Figura 49: Tecnologia cdte - previsão de todos os dados com ajuste no eixo X.**

A figura acima (Figura 49) ilustra o resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Reflete resultados completamente insatisfatórios.

Com dados apenas das primeiras 24 horas:



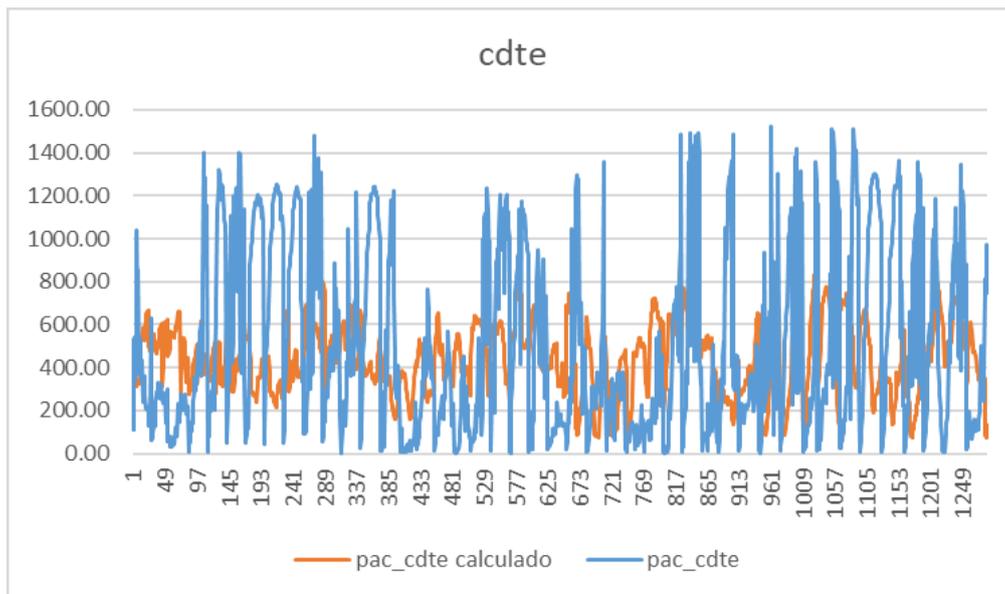
**Figura 50: Tecnologia cdte - primeiras 24h com ajuste no eixo X.**

Já a figura acima (Figura 50) representa a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Reflete resultados completamente insatisfatórios.

Mesmo possuindo uma linha de formato diferente, não há a possibilidade de afirmar que este é um modelo preditivo superior ao anterior, pois falha na previsão de diversos picos ao longo do dia.

Ao implementar a equação da regressão linear que desconsidera os valores nulos, os resultados obtidos foram representados a seguir:

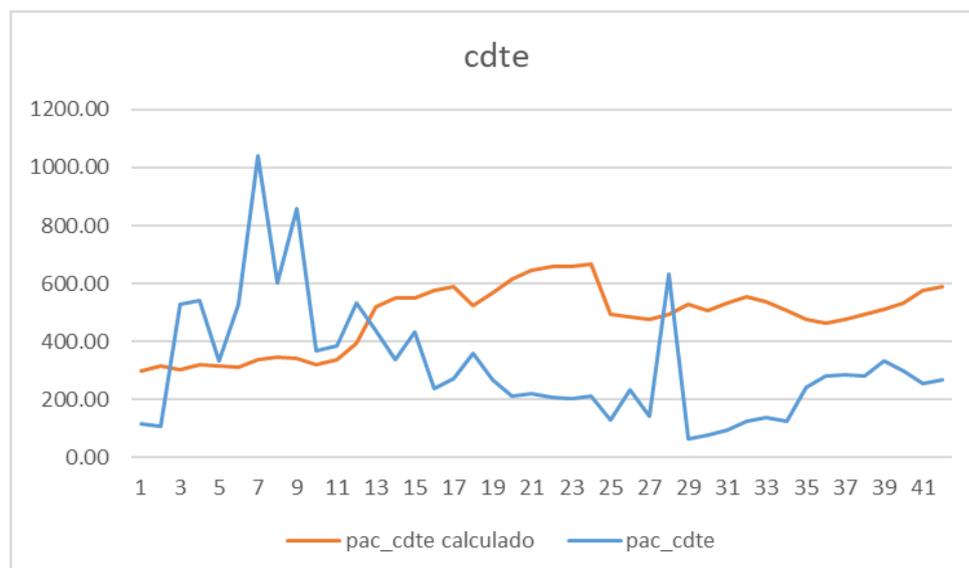
Considerando o eixo x sequencial:



**Figura 51: Tecnologia cdte - previsão de todos os dados com eixo X sequencial.**

De forma semelhante, a figura acima (Figura 51) representa o resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Reflete resultados insatisfatórios.

Demonstração apenas das primeiras 24 horas:

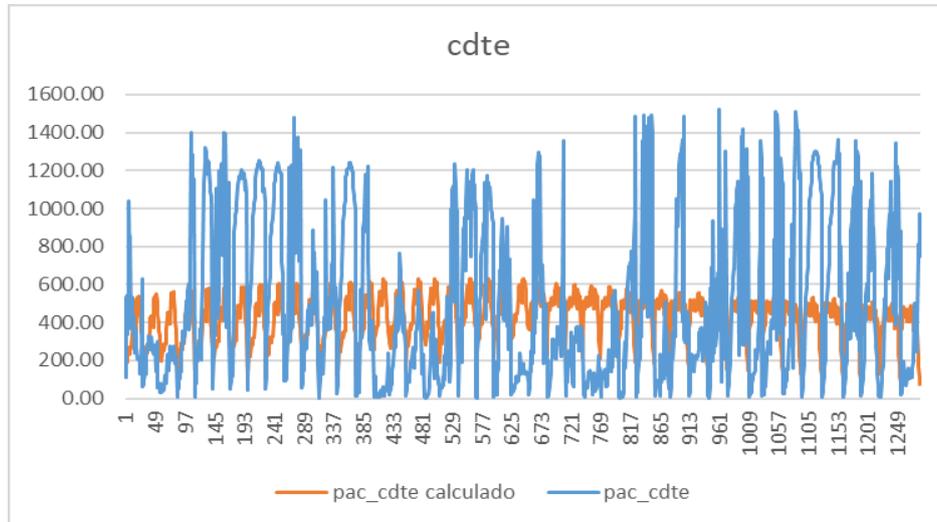


**Figura 52: Tecnologia cdte - primeiras 24 horas com eixo X sequencial.**

A figura acima (Figura 52) ilustra a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Reflete resultados insatisfatórios.

O sistema permanece impreciso, mesmo após utilizar a equação de regressão com dados nulos filtrados.

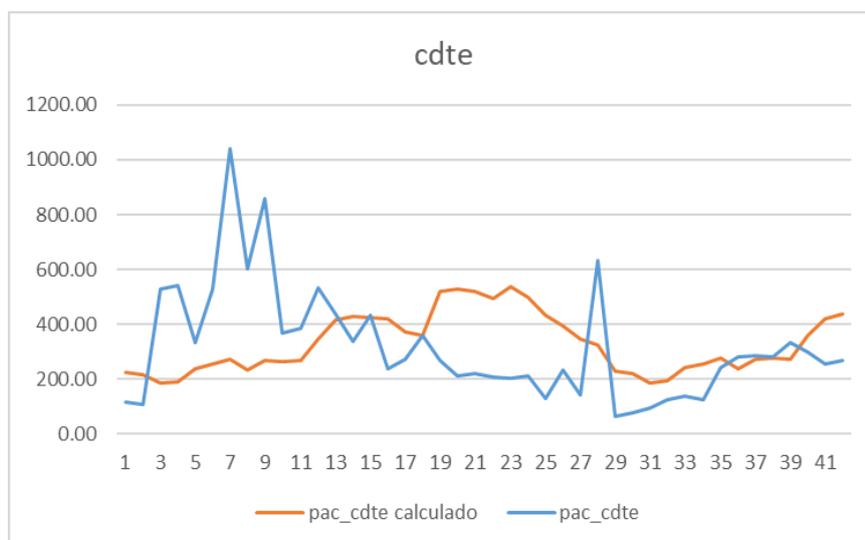
Após os dados do eixo x sequencial, foi implementado o modelo com o eixo x ajustado:



**Figura 53: Tecnologia cdte - exposição da previsão de todos os dados após ajuste no eixo X.**

O Resultado da integração das previsões à equação de regressão linear é representado pela figura acima (Figura 53), onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Também reflete resultados insatisfatórios.

Dados das primeiras 24 horas:



**Figura 54: Tecnologia cdte - exposição da previsão das primeiras 24 horas após ajuste no eixo X.**

Visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear também é representada acima (Figura 54), onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado

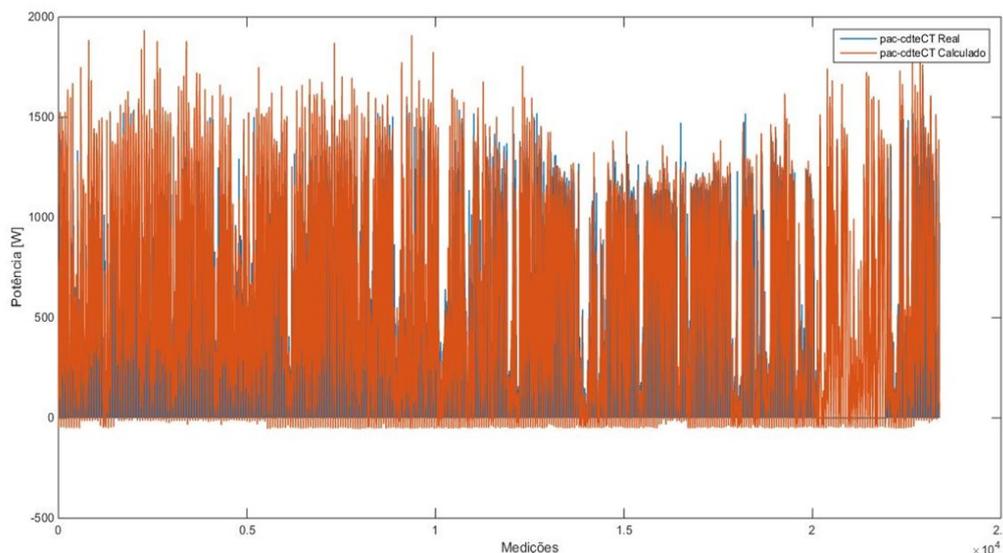
calculado da previsão com a equação da regressão linear. Reflete resultados insatisfatórios.

É possível observar uma suavização quando comparado com os gráficos do eixo x sequencial, porém, como sistema preditivo, os valores ainda diferem por uma margem elevada.

### 3.4.1.6 Segunda Ampliação do Conjunto de Dados

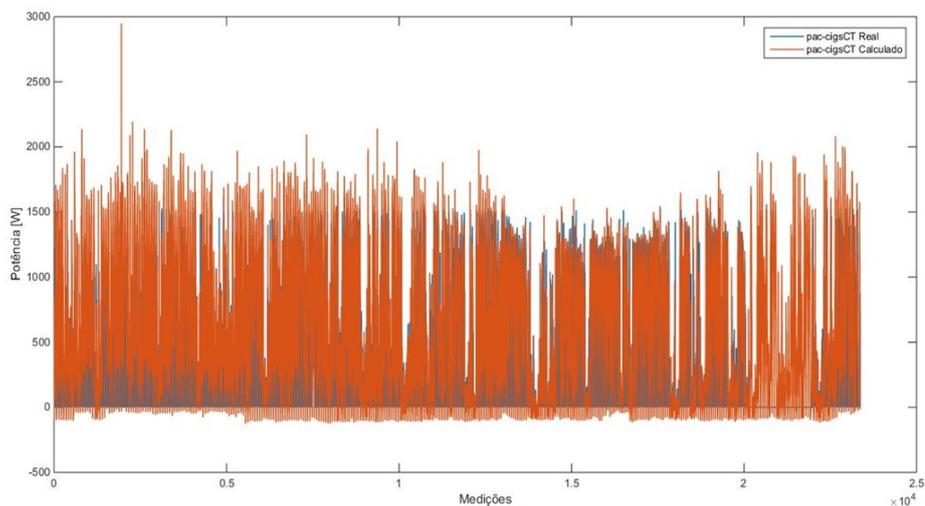
Após a realização da abordagem descrita acima (com dados do período de novembro do ano de 2021 até outubro do ano de 2022) e apreciação dos resultados, foi decidido então testar a regressão anterior neste novo conjunto de dados: Aplicou-se os resultados da primeira regressão (utilizando a fórmula feita a partir dos coeficientes calculados através dos dados de junho de 2021) no conjunto de dados de novembro/2021 ~ outubro/2020.

Novamente, comparando os valores de potência teóricos (calculados) com os valores de potência reais (medidos), vemos os resultados nos gráficos abaixo:



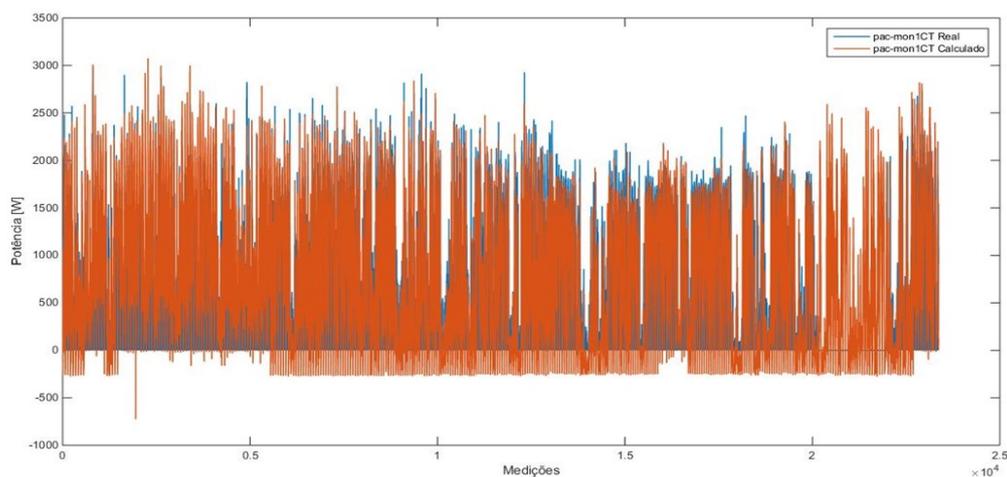
**Figura 55: Segunda Ampliação - Gráfico comparativo da variável pac-cdteCT**

Na figura acima (Figura 55), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se desta vez que ambas as linhas estão muito mais próximas, indicando que a equação de regressão do mês de Junho de 2021 calcula com mais precisão o valor da potência.



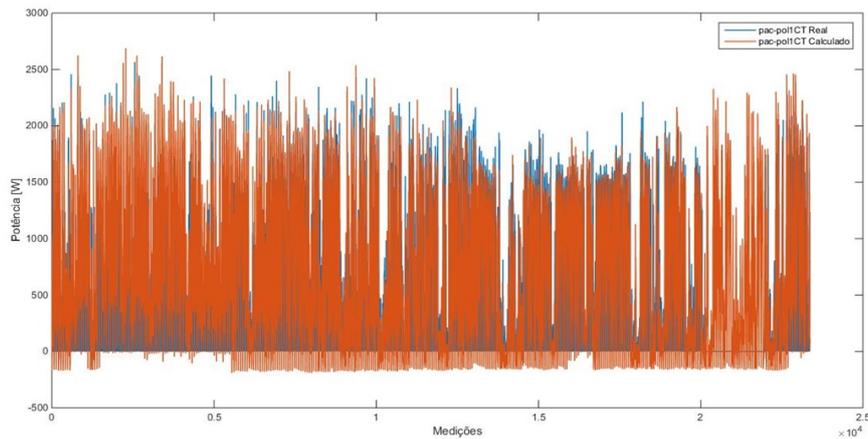
**Figura 56: Segunda Ampliação - Gráfico comparativo da variável *pac-cigsCT*.**

Também na figura acima (Figura 56), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se desta vez que ambas as linhas estão muito mais próximas, indicando que a equação de regressão do mês de Junho de 2021 calcula com mais precisão o valor da potência.



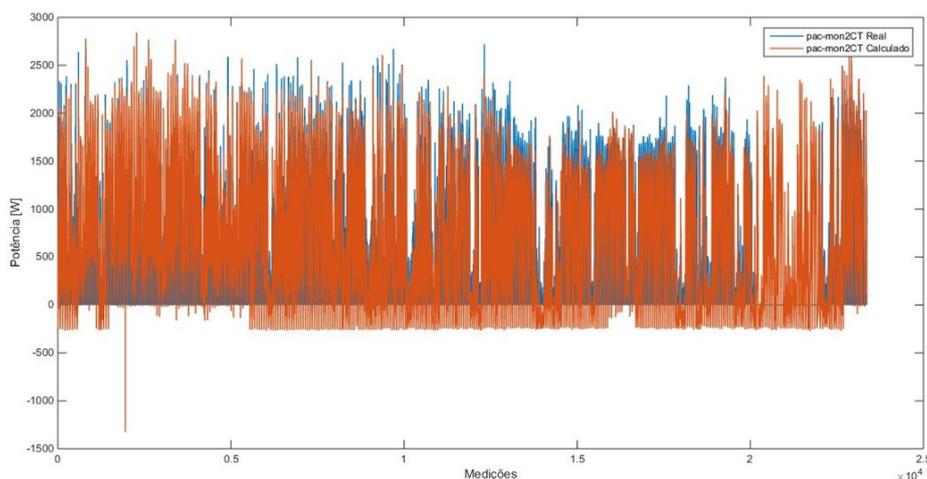
**Figura 57: Segunda Ampliação - Gráfico comparativo da variável *pac-mon1CT*.**

Da mesma forma, na figura acima (Figura 57), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se desta vez que ambas as linhas estão muito mais próximas, indicando que a equação de regressão do mês de Junho de 2021 calcula com mais precisão o valor da potência.



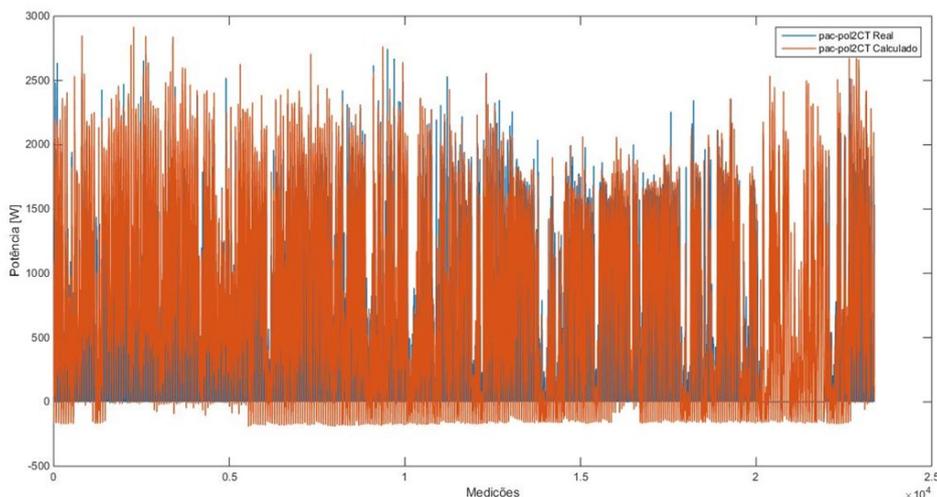
**Figura 58: Segunda Ampliação - Gráfico comparativo da variável *pac-pol1CT*.**

Também, na figura acima (Figura 58), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se desta vez que ambas as linhas estão muito mais próximas, indicando que a equação de regressão do mês de Junho de 2021 calcula com mais precisão o valor da potência.



**Figura 59: Segunda Ampliação - Gráfico comparativo da variável *pac-mon2*.**

Da mesma forma, na figura acima (Figura 59), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se desta vez que ambas as linhas estão muito mais próximas, indicando que a equação de regressão do mês de Junho de 2021 calcula com mais precisão o valor da potência.



**Figura 60: Segunda Ampliação - Gráfico comparativo da variável pac-pol2.**

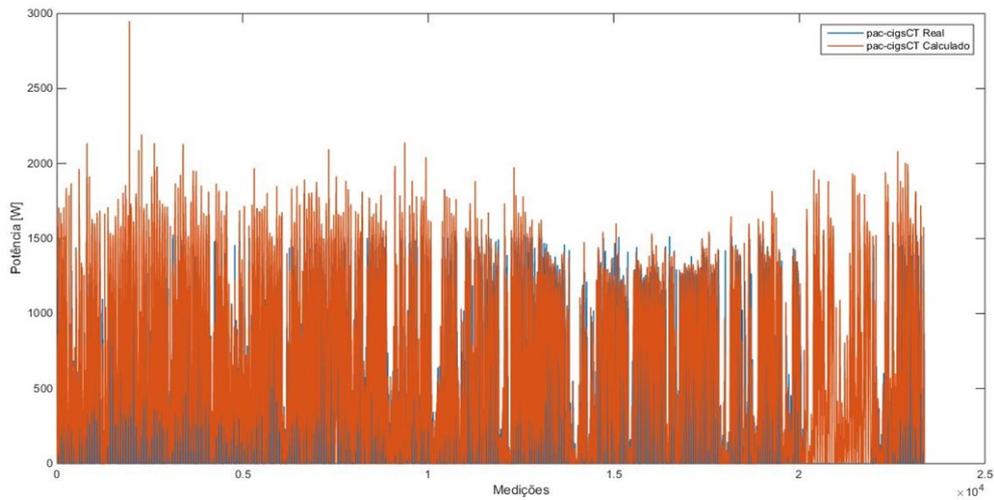
Por último, a figura acima (Figura 60), a linha azul corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Nota-se desta vez que ambas as linhas estão muito mais próximas, indicando que a equação de regressão do mês de Junho de 2021 calcula com mais precisão o valor da potência

Surpreendentemente, os resultados acima observados mostram-se melhores que a última análise realizada (regressão linear e cálculo de potência com os dados de novembro/2021~outubro/2022), mesmo após a remoção dos dados nulos, provando que a equação para estimativa de potência adquirida através dos dados meteorológicos do mês de junho de 2021 é mais precisa e confiável que a equação adquirida através dos dados de novembro/2021~outubro/2022.

Notamos ainda que existem valores negativos de potência calculados, resultados da aplicação da equação. Corrigindo estes valores, pode-se obter um resultado ainda mais satisfatório.

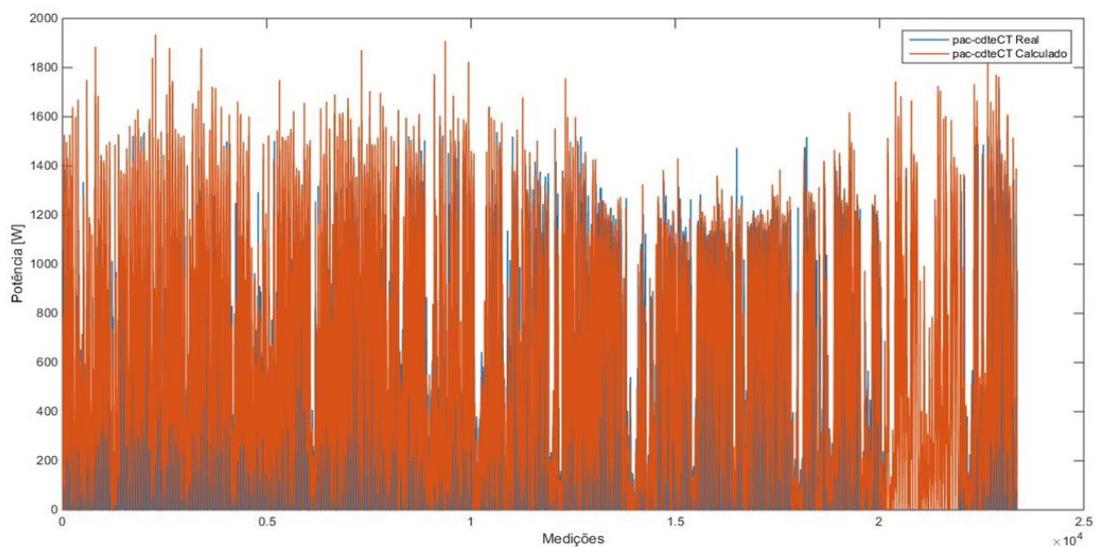
### 3.4.1.6.1 Filtragem de Valores Negativos

Aplicando aos mesmos resultados acima, uma filtragem simples (substituindo os valores calculados negativos pelo valor zero), temos agora os seguintes resultados:



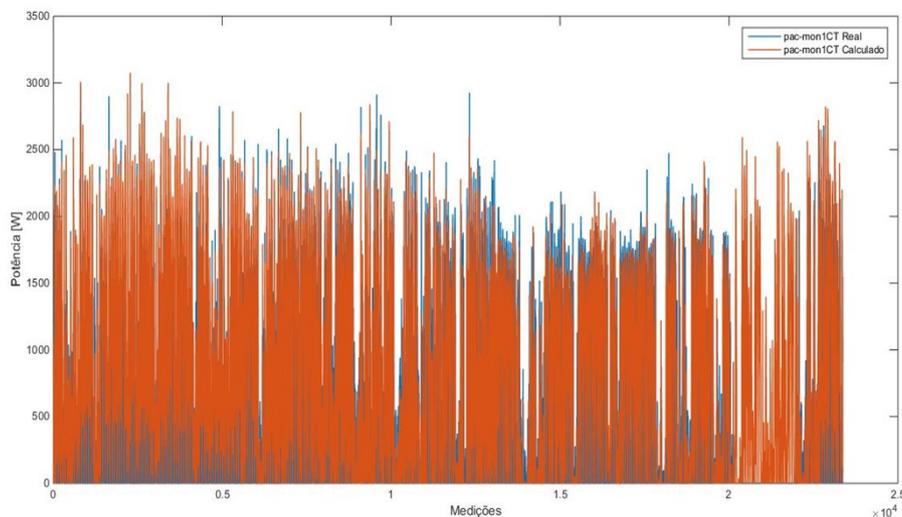
**Figura 61: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável *pac-cdteCT*.**

A linha azul presente na figura acima (Figura 61) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Ao removermos os valores negativos do gráfico, temos um resultado ainda mais satisfatório.



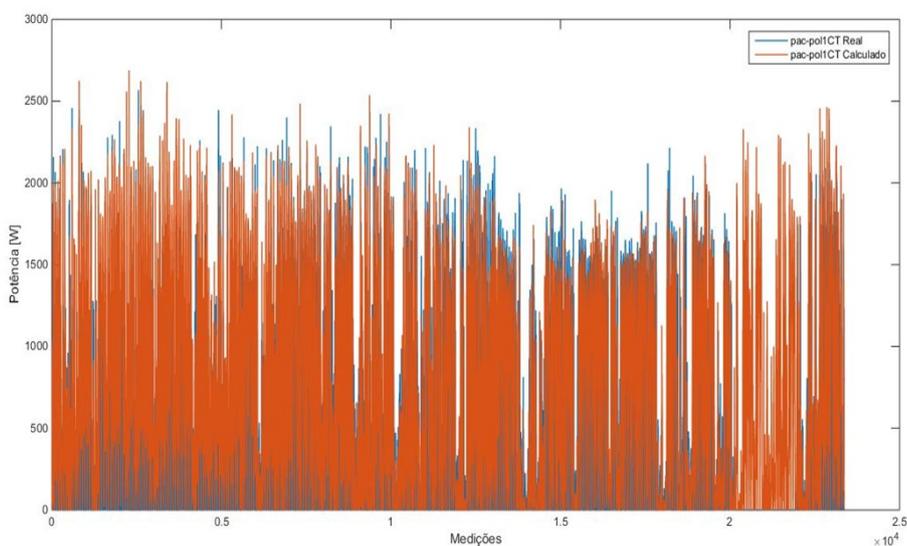
**Figura 62: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável *pac-cigsCT*.**

Da mesma forma, a linha azul da figura representada anteriormente (Figura 62) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Ao removermos os valores negativos do gráfico, temos um resultado ainda mais satisfatório.



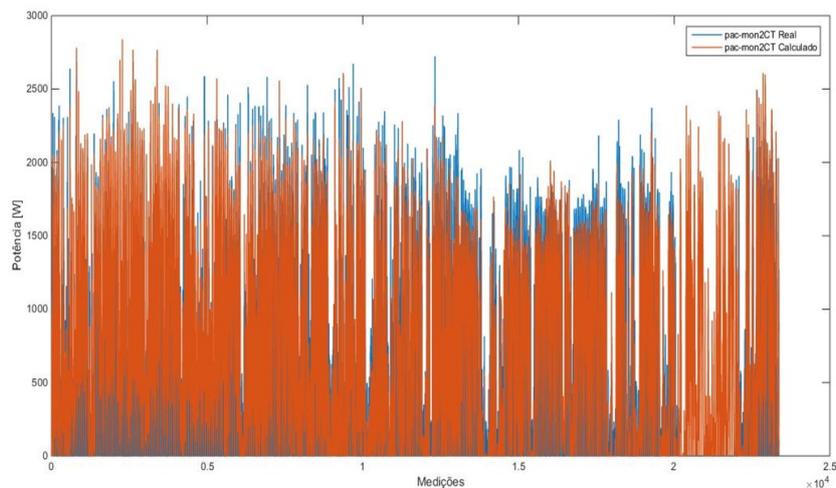
**Figura 63: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-mon1CT.**

De modo semelhante, a linha azul presente na figura acima (Figura 63) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Ao removermos os valores negativos do gráfico, temos um resultado ainda mais satisfatório.



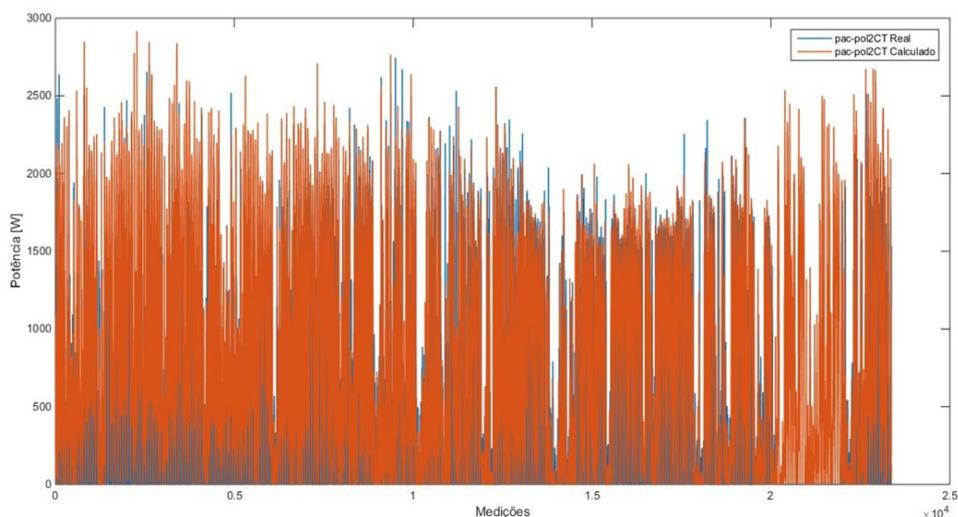
**Figura 64: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-pol1CT.**

A linha azul presente na figura anterior (Figura 64) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Ao removermos os valores negativos do gráfico, temos um resultado ainda mais satisfatório.



**Figura 65: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-mon2CT.**

Novamente, a linha azul presente na figura acima (Figura 65) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Ao removermos os valores negativos do gráfico, temos um resultado ainda mais satisfatório



**Figura 66: Segunda Ampliação sem valores negativos - Gráfico comparativo da variável pac-pol2CT.**

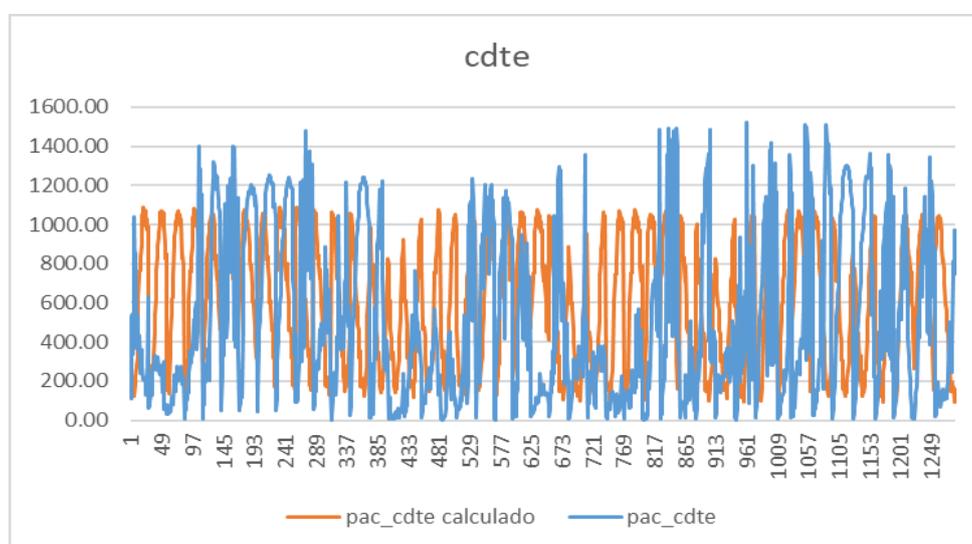
Por último, a linha azul presente na figura anterior (Figura 66) corresponde ao valor de potência real, enquanto que a linha vermelha corresponde ao valor de potência calculado. Ao removermos os valores negativos do gráfico, temos um resultado ainda mais satisfatório

É evidente a melhoria dos resultados após a filtragem dos dados negativos, podendo ser observada nos gráficos a forma com que os valores das potências calculadas acompanham os valores das potências reais.

### 3.4.1.7 Resultados da Previsão após Segunda Ampliação de Dados

Uma vez que as entradas não possuíam dados nulos como as saídas, os mesmos 2 sistemas demonstrados anteriormente poderão ser usados nesta nova equação de regressão linear.

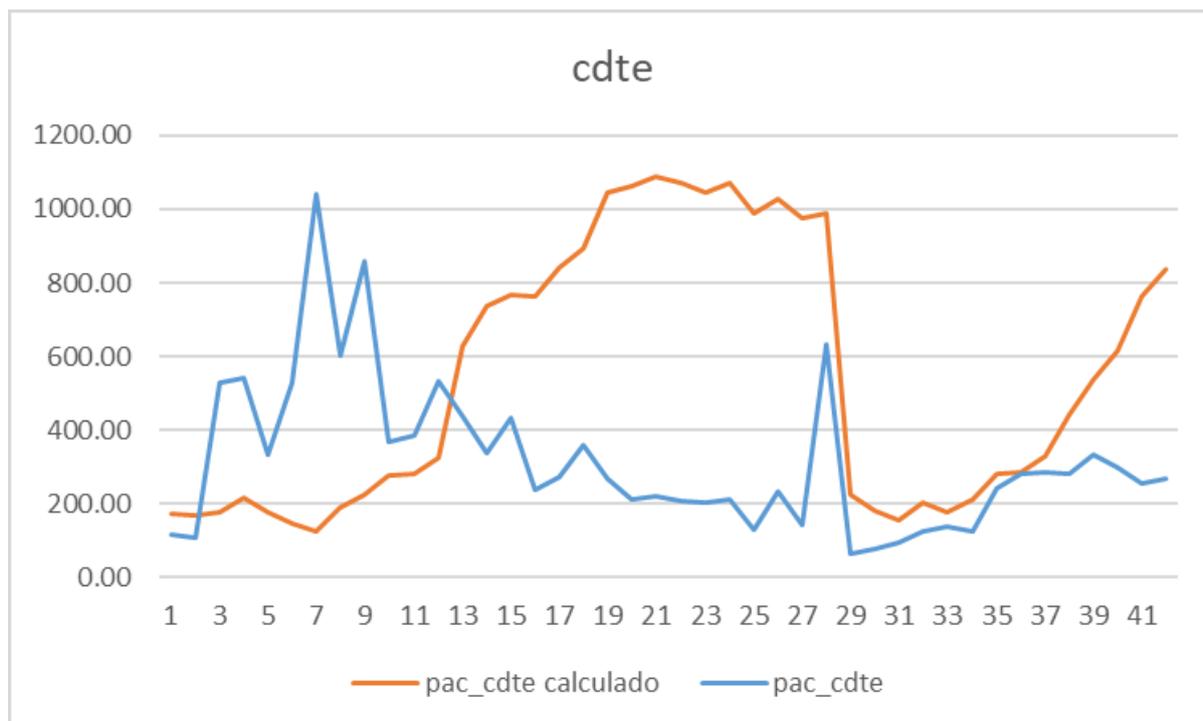
Eixo x sequencial:



**Figura 67: Tecnologia cdte - exposição da previsão de todos os dados após 2ª ampliação.**

A figura acima (Figura 67) ilustra o resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Com a limpeza dos valores nulos, os resultados apresentam padrões factíveis.

Primeiras 24 horas:

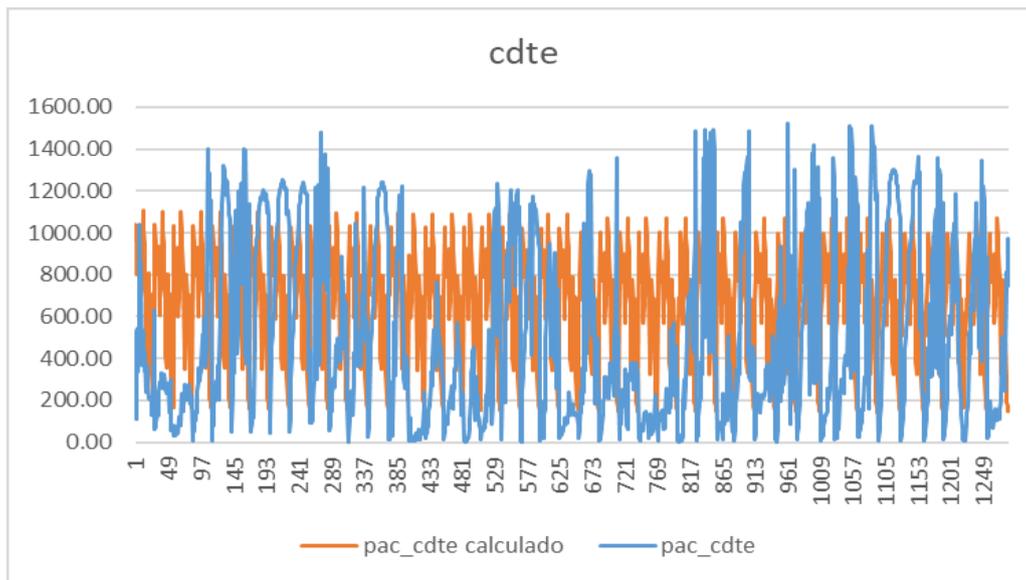


**Figura 68: Tecnologia cdte - exposição da previsão das primeiras 24h após 2ª ampliação.**

De modo semelhante, a figura acima (Figura 68) ilustra a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Há similaridades na amplitude, porém, o padrão diário foi insatisfatório.

Através dos gráficos, observa-se uma melhora no sistema, principalmente considerando a terça parte final dos mesmos.

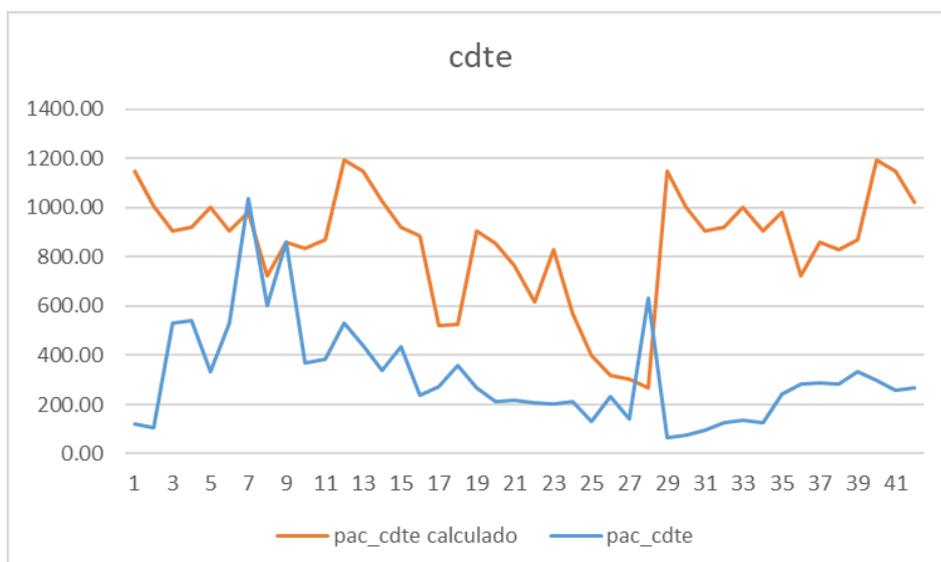
Eixo x ajustado:



**Figura 69: Tecnologia cdte - exposição da previsão de todos os dados após 2ª ampliação e ajuste em X.**

Já a figura acima (Figura 69) ilustra o resultado da integração das previsões à equação de regressão linear, onde a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Com a limpeza dos valores nulos, os resultados apresentam padrões factíveis.

Para as primeiras 24 horas:



**Figura 70: tecnologia cdte - exposição da previsão das primeiras 24h após 2ª ampliação e ajuste em X.**

Da mesma forma, a figura acima (Figura 70) representa a visão das primeiras 24 horas do resultado da integração das previsões à equação de regressão linear, onde

a linha azul corresponde aos dados medidos e a linha laranja corresponde ao resultado calculado da previsão com a equação da regressão linear. Há similaridades na amplitude, porém, o padrão diário foi insatisfatório.

Ao contrário do sistema com  $x$  sequencial, este modelo se mostra mais eficiente nos primeiros momentos da previsão.

### **3.4.2 MatLab**

Paralelamente ao desenvolvimento do sistema de previsão através do Excel, foi também estudado e criado um código para a implementação no software MATLAB, pois também existe uma função pronta para a geração de dados futuros. O nome desta função é `forecast`, que utiliza de um modelo matemático e dados passados para efetuar a previsão. O modelo utilizado será o ARMAX, explicado anteriormente.

#### **3.4.2.1 Modelagem Caixa-Preta**

Como mencionado na fundamentação teórica (Tópico 2.6), uma forma eficiente de se obter um modelo quando não há conhecimento pleno das relações entre todas as variáveis de entrada e saída, é a utilização de um modelo caixa-preta (o qual não se conhece o modelo em si, no entanto trabalha-se somente com o conjunto de dados de entradas e saídas).

Além disso, anteriormente, as dinâmicas meteorológicas são inevitavelmente caóticas, tornando muito onerosa a tarefa de se obter um modelo que descreva com uma precisão aceitável as dinâmicas meteorológicas locais. Por isso, decidiu-se utilizar um modelo caixa-preta.

Foi utilizado, para realizar a modelagem do sistema, um modelo Autorregressivo com entradas Exógenas (ARX), pelo fato do mesmo demandar pouca memória e atender bem às necessidades desse projeto. (ICPS, 2017)

Será descrito a seguir a metodologia utilizada para se obter os modelos ARX que descrevem o problema, tal qual foram gerados seis modelos MISO e seis modelos SISO, cada um descrevendo a relação entre as entradas (descritas a seguir) e uma única saída, sendo ela a potência em um dos seis inversores estudados nesse projeto.

### **3.4.2.1.1 Escolha de Variáveis de Entrada e Saída**

Para a realização dos testes, foi utilizada a base de dados do LABENS, com os dados referentes ao mês de junho de 2021. Essa base possui várias medidas referentes à estação meteorológica instalada em 09/12/2019. Foram verificados os possíveis dados, tais quais escolheu-se 6 variáveis de entrada e 6 variáveis de saída (ICPS, 2017).

Como mencionado no tópico 1.3, este trabalho possui como intenção realizar a predição da eficiência de um painel fotovoltaico, logo foram escolhidas como saídas do sistema as potências em 6 inversores diferentes, conectados à 6 tipos diferentes de módulos fotovoltaicos. Como entradas, foram utilizadas os dados referentes à 6 medidas meteorológicas, sendo elas a umidade relativa do ar, pressão atmosférica local, índice de chuva, temperatura superficial da estação, velocidade do vento e radiação média no plano inclinado na latitude.

#### **3.4.2.1.1.1 Verificação de Linearidade**

Antes de iniciar os testes, foi necessário verificar relação entre as variáveis de entrada e de saída, uma vez que o modelo ARX é um modelo linear. A utilização de variáveis de entrada e saída que possuem uma relação não linear significaria a necessidade da utilização de um modelo que leve em consideração tal não linearidade. Um exemplo de possível sistema seria um NARMAX, ou NARX, que possuem um grau de complexidade superior ao ARX ou ARMA, saindo do escopo do trabalho.

Para verificar a linearidade, foi utilizada a função “corrcoef(A)” do Matlab, que gera uma matriz de correlação entre  $n$  variáveis. A interpretação dessa matriz se dá pela verificação dos valores que não estão na diagonal principal. Quanto mais próximos de um, maior é a relação linear das duas variáveis, enquanto se o valor for próximo de zero, não há relação linear entre as variáveis. Um valor próximo a um negativo indica uma relação antilinear entre as variáveis.

Uma vez definidas as variáveis que possuem uma relação linear com as saídas, decidiu-se realizar os testes obtendo dois modelos: um sistema SISO utilizando apenas as variáveis com relação linear com a saída (radiação no plano inclinado) e um sistema MISO considerando as seis variáveis de entrada como lineares.

### 3.4.2.1.2 Modelagem ARX

Cada modelo ARX, tanto SISO quando MISO, foi gerado duas vezes: uma vez utilizando um método recursivo e outra utilizando as funções nativas do Matlab, que definem quais repressores na entrada e saída que zeram o **Critério de informação de Akaike**.

#### 3.4.2.1.2.1 Método Recursivo

O primeiro método utilizado foi utilizando uma função recursiva criada no matlab. A mesma utilizava 3 laços tipo “for”, alterando três variáveis na, nb e nk de um valor inicial até um final, determinado pelo operador. Para o sistema SISO os valores de na nb e nk variavam entre 1 e 5, enquanto no sistema MISO na variava entre 1 e 5, enquanto nb e nk variavam entre 3 e 5. Os valores de na, nb e nk foram inseridos na função arx, do System Identification Toolbox. Em tal função, os valores na, nb e nk representam a ordem do polinômio  $A(q)$ ,  $B(q)$  e o atraso de transporte entrada-saída, tal qual nb e nk são vetores com a quantidade de elementos referentes ao número de entradas. Foi feita uma simplificação no sistema MISO, na qual considerou-se as que nb e nk são vetores compostos da mesmas ordens para todos elementos, tal qual os vetores de 6 elementos se tonam [nb nb nb nb nb nb] e [nk nk nk nk nk nk], respectivamente.

Tal simplificação se deu pelo fato de que, ao realizar o teste utilizando-se de um laço for para cada elemento do vetor, obteve-se uma quantidade de possíveis combinações na casa dos bilhões, tornando o teste extremamente oneroso computacionalmente, levando cerca de quatro dias para se completar.

A função utilizada para gerar o modelo ARX (arx, System Identification Toolbox) possui como saída um objeto tipo idpoly que contém o modelo, gerado a partir de um conjunto de dados de entrada, ordens de sistema na e nb e atraso no tempo nk. Foi verificado o fit de cada modelo sobre o conjunto de dados do modelo, e salvo o modelo com melhor fit. O modelo foi comparado com o conjunto de dados de teste utilizando a função compare.

Foram gerados modelos utilizando como conjunto de dados a base de dados do LABENE, tal qual para o sistema SISO foi utilizada um tempo de amostragem de um minuto, enquanto para o sistema MISO, um tempo de amostragem de dez minutos.

Foram geradas previsões de séries no tempo utilizando os dados de modelagem, teste e modelos ARX utilizando a função forecast.

#### **3.4.2.1.2.2 Método TOOLBOX**

O segundo método utilizado para obtenção do modelo ARX foi a utilização das funções da biblioteca nativa do Matlab, System Identification Toolbox.

Foram utilizadas funções de seleção de ordem de modelo ARX, que dentro de um certo escopo pré determinado, encontra o conjunto de dados na, nb e nk que zeram a função de comparação de Akaike. Foram gerados modelos ARX SISO e MISO utilizando tais funções.

Foram geradas previsões de séries no tempo utilizando os dados de modelagem, teste e modelos ARX utilizando a função forecast.

## 4 RESULTADOS E DISCUSÕES

Edward Lorenz, matemático americano, descobriu o caos de um sistema durante pesquisa feita no Instituto Tecnológico do Massachussets (MIT), no final da década de 1950, onde ele modelou um sistema climático utilizando doze equações diferenciais. Lorenz deduziu que, a menor das modificações nas condições iniciais, podem modificar drasticamente o comportamento a longo termo de um sistema meteorológico (ENCYCLOPEDIA, 2021).

Com estas informações obtidas, junto da revisão das referências bibliográficas, foi concluído que a busca por uma base de dados mais robusta, junto da aplicação da linearização desses dados, são capazes de auxiliar na obtenção de um modelo preditivo que consiga estipular valores para até uma hora após determinado momento, extremamente distante dos sete dias simulados nos modelos previamente discutidos.

Tendo em mãos o conjunto de dados a ser utilizado para a preparação do método, analisou-se as variáveis disponíveis e dentre elas, optou-se pela utilização das listadas abaixo:

- humid-so10CT (%) - umidade;
- press-so10CT (mb) - pressão atmosférica;
- rain-so10CT (mm/h) - precipitação;
- tp\_sfc-so10CT (°C) - temperatura do ar superficial;
- ws\_max-so10CT (m/s) - velocidade máxima do vento;
- tilt\_avg-so01CT (W/m<sup>2</sup>) - média da radiação;

### 4.1 SELEÇÃO DO MÉTODO A SER APLICADO

Após diversas tentativas, a abordagem que mais se adequou à proposta do trabalho foi a aplicação de regressão linear no conjunto de dados disponíveis, mostrando-se a mais promissora, sendo este método o mais simples encontrado capaz de estimar a potência gerada através da equação obtida composta pelas demais variáveis utilizadas como entrada e seus respectivos pesos na equação.

### 4.1.1 Primeira Aplicação

Como dentre as variáveis escolhidas, a média da radiação é a que está mais fortemente ligada à geração fotovoltaica, inicialmente buscou-se encontrar uma relação entre as demais variáveis para com a média da radiação, considerando neste caso as 5 primeiras variáveis como entrada e a média da radiação (última) como saída.

Contudo, após simulações e considerações, esta abordagem não se mostrou viável e com grande aleatoriedade nos resultados, causada pelo fato das variáveis “de entrada” possuírem muita influência entre si, e não somente na radiação. Foi constatado então que nestas considerações (com tais entradas e saída) o sistema é muito dinâmico, cuja tentativa de linearização se mostrou ineficaz.

### 4.1.2 Segunda Aplicação

Após a primeira tentativa falha descrita acima, optou-se por uma nova abordagem: considerar todas as variáveis tratadas até o momento como sendo variáveis de entrada (incluindo a média da radiação), e como saída, utilizar os valores de potência gerada:

- pac-cdteCT (W)- Potência da rede AC - inversor de 1,5kW;
- pac-cigsCT (W)- Potência da rede AC - inversor de 1,5kW;
- pac-mon1CT (W)- Potência da rede AC - inversor de 3kW (monocristalino 1);
- pac-mon2CT (W)- Potência da rede AC - inversor de 3kW (monocristalino 2);
- pac-pol1CT (W)- Potência da rede AC - inversor de 3kW (policristalino 1);
- pac-pol2CT (W)- Potência da rede AC - inversor de 3kW (policristalino 2);

Nesta segunda abordagem obteve-se um resultado muito mais satisfatório. No modelo de regressão utilizado, pode-se ter inúmeras variáveis de entrada “X”, mas apenas uma variável de saída “Y”. Sendo assim, foram realizadas 6 regressões lineares, uma para cada variável de potência.

### 4.1.3 Definição do Método Para Excel

A aplicação da regressão linear retorna, como já demonstrado no desenvolvimento do trabalho, os valores dos coeficientes para cada variável de entrada, o que permite verificar qual o grau de influência de cada uma na potência gerada. Com posse dos coeficientes, montou-se a equação para determinar a

potência, calcular, e obter valores seus teóricos para então comparar estes com os valores de potência reais. Comparando os valores de potência calculados e reais para cada uma das medições (medições de Junho de 2021, com intervalos de 10 minutos), obteve-se um resultado significativamente satisfatório, pois pôde-se notar a grande semelhança entre os valores calculados (teóricos) e os valores medidos (reais), para todas as 6 potências analisadas, reafirmando a escolha da utilização do cálculo de regressão linear do excel para este objetivo.

#### **4.1.4 Ampliação do Conjunto de Dados e Análise de Resultados Obtidos**

Assim como na análise anterior, fazendo uso dos coeficientes gerados pela aplicação da regressão, montou-se a equação capaz de calcular a(s) potência(s) gerada(s), obtendo seus valores teóricos para então comparar com os valores de potência reais.

Desta vez, comparando os valores de potência calculados e reais para cada uma das medições (medições de Novembro de 2021 até Outubro de 2022, com intervalos de 10 minutos, somente de períodos diurnos), obteve-se um resultado menos satisfatório que o anterior, principalmente com relação aos módulos das potências calculadas.

Entretanto, a principal razão se deve (conforme já detalhado acima na metodologia do presente trabalho) devido à quantidade de medições “nulas” presentes neste conjunto de dados ser muito maior (mesmo proporcionalmente) do que no primeiro conjunto analisado. Porém, os resultados obtidos neste caso não desvalorizam nem modificam a escolha do emprego desta abordagem de aplicação da regressão linear, já que apesar das diferenças observadas na intensidade, apresentou boa resposta referente à sazonalidade (mesmo para um período estendido), sendo capaz de acompanhar precisamente as variações das potências geradas ao longo de todo o período analisado.

Conforme demonstrado no desenvolvimento, na tentativa de melhorar os resultados obtidos nesta segunda análise, foi realizada uma limpeza nos dados utilizados, removendo as medições onde haviam valores nulos de potência, para melhorar a qualidade do conjunto de dados e conseqüentemente da equação

resultante deste. Como pôde ser observado, embora isso tenha melhorado os resultados, estes ainda apresentaram-se insatisfatórios.

Em seguida, em uma última tentativa de melhorar os resultados da análise deste conjunto de dados (novembro/2021~outubro/2022), foi aplicada a equação baseada nos coeficientes adquiridos na primeira regressão linear realizada (com dados de junho de 2021) e, ao contrário do esperado, os resultados desta vez mostraram-se melhores que os anteriores (para este conjunto de dados maior), provando que a equação para cálculo de potência feita através da análise dos dados meteorológicos do mês de junho de 2021 se mostrou a mais eficaz para este tipo de abordagem.

## **4.2 RESULTADOS POR REGRESSÃO LINEAR**

Após inúmeras análises e tentativas, a aplicação de uma regressão linear com os dados disponíveis se mostrou a mais promissora para o objetivo e proposta do trabalho, de apresentar um método ou cálculo, sendo este o mais simplificado possível, de estimar a potência gerada num futuro próximo fazendo uso de dados observações passadas.

E ainda, seguindo neste raciocínio, a definição do modelo onde se definiu as seis variáveis de entrada (umidade, pressão, precipitação, velocidade do vento, temperatura e radiação) e as potências como saídas, se mostrou o mais satisfatório para esta abordagem. Muitos pontos observados nos resultados obtidos reforçam a ideia: como esperado o fator que mais ligado à potência gerada e conseqüentemente mais influente sobre esta, é a média da radiação incidente nos módulos. Como os demais fatores climáticos afetam, não só a radiação, mas também o próprio funcionamento do sistema fotovoltaico, a decisão de considerar também estes fatores na análise da potência agrega robustez à estimativa apresentada, tanto na aplicação da regressão linear neste conjunto de dados quanto na aplicação dos métodos preditivos.

## 4.3 RESULTADOS DA PREVISÃO POR EXCEL

### 4.3.1 Considerando Apenas Dados de JUN/2021

Ao implementar o modelo que considera apenas os dados de Junho de 2021, tanto para a regressão quanto para a previsão, desconsideramos o dia 28 de junho, pois as saídas revelam um possível lapso nas medições, com diversos valores nulos. o erro médio das previsões são representados na tabela a seguir:

| Tempo             | cdte   | cigs    | mon1   | mon2    | pol1   | pol2    |
|-------------------|--------|---------|--------|---------|--------|---------|
| <b>10 minutos</b> | 16,60% | 4,70%   | 63,30% | 70,20%  | 68,30% | 72,90%  |
| <b>1 hora</b>     | 58,70% | 57,70%  | 54,10% | 101,60% | 55,90% | 51,70%  |
| <b>1 dia</b>      | 48,50% | 61,70%  | 64,00% | 66,90%  | 60,30% | 74,30%  |
| <b>2 dias</b>     | 95,20% | 149,10% | 89,60% | 78,20%  | 94,80% | 104,80% |

*Tabela 13: Erro médio de previsões - JUN/2021.*

Através da análise dos gráficos gerados, os valores previstos podem ser observados como uma linha de tendência para os valores que serão obtidos durante o dia, servindo como um guia, pois o erro foi demasiado. Algo a se tomar disso é o fato de que a precisão é muito inferior quando tratamos de valores da manhã, especificamente os primeiros dados, que se mostram com a maior taxa de erro, ou prevendo valores em diversas ordens de magnitude superior ou inferior. Valores estes que aparentam estabilizar em um erro menor em torno das 10 horas da manhã, onde o algoritmo tem suas melhores previsões.

Embora não tenha se provado como uma ferramenta com confiabilidade indubitável, é possível observar, de certa forma, uma estimativa válida, que pode ser refinada e polida, a fim de auxiliar usuários de usinas solares residenciais a estimar a obtenção de potência ao longo do dia.

## 4.3.2 Considerando Dados de 1 Ano e Previsão Sequencial

### 4.3.2.1 Valores Nulos Presentes

#### 4.3.2.1.1 Eixo X Sequencial

Tendo este como o modelo mais rudimentar, e primeira tentativa realizada, foram consideradas as 23 mil linhas de dados, que levam em consideração momentos em que os painéis estão em funcionamento. Para a equação da regressão linear, foi considerada a primeira do tópico 3.4.1.5 com os seguintes erros médios, representados na tabela a seguir:

| Tempo             | cdte   | cigs    | mon1   | mon2   | pol1   | pol2   |
|-------------------|--------|---------|--------|--------|--------|--------|
| <b>10 minutos</b> | 18,40% | 4,10%   | 49,20% | 70,80% | 66,50% | 60,10% |
| <b>1 hora</b>     | 51,00% | 50,10%  | 52,90% | 56,20% | 55,70% | 57,20% |
| <b>1 dia</b>      | 47,00% | 54,30%  | 49,40% | 50,30% | 50,80% | 52,20% |
| <b>2 dias</b>     | 56,80% | 194,50% | 49,40% | 49,90% | 54,90% | 50,30% |

*Tabela 14: Tabela de erros médios de previsões – Período de 1 ano e eixo sequencial.*

Devido à presença de 12 vezes mais dados, havia a expectativa de um sistema mais assertivo, porém, o que é observado é um sistema mais consistente no erro, com menor variação, exceto pela tecnologia cigs.

Em qualquer sistema que considera os valores nulos, a previsão estabelece um padrão sazonal de valores durante o dia e, quando comparado com valores nulos, o erro tende ao infinito.

#### 4.3.2.1.2 Eixo X Ajustado

Ainda na equação regressiva mais rudimentar, os dados de previsão do sistema que considera o eixo x ajustado são implementados na base de dados com 23 mil linhas, que levam em consideração momentos em que os painéis estão em funcionamento. Para a equação da regressão linear, foi considerada a primeira do tópico 3.4.1.5 com os seguintes erros médios:

| Tempo             | cdte    | cigs    | mon1   | mon2   | pol1   | pol2   |
|-------------------|---------|---------|--------|--------|--------|--------|
| <b>10 minutos</b> | 162,20% | 168,60% | 44,70% | 17,70% | 6,10%  | 13,00% |
| <b>1 hora</b>     | 108,00% | 126,40% | 75,00% | 62,80% | 55,20% | 57,70% |
| <b>1 dia</b>      | 52,20%  | 64,70%  | 47,80% | 46,60% | 46,40% | 45,70% |
| <b>2 dias</b>     | 88,30%  | 263,40% | 83,60% | 82,80% | 96,70% | 96,20% |

*Tabela 15: Tabela de erros médios de previsões – Período de 1 ano e eixo ajustado.*

Têm-se, pelas tabelas apresentadas anteriormente, que os valores apresentados por esse sistema mostram grande inconsistência.

### 4.3.2.2 Valores Nulos Filtrados

#### 4.3.2.2.1 Eixo Sequencial

Agora com os valores nulos filtrados, foram considerados momentos em que os painéis geram valores observáveis, com o intuito de evitar as peculiaridades do modelo apresentado anteriormente. Para a equação da regressão linear, foi considerada a segunda do tópico 3.4.1.5 com os seguintes erros médios:

| Tempo             | cdte    | cigs    | mon1    | mon2    | pol1    | pol2    |
|-------------------|---------|---------|---------|---------|---------|---------|
| <b>10 minutos</b> | 152,60% | 131,50% | 34,60%  | 22,30%  | 11,90%  | 7,30%   |
| <b>1 hora</b>     | 79,50%  | 85,40%  | 60,90%  | 55,20%  | 51,70%  | 51,00%  |
| <b>1 dia</b>      | 123,80% | 272,00% | 111,90% | 110,00% | 113,50% | 109,50% |
| <b>2 dias</b>     | 201,40% | 421,40% | 159,90% | 162,50% | 182,90% | 164,80% |

*Tabela 16: Tabela de erros médios – Período de 1 ano, eixo sequencial e valores nulos filtrados.*

O motivo do erro exacerbado neste modelo se deve a momentos onde a previsão estima valores realistas para o futuro, porém, nos dados de teste de saída, algumas potências apresentaram valores inferiores aos esperados.

#### 4.3.2.2.2 Eixo x Ajustado

Utilizando os mesmos dados de regressão demonstrados anteriormente, apenas com a diferença do sistema preditivo utilizado, onde o eixo x foi ajustado:

| Tempo             | cdte    | cigs    | mon1    | mon2    | pol1    | pol2    |
|-------------------|---------|---------|---------|---------|---------|---------|
| <b>10 minutos</b> | 458,70% | 434,40% | 188,60% | 64,30%  | 87,40%  | 126,30% |
| <b>1 hora</b>     | 190,00% | 205,60% | 124,50% | 103,70% | 102,00% | 99,00%  |
| <b>1 dia</b>      | 179,80% | 369,50% | 147,10% | 144,90% | 152,90% | 148,40% |
| <b>2 dias</b>     | 275,80% | 563,80% | 215,40% | 218,70% | 243,10% | 223,20% |

*Tabela 17: Tabela de erros médios – Período de 1 ano, eixo ajustado e valores nulos filtrados.*

Embora haja boa previsão temporal de variáveis neste sistema, quando implementado junto com a fórmula de regressão linear, os resultados apresentados provam-se inaptos a prever resultados futuros corretamente.

### 4.3.2.3 Regressão de Junho/2021

#### 4.3.2.3.1 Eixo X Sequencial

Para os testes a seguir, a equação gerada a partir dos dados de Junho de 2021 foi utilizada. Com os seguintes erros médios:

| Tempo             | cdte    | cigs    | mon1    | mon2    | pol1    | pol2    |
|-------------------|---------|---------|---------|---------|---------|---------|
| <b>10 minutos</b> | 44,90%  | 9,50%   | 65,20%  | 82,40%  | 69,20%  | 58,50%  |
| <b>1 hora</b>     | 58,10%  | 55,60%  | 73,30%  | 77,80%  | 66,60%  | 63,40%  |
| <b>1 dia</b>      | 191,70% | 250,60% | 164,50% | 161,30% | 170,60% | 178,90% |
| <b>2 dias</b>     | 260,30% | 306,40% | 191,50% | 189,60% | 226,70% | 214,10% |

*Tabela 18: Tabela de erros médios - Regressão JUN/2021 e eixo sequencial.*

Para períodos mais recentes, especificamente 10 minutos e 1 hora, o sistema apresenta resultados mais favoráveis, quando comparado aos modelos revelados anteriormente.

#### 4.3.2.3.2 Eixo X Ajustado

Utilizando os mesmos dados de regressão demonstrados anteriormente, apenas com a diferença do sistema preditivo utilizado, onde o eixo x foi ajustado:

| Tempo             | cdte   | cigs   | mon1   | mon2   | pol1   | pol2   |
|-------------------|--------|--------|--------|--------|--------|--------|
| <b>10 minutos</b> | 865,5% | 844,8% | 353,5% | 148,1% | 193,4% | 273,8% |
| <b>1 hora</b>     | 350,8% | 82,0%  | 200,1% | 160,1% | 167,9% | 180,8% |
| <b>1 dia</b>      | 243,7% | 571,0% | 160,0% | 143,9% | 166,2% | 173,9% |
| <b>2 dias</b>     | 350,5% | 801,3% | 223,9% | 210,8% | 244,0% | 252,1% |

*Tabela 19: Tabela de erros médios - Regressão JUN/2021 e eixo ajustado.*

Embora haja boa previsão temporal de variáveis neste sistema, quando implementado junto com a fórmula de regressão linear, assim como demonstrado anteriormente com outra equação regressiva, o sistema se mostrou inapto a prever resultados futuros.

## 4.4 RESULTADOS DA PREVISÃO POR MATLAB

### 4.4.1 ESTUDO DE LINEARIDADE

Foi feito o estudo de linearidade entre cada uma das variáveis de entrada com a variável de saída, resultando em apenas uma variável de entrada como tendo relação linear com a saída, como mostra a tabela Tabela 20.

|           | humidso10CT | pressso10CT | rainso10CT | tp_sfcso10CT | ws_maxso10CT | tilt_avgso01CT |
|-----------|-------------|-------------|------------|--------------|--------------|----------------|
| pacdteCT  | -0.6727     | 0.0030      | -0.0819    | 0.4867       | 0.1518       | 0.9818         |
| paccigsCT | -0.6675     | 0.0068      | -0.0838    | 0.4812       | 0.1492       | 0.9807         |
| pacmon1CT | -0.6751     | 0.0017      | -0.0791    | 0.4949       | 0.1610       | 0.9768         |
| pacmon2CT | -0.6707     | 0.0020      | -0.0771    | 0.4922       | 0.1614       | 0.9754         |
| pacpol1CT | -0.6698     | 0.0018      | -0.0790    | 0.4853       | 0.1563       | 0.9759         |
| pacpol2CT | -0.6797     | 0.0019      | -0.0799    | 0.4934       | 0.1584       | 0.9892         |

Tabela 20: Estudo de linearidade entre variáveis.

### 4.4.2 MODELO SISO REGRESSIVO

Ao gerar o sistema SISO utilizando o método regressivo, obteve-se os modelos apresentados no ponto 1) do **ANEXO 2**, tais quais a ordem do polinômio  $A(q)$ ,  $B(q)$  e ordem do atraso de entrada-saída, para o modelo com o melhor fit considerando o conjunto de dados de modelagem, é mostrado na Figura 71.

|         | SISO<br>Regressivo | SISO<br>Toolbox | MISO<br>Regressivo | MISO<br>Toolbox |
|---------|--------------------|-----------------|--------------------|-----------------|
| pacdte  | 76,65%             | 76,94%          | 71,67%             | 33,58%          |
| paccgis | 67,61%             | 70,5%           | 70,2%              | 33,66%          |
| pacmon1 | 69,64%             | 69,24%          | 69,94%             | 50,36%          |
| pacmon2 | 69,8%              | 71,22%          | 69,98%             | 48,72%          |
| pacpol1 | 67,22%             | 67,28%          | 67,56%             | 56,34%          |
| pacpol2 | 63,45%             | 72,87%          | 70,42%             | 29,91%          |

Figura 71: FIT sobre dados dos testes.

| <b>SISO REGRESSIVO</b> |           |           |           |
|------------------------|-----------|-----------|-----------|
|                        | <b>na</b> | <b>nb</b> | <b>nk</b> |
| <b>pacdte</b>          | 7         | 7         | 1         |
| <b>pacgis</b>          | 7         | 7         | 7         |
| <b>pacmon1</b>         | 7         | 7         | 7         |
| <b>pacmon2</b>         | 7         | 7         | 7         |
| <b>pacpol1</b>         | 7         | 7         | 7         |
| <b>pacpol2</b>         | 7         | 7         | 7         |

*Figura 72: Ordem do Sistema SISO Regressivo.*

| <b>MISO REGRESSIVO</b> |           |             |             |
|------------------------|-----------|-------------|-------------|
|                        | <b>na</b> | <b>nb</b>   | <b>nk</b>   |
| <b>pacdte</b>          | 5         | 5 5 5 5 5 5 | 1 1 1 1 1 1 |
| <b>pacgis</b>          | 5         | 5 5 5 5 5 5 | 1 1 1 1 1 1 |
| <b>pacmon1</b>         | 5         | 5 5 5 5 5 5 | 1 1 1 1 1 1 |
| <b>pacmon2</b>         | 5         | 5 5 5 5 5 5 | 1 1 1 1 1 1 |
| <b>pacpol1</b>         | 5         | 5 5 5 5 5 5 | 1 1 1 1 1 1 |
| <b>pacpol2</b>         | 5         | 5 5 5 5 5 5 | 1 1 1 1 1 1 |

*Figura 73: Ordem do Sistema MISO Regressivo.*

| <b>SISO TOOLBOX</b> |           |           |           |
|---------------------|-----------|-----------|-----------|
|                     | <b>na</b> | <b>nb</b> | <b>nk</b> |
| <b>pacdte</b>       | 3         | 7         | 1         |
| <b>pacgis</b>       | 6         | 7         | 7         |
| <b>pacmon1</b>      | 7         | 4         | 7         |
| <b>pacmon2</b>      | 5         | 4         | 7         |
| <b>pacpol1</b>      | 6         | 1         | 7         |
| <b>pacpol2</b>      | 6         | 1         | 1         |

*Figura 74: Ordem do Sistema SISO TOOLBOX.*

| MISO TOOLBOX   |    |             |             |
|----------------|----|-------------|-------------|
|                | na | nb          | nk          |
| <b>pacdte</b>  | 3  | 3 4 3 3 3 3 | 3 5 3 5 5 5 |
| <b>pacgis</b>  | 3  | 3 4 3 3 3 3 | 3 5 3 5 5 5 |
| <b>pacmon1</b> | 3  | 3 4 3 3 3 3 | 3 5 4 5 5 4 |
| <b>pacmon2</b> | 3  | 3 3 3 3 3 3 | 3 5 4 5 5 4 |
| <b>pacpol1</b> | 3  | 3 3 3 3 3 3 | 5 5 5 5 5 3 |
| <b>pacpol2</b> | 3  | 3 4 3 3 3 3 | 3 5 4 5 5 5 |

*Figura 75: Ordem do Sistema MISO TOOLBOX.*

Foram obtidos dados de comparação entre as saídas reais do modelo de teste e as saídas geradas pelo modelo ARX, considerando as entradas do modelo de teste. Os resultados são mostrados na figura 3 do **ANEXO 1**, tal qual o fit dos resultados são mostrados na Tabela 20. Observa-se que o fit do modelo ficou entre 76,65% e 63,45% quando testados contra os conjuntos de dados de testes. Ao gerar os gráficos de previsão, obteve os resultados mostrados na figura 1 do **ANEXO 1**.

#### 4.4.3 MODELO MISO REGRESSIVO

Realizou-se o mesmo procedimento ao descrito para o método SISO regressivo, mudando apenas a quantidade de entradas para seis. O conjunto de dados utilizado para gerar o sistema MISO possui um tempo de amostragem de 10 minutos, diferente do conjunto de dados SISO, que possui um tempo de amostragem de 1 minuto.

Após a execução, foi salva em memória cada um dos conjuntos de ordens dos polinômios  $A(q)$ ,  $B(q)$  e quantidade de repressores que gerou o melhor modelo dentro dos parâmetros do programa. As ordens, assim como a quantidade de regressores é mostrada na Figura 73. Os modelos gerados são mostrados no tópico 3) do **ANEXO 2**. Os modelos foram comparados com os dados de teste, como no item acima, tal qual o fit dos sistemas é mostrado na figura Figura 71. Os resultados do modelo sobre os dados de teste e o forecast do modelo são mostrados nas figuras 6 e 8 do **ANEXO 1**, respectivamente.

#### 4.4.4 Método TOOLBOX

Como descrito na metodologia, foram utilizadas as funções nativas da Toolbox de identificação de sistemas para gerar os modelos ARX. A figura 2 do **ANEXO 1** mostra a comparação dos modelos gerados pela biblioteca com os dados de teste, para os modelos SISO. Verifica-se que o erro é levemente mais alto do que o modelo gerado pelo método regressivo, uma vez que a função de Akaike possui como premissa encontrar um modelo que não cause overfitting.

Na figura 7 do **ANEXO 1**, assim como na figura Figura 71 podemos observar que, ao contrário dos sistemas SISO, no sistema MISO o erro é extremamente grande, havendo vários valores de potências negativas. Isso se dá pelo fato de a toolbox ter sido desenvolvida para funcionar, primariamente, com sistemas SISO, fazendo com que os erros do sistema MISO sejam extremamente altos. Nas figuras 4 e 5 do **ANEXO 1** observa-se o forecast dos sistemas SISO e MISO gerados pelas funções nativas da Toolbox, respectivamente. As figuras 74 e 75 possuem as ordens de sistema para os modelos SISO toolbox e MISO toolbox, respectivamente.

## 5 CONCLUSÃO

### 5.1 APLICAÇÃO EM EXCEL

A escolha do modelo onde se definiu as seis variáveis de entrada (umidade, pressão, precipitação, velocidade do vento, temperatura e radiação) e as potências como saídas, se mostrou o mais satisfatório para esta abordagem. Muitos pontos observados nos resultados obtidos reforçaram ideias iniciais (como da média da radiação ser o fator que mais influencia a potência gerada). Como os demais fatores climáticos afetam, não só a radiação, mas também o próprio funcionamento do sistema fotovoltaico, a decisão de considerar também estes fatores na análise da potência agrega robustez à estimativa apresentada, tanto na aplicação da regressão linear neste conjunto de dados quanto na aplicação dos métodos preditivos.

Vale destacar também a importância da qualidade do conjunto de dados a ser estudado quando na aplicação deste método, pois valores nulos e/ou poluídos afetam drasticamente o resultado da regressão linear, o que conseqüentemente aumenta a incerteza e diminui a precisão dos coeficientes (pesos) das variáveis consideradas na equação do cálculo de potência.

Quando se trata da previsão analisada isoladamente uma a uma, o sistema com mais dados mostrou-se mais robusto, quando comparado com o sistema que possuía apenas dados de Junho de 2021. E dentro dos sistemas com mais dados, o que levou em consideração a data/hora como eixo x revelou-se a mais precisa. O motivo de haver a realização das duas se deve à dificuldade de trabalhar com a ferramenta, que não gera dados a menos que 100% das timestamps estejam corretas, portanto, uma parte da solução de problemas envolve dedicar tempo à análise de dados faltantes. Após o sucesso na realização do mesmo, os gráficos foram gerados e provaram-se satisfatórios.

Para a implementação da fórmula de regressão junto com os dados de previsão, o sistema que leva em consideração apenas os dados de Junho de 2021 se mostrou o melhor e mais apto a prever valores futuros, já que o sistema que considerou o conjunto com mais dados apresentou uma grande quantidade de medições nulas. Este resultado se deve ao fato de Junho é uma das épocas do ano com a menor incidência de chuvas (CLIMATE DATA, 2021), que são precedidas por nuvens e que atrapalham, ou até mesmo interrompem, o funcionamento dos painéis solares. Portanto, utilizar

Junho como parâmetro, prova-se como a estratégia mais eficaz para tratamento de dados, pois é o mês que terá maior quantidade de dados mensuráveis.

|                         | Janeiro | Fevereiro | Março | Abril | Maior | Junho | Julho | Agosto | Setembro | Outubro | Novembro | Dezembro |
|-------------------------|---------|-----------|-------|-------|-------|-------|-------|--------|----------|---------|----------|----------|
| Temperatura média (°C)  | 20.4    | 20.6      | 19.8  | 18.1  | 15    | 13.9  | 13.4  | 14.5   | 15.8     | 17.4    | 18.1     | 19.8     |
| Temperatura mínima (°C) | 17.6    | 17.8      | 17.1  | 15.2  | 12    | 10.7  | 9.9   | 10.6   | 12.2     | 14      | 15       | 16.6     |
| Temperatura máxima (°C) | 24.6    | 24.7      | 23.8  | 22.3  | 18.9  | 18.2  | 18    | 19.7   | 20.9     | 22.2    | 22.6     | 24.2     |
| Chuva (mm)              | 233     | 199       | 137   | 84    | 100   | 101   | 104   | 84     | 144      | 142     | 134      | 168      |
| Umidade(%)              | 85%     | 86%       | 85%   | 83%   | 83%   | 84%   | 82%   | 80%    | 80%      | 83%     | 83%      | 83%      |
| Dias chuvosos (d)       | 18      | 16        | 15    | 9     | 8     | 7     | 6     | 6      | 9        | 12      | 13       | 15       |
| Horas de sol (h)        | 5.6     | 6.1       | 5.2   | 4.9   | 4.7   | 5.5   | 5.8   | 6.3    | 5.3      | 4.8     | 5.0      | 5.7      |

Figura 76: gráfico com dados climatológicos para Curitiba 1991-2021

## 5.2 MODELAGEM ARX

Verificou-se através da utilização da função *corrcoef* (Matriz de coeficientes de Correlação) do Matlab que a relação linear entre as entradas e saídas não são lineares, exceto pela radiação global horizontal. Por conta disso, realizou-se dois testes: um considerando a global horizontal como entrada (SISO) e um considerando todas as seis variáveis como descrito anteriormente (MISO).

Foi escolhido o método ARX como modelo regressivo, seguindo o trabalho de Da Silva (2022), que também utilizou deste modelo para realizar o cálculo de diferentes tipos de radiação a partir da radiação global horizontal. Tal modelo (ARX), por considerar que a relação entre as variáveis é linear e não possuir médias móveis, requer um poder computacional relativamente baixo, comparado com sistemas ARMAX e NARX.

Foram gerados modelos ARX de duas maneiras diferentes, e utilizando dois conjuntos de dados, totalizando quatro testes:

1. Modelo SISO regressivo;
2. Modelo SISO toolbox;
3. Modelo MISO regressivo;
4. Modelo MISO toolbox.

Utilizou-se 20% dos dados do mês de junho de 2021 para realizar os testes dos modelos, e 80% dos dados para realizar o modelo em si (MORENO, 2007). O modelo regressivo variava a ordem de cada um dos elementos do sistema para tentar encontrar o modelo que trouxesse a menor taxa de erro médio quadrático (fit % no Matlab), enquanto o método Toolbox utiliza do critério de comparação de Akaike para encontrar o melhor modelo para um dado vetor de ordens de sistema.

Encontrou-se uma certa dificuldade para realizar os testes do modelo MISO, uma vez que, entre as ordens de  $n_a$ ,  $n_b$  e atraso  $n_k$ , há 13 variáveis a serem consideradas, tornando o script extremamente oneroso computacionalmente. Para o modelo 3, MISO regressivo, a grande quantidade de variáveis faz com que o script tenha que ser rodado algumas milhões de vezes afim de encontrar o melhor modelo, fazendo com que o programa demorasse cerca de cinco dias para executar, tornando o programa mais lento do que o tempo que estava sendo previsto na série temporal (três dias), tornando o modelo inutilizável sem fazer algumas simplificações. Com as simplificações, o script passou a ter um tempo de execução de cerca de cinco horas. Foram feitas as seguintes simplificações:

- As ordens de  $n_b$  e  $n_k$  foram limitadas de 3 a 5, diminuindo a quantidade de execuções do script regressivo;
- Considerou-se o vetor de ordens  $n_b$  como possuindo todos elementos de mesma ordem, diminuindo a quantidade de laços necessários de seis para um;
- Considerou-se o vetor de ordens  $n_k$  como possuindo todos elementos de mesma ordem, diminuindo a quantidade de laços necessários de seis para um.

O modelo 4, MISO Toolbox, também teve que ser simplificado, aplicando-se somente a primeira simplificação, limitando a ordem dos vetores  $n_b$  e  $n_k$ . O motivo da simplificação é que o modelo toolbox gera todas as possíveis combinações de ordens  $n_a$ ,  $n_b$  e  $n_k$  em uma matriz, checando uma a uma. Sem a simplificação o Matlab acusava erro de memória, pois não é possível gerar uma matriz com bilhões de linhas e treze colunas.

Ao realizar os testes utilizando o método ARX, observou-se uma melhora significativa no modelo ao utilizar o modelo SISO em relação do MISO, uma vez que as variáveis de umidade, vento, temperatura, chuva e pressão não possuem relação linear com as saídas escolhidas (as potências nos inversores).

Utilizando os quatro modelos gerados, foi possível realizar uma prova de conceito, realizando uma previsão de série temporal a partir de cada um dos modelos, para cada uma das saídas, utilizando a função **forecast** do Matlab.

Observou-se que a modelagem ARX descreve de forma satisfatória um sistema meteorológico local de curto prazo, principalmente considerando uma modelagem SISO (um fit de 63% a 76% para o sistema regressivo). Para trabalhos futuros, sugere-se realizar a previsão MISO utilizando um sistema não linear, como um NARX ou NARMAX, tendo a possibilidade de se obter um modelo que descreva os fenômenos estudados utilizando todas as seis variáveis de entrada. Outra sugestão seria utilizar os modelos estudados para aplicação em um universo de dados maior tentando fazer um modelo que seja capaz de prever em médio e longo prazo. Para tal, seria necessário um sistema computacional extremamente otimizado, possibilitando a geração de sistemas MISO com mais rapidez, gerando assim modelos relevantes dentro do tempo estudado.

## REFERÊNCIAS

ARISTONE, Flavio; MELLO, G. O. L.; OLIVEIRA, A. P. G.; RODRIGUES, V. S.; SOUZA, Amaury. **EFEITOS DA TEMPERATURA AMBIENTE E DA VELOCIDADE DO VENTO NO DESEMPENHO DE MÓDULOS SOLARES FOTOVOLTAICOS**. Revista de Geografia e Interdisciplinaridade, InterEspaço. Grajaú, 2018.

BEIGELMAN, B. B. **A Energia Solar Fotovoltaica E A Aplicação Na Usina Solar De Tauá**. Trabalho de Conclusão de Curso – Universidade Federal do Rio de Janeiro, Escola Politécnica. Rio de Janeiro, 2013.

CALDAS, R. P.; MELO, A. C. G.; PESSANHA, J. F. M.; FALCÃO, D. M.. **UM MODELO DE PREVISÃO DE CURTO PRAZO DE ENERGIA SOLAR FOTOVOLTAICA, USANDO MÁQUINA DE VETORES DE SUPORTE (SVM)**. XXV Seminário Nacional de Produção e Transmissão de Energia Elétrica. Belo Horizonte. Disponível em [https://www.researchgate.net/publication/337289573\\_UM\\_MODELO\\_DE\\_PREVISA\\_O\\_DE\\_CURTO\\_PRAZO\\_DE\\_ENERGIA\\_SOLAR\\_FOTOVOLTAICA\\_USANDO\\_MAQUINA\\_DE\\_VETORES\\_DE\\_SUPORTE\\_SVM](https://www.researchgate.net/publication/337289573_UM_MODELO_DE_PREVISA_O_DE_CURTO_PRAZO_DE_ENERGIA_SOLAR_FOTOVOLTAICA_USANDO_MAQUINA_DE_VETORES_DE_SUPORTE_SVM). Acesso em 12 de out de 2020.

Centro de Referência para Energia Solar e Eólica Sérgio de Salvo Brito (CRESESB). **ENERGIA SOLAR: PRINCÍPIOS E APLICAÇÕES**. Rio de Janeiro, 2006. Disponível em [http://www.cresesb.cepel.br/download/tutorial/tutorial\\_solar\\_2006.pdf](http://www.cresesb.cepel.br/download/tutorial/tutorial_solar_2006.pdf). Acesso em 20 de nov de 2021.

CREF UFRGS, 2019. **QUEDA DA PRESSÃO ATMOSFÉRICA ANTES DA CHUVA?** Disponível em: <http://cref.if.ufrgs.br/?contact-pergunta=queda-da-pressao-atmosferica-antes-da-chuva>. Acesso em set de 2022.

DA SILVA, Vinícius Leonardo Gadioli et al. An Approach to Solar Radiation Prediction Using ARX and ARMAX Models. **Frontiers in Energy Research**, v. 10, 2022.

DANTAS, Stefano G.; POMPERMAYER, Faniano M. **VIABILIDADE ECONÔMICA DE SISTEMAS FOTOVOLTAICOS NO BRASIL E POSSÍVEIS EFEITOS NO SETOR ELÉTRICO**, IPEA, Rio de Janeiro, 2018.

Emiliano, Paulo C., et al. "**Critérios de informação de akaike versus bayesiano: análise comparativa.**" *19º Simpósio Nacional de Probabilidade e Estatística* (2010).

FADIGAS, Eliane A. F. **Amaral. Energia Solar Fotovoltaica: Fundamentos, Conversão e Viabilidade técnico-econômica**. GEPEA - Grupo de Energia Escola Politécnica Universidade de São Paulo, 2004.

HAYANI, R.A.; OLIVEIRA, M. A. G.; CAMARGO, I. M. T. **Comparação do Custo entre Energia Solar Fotovoltaica e Fontes Convencionais**. V Congresso Brasileiro de Planejamento Energético. Brasília. Disponível em: [https://edisciplinas.usp.br/pluginfile.php/3427159/mod\\_resource/content/1/solar.pdf](https://edisciplinas.usp.br/pluginfile.php/3427159/mod_resource/content/1/solar.pdf). Acesso em 12 de out de 2020

INSTITUTO NACIONAL DE METEOROLOGIA. **Dados Históricos Anuais**. Disponível em <https://portal.inmet.gov.br/dadoshistoricos>. Acesso em: 22 de out 2020.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (São José dos Campos). **Atlas Brasileiro de Energia Solar**. 2.ed. São José dos Campos: 2017. Disponível em [http://labren.ccst.inpe.br/atlas\\_2017.html](http://labren.ccst.inpe.br/atlas_2017.html). Acesso em: 20 de out 2020.

JUNIOR, J.U. **Sistemas fotovoltaicos conectados a redes de distribuição urbanas: sua influência na qualidade da energia elétrica e análise dos parâmetros que possam afetar a conectividade**. Tese (Doutorado - Programa de Pós-Graduação em Engenharia Civil) – Universidade Federal de Santa Catarina, Centro Tecnológico. Florianópolis, 2010.

KULKARNI, Rohan S.; TALANGE, Dhananjay B.. **Modeling of Solar Photovoltaic module using system identification**. 7<sup>th</sup> International Conference of Power Systems (ICPS), College of Engineering Pune, India. 2017.

LAZZARETTI, AE; COSTA; CHD, RODRIGUES, MP; YAMADA, GD; LEXINOSKI, G; MORITZ, GL; OROSKI, E; GOES, RE; LINHARES, RR; STADZISZ, PC; OMORI, JS; SANTOS, RBD; **A Monitoring System for Online Fault Detection and Classification in Photovoltaic Plants**. Sensors (Basel). 2020 Aug;2017.

MARIANO, Glauber L. **PRESSÃO E MOVIMENTOS ATMOSFÉRICOS**. UFPEL, 2014. Disponível em: <https://wp.ufpel.edu.br/glaubermariano/files/2014/01/Unidade-III-Press%C3%A3o-e-movimentos-atmosf%C3%A9ricos.pdf> Acesso em: jul de 2021.

MICROSOFT. **Usar Ferramentas de Análise Para Executar Análises de Dados Complexas**. Disponível em <https://support.microsoft.com/pt-br/office/usar-ferramentas-de-an%C3%A1lise-para-executar-an%C3%A1lises-de-dados-complexas-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6>. Acesso em 8 de jul de 2022.

MONDOC B.; POP F. **Factors Influencing the Performance of a Photovoltaic Power Plant, 2010**. 3rd International Conference On Modern Power Systems. Romania. Disponível em: [https://ie.utcluj.ro/files/acta/2010/Number5/MPS\\_2010\\_Mondoc.pdf](https://ie.utcluj.ro/files/acta/2010/Number5/MPS_2010_Mondoc.pdf). Acesso em 15 de out 2020.

MORENO, Raquel Salazar et al. **Modelos ARX para predecir la temperatura del aire de un invernadero: una metodología**. Agrocienca, v. 41, n. 2, p. 181-192, 2007.

NEVES, R. N. F. **Desenvolvimento de Modelos de Previsão de Produção de Centrais Solares Fotovoltaicas**. Dissertação (Mestrado Integrado em Engenharia Electrotécnica e de Computadores Major Energia) - Faculdade de Engenharia da Universidade do Porto. Porto 2010.

O SETOR ELÉTRICO. **Avaliação econômica de diferentes tecnologias fotovoltaicas com banco de baterias**. Disponível em: <https://www.osestoreletrico.com.br/tecnologia-fotovoltaica-avaliacao-economica-de-diferentes-tecnologias-com-banco-de-baterias>. Acesso em: 19 de out. 2011.

PORTALENERGIA. **Energia Fotovoltaica: Manual sobre tecnologias, projecto e instalação**. Disponível em <https://www.portal-energia.com/category/energiasolar/documentacao-energia-solar/>. Acesso em 19 de nov de 2021.

PORTALSOLAR. **ENERGIA FOTOVOLTAICA**. 2014 a 2022. Disponível em: <https://www.portalsolar.com.br/energia-fotovoltaica.html>. Acesso em jan 2022.

SAUAIA, R. L. **Energia Solar Fotovoltaica: Panorama, Oportunidades e Desafios**. Seminário Desafios da Geração de Energia Elétrica no Brasil. Brasília. Disponível em <https://www.aneel.gov.br/documents/10184/15266087/painel+3+ap+7+2017.10.19+ABSOLAR+-+Energia+Solar+Fotovoltaica+-+Dr.+Rodrigo+Lopes+Sauaia.pdf/54f8b161-751b-0639-bd04-77a60cac45c3>. Acesso em 28 de out de 2020.

SILVA, Vinicius Leonardo Gadioli da. **Predição da radiação solar por meio de modelagem linear (ARX e ARMAX) e modelagem não-linear (Redes Neurais)**. 2016.

TIEPOLO G. M.; JUNIORB, O. C.; JUNIORA, J. U. Analysis of the Electricity Generation Potential by Solar Photovoltaic Source in the State of Paraná – Brazil. **Chemical Engineering Transactions**. Volume 35, 2013. Disponível em: <https://www.aidic.it/cet/13/35/100.pdf>. Acesso em 28 de out 2020.

TIEPOLO, G. M; PEREIRA, E.B.; JUNIOR, J. U.; PEREIRA, S. V.; RODRIGUES, A.; LIMA, F. J. L. L.; COSTA, R. S.; ALVES, A. R.; Atlas De Energia Solar Do Estado Do Paraná – Resultados. **Revista Brasileira de Energia Solar** Ano 9 Volume IX Número 1 Julho de 2018 p. 01-10. Disponível <https://anaiscbens.emnuvens.com.br/cbens/article/view/671>. Acesso em 21 de out 2020

UFRGS; REAMAT. Algebra Linear - Um Livro Colaborativo. **Método dos Mínimos Quadrados**. 2020. Disponível em [https://www.ufrgs.br/reatmat/AlgebraLinear/livro/s14-mx00e9todo\\_dos\\_mx00ednimos\\_quadrados.html](https://www.ufrgs.br/reatmat/AlgebraLinear/livro/s14-mx00e9todo_dos_mx00ednimos_quadrados.html). Acesso em 2 de fev de 2022.

CASTRO, R.M.G; **Energias Renováveis e Produção Descentralizada**. INTRODUÇÃO À ENERGIA FOTOVOLTAICA Edição 0 Novembro de 2002.

SANDIA NATIONAL LABORATORIES; **Irradiance & Insolation**. PVPerformance Modeling Colaborative. Disponível <https://pvpmc.sandia.gov/modeling-steps/1-weather-design-inputs/irradiance-and-insolation-2>. Acesso em 20 de mai 2021

BILLINGS, Stephen A. **Nonlinear System Identification: NARMAX Methods in the Time, frequency, and Spatio-Temporal Domains**. Chichester, Reino Unido: John Wiley & Sons, 2013.

DUARTE, Marcelo Machado. **Modelagem Paramétrica De Sistemas Dinâmicos Não-estacionários**. Dezembro de 2006, 134 pg. Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

BROCKWELL, Peter J.; DAVIS, Richard A. **Introduction to Time Series and Forecasting**. Segunda edição. Fort Collins, Estados Unidos: Springer-Verlag, 2002.  
ISERMANN, Rolf; MÜNCHHOF, Marco. **Identification of Dynamic Systems: An introduction with Applications**. Nova Iorque, Estados Unidos. Springer-Verlag, 2011.

YOUNG, Peter C. **Recursive Estimation and Time-Series Analysis**. Segunda edição. Nova Iorque, Estados Unidos. Springer-Verlag, 2011.

RACHAD, Sofia; et. Al. **System Identification of Inventory System Using ARX and ARMAX Models**. International Journal of Control and Automation, Casablanca, Marrocos. Vol.8, No.12, pp.283-294, 2015.

ENCYCLOPEDIA; **Chaos Theory And Meteorological Predictions**. Disponível <https://www.encyclopedia.com/environment>. Acesso em 20 de mai 2021

YALE; **Linear Regression**, 1997-98. Disponível <http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm>. Acesso em 20 de mai 2021

SOLARVIEW; **O que é irradiância solar?** Disponível <https://solarview.zendesk.com/hc>. Acesso em 20 de mai 2021

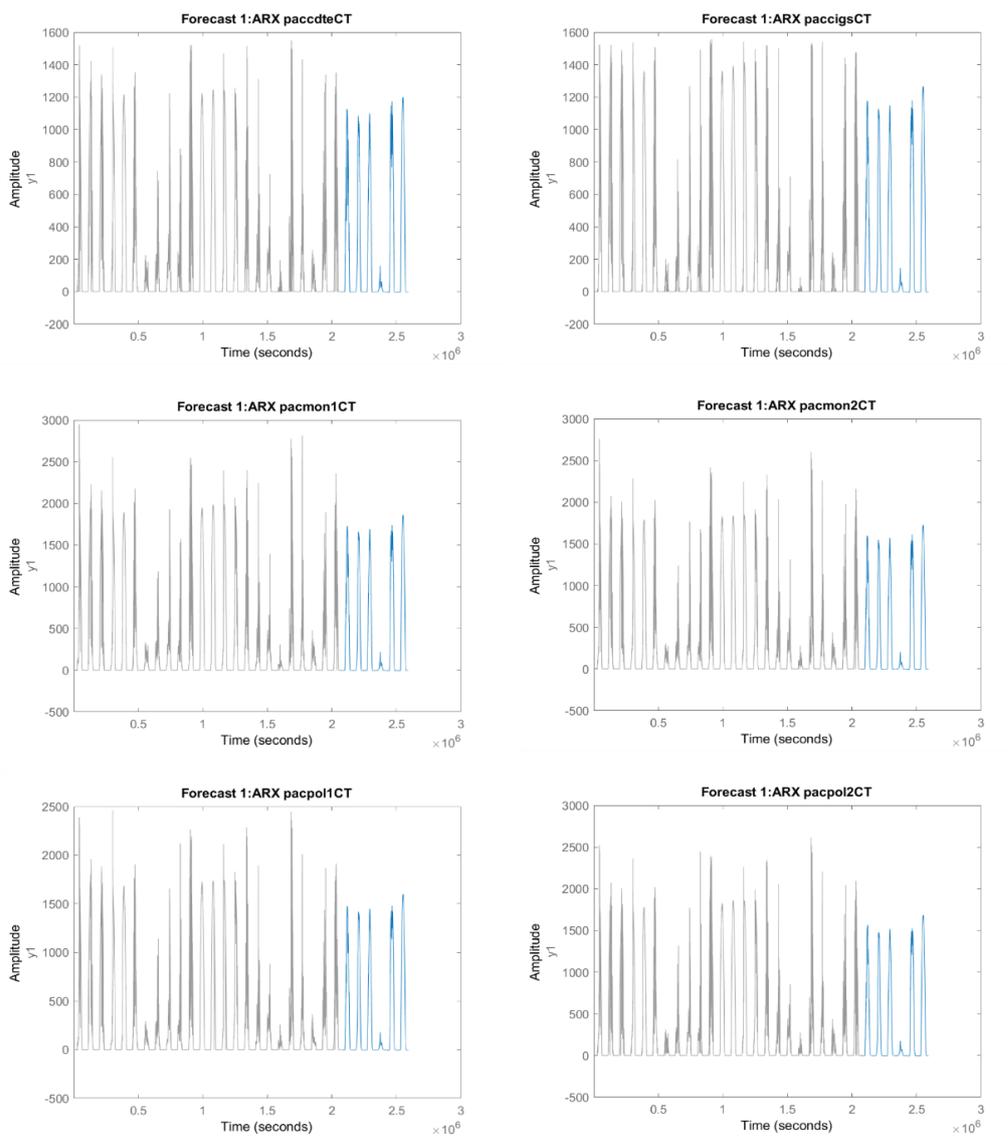
AGUIRRE, Luis. **Introdução à identificação de sistemas: Técnicas lineares e não-lineares aplicadas a sistemas reais**. Segunda edição. Editora UFMG, Minas Gerais, 2004.

HYNDMAN, Rob J; **Forecasting: Principles and Practice**. Segunda edição, editora Texts, 2018.

## ANEXO 1

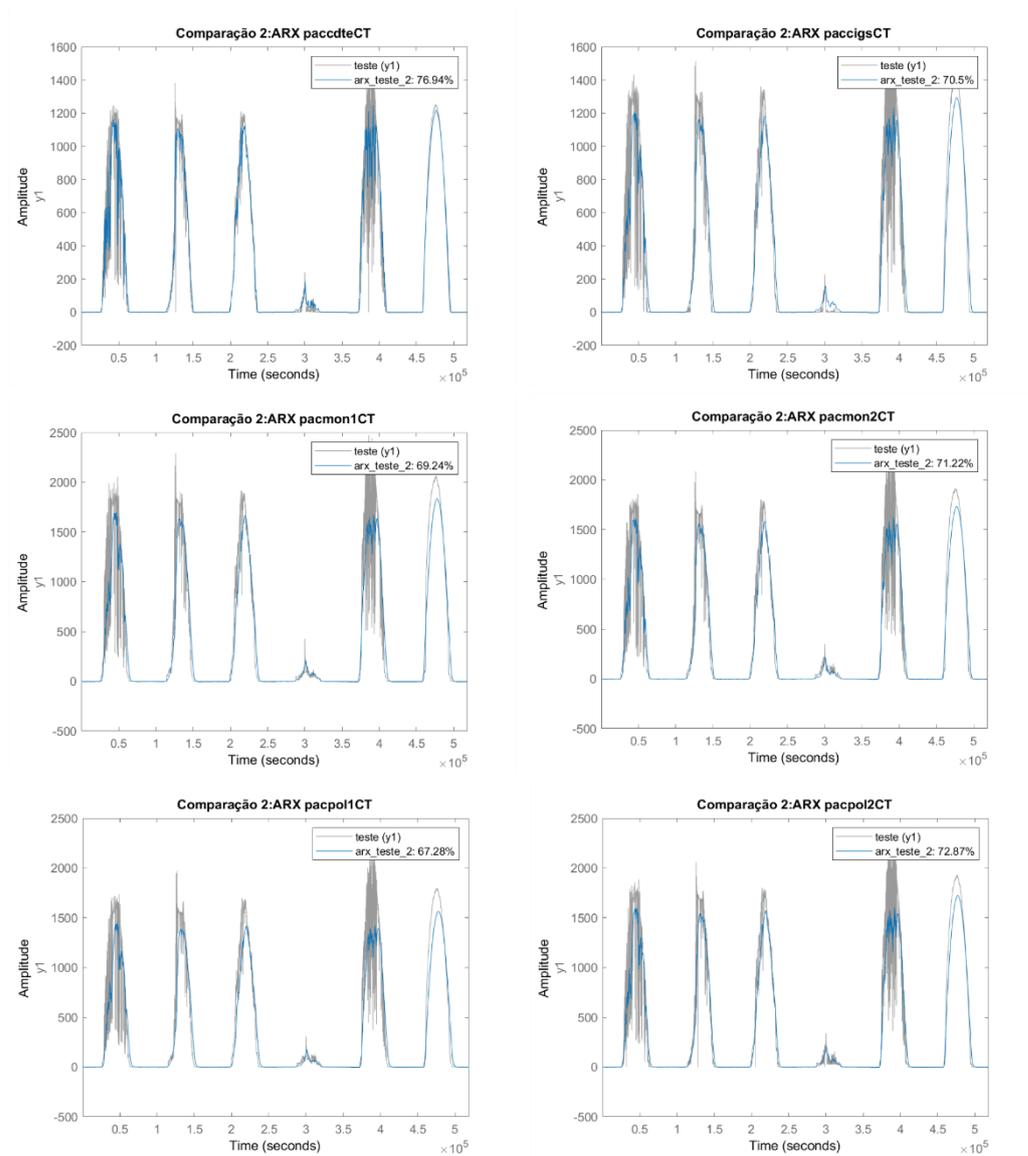
Este anexo contém os resultados obtidos através da modelagem ARX, tal qual há a comparação entre os modelos e dados reais, e a previsão em série temporal de cada um dos modelos. Em ambos casos, o eixo X possui o tempo em segundos. Nas comparações de modelos o eixo X possui a potência em cada uma das tecnologias, em cinza o medido e em azul o modelado. Na previsão de série temporal as potências são mostradas em Y, onde em cinza temos os dados reais e em azul os dados previstos.

### 1) FORECAST DO SISTEMA SISO REGRESSIVO



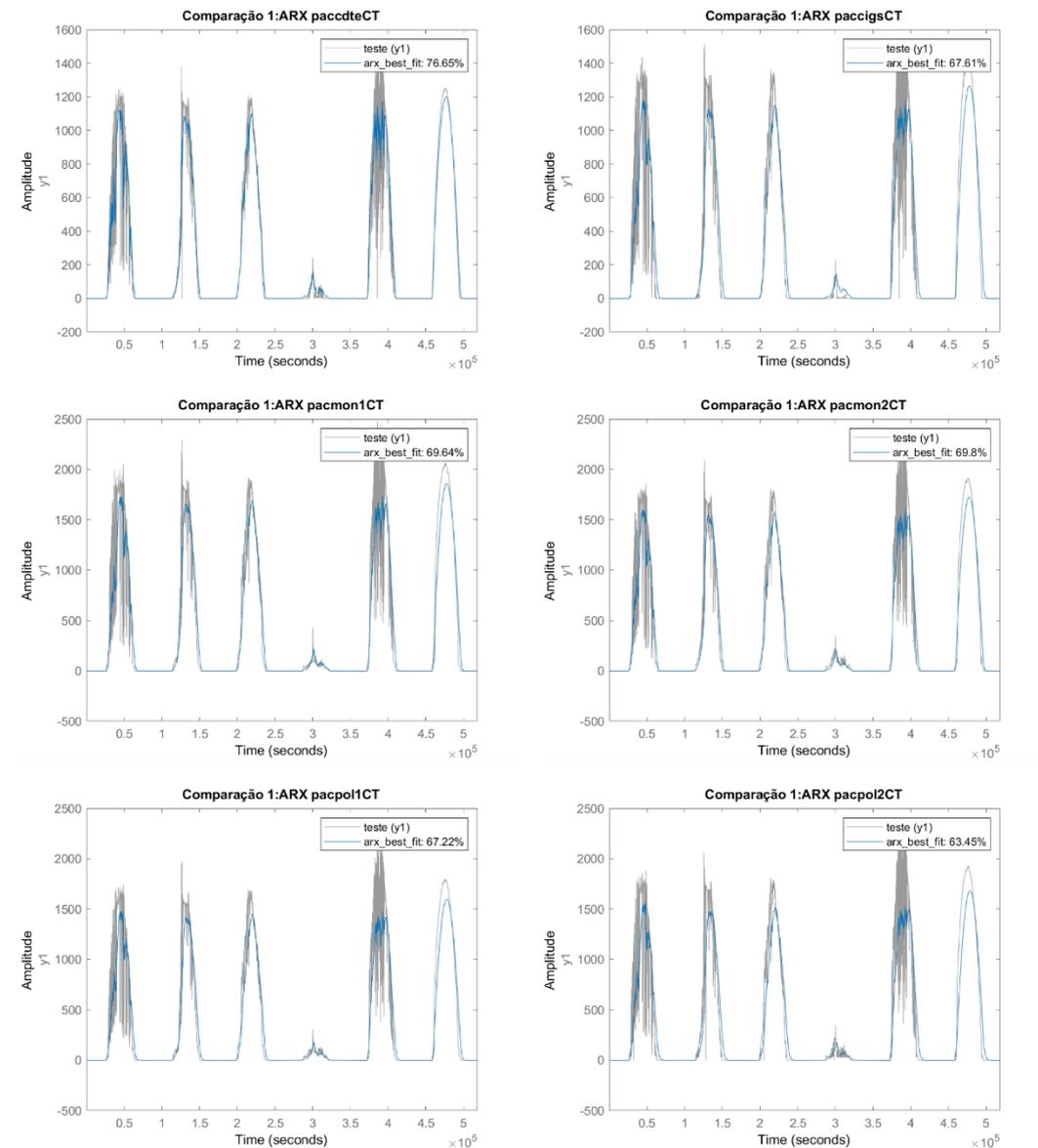
**Figura 77: ANEXO 1 Resultados - Forecast SISO Regressivo.**

## 2) COMPARATIVO: MODELO TOOLBOX (SISO) X DADOS REAIS



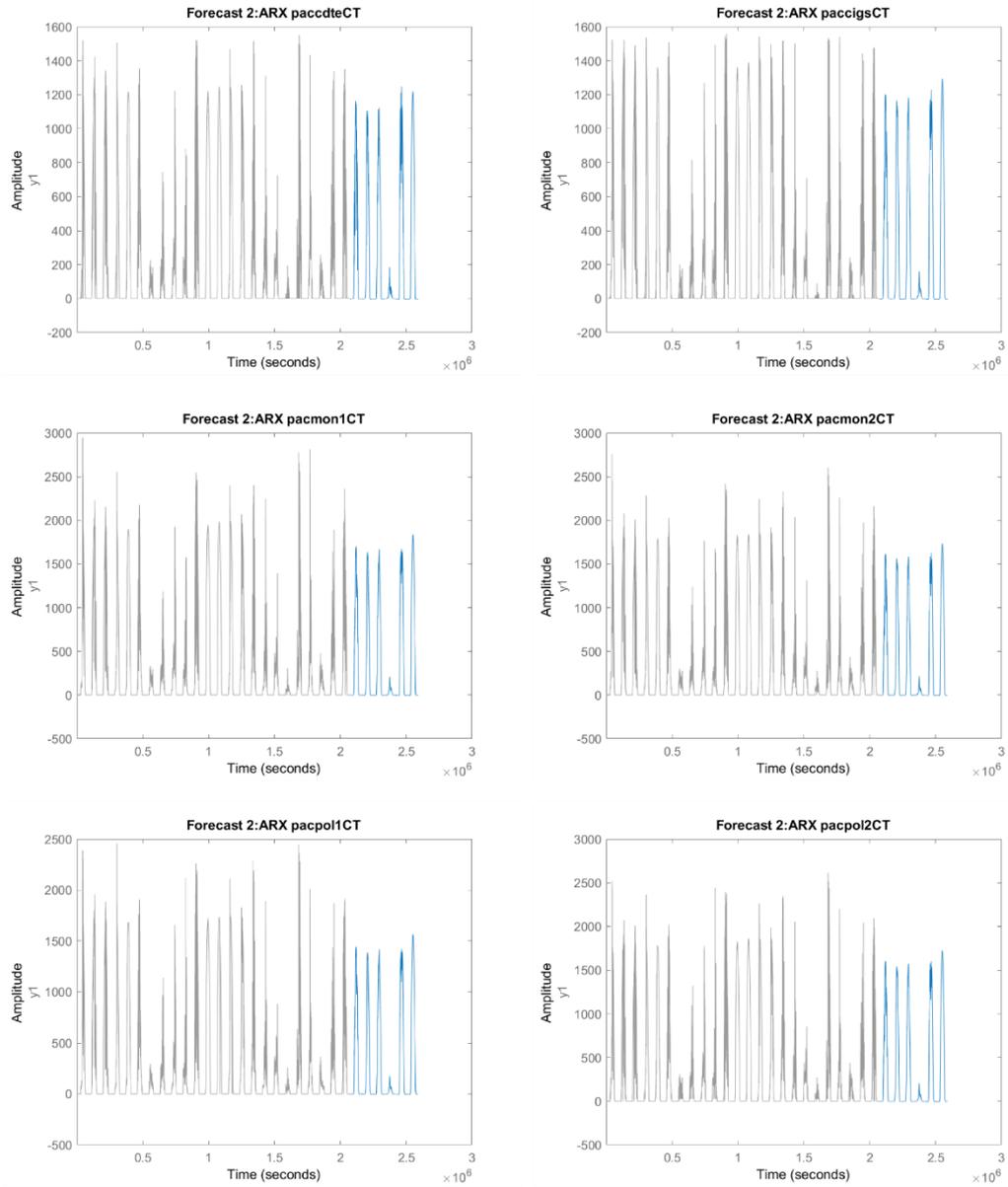
**Figura 78: ANEXO 1 Resultados – Modelo Toolbox MISO x Dados Reais.**

### 3) COMPARATIVO: MÉTODO REGRESSIVO (SISO) X DADOS REAIS



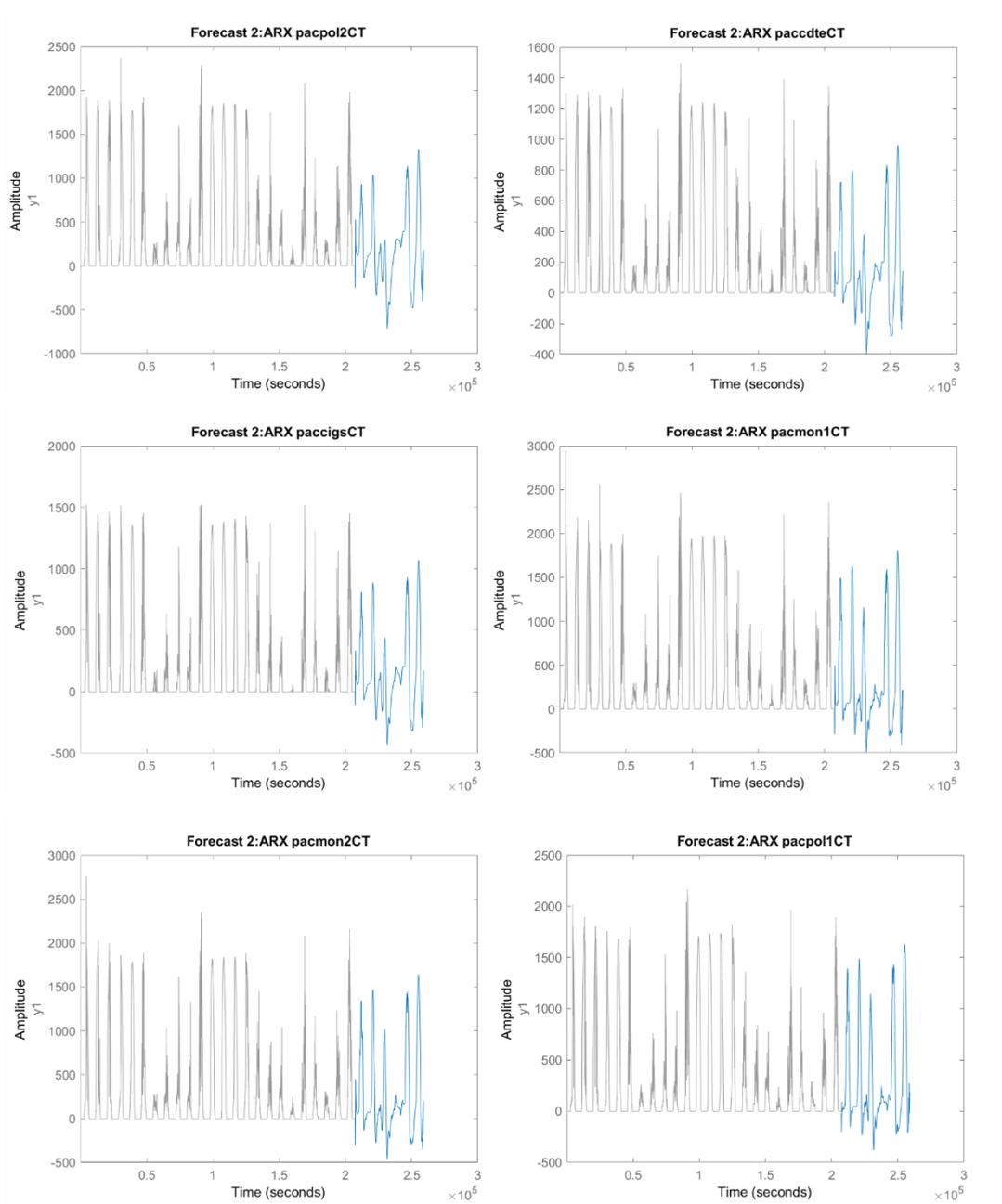
**Figura 79: ANEXO 1 Resultados – Método Regressivo SISO x Dados Reais.**

## 4) FORECAST DO SISTEMA SISO TOOLBOX



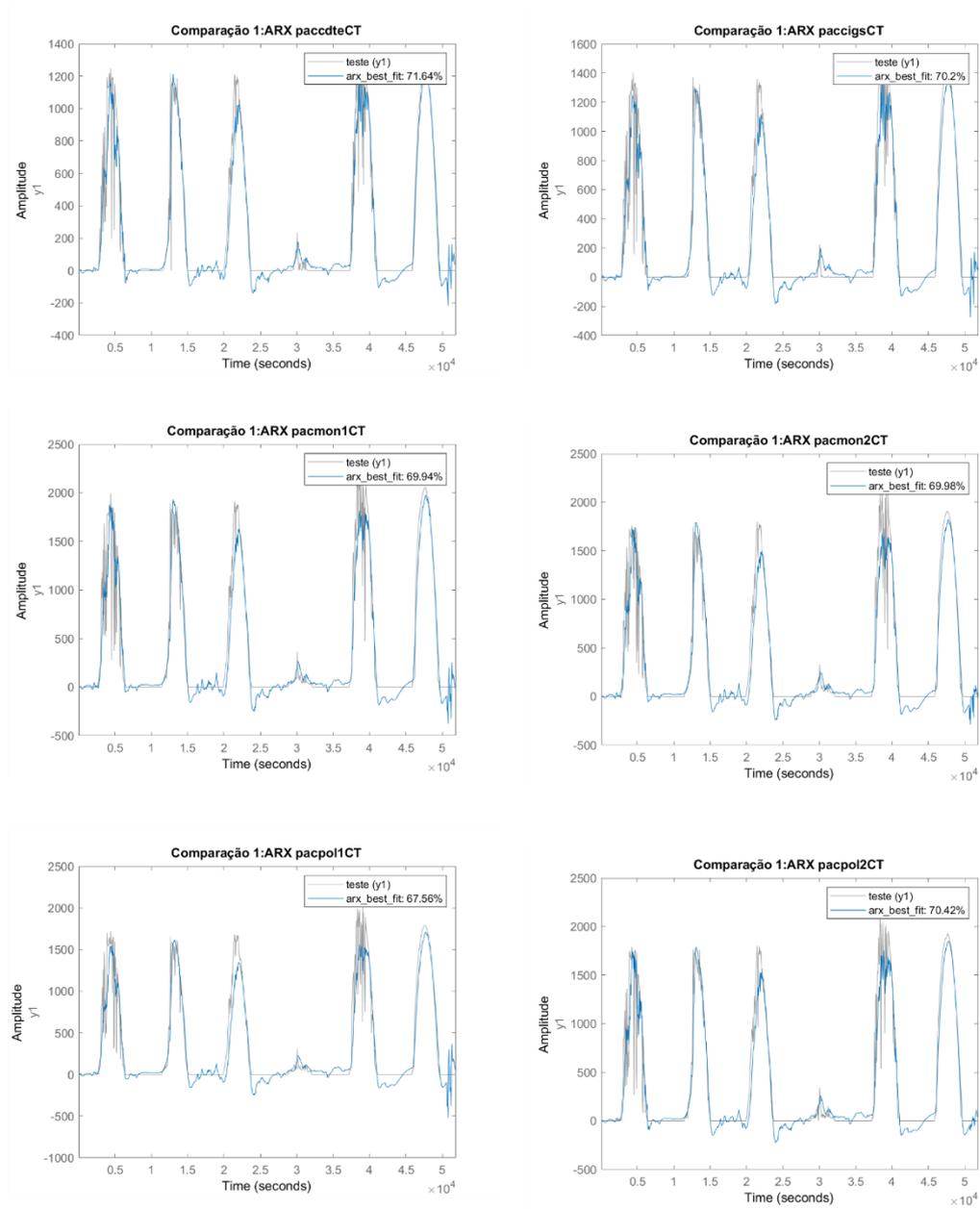
**Figura 80: ANEXO 1 Resultados - Forecast SISO Toolbox.**

## 5) FORECAST DO SISTEMA MISO TOOLBOX



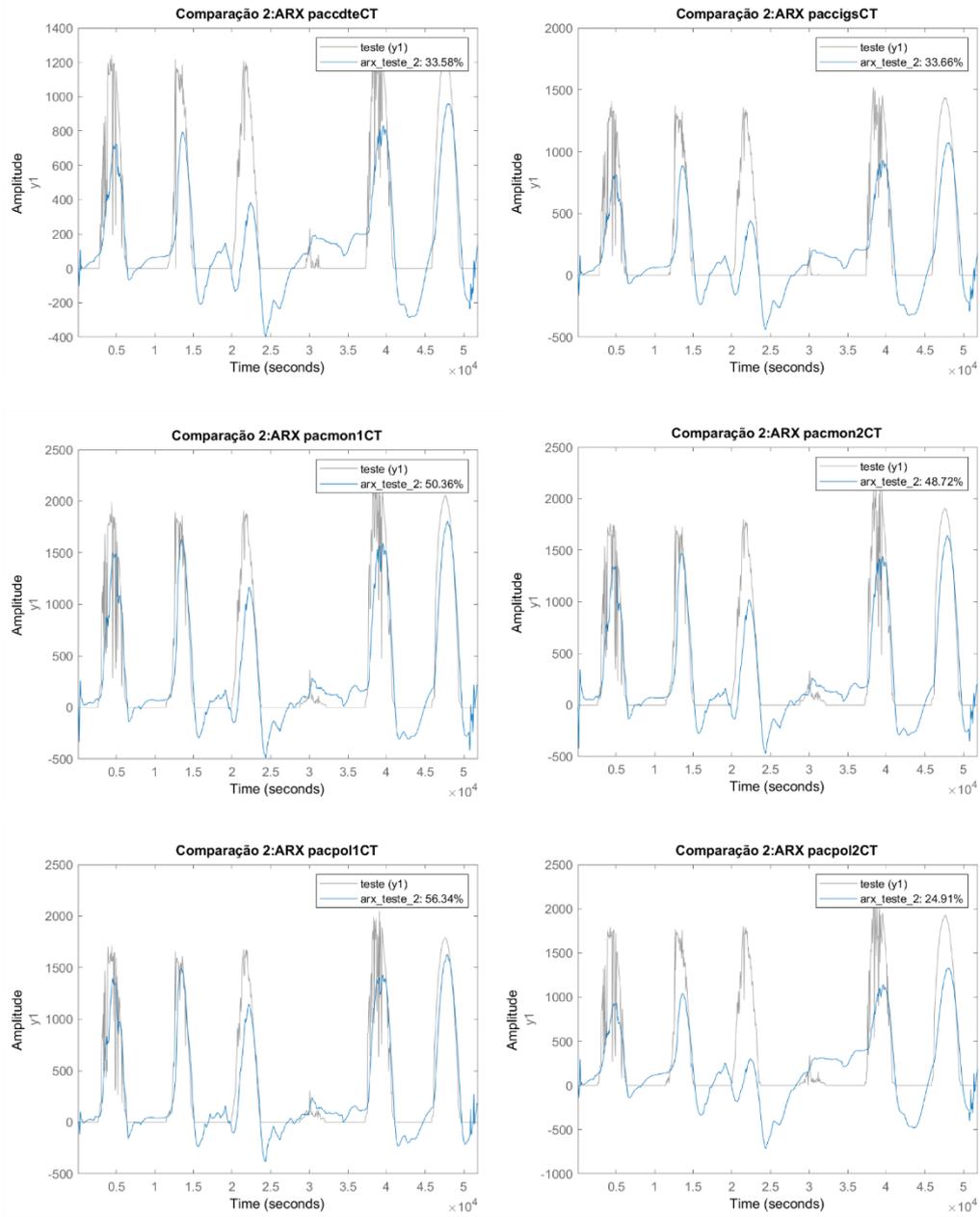
**Figura 81: ANEXO 1 Resultados - Forecast MISO Toolbox.**

## 6) COMPARATIVO: MÉTODO REGRESSIVO (MISO) X DADOS REAIS



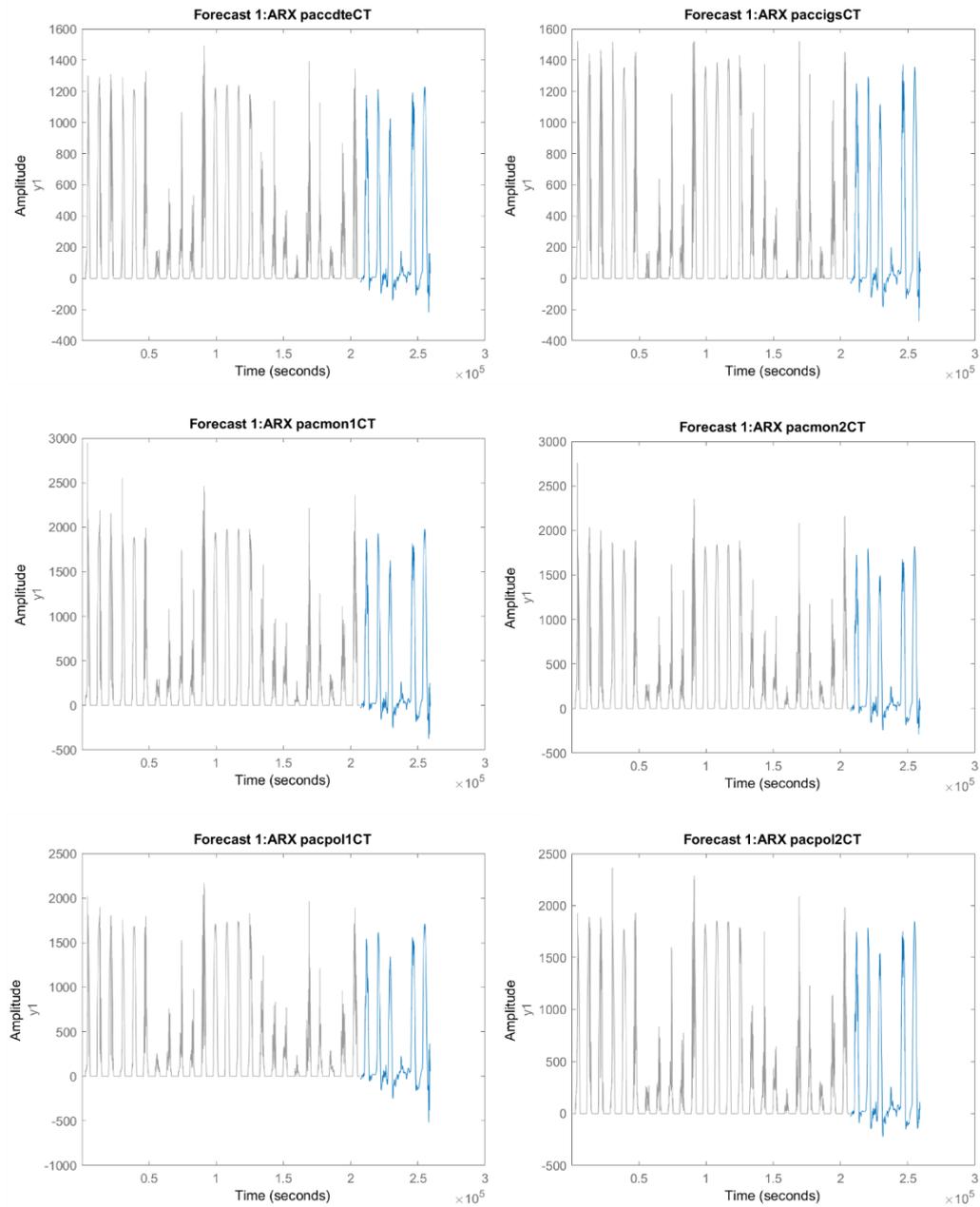
**Figura 82: ANEXO 1 Resultados – Método Regressivo MISO x Dados Reais.**

## 7) COMPARATIVO: MODELO TOOLBOX (MISO) X DADOS REAIS



**Figura 83: ANEXO 1 Resultados – Modelo Toolbox MISO x Dados Reais**

## 8) FORECAST DO SISTEMA MISO REGRESSIVO



**Figura 84: ANEXO 1 Resultados - Forecast MISO Regressivo.**

## ANEXO 2

Este anexo possui os modelos ARX gerados via Matlab. São mostrados os polinômios arbitrários A e B, assim como as configurações, o tempo de amostragem e o fit sobre o conjunto de dados de modelagem.

### 1) SISO REGRESSIVO

#### **arx\_paccdteCT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.5801 z^{-1} - 0.03323 z^{-2} - 0.01769 z^{-3} - 0.03967 z^{-4} - 0.04876 z^{-5} - 0.0674 z^{-6}$

$- 0.0203 z^{-7}$

$B(z) = 0.08224 z^{-1} + 0.06616 z^{-2} + 0.1232 z^{-3} - 0.05681 z^{-4} - 0.07372 z^{-5} + 0.06207 z^{-6}$

$+ 0.04 z^{-7}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=7 nb=7 nk=1

Number of free coefficients: 14

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 80.13% (prediction focus)

FPE: 4343, MSE: 4338

#### **arx\_paccigsCT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.6452 z^{-1} - 0.07762 z^{-2} - 0.07926 z^{-3} - 0.04487 z^{-4} - 0.008293 z^{-5} - 0.01539 z^{-6}$

$- 0.04674 z^{-7}$

$B(z) = 0.05563 z^{-7} + 0.02302 z^{-8} - 0.1013 z^{-9} + 0.1101 z^{-10} - 0.09073 z^{-11} + 0.07071 z^{-12}$

$+ 0.04308 z^{-13}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=7 nb=7 nk=7

Number of free coefficients: 14

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 80.26% (prediction focus)

FPE: 5370, MSE: 5364

#### **arx\_pacmon1CT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.659 z^{-1} - 0.05182 z^{-2} - 0.08063 z^{-3} - 0.05099 z^{-4} - 0.01361 z^{-5} - 0.03763 z^{-6}$

$- 0.01704 z^{-7}$

$B(z) = 0.1203 z^{-7} - 0.04997 z^{-8} - 0.07524 z^{-9} + 0.09034 z^{-10} - 0.04651 z^{-11} + 0.1565 z^{-12}$

$- 0.02012 z^{-13}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=7 nb=7 nk=7

Number of free coefficients: 14

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 80.57% (prediction focus)

FPE: 1.064e+04, MSE: 1.063e+04

**arx\_pacmon2CT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.6501 z^{-1} - 0.0363 z^{-2} - 0.08191 z^{-3} - 0.05191 z^{-4} - 0.03172 z^{-5} - 0.03983 z^{-6} - 0.0113 z^{-7}$

$B(z) = 0.1352 z^{-7} - 0.1144 z^{-8} + 0.06059 z^{-9} + 0.006869 z^{-10} - 0.04251 z^{-11} + 0.1661 z^{-12} - 0.03553 z^{-13}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=7 nb=7 nk=7

Number of free coefficients: 14

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 79.78% (prediction focus)

FPE: 9879, MSE: 9867

**arx\_pacpol1CT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.6566 z^{-1} - 0.02285 z^{-2} - 0.1054 z^{-3} - 0.04126 z^{-4} - 0.0183 z^{-5} - 0.04577 z^{-6} - 0.03079 z^{-7}$

$B(z) = 0.09812 z^{-7} - 0.1214 z^{-8} + 0.07354 z^{-9} + 0.03032 z^{-10} - 0.07683 z^{-11} + 0.1919 z^{-12} - 0.06204 z^{-13}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=7 nb=7 nk=7

Number of free coefficients: 14

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 80.1% (prediction focus)

FPE: 8532, MSE: 8522

**arx\_pacpol2CT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.6435 z^{-1} - 0.03482 z^{-2} - 0.08639 z^{-3} - 0.06557 z^{-4} - 0.006169 z^{-5} - 0.03973 z^{-6} - 0.05937 z^{-7}$

$B(z) = -0.03894 z^{-7} + 0.07109 z^{-8} - 0.07776 z^{-9} + 0.06667 z^{-10} - 0.05516 z^{-11} + 0.02499 z^{-12} + 0.1245 z^{-13}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=7 nb=7 nk=7

Number of free coefficients: 14

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 79.83% (prediction focus)

FPE: 1.011e+04, MSE: 1.01e+04

## 2) SISO TOOLBOX

### arx\_paccdteCT\_2 =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$   
 $A(z) = 1 - 0.5898 z^{-1} - 0.05161 z^{-2} - 0.04022 z^{-3}$   
 $B(z) = 0.06093 z^{-1} + 0.03179 z^{-2} + 0.1552 z^{-3} + 0.001832 z^{-4} + 0.009219 z^{-5} + 0.1023 z^{-6} + 0.04542 z^{-7}$   
 Sample time: 60 seconds  
 Parameterization:  
 Polynomial orders: na=3 nb=7 nk=1  
 Number of free coefficients: 10  
 Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.  
 Status:  
 Estimated using ARX on time domain data "modelo".  
 Fit to estimation data: 80.08% (prediction focus)  
 FPE: 4365, MSE: 4361

### arx\_paccigsCT\_2 =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$   
 $A(z) = 1 - 0.6455 z^{-1} - 0.078 z^{-2} - 0.08042 z^{-3} - 0.04586 z^{-4} - 0.009399 z^{-5} - 0.03298 z^{-6}$   
 $B(z) = 0.1022 z^{-7} + 0.009553 z^{-8} - 0.09669 z^{-9} + 0.1111 z^{-10} - 0.0916 z^{-11} + 0.06806 z^{-12} + 0.0443 z^{-13}$   
 Sample time: 60 seconds  
 Parameterization:  
 Polynomial orders: na=6 nb=7 nk=7  
 Number of free coefficients: 13  
 Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.  
 Status:  
 Estimated using ARX on time domain data "modelo".  
 Fit to estimation data: 80.24% (prediction focus)  
 FPE: 5377, MSE: 5371

### arx\_pacmon1CT\_2 =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$   
 $A(z) = 1 - 0.6616 z^{-1} - 0.05285 z^{-2} - 0.0791 z^{-3} - 0.05113 z^{-4} - 0.01621 z^{-5} - 0.03972 z^{-6} - 0.0174 z^{-7}$   
 $B(z) = 0.1267 z^{-7} - 0.03702 z^{-8} - 0.06829 z^{-9} + 0.1376 z^{-10}$   
 Sample time: 60 seconds  
 Parameterization:  
 Polynomial orders: na=7 nb=4 nk=7  
 Number of free coefficients: 11  
 Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.  
 Status:  
 Estimated using ARX on time domain data "modelo".  
 Fit to estimation data: 80.54% (prediction focus)  
 FPE: 1.068e+04, MSE: 1.067e+04

**arx\_pacmon2CT\_2 =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.6544 z^{-1} - 0.03887 z^{-2} - 0.08531 z^{-3} - 0.05171 z^{-4} - 0.06 z^{-5}$

$B(z) = 0.1774 z^{-7} - 0.1072 z^{-8} + 0.07286 z^{-9} + 0.05706 z^{-10}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=4 nk=7

Number of free coefficients: 9

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 79.72% (prediction focus)

FPE: 9933, MSE: 9925

**arx\_pacpol1CT\_2 =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.6593 z^{-1} - 0.02472 z^{-2} - 0.1087 z^{-3} - 0.0442 z^{-4} - 0.02261 z^{-5} - 0.06699 z^{-6}$

$B(z) = 0.1216 z^{-7}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=6 nb=1 nk=7

Number of free coefficients: 7

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 80.01% (prediction focus)

FPE: 8600, MSE: 8594

**arx\_pacpol2CT\_2 =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$A(z) = 1 - 0.6142 z^{-1} + 0.002154 z^{-2} - 0.0934 z^{-3} - 0.07735 z^{-4} - 0.01358 z^{-5} - 0.09956 z^{-6}$

$B(z) = 0.1893 z^{-1}$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=6 nb=1 nk=1

Number of free coefficients: 7

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 79.7% (prediction focus)

FPE: 1.023e+04, MSE: 1.023e+04

### 3) MISO REGRESSIVO

#### arx\_paccdteCT =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.2535 z^{-1} - 0.1404 z^{-2} + 0.03925 z^{-3} - 0.1514 z^{-4} + 0.152 z^{-5}$$

$$B1(z) = 1.137 z^{-1} + 3.823 z^{-2} - 6.019 z^{-3} + 3.112 z^{-4} - 0.05911 z^{-5}$$

$$B2(z) = -0.04788 z^{-1} - 0.3696 z^{-2} + 0.7289 z^{-3} - 0.4467 z^{-4} - 0.03244 z^{-5}$$

$$B3(z) = -4.38 z^{-1} - 2.491 z^{-2} - 9.829 z^{-3} + 4.014 z^{-4} - 10.62 z^{-5}$$

$$B4(z) = -2.684 z^{-1} + 15.43 z^{-2} - 3.153 z^{-3} + 10.59 z^{-4} - 21.92 z^{-5}$$

$$B5(z) = 0.05069 z^{-1} - 0.000999 z^{-2} - 0.03472 z^{-3} + 0.01571 z^{-4} - 0.07859 z^{-5}$$

$$B6(z) = 0.4243 z^{-1} + 0.1405 z^{-2} + 0.1414 z^{-3} - 0.1118 z^{-4} + 0.2847 z^{-5}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=[5 5 5 5 5] nk=[1 1 1 1 1]

Number of free coefficients: 35

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 69.16% (prediction focus)

FPE: 1.062e+04, MSE: 1.038e+04

#### arx\_paccigsCT =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.3094 z^{-1} - 0.05669 z^{-2} - 0.07325 z^{-3} - 0.155 z^{-4} + 0.1363 z^{-5}$$

$$B1(z) = 1.618 z^{-1} + 3.648 z^{-2} - 5.949 z^{-3} + 2.957 z^{-4} + 0.02755 z^{-5}$$

$$B2(z) = -0.08053 z^{-1} - 0.5433 z^{-2} + 1.03 z^{-3} - 0.5896 z^{-4} - 0.01196 z^{-5}$$

$$B3(z) = -2.623 z^{-1} - 4.405 z^{-2} - 9.569 z^{-3} - 4.238 z^{-4} - 12.83 z^{-5}$$

$$B4(z) = -5.08 z^{-1} + 33.77 z^{-2} - 34.07 z^{-3} + 27.01 z^{-4} - 23.74 z^{-5}$$

$$B5(z) = 0.05155 z^{-1} - 0.01053 z^{-2} - 0.02788 z^{-3} + 0.04572 z^{-4} - 0.1158 z^{-5}$$

$$B6(z) = 0.3719 z^{-1} + 0.287 z^{-2} - 0.0126 z^{-3} - 0.1242 z^{-4} + 0.3128 z^{-5}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=[5 5 5 5 5] nk=[1 1 1 1 1]

Number of free coefficients: 35

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 67.5% (prediction focus)

FPE: 1.488e+04, MSE: 1.454e+04

#### arx\_pacmon1CT =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.2725 z^{-1} - 0.2301 z^{-2} + 0.09363 z^{-3} - 0.07256 z^{-4} + 0.07459 z^{-5}$$

$$B1(z) = 1.516 z^{-1} + 5.1 z^{-2} - 4.201 z^{-3} + 3.108 z^{-4} - 1.944 z^{-5}$$

$$B2(z) = 0.1511 z^{-1} - 1.039 z^{-2} + 1.031 z^{-3} - 0.5047 z^{-4} + 0.01954 z^{-5}$$

$$B3(z) = -5.541 z^{-1} - 2.783 z^{-2} - 7.175 z^{-3} - 8.362 z^{-4} - 5.815 z^{-5}$$

$$B4(z) = -19.11 z^{-1} + 84.76 z^{-2} - 53.56 z^{-3} + 10.93 z^{-4} - 25.26 z^{-5}$$

$$B5(z) = 0.1 z^{-1} + 0.005222 z^{-2} - 0.01822 z^{-3} + 0.07545 z^{-4} - 0.1587 z^{-5}$$

$$B6(z) = 0.614 z^{-1} - 0.02098 z^{-2} + 0.2832 z^{-3} + 0.08409 z^{-4} + 0.3556 z^{-5}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=[5 5 5 5 5] nk=[1 1 1 1 1]

Number of free coefficients: 35

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 68.49% (prediction focus)

FPE: 2.884e+04, MSE: 2.818e+04

**arx\_pacmon2CT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.2962 z^{-1} - 0.2261 z^{-2} + 0.06245 z^{-3} - 0.08767 z^{-4} + 0.115 z^{-5}$$

$$B1(z) = 0.7286 z^{-1} + 5.823 z^{-2} - 3.938 z^{-3} + 1.188 z^{-4} - 0.4124 z^{-5}$$

$$B2(z) = 0.2193 z^{-1} - 1.171 z^{-2} + 0.8989 z^{-3} - 0.3828 z^{-4} + 0.1059 z^{-5}$$

$$B3(z) = -5.297 z^{-1} - 0.9368 z^{-2} - 8.471 z^{-3} - 3.689 z^{-4} - 1.607 z^{-5}$$

$$B4(z) = -16.71 z^{-1} + 71.46 z^{-2} - 46.82 z^{-3} + 12.57 z^{-4} - 22.59 z^{-5}$$

$$B5(z) = 0.09442 z^{-1} + 0.03322 z^{-2} - 0.03124 z^{-3} + 0.06535 z^{-4} - 0.1428 z^{-5}$$

$$B6(z) = 0.5441 z^{-1} - 0.04385 z^{-2} + 0.2489 z^{-3} + 0.04112 z^{-4} + 0.3779 z^{-5}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=[5 5 5 5 5] nk=[1 1 1 1 1]

Number of free coefficients: 35

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 68.42% (prediction focus)

FPE: 2.471e+04, MSE: 2.415e+04

**arx\_pacpol1CT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.3539 z^{-1} - 0.1994 z^{-2} + 0.03766 z^{-3} - 0.06644 z^{-4} + 0.03975 z^{-5}$$

$$B1(z) = 2.543 z^{-1} + 5.558 z^{-2} - 7.606 z^{-3} + 3.17 z^{-4} - 0.7875 z^{-5}$$

$$B2(z) = 0.1191 z^{-1} - 0.836 z^{-2} + 1.366 z^{-3} - 0.6964 z^{-4} - 0.2007 z^{-5}$$

$$B3(z) = -4.151 z^{-1} - 1.558 z^{-2} - 9.172 z^{-3} - 4.046 z^{-4} - 4.901 z^{-5}$$

$$B4(z) = -8.99 z^{-1} + 63.76 z^{-2} - 43.49 z^{-3} + 5.686 z^{-4} - 19.29 z^{-5}$$

$$B5(z) = 0.132 z^{-1} - 0.05494 z^{-2} - 0.1134 z^{-3} + 0.04349 z^{-4} - 0.06142 z^{-5}$$

$$B6(z) = 0.4579 z^{-1} + 0.02226 z^{-2} + 0.1528 z^{-3} + 0.06172 z^{-4} + 0.1979 z^{-5}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=[5 5 5 5 5] nk=[1 1 1 1 1]

Number of free coefficients: 35

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 69.35% (prediction focus)

FPE: 2.066e+04, MSE: 2.019e+04

**arx\_pacpol2CT =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.2652 z^{-1} - 0.2212 z^{-2} + 0.05967 z^{-3} - 0.1808 z^{-4} + 0.1616 z^{-5}$$

$$B1(z) = 4.873 z^{-1} + 1.113 z^{-2} - 4.142 z^{-3} + 0.2493 z^{-4} + 0.7757 z^{-5}$$

$$B2(z) = -0.6316 z^{-1} - 0.08953 z^{-2} + 0.3496 z^{-3} - 0.07053 z^{-4} + 0.1927 z^{-5}$$

$$B3(z) = -3.269 z^{-1} - 4.449 z^{-2} - 18.03 z^{-3} + 3.81 z^{-4} - 5.821 z^{-5}$$

$$B4(z) = 14.71 z^{-1} + 2.148 z^{-2} + 9.919 z^{-3} - 2.162 z^{-4} - 27.22 z^{-5}$$

$$B5(z) = 0.1263 z^{-1} - 0.05941 z^{-2} - 0.09647 z^{-3} + 0.08904 z^{-4} - 0.0862 z^{-5}$$

$$B6(z) = 0.5492 z^{-1} + 0.05423 z^{-2} + 0.3427 z^{-3} - 0.2531 z^{-4} + 0.4453 z^{-5}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=[5 5 5 5 5] nk=[1 1 1 1 1]

Number of free coefficients: 35

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 69.89% (prediction focus)

FPE: 2.289e+04, MSE: 2.236e+04

## 4) MISO TOOLBOX

### arx\_paccdteCT\_2 =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$   
 $A(z) = 1 - 0.511 z^{-1} - 0.2839 z^{-2} - 0.08658 z^{-3}$   
 $B1(z) = 0.6382 z^{-3} + 0.2894 z^{-4} + 1.029 z^{-5}$   
 $B2(z) = -0.3763 z^{-5} + 0.08297 z^{-6} + 0.1361 z^{-7} + 0.002682 z^{-8}$   
 $B3(z) = -12.3 z^{-3} + 3.212 z^{-4} - 10.46 z^{-5}$   
 $B4(z) = 17.62 z^{-5} - 16.15 z^{-6} - 3.134 z^{-7}$   
 $B5(z) = -0.008751 z^{-5} - 0.005701 z^{-6} - 0.05317 z^{-7}$   
 $B6(z) = 0.1473 z^{-5} + 0.05849 z^{-6} - 0.02454 z^{-7}$   
Sample time: 60 seconds  
Parameterization:  
Polynomial orders: na=3 nb=[3 4 3 3 3 3] nk=[3 5 3 5 5 5]  
Number of free coefficients: 22  
Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.  
Status:  
Estimated using ARX on time domain data "modelo".  
Fit to estimation data: 68.17% (prediction focus)  
FPE: 1.122e+04, MSE: 1.106e+04

### arx\_paccigsCT\_2 =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$   
 $A(z) = 1 - 0.511 z^{-1} - 0.2331 z^{-2} - 0.1294 z^{-3}$   
 $B1(z) = 1.01 z^{-3} + 0.2862 z^{-4} + 0.9828 z^{-5}$   
 $B2(z) = -0.3709 z^{-5} + 0.03631 z^{-6} + 0.1435 z^{-7} + 0.0113 z^{-8}$   
 $B3(z) = -10.98 z^{-3} - 6.442 z^{-4} - 8.72 z^{-5}$   
 $B4(z) = 19.98 z^{-5} - 17.3 z^{-6} - 4.59 z^{-7}$   
 $B5(z) = -0.01938 z^{-5} - 0.00496 z^{-6} - 0.05961 z^{-7}$   
 $B6(z) = 0.1845 z^{-5} + 0.03657 z^{-6} - 0.004915 z^{-7}$   
Sample time: 60 seconds  
Parameterization:  
Polynomial orders: na=3 nb=[3 4 3 3 3 3] nk=[3 5 3 5 5 5]  
Number of free coefficients: 22  
Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.  
Status:  
Estimated using ARX on time domain data "modelo".  
Fit to estimation data: 66.66% (prediction focus)  
FPE: 1.554e+04, MSE: 1.531e+04

### arx\_pacmon1CT\_2 =

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$   
 $A(z) = 1 - 0.4746 z^{-1} - 0.2927 z^{-2} - 0.01969 z^{-3}$   
 $B1(z) = 2.318 z^{-3} + 0.2886 z^{-4} + 0.6264 z^{-5}$   
 $B2(z) = -0.6993 z^{-5} + 0.06635 z^{-6} + 0.34 z^{-7} + 0.03161 z^{-8}$   
 $B3(z) = -16.56 z^{-4} + 0.7352 z^{-5} - 14.62 z^{-6}$   
 $B4(z) = 42.33 z^{-5} - 23.27 z^{-6} - 21.66 z^{-7}$   
 $B5(z) = -0.09402 z^{-5} - 0.04674 z^{-6} + 0.04321 z^{-7}$   
 $B6(z) = 0.3036 z^{-4} + 0.224 z^{-5} - 0.04191 z^{-6}$   
Sample time: 60 seconds  
Parameterization:  
Polynomial orders: na=3 nb=[3 4 3 3 3 3] nk=[3 5 4 5 5 4]  
Number of free coefficients: 22  
Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.  
Status:  
Estimated using ARX on time domain data "modelo".  
Fit to estimation data: 67.66% (prediction focus)  
FPE: 3.014e+04, MSE: 2.969e+04

**arx\_pacmon2CT\_2 =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.4808 z^{-1} - 0.285 z^{-2} - 0.03611 z^{-3}$$

$$B1(z) = 1.637 z^{-3} + 0.08231 z^{-4} + 1.227 z^{-5}$$

$$B2(z) = -0.4989 z^{-5} - 0.008398 z^{-6} + 0.2675 z^{-7}$$

$$B3(z) = -12.79 z^{-4} + 2.026 z^{-5} - 13.24 z^{-6}$$

$$B4(z) = 36.03 z^{-5} - 17.89 z^{-6} - 20.54 z^{-7}$$

$$B5(z) = -0.09088 z^{-5} - 0.04565 z^{-6} + 0.05471 z^{-7}$$

$$B6(z) = 0.2421 z^{-4} + 0.2169 z^{-5} - 0.04295 z^{-6}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=3 nb=[3 3 3 3 3 3] nk=[3 5 4 5 5 4]

Number of free coefficients: 21

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 67.58% (prediction focus)

FPE: 2.58e+04, MSE: 2.545e+04

**arx\_pacpol1CT\_2 =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.5305 z^{-1} - 0.2417 z^{-2} + 0.04488 z^{-3} - 0.05943 z^{-4} + 0.03741 z^{-5}$$

$$B1(z) = 3.559 z^{-5} + 0.8361 z^{-6} - 1.981 z^{-7}$$

$$B2(z) = -0.9298 z^{-5} + 0.339 z^{-6} + 0.4097 z^{-7}$$

$$B3(z) = -2.032 z^{-5} - 10.64 z^{-6} - 10.68 z^{-7}$$

$$B4(z) = 41.95 z^{-5} - 31.85 z^{-6} - 12.42 z^{-7}$$

$$B5(z) = -0.01463 z^{-5} - 0.02171 z^{-6} - 0.05952 z^{-7}$$

$$B6(z) = 0.2436 z^{-3} + 0.09897 z^{-4} + 0.1367 z^{-5}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=5 nb=[3 3 3 3 3 3] nk=[5 5 5 5 5 3]

Number of free coefficients: 23

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 68.81% (prediction focus)

FPE: 2.125e+04, MSE: 2.09e+04

**arx\_pacpol2CT\_2 =**

Discrete-time ARX model:  $A(z)y(t) = B(z)u(t) + e(t)$

$$A(z) = 1 - 0.5069 z^{-1} - 0.2995 z^{-2} - 0.1057 z^{-3}$$

$$B1(z) = 1.473 z^{-3} - 0.3534 z^{-4} + 1.631 z^{-5}$$

$$B2(z) = -0.3604 z^{-5} - 0.1872 z^{-6} + 0.1999 z^{-7} + 0.1338 z^{-8}$$

$$B3(z) = -6.454 z^{-4} - 2.312 z^{-5} - 14.66 z^{-6}$$

$$B4(z) = 24 z^{-5} - 14.96 z^{-6} - 11.53 z^{-7}$$

$$B5(z) = 0.03241 z^{-5} + 0.01163 z^{-6} - 0.1389 z^{-7}$$

$$B6(z) = 0.1772 z^{-5} + 0.03462 z^{-6} - 0.004123 z^{-7}$$

Sample time: 60 seconds

Parameterization:

Polynomial orders: na=3 nb=[3 4 3 3 3 3] nk=[3 5 4 5 5 5]

Number of free coefficients: 22

Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:

Estimated using ARX on time domain data "modelo".

Fit to estimation data: 69.28% (prediction focus)

FPE: 2.363e+04, MSE: 2.327e+04