



FUNDAÇÃO UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL

FACULDADE DE ENGENHARIAS, ARQUITETURA E URBANISMO E GEOGRAFIA

PROGRAMA DE PÓS-GRADUAÇÃO EM EFICIÊNCIA ENERGÉTICA E SUSTENTABILIDADE

CURSO DE MESTRADO PROFISSIONAL EM EFICIÊNCIA ENERGÉTICA E SUSTENTABILIDADE

# **ANÁLISE DO FATOR DE SOBREDIMENSIONAMENTO PARA SISTEMAS FOTOVOLTAICOS COM O USO DE REDE NEURAL ARTIFICIAL**

**GABRIEL EDGAR HERMANN**

**2024**

**CAMPO GRANDE/MS**

**FUNDAÇÃO UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL**

**FACULDADE DE ENGENHARIAS, ARQUITETURA E URBANISMO E GEOGRAFIA**

**MESTRADO PROFISSIONAL EM EFICIÊNCIA ENERGÉTICA E SUSTENTABILIDADE**

**ANÁLISE DO FATOR DE SOBREDIMENSIONAMENTO PARA SISTEMAS  
FOTOVOLTAICOS COM O USO DE REDE NEURAL ARTIFICIAL**

**GABRIEL EDGAR HERMANN**

Dissertação apresentada ao Mestrado Profissional da Faculdade de Engenharias, Arquitetura e Urbanismo e Geografia da Universidade Federal de Mato Grosso do Sul, para obtenção do título de Mestre em Eficiência Energética e Sustentabilidade, na área de concentração Eficiência Energética.

**Orientadora: Profa. Dra. Andréa Teresa Riccio Barbosa**

CAMPO GRANDE  
Junho/2024



Serviço Público Federal  
Ministério da Educação  
Fundação Universidade Federal de Mato Grosso do Sul



## FOLHA DE APROVAÇÃO

**Gabriel Edgar Hermann**

### **Análise de fator de sobre dimensionamento em sistemas fotovoltaicos com o uso de rede neural artificial**

Redação final do Trabalho de Conclusão de Curso, aprovado pela Banca Examinadora em 19 de junho de 2024, na Faculdade de Engenharias, Arquitetura e Urbanismo e Geografia da Universidade Federal de Mato Grosso do Sul para obtenção do título de Mestre em Eficiência Energética e Sustentabilidade.

#### **Banca examinadora:**

Dra. Andrea Teresa Riccio Barbosa (Presidente - Interno)

Dr. Frederico Silva Moreira (Interno) (Suplente)

Dr. Flavio Aristone (Externo)

Dr. Jair de Jesus Fiorentino (Externo) (Suplente)

Dr. Ramon Eduardo Pereira Silva (Interno)

NOTA  
MÁXIMA  
NO MEC

UFMS  
É 10!!!



Documento assinado eletronicamente por **Andrea Teresa Riccio Barbosa, Coordenador(a) de Curso de Pós-graduação**, em 22/07/2024, às 17:06, conforme horário oficial de Mato Grosso do Sul, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Flavio Aristone, Professor do Magisterio Superior**, em 24/07/2024, às 16:55, conforme horário oficial de Mato Grosso do Sul, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ramon Eduardo Pereira Silva, Professor Permanente**, em 25/07/2024, às 14:37, conforme horário oficial de Mato Grosso do Sul, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufms.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufms.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4976564** e o código CRC **38D2BB6E**.

## COLEGIADO DE PÓS-GRADUAÇÃO EM EFICIÊNCIA ENERGÉTICA E SUSTENTABILIDADE

Av Costa e Silva, s/nº - Cidade Universitária

Fone:

CEP 79070-900 - Campo Grande - MS

Referência: Processo nº 23104.016210/2024-32

SEI nº 4976564

## RESUMO

Sistemas de geração fotovoltaica são hoje uma alternativa para quem quer investir em gerar energia limpa. Sendo assim, é importante que cada vez mais seja discutido, em meio acadêmico e profissional, como fazer o correto dimensionamento desses sistemas, visando uma análise que se aproxime da real geração pretendida. Não há consenso no meio acadêmico sobre como a diminuição da visibilidade solar em períodos de estiagem reduz a energia gerada em painéis instalados em áreas rurais e urbanas. Uma revisão bibliográfica foi feita acerca do tema proposto, e estudos de diversos lugares do mundo foram agrupados nesse trabalho elucidando o problema. Sendo assim, este estudo apresenta como determinar o fator de sobredimensionamento para correção da potência de geração nos projetos de sistemas fotovoltaicos, em regiões de longos períodos de estiagem, além de determinar os fatores (parâmetros) que influenciam na geração fotovoltaica nesses locais considerados críticos. Para isso, foram coletados os dados de geração de dois sistemas fotovoltaicos, um instalado em uma área urbana em Campo Grande/MS e um na zona rural na cidade de Bela Vista/MS, que foram posteriormente compilados em conjunto com dados meteorológicos coletados do INMET. Esses dados foram implementados em uma Rede Neural Artificial (RNA) do tipo *MultiLayer Perceptron* (MLP), com auxílio do simulador WEKA e, como resultado esperou-se obter com o menor erro possível, o fator de correção em projeto para os sistemas instalados em regiões com períodos de estiagem, além de analisar a diferença em sistemas rurais e urbanos quanto a sujidade. Os resultados foram satisfatórios, ao passo que valores de erro em aproximadamente 1% mostraram a eficiência na explicação e modelagem do problema usando as RNAs do tipo MLP. Outro resultado importante foi o fator de sobre dimensionamento na ordem de 4% para áreas urbanas e 10% para áreas rurais, em sistemas sem manutenção periódica. O estudo também mostra os principais fatores meteorológicos que influenciam na geração fotovoltaica. Por fim, encontraram-se os pesos das Redes Neurais Artificiais que estabelecem o conhecimento da rede.

**Palavras-chave:** Energia solar; Módulos fotovoltaicos; Estiagem; Fator de sobredimensionamento.

## ABSTRACT

Photovoltaic generation systems are today an alternative for those who want to invest in generating clean energy. Therefore, it is important that it is increasingly discussed, in academic and professional circles, how to correctly size these systems, generating an analysis that comes closer to the intended real generation. There is no consensus in academia on how the decrease in solar visibility during estimated periods reduces the energy generated in panels installed in rural and urban areas. A bibliographical review was carried out on the proposed topic, and studies from different places around the world were grouped in this work, elucidating the problem. Therefore, this study presents how to determine the oversizing factor to correct generation power in photovoltaic system projects, in regions with a long estimation period, in addition to determining the factors (parameters) that influence photovoltaic generation in these locations considered critical. For this, generation data from two photovoltaic systems were found, one installed in an urban area in Campo Grande/MS and one in a rural area in the city of Bela Vista/MS, which were later compiled together with meteorological data found from INMET. These data were implemented in an Artificial Neural Network (ANN) of the MultiLayer Perceptron (MLP) type, with the help of the WEKA simulator and, as a result, I hope that you obtain, with the smallest possible error, the correction factor in design for systems installed in regions with periods of drought, in addition to analyzing the difference between rural and urban systems in terms of dirt. The results were overwhelming, while error values of approximately 1% showed the efficiency in explaining and modeling the problem using MLP-type ANNs. Another important result was the scaling factor of around 4% for urban areas and 10% for rural areas, in systems without periodic maintenance. The study also shows the main meteorological factors that influence photovoltaic generation. Finally, we found the weights of Artificial Neural Networks that establish the knowledge of the network.

**Keywords:** Solar energy; Photovoltaic modules; Drought; Oversizing factor.

## LISTA DE FIGURAS

Figura 1 – Brasil: incidência solar média diária (Em kWh/m <sup>2</sup> ). .....	16
Figura 2 - 1ª etapa de buscas de artigos científicos com quantitativos dos trabalhos encontrados.....	21
Figura 3 – 2ª etapa de buscas de artigos científicos com quantitativos dos trabalhos encontrados.....	22
Figura 4 - 3ª etapa de buscas de artigos científicos com quantitativos dos trabalhos encontrados.....	22
Figura 5 - 3ª etapa de buscas de artigos científicos ano a ano com quantitativos dos trabalhos encontrados.....	23
Figura 6 - Curvas <i>I-V</i> medidas durante um dia de sol em 2006 e 2017 e transladadas para a condição $G = 1.000 \text{ W/m}^2$ e $T_{cel} = 60 \text{ }^\circ\text{C}$ . .....	26
Figura 7 - Curvas <i>P-V</i> medidas durante um dia de sol em 2006 e 2017 e transladadas para a condição $G = 1.000 \text{ W/m}^2$ e $T_{cel} = 60 \text{ }^\circ\text{C}$ . .....	27
Figura 8 – <i>Principais defeitos encontrados nos módulos inspecionados</i> . .....	27
Figura 9 – Comparação da geração de energia entre as três strings no mês de julho. ....	28
Figura 10 – Curvas dos painéis antes da limpeza. ....	30
Figura 11 – Curvas dos painéis antes e após a limpeza.....	30
Figura 12 – Exemplo de Pacote corrplot no <i>software RStudio</i> . .....	37
Figura 13 – Arquitetura MLP com duas camadas intermediárias. ....	39
Figura 14 – Aprendizado supervisionado. ....	41
Figura 15 - Fluxograma das etapas de trabalho. ....	44
Figura 16 – Imagem aérea da Instalação A.....	46
Figura 17 – Imagem aérea da Instalação B.....	46

Figura 18 – Variáveis pesquisadas e consideradas possivelmente relevantes para a solução do problema. ....	48
Figura 19 – Geração mensal estimada em projeto para o sistema A. ....	52
Figura 20 – Sistema A - Inversor e <i>Stringbox</i> . ....	53
Figura 21 – Geração mensal estimada em projeto para o sistema B. ....	54
Figura 22 – Sistema B - Inversor e <i>Stringbox</i> . ....	54
Figura 23 – Histograma de chuva em (mm). ....	57
Figura 24 - Geração estimada em (kWh/dia). ....	57
Figura 25 - Histograma da geração real em (kWh/dia). ....	58
Figura 26 - Histograma de rajada de vento (km/h). ....	58
Figura 27 - Histograma de temperatura Instantânea (°C). ....	59
Figura 28 - Histograma de temperatura Máxima (°C). ....	59
Figura 29 - Histograma de temperatura Mínima (°C). ....	60
Figura 30 - Histograma de umidade Instantânea (%). ....	60
Figura 31 - Histograma de umidade Máxima (%). ....	61
Figura 32 - Histograma de umidade mínima (%). ....	61
Figura 33 - Histograma de velocidade do vento (km/h). ....	62
Figura 34 - Gráfico de correlação entre as variáveis do sistema A. ....	63
Figura 35 – Histograma de chuva em (mm). ....	65
Figura 36 - Geração estimada em (kWh/dia). ....	66
Figura 37 - Geração real em (kWh/dia). ....	66
Figura 38 - Histograma de rajada de vento (km/h). ....	67
Figura 39 - Histograma de temperatura Instantânea (°C). ....	68
Figura 40 - Histograma de temperatura máxima (°C). ....	68

Figura 41 - Histograma de temperatura mínima (°C).....	69
Figura 42 - Histograma de umidade Instantânea (%).....	70
Figura 43 - Histograma de umidade máxima (%).....	70
Figura 44 - Histograma de umidade mínima (%).....	71
Figura 45 - Histograma de velocidade do vento (km/h).....	71
Figura 46 - Gráfico de correlação entre as variáveis do sistema B. ....	73
Figura 47 - Imagem da Rede Neural do Sistema A, com as 8 variáveis de entrada, e a saída GEREAL.....	76
Figura 48 – Imagem da Rede Neural do Sistema B, com as 8 variáveis de entrada, e a saída GEREAL.....	78
Figura 49 - Exemplo de 10 linhas da planilha de dados contendo as 8 variáveis e a variável de saída GEREAL. ....	81
Figura 50 - Exemplo de 10 linhas da planilha de dados contendo as 8 variáveis e a variável de saída GEREAL. ....	81
Figura 51 – Gráfico de geração x valores preditos pelo <i>software WEKA</i> , em kWh/dia.....	89

## LISTA DE TABELAS

Tabela 1 – Principais trabalhos relacionados ao atual trabalho. ....	32
Tabela 2 – Resumo de dados das variáveis do sistema A. ....	56
Tabela 3 - Resumo de dados das variáveis do sistema B. ....	64
Tabela 4 – Comparativo entre uso com e sem a variável UMIDMÍN para resolução do problema no sistema A. ....	75
Tabela 5 - Comparativo entre uso com, sem UMIDMÍN para resolução do problema no sistema B. ....	77
Tabela 6 – Resultados de erro para o TESTE 0 em diferentes épocas de treinamento. ....	80
Tabela 7 – Resultados de erro para o TESTE 1 em diferentes épocas de treinamento. ....	82
Tabela 8 – Resultados de erro para o TESTE 2 em diferentes épocas de treinamento. ....	83
Tabela 9 – Resultados de erro para o TESTE 3 em diferentes épocas de treinamento. ....	84
Tabela 10 – Resultados de erro para o TESTE 4 em diferentes épocas de treinamento. ....	85
Tabela 11 – Resultados de erro para o TESTE 5 em diferentes épocas de treinamento. ....	86
Tabela 12 – Resultados de erro para o TESTE 6 em diferentes épocas de treinamento. ....	87
Tabela 13 – Resultados de erro para o TESTE 7 em diferentes épocas de treinamento. ....	88
Tabela 14 – Comparativo entre as perdas de geração para os trabalhos citados e o trabalho atual. ....	92
Tabela 15 – Resultado dos neurônios intermediários para o neurônio de saída da MLP para os sistemas rural e urbano combinados. ....	101
Tabela 16 - Valor dos Pesos das entradas para os neurônios intermediários da MLP para os sistemas misturados. ....	102
Tabela 17 – Resultado dos neurônios intermediários para o neurônio de saída da MLP para o sistema A. ....	103

Tabela 18 - Valor dos Pesos das entradas para os neurônios intermediários da MLP para o sistema A. ....	104
Tabela 19 – Resultado dos neurônios intermediários para o neurônio de saída da MLP para o sistema B.....	105
Tabela 20 - Valor dos Pesos das entradas para os neurônios intermediários da MLP para o sistema B.....	106

## LISTA DE SIGLAS E ABREVIATURAS

ANEEL – Agência Nacional de Energia Elétrica

CEMTEC – Centro de Monitoramento do Tempo e do Clima de MS

CPTEC – Centro de Previsão de Tempo e Estudos Climáticos

CSV – *Comma separated values*

EMC – Escola de Engenharia Elétrica, Mecânica e de Computação

FV – Fotovoltaica

GD – Geração distribuída

h – hora

IA – Inteligência artificial

INMET – Instituto Nacional de Meteorologia

INPE – Instituto Nacional de Pesquisas Espaciais

k – quilo

m – metro

MLP – *Multiplayer Perceptron*

NTP – *Network Time Protocol*

p – pico

RNA – Rede Neural Artificial

RNAs – Redes Neurais Artificiais

SFV – Sistema Fotovoltaico

UFG – Universidade Federal de Goiás

UFMS – Universidade Federal de Mato Grosso do Sul

UFRGS – Universidade Federal do Rio Grande do Sul

UNIPAM – Centro Universitário de Patos de Minas

UTFPR – Universidade Tecnológica Federal do Paraná

W – watt

WEKA – *Waikato Environment for Knowledge Analysis*

XLS – *Extensible Style Language*

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>15</b>
1.1	<i>Objetivo Geral .....</i>	17
1.2	<i>Objetivos Específicos .....</i>	17
1.3	<i>Justificativa .....</i>	17
1.4	<i>Estrutura do Trabalho .....</i>	19
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA .....</b>	<b>20</b>
2.1	<i>Pesquisa bibliométrica .....</i>	20
2.2	<i>Resultados bibliométricos .....</i>	21
2.3	<i>Trabalhos relacionados .....</i>	24
2.3.1	<i>Estudo estatístico .....</i>	35
2.4	<i>Sistemas computacionais .....</i>	38
2.4.1	<i>Redes MLP .....</i>	38
2.4.2	<i>Aprendizado Supervisionado .....</i>	40
2.4.3	<i>Backpropagation .....</i>	41
<b>3</b>	<b>METODOLOGIA .....</b>	<b>43</b>
	<i>Etapa 1: Local de projeto .....</i>	45
	<i>Etapa 2: Levantamento e Coleta de dados .....</i>	47
	<i>Etapa 3: Implementação da Rede Neural Artificial do tipo MLP .....</i>	49
	<i>Etapa 4: Análise dos resultados .....</i>	51
<b>4</b>	<b>RESULTADOS .....</b>	<b>52</b>
4.1	<i>Dimensionamento do sistema .....</i>	52
4.2	<i>Visualização de dados .....</i>	55
4.2.1	<i>Filtragem de dados .....</i>	55
4.2.2	<i>Histograma .....</i>	55
4.3	<i>Seleção dos atributos .....</i>	73
4.4	<i>Implementação da RNA do tipo MLP .....</i>	78

4.5	<i>Predição</i> .....	89
4.6	<i>Discussão</i> .....	90
<b>5</b>	<b>CONCLUSÃO</b> .....	<b>93</b>
5.1	<i>Trabalhos futuros</i> .....	95
	<b>REFERÊNCIAS</b> .....	<b>96</b>

## 1 INTRODUÇÃO

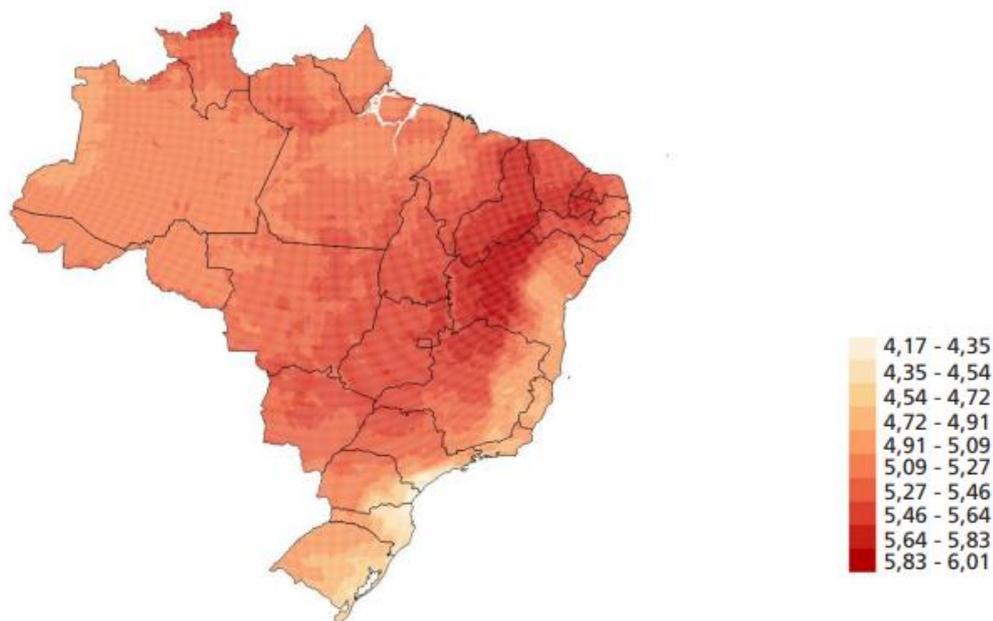
Com o passar dos anos, o aumento crescente da demanda tem acarretado a busca por novos métodos de produção de energia elétrica. A energia Solar Fotovoltaica (FV) já é um dos meios mais procurados devido à sua facilidade de instalação, a vantagem de não produzir poluição em sua geração de eletricidade e pelo alto investimento por novas tecnologias e estudos acerca de métodos e materiais mais eficazes que proporcionam ao sistema melhoras de eficiência a cada ano. Células solares convertem diretamente a energia do sol em energia elétrica de forma estática, silenciosa, não poluente e renovável (RUTHER, 2004).

Em outubro de 2020, o Brasil atingiu 7 GW de potência instalada em energia solar, sendo mais de 4 GW correspondentes à porcentagem de 99,9% de toda micro e minigeração distribuída em residências, comércios, indústrias e propriedades rurais. Já na geração centralizada, o Brasil obteve 2,955 GW gerados pela energia solar, sendo 1,7% de toda a matriz energética brasileira, com previsão de 2,7% até o começo de 2022 (ANEEL, 2020).

Devido a essas condições, portanto, é de suma importância haver estudos que mostrem as peculiaridades dos sistemas, bem como seu correto dimensionamento, levando em consideração fatores como o clima e as intempéries. Além do dimensionamento correto deve ser citado que a manutenção preventiva e a corretiva de Sistemas Fotovoltaicos (SFV) é algo crítico e que merece fundamental atenção (PINHO e GALDINO, 2014).

Alguns estados do Brasil se destacam na geração solar devido ao alto índice de irradiação recebida do sol. Juntos, eles formam o chamado cinturão solar brasileiro possuindo médias de irradiação diária maiores que 5,26 kWh/m<sup>2</sup> (DANTAS, 2018). Um desses estados é o Mato Grosso do Sul, que possui uma média de irradiação anual maior que a maioria do país, conforme mostrado na Figura 1.

Figura 1 – Brasil: incidência solar média diária (Em kWh/m<sup>2</sup>).



Fonte: Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) e Instituto Nacional de Pesquisas Espaciais (INPE), 2016.

O estado do Mato Grosso do Sul tem a sua economia diretamente ligada ao agronegócio, e muitos produtores rurais aderiram à geração fotovoltaica nos últimos anos, visando reduzir os custos de produção com energia elétrica. No entanto, eles relatam que em períodos de estiagem, que duram mais de 30 dias, a geração solar é reduzida consideravelmente.

Um dos fatores que comprometem a geração nesses períodos é o grande número de material particulado no ar, que para tais áreas é levantado do solo não apenas pelo vento, mas também pelo trânsito de caminhões, tratores, colheitadeiras e outras máquinas que trabalham perto de onde a usina fotovoltaica foi instalada.

Essas usinas, quando em áreas rurais, podem ser instaladas no telhado de barracões, mangueiros, sobre casas ou até mesmo em solo, com os devidos cuidados de esforços para ventos. Quando em solo, sua manutenção é feita de uma forma mais simples, pois evita a necessidade de o técnico responsável pela manutenção subir em um telhado para fazer a limpeza, reaperto ou troca de conexões. No entanto, quando instalada no telhado, devido à maior complexidade, muitas vezes a manutenção é negligenciada, passando anos sem uma correta limpeza.

Esse trabalho tem como objetivo, portanto, avaliar as taxas de perda em sistemas fotovoltaicos em períodos de estiagem, tanto urbanos quanto rurais, e propor um fator de sobredimensionamento de projeto que possa ser utilizado por engenheiros e técnicos da área para corrigir a potência dos SFV em áreas cuja sujidade é agravada, devido aos períodos secos.

### **1.1 Objetivo Geral**

Determinar o fator de sobredimensionamento para correção da potência de geração nos projetos de Sistemas Fotovoltaicos, em regiões de longo período de estiagem.

### **1.2 Objetivos Específicos**

- Obter dados de geração fotovoltaica em sistemas instalados no Mato Grosso do Sul;
- Avaliar a importância de dados meteorológicos na geração fotovoltaica em locais de estiagem;
- Determinar os fatores (parâmetros) que influenciam na geração fotovoltaica em locais de estiagem;
- Identificar melhorias de geração fotovoltaica em sistemas que realizam limpeza periódica;
- Definir um método e uma ferramenta adequada para a solução da determinação do fator de sobredimensionamento para correção de geração fotovoltaica;
- Analisar a diferença em sistemas rurais e urbanos quanto a sujidade.

### **1.3 Justificativa**

Sistemas fotovoltaicos vêm ganhando espaço a cada ano no mundo. No Brasil, onde os custos pagos pela energia são uma parcela significativa nas despesas mensais dos consumidores, muitos já aderiram à instalação desses sistemas como uma fonte de economia mensal, reduzindo seus gastos com energia elétrica. Portanto, faz-se necessário o correto dimensionamento por parte dos engenheiros e técnicos responsáveis pelo projeto e instalação dessas usinas.

Esta pesquisa visa não apenas incorporar um novo conhecimento ao dia a dia dos profissionais que atuam no ramo, mas também identificar quais são os fatores que influenciam na geração de energia em SFV, tais como precipitação, temperatura, vento, humidade, entre outros.

Dessa forma, a pesquisa acarretará um impacto econômico com o correto dimensionamento de sistemas em áreas de períodos secos. Também organizacional, mostrando aos responsáveis pelo dimensionamento, quando e onde utilizar o fator de sobredimensionamento, visando uma correta geração fotovoltaica mensal e melhorando a eficiência energética global do sistema.

A contribuição do trabalho não ocorre, portanto, somente quando se deve aumentar a potência de projeto para alcançar determinada geração requisitada, mas também poderá contribuir em identificar situações que tal fator não se faz necessário. Reduz-se assim, os gastos iniciais com equipamentos, tornando o SFV ainda mais viável para os clientes e consumidores que querem investir na geração solar.

Quanto a contribuição acadêmica, refere-se à disponibilização dos dados que mostrem quais características naturais influenciam na geração solar no estado de Mato Grosso do Sul. Tais dados podem ser utilizados em pesquisas futuras tanto na própria UFMS, quanto em outras universidades pelo país, por futuros pesquisadores que venham a estudá-los. Além do fator de sobredimensionamento que é o alvo principal desse trabalho, que posteriormente disponibilizado, pode ajudar futuros projetos de pesquisa.

Outra contribuição do trabalho é que ele visa a eficiência energética que hoje é pautada na ODS 7 (Objetivo de Desenvolvimento Sustentável), que tem como principal objetivo garantir o acesso universal, confiável, moderno e a preços acessíveis a serviços de energia. Outra ODS que esse trabalho engloba é a ODS 13 que visa combater as mudanças climáticas, visto que, a energia solar é uma fonte limpa e inesgotável de energia.

A inserção do pesquisador deste trabalho no mercado da geração fotovoltaica em Campo Grande – MS, possibilitou a identificação do problema principal da pesquisa e, foi o que motivou o interesse e inquietude por dados que resultem em um melhor dimensionamento para os SFVs.

Ao ingressar no Mestrado Profissional em Eficiência Energética e Sustentabilidade, as disciplinas cursadas e conhecimentos agregados, permitiram ao pesquisador esclarecer

algumas destas inquietações, bem como traçar estratégias que possam levar a resultados significativos, direcionado à possíveis melhorias para o serviço prestado, e agregar contribuições para a sociedade e para o meio acadêmico como um todo.

#### **1.4 Estrutura do Trabalho**

Esse trabalho encontra-se dividido em cinco sessões. A primeira é a introdução que retrata a problematização do tema, e contém também os objetivos gerais e específicos da pesquisa, bem como a justificativa para tal. A segunda sessão trata de uma revisão de literatura onde foram alocados dados de um apanhado de artigos e dissertações que embasam a pesquisa atual. A terceira sessão corresponde à metodologia aplicada na coleta, no tratamento dos dados e na análise matemática e computacional. Na quarta sessão, constam os resultados obtidos com este estudo e o cronograma de atividades desde a elaboração do projeto, no qual também resultou na submissão de artigo em periódico. A quinta e última sessão mostra a conclusão do trabalho com propostas de trabalhos futuros.

## 2 REVISÃO BIBLIOGRÁFICA

Nesta seção são apresentadas teorias consideradas relevantes para o acompanhamento do desenvolvimento deste trabalho e artigos que abordam alguns temas significativos e suas relações com o estudo proposto.

### 2.1 Pesquisa bibliométrica

Fazendo uma verificação de trabalhos relacionados a esse projeto, foi realizado um levantamento bibliométrico com o tema principal de: “interferência da sujidade em sistemas fotovoltaicos”. Foi utilizado o *software* livre *Harzing's Publish or Perish* versão 8.6.4214.8378. Também foram feitas pesquisas nos periódicos eletrônicos da Capes.

A pesquisa consistiu em 3 etapas que foram limitadas a alguns fatores importantes para a atuação dos mecanismos de pesquisa, todas as etapas aconteceram em dezembro de 2022. Por conta das limitações dos resultados máximos do *Google Scholar* (1000 resultados), a pesquisa necessitou ser dividida em diversos períodos de publicação, que posteriormente foram consolidados.

Para a 1ª etapa da pesquisa nos *sites* foi realizada a busca utilizando como palavras-chave os termos “*losses, photovoltaic, dirt e drought*” que em português refere-se a “perdas, fotovoltaicos, sujeira, estiagem (seca)”.

Na 2ª etapa da pesquisa a prioridade foram os trabalhos apresentados a partir de 2013, delimitando os artigos. O fato de selecionar os artigos mais novos, dos últimos dez anos, leva em consideração a evolução dos materiais que compõem os módulos fotovoltaicos, e com isso, a melhoria do desempenho deles. Também pode-se citar o aumento do número de usinas fotovoltaicas instaladas pelo mundo, o que gera mais informação sobre o assunto.

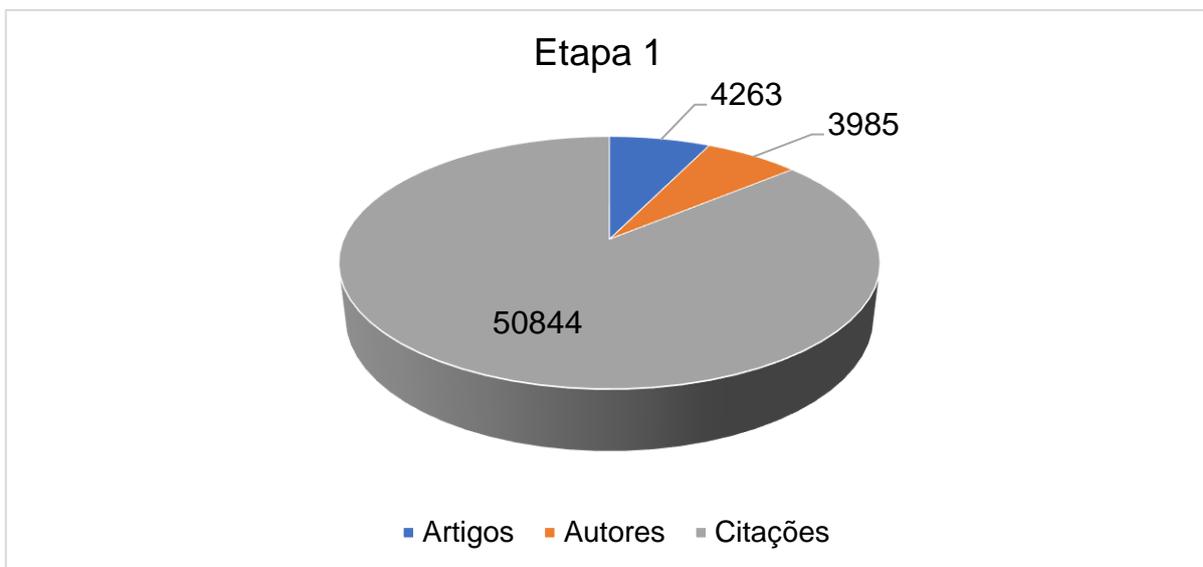
Para a 3ª etapa da pesquisa foi efetuada uma procura pelos títulos, onde foi possível melhorar a seleção dos artigos, utilizando um filtro rejeitando palavras, as quais foram “*heating, shading, biology, microbiology, medicine, hydraulic, physiological agriculture e vegetation*” que em português são “*aquecimento de água, sombreamento, biologia, microbiologia, medicina, hidráulica, fisiológico e vegetação*”. Isto devido ao fato de que estas

palavras não são comumente utilizadas em artigos científicos para tratar do assunto desse projeto.

## 2.2 Resultados bibliométricos

Na 1ª etapa da pesquisa foi possível encontrar um total de 4.263 artigos, com 3.985 autores e um total de 50.844 citações, como pode ser observado na Figura 2.

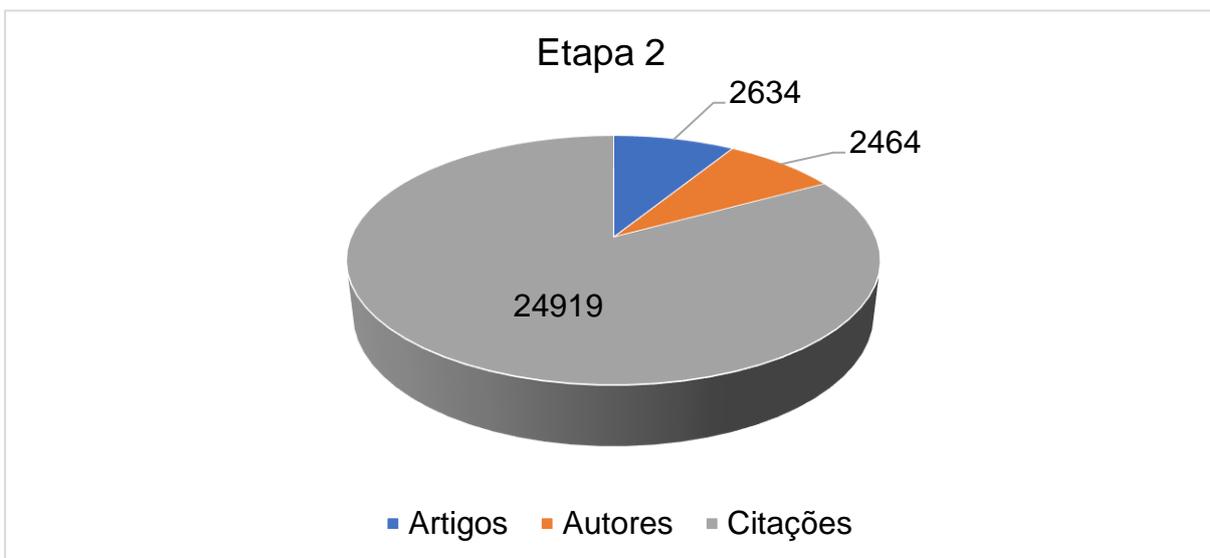
Figura 2 - 1ª etapa de buscas de artigos científicos com quantitativos dos trabalhos encontrados.



Fonte: próprio autor.

Na 2ª etapa da pesquisa foi aplicado o filtro de data buscando os artigos com tecnologias mais recentes, desde o ano de 2013 até o ano de 2022, resultando em 2.634 artigos, com 2.464 autores e um total de 24.919 citações, conforme a Figura 3.

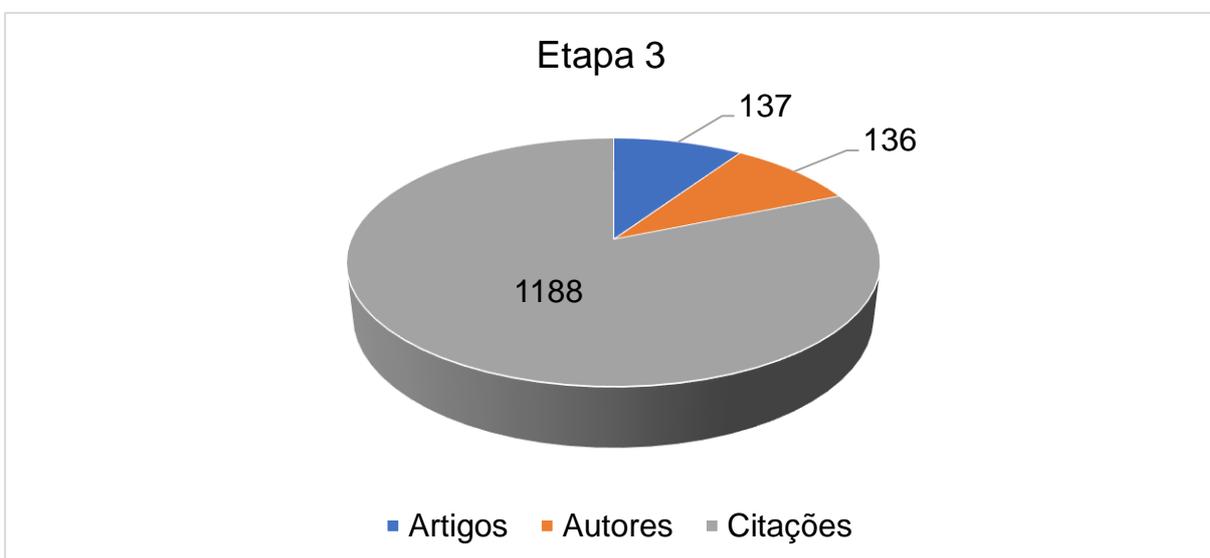
Figura 3 – 2ª etapa de buscas de artigos científicos com quantitativos dos trabalhos encontrados.



Fonte: próprio autor.

Já na 3ª etapa da pesquisa, ao analisar os títulos utilizando o filtro rejeitando palavras, foi possível filtrar para um total de 137 artigos, com 136 autores e um total de 1.188 citações, como pode ser observado na Figura 4.

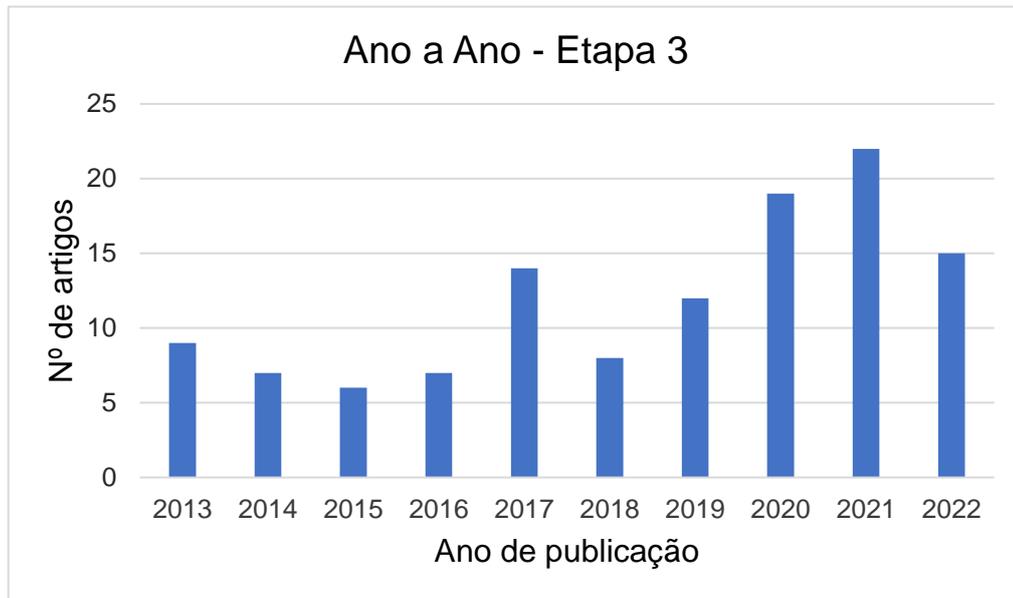
Figura 4 - 3ª etapa de buscas de artigos científicos com quantitativos dos trabalhos encontrados.



Fonte: próprio autor.

Como a Etapa 3 elenca artigos de maior interesse para essa pesquisa, uma análise ano a ano foi feita, bem como a quantidade de artigos publicados por ano conforme a Figura 5.

Figura 5 - 3ª etapa de buscas de artigos científicos ano a ano com quantitativos dos trabalhos encontrados.



Fonte: próprio autor.

Na seção seguinte alguns desses trabalhos mais relevantes serão apresentados com maiores detalhes.

### 2.3 Trabalhos relacionados

Araújo (2016) realizou uma pesquisa para analisar os principais fatores de perdas em sistemas fotovoltaicos. Ela elenca quatro, que são: temperatura, sujidade, sombreamento e perdas no inversor. Para a temperatura, seus estudos mostram uma diferença de 10 ° C de um sistema para outro, instalado em lugares próximos, mas de maneiras diferentes. Quanto à sujidade, a limpeza periódica resultou em melhorias de até 10 % em suas análises. Quanto ao sombreamento, a autora cita que é de suma importância o planejamento, antes da execução da instalação, visando o melhor local para instalação. Por fim, a autora mostra que em sistemas que possuem inversores mais atualizados e com materiais mais novos, a diferença nas perdas para modelos mais antigos pode chegar a 6%.

Costa (2011) identificou perdas em torno de 30% do rendimento de um sistema fotovoltaico instalado em Belo Horizonte, comparando a energia produzida antes e depois da ocorrência de precipitação que promoveu a limpeza natural dos módulos FV, eliminando a sujidade acumulada. Ele levanta dados que apresentam que a redução da irradiância solar absorvida ainda fica mais comprometida em regiões com longos períodos secos e com elevado índice de material particulado sobre os módulos fotovoltaicos. Isso ocorre em regiões com elevada poluição do ar, fator esse registrado com frequência nas grandes áreas urbanas, onde há um grande aumento anual da Geração Distribuída (GD).

Al-Sabounchi (2013) avaliou um sistema de 36 kWp, interligado à rede e instalado em Abu Dhabi, nos Emirados Árabes. O estudo mostrou uma redução em torno de 27% de produção de energia, nos meses vizinhos a julho, com o foco no próprio mês de julho, verão no Hemisfério Norte.

Segundo Duarte (2020) a deposição de sujidades tem impacto negativo significativo no desempenho de módulos Fotovoltaicos (FV). Para tal, ele apresenta uma análise sobre a taxa de sujidade em módulos fotovoltaicos e as principais interferências na geração de energia solar. Esse autor também utilizou cupons de vidro, que foram colocados em uma área externa perto do laboratório da GREEN PUC Minas, em Belo Horizonte/MG onde a densidade gravimétrica foi medida em períodos e a taxa de sujidade foi estudada e quantificada, a partir da pesagem com balança de precisão.

Ainda no trabalho de Duarte dados obtidos foram estudados e comparados com os de estações meteorológicas próximas. Além de mostrar uma grande concordância entre os dois métodos observados, o autor revela que a sujidade reduz a geração de energia em até 14,02% após um período de 75 dias de exposição, onde após aferições observa uma deposição de 2,57g de sujeira, por metro quadrado, quando comparado ao sistema recém lavado. Por fim, o autor cita que tanto o estudo baseado *in loco* (vidro) quanto o de estações gravimétricas apresentam valores muito parecidos.

Ghazi *et al.* (2014) relatam em seus estudos que esta sujidade pode ser composta por uma grande quantidade de substâncias, sendo a mais comum delas a poeira (por poeira entende-se qualquer agregado de partículas com menos de 500  $\mu\text{m}$  de diâmetro). A composição da poeira também é diversa, podendo conter pólen, material têxtil e frequentemente minerais orgânicos, como areia e argila.

Romanholo (2019) utilizou para sua análise um sistema conectado à rede e instalado no campus da Universidade Federal de Goiás (UFG) em Goiânia, com potência de 34 kWp. Vizinho a esse, encontra-se uma central meteorológica capaz de mensurar e armazenar os valores de velocidade e direção do vento, temperatura ambiente, umidade relativa do ar, índice pluviométrico, irradiância solar e pressão atmosférica. Para tal, uma análise completa da morfologia das partículas foi feita, tanto visualmente quando em um micrografo. Com a Espectroscopia de Energia Dispersiva de Raios-X (EDS), puderam analisar a composição química das partículas depositadas. As amostras coletadas mostraram uma grande presença de silício, indicando possível participação de areia ou de outra estrutura do dióxido de silício.

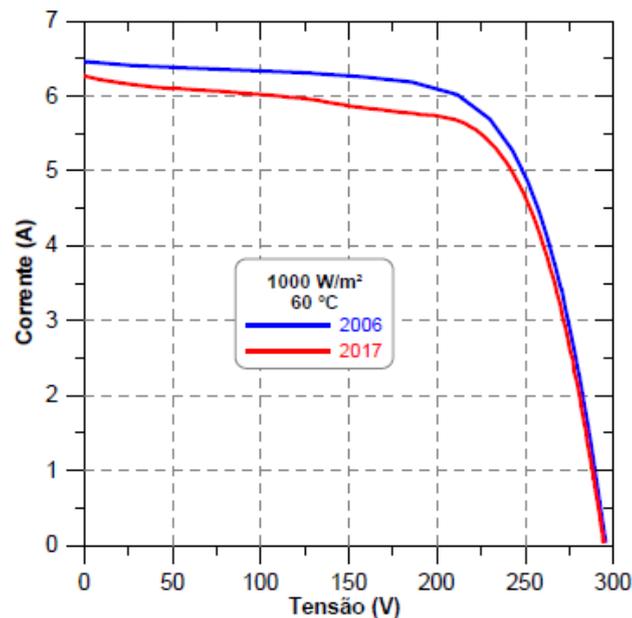
Além disso, Ndiaye *et al.* (2013) apresentam uma revisão de trabalhos publicados com um compilado dos principais problemas acerca da degradação de módulos fotovoltaicos. São eles:

- Interconexões quebradas;
- Células quebradas;
- Corrosão;
- Delaminação;

- Descoloração (*browning* e *yellowing*);
- Vidro quebrado;
- Falhas no diodo de *bypass*;
- Falhas na fita de solda.

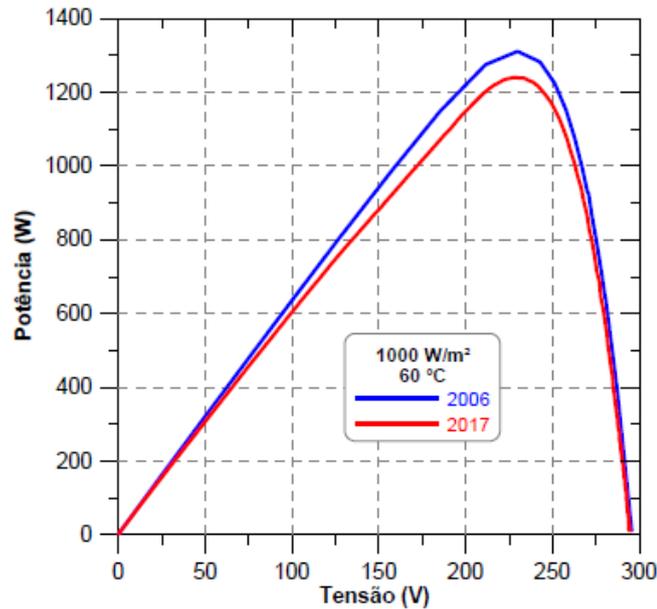
De Oliveira (2018) em seu trabalho, analisou um sistema fotovoltaico conectado à rede de 48 módulos de 100 Wp que foi instalado na fachada norte do Laboratório de Energia Solar da UFRGS (LABSOL) em 2004. Em seu trabalho, ele avalia o sistema após 11 anos de funcionamento, levanta curvas de tensão e corrente e explica que os módulos fotovoltaicos ao envelhecerem diminuem sua capacidade de fornecer potência. Comparando as curvas de parâmetros elétricos, nota-se uma redução de 5% na potência gerada, conforme mostra a Figura 6 e a Figura 7.

Figura 6 - Curvas I-V medidas durante um dia de sol em 2006 e 2017 e transladadas para a condição  $G = 1.000 \text{ W/m}^2$  e  $T_{cel} = 60 \text{ }^\circ\text{C}$ .



Fonte: De Oliveira (2018).

Figura 7 - Curvas P-V medidas durante um dia de sol em 2006 e 2017 e transladadas para a condição  $G = 1.000 \text{ W/m}^2$  e  $T_{cel} = 60 \text{ }^\circ\text{C}$ .



Fonte: De Oliveira (2018).

O mesmo autor relata ainda os principais problemas que ocorreram nos painéis em questão, através de inspeção visual, conforme a Figura 8.

Figura 8 – Principais defeitos encontrados nos módulos inspecionados.

Tipo de defeito	Módulos afetados	Módulos afetados (%)
Descoloração	16	100
Delaminação	13	81
Célula quebrada	2	12
Infiltração	4	25

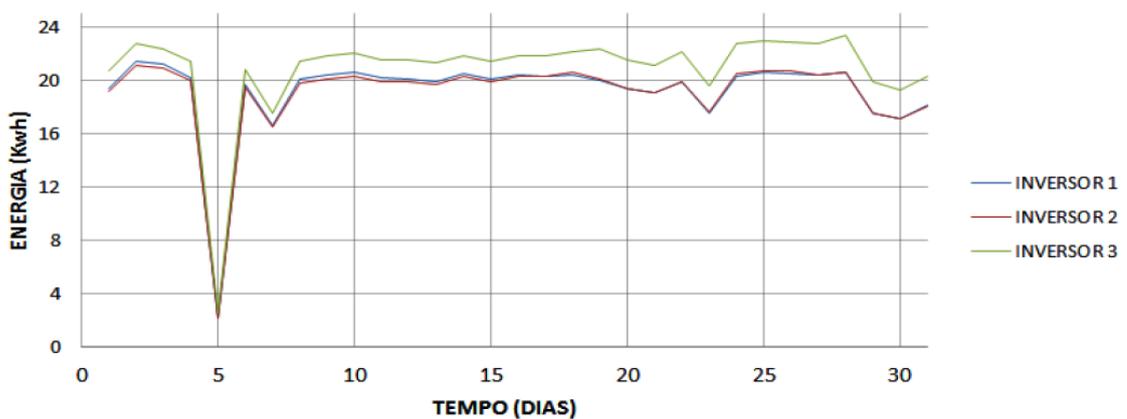
Fonte: De Oliveira (2018).

Barbosa (2019) define que a sujidade pode ser composta por diferentes partículas, sendo a poeira uma delas. Por poeira ele define “como qualquer material particulado de diâmetro inferior a  $500 \mu\text{m}$  e composto por pequenas quantidades de pólen (vegetação, fungos, bactérias), células humanas/animais, fibras de carpete/materiais têxteis e outras micro areias, sendo mais comumente, minerais geomorfológicos (areia, argila, calcário erodido)”.

O mesmo autor tece também um estudo sobre a influência da sujidade em um sistema solar, conectado à rede e instalado na cidade de Patos de Minas, mais especificamente no

Centro Universitário de Patos de Minas (UNIPAM). Em seu trabalho, ele analisou durante dois meses a diferença entre a geração de uma *string* que estava sendo limpa a cada 7 dias e duas outras que estavam a mais de um ano sem nenhuma limpeza. Os dois meses, como ressalta o autor, foram de mesmo clima e não apresentaram chuva. Para o ensaio o autor mostra diferença de mais de 10% a mais na geração dos painéis que possuíam a limpeza regular, conforme a Figura 9. Cita ainda que para uma melhor eficiência em sistemas fotovoltaicos, a manutenção preventiva deve ser feita periodicamente.

Figura 9 – Comparação da geração de energia entre as três strings no mês de julho.



Fonte: Barbosa (2019).

Mejia (2014) traz um estudo sobre a sujidade de um arranjo de painéis de um Sistema Fotovoltaico Concentrado (CPV) conectado à rede, que foi instalado em Santa Clara, Califórnia (USA), no ano de 2010. O autor cita que o estudo foi feito naquele lugar específico, devido à proximidade de uma estação meteorológica que fornece dados precisos que são usados para esta análise. Coletando dados de geração a cada 15 minutos, foi possível obter informações que demonstram uma perda de -0,021% ao dia e uma média anual de eficiência em torno de 6,8%. Relata também que, durante o período de estiagem o sistema perdeu 1,6% de eficiência, e que essa teve uma grande melhora após a ocorrência de chuva. Posteriormente, faz um comparativo de valores entre um sistema com e outro sem processo de lavagem, observando ganhos de 11,1% da energia anual, calculado como a integral entre a eficiência do painel limpo e a eficiência real observada. Por fim, finaliza seu estudo resumindo que as perdas por sujeiras têm seu maior impacto durante os longos verões secos da Califórnia.

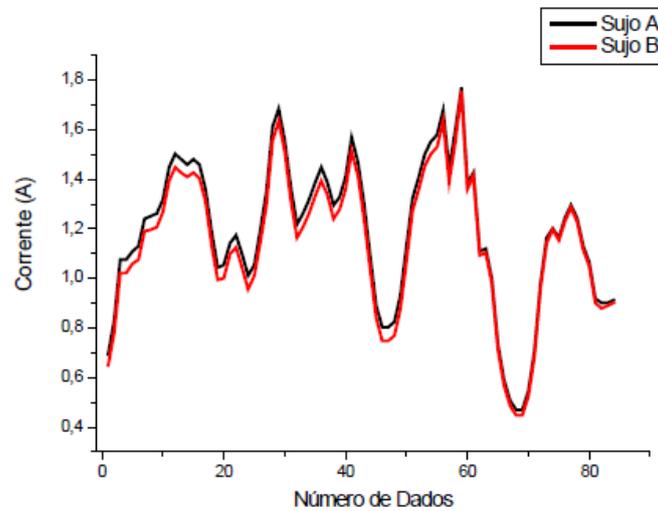
Monto (2010) e El-Nashar (2003), em suas pesquisas no Oriente Médio, obtiveram resultados parecidos com Garg (1974) na Índia. Assim como em Santa Clara nos Estados Unidos, foram observadas perdas de geração médias de 0,2% ao dia para sistemas conectados à rede, em longos períodos de estiagem.

Vera (2006) apresenta uma comparação do desempenho de módulos fotovoltaicos após seis anos de operação. Para tal, apresenta uma análise dos dados adquiridos através do comportamento das curvas características ( $I-V$ ) e da variação nos valores da potência máxima ( $P_m$ ), corrente de curto-circuito ( $I_{sc}$ ), tensão de circuito aberto ( $V_{oc}$ ), para 8 módulos fotovoltaicos, antes e depois de serem lavados.

Em seu estudo, Vera exalta que os valores de  $I_{sc}$  e  $P_m$  sofreram uma melhoria da ordem de 5 % e 4 %, respectivamente, após a lavagem dos módulos e que a variação de  $V_{oc}$  é praticamente desprezível. Para tal, o autor mostra que a poeira não introduziu distorções no formato das curvas  $I-V$ , mesmo em módulos nos quais a deposição de sujeira não foi totalmente uniforme. Mostra, também, a evidente necessidade da limpeza periódica dos módulos, a fim de evitar a perda de potência por sujeira e a consequente queda no desempenho geral do sistema. Por fim, o autor apresenta variações no  $I_{sc}$  e  $P_m$  da ordem de 2 % e 6 %, respectivamente.

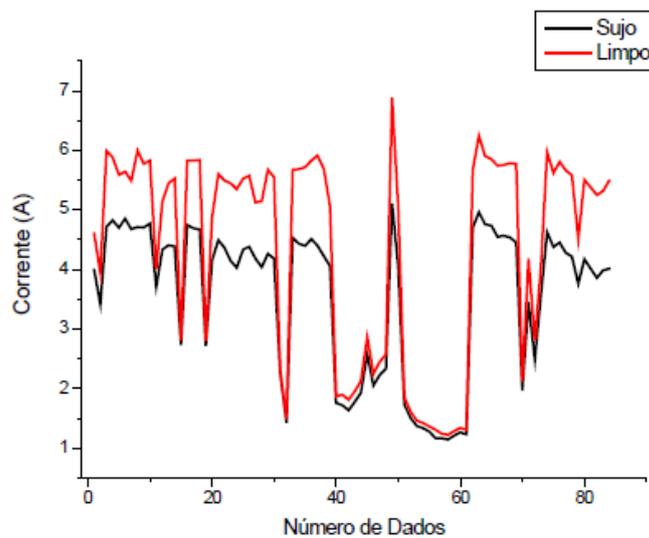
Michels (2009) buscou quantificar a diferença na eficiência, entre dois conjuntos de painéis fotovoltaicos, um limpo e outro sujo, na cidade de Medianeira/PR. Um sistema de 10 módulos fotovoltaicos foi analisado, após ser instalado nas dependências da Universidade Tecnológica Federal do Paraná (UTFPR). Sendo assim, dos 10 módulos, 4 foram analisados, 2 sujos e 2 limpos, e levantou-se os gráficos de corrente antes e depois da limpeza, conforme a Figura 10 e Figura 11.

Figura 10 – Curvas dos painéis antes da limpeza.



Fonte: Michels (2009).

Figura 11 – Curvas dos painéis antes e após a limpeza.



Fonte: Michels (2009).

Segundo o autor, verificou-se, após análise e comparação dos dados, que a configuração com os painéis limpos apresentou uma eficiência 16,26% maior do que com a configuração de painéis sujos. Por fim, o autor cita que é de suma importância fazer limpezas periódicas nos painéis fotovoltaicos, para não ocorrer perda de eficiência.

Já Tonolo (2018) usa o *software* RADIASOL para analisar o impacto da sujidade no rendimento de dois sistemas fotovoltaicos instalados na sede da Universidade Tecnológica Federal do Paraná (UTFPR), campus Curitiba. Para tal, foi feito um levantamento em todos os

dias dos meses de junho até outubro, comparando a performance dos sistemas fotovoltaicos pré-limpeza programada, com o pós-limpeza. Como o autor cita, após a limpeza houve um período de estiagem, que durou 30 dias e posteriormente um período de grande quantidade de chuva, ao qual, a autolimpeza dos painéis foi de grande eficiência.

Tonolo demonstra, em seu trabalho, que o período de seca apresentou alta influência sobre o desempenho dos sistemas fotovoltaicos, diminuindo muito a geração devido ao acúmulo de sujeira. O autor apresentou através de tabelas que tanto a limpeza programada quanto a autolimpeza mostraram praticamente os mesmos resultados, e que quanto maior a inclinação dos painéis, menor a interferência da geração devido ao acúmulo de sujeira.

Tonolo (2019) comparou ainda os mesmos sistemas anteriores, apresentando aumento de geração fotovoltaica de aproximadamente 10% para a limpeza e 6% para a autolimpeza, quando comparados com dados em período anterior a limpeza.

Araújo (2020) estudou um sistema fotovoltaico de 3,9 kWp instalado no Laboratório de Energias Alternativas da Universidade Federal do Ceará (LEA-UFC). Para tal, foi realizada a observação de dados pluviométricos, de temperatura e de irradiação, durante 13 meses seguidos. Dois períodos de estiagem foram constatados, os valores de geração durante o período de estiagem decaíram, atingindo valores próximos de 75% do valor de projeto (22 dias sem chuva), enquanto no período chuvoso aumentaram, atingindo valores próximos de 100%.

Uma outra análise para um período de 15 dias sem chuva, foi constatado um valor de decaimento de 12% de geração, sendo então recomendado limpezas periódicas entre 10 e 15 dias para módulos FV instalados em condições ambientais e de instalação do presente estudo.

A Tabela 1 contém os principais trabalhos que contribuíram para o desenvolvimento deste projeto, bem como o ano, autores, local de publicação, área e principais relevâncias.

Tabela 1 – Principais trabalhos relacionados ao atual trabalho.

<i>Nome do trabalho</i>	<i>Autores</i>	<i>Citação</i>	<i>Local de publicação</i>	<i>Área de concentração</i>	<i>Importância para a atual pesquisa</i>
Análise dos fatores de perdas em sistemas fotovoltaicos da UTFPR campus Curitiba	Édwin Augusto Tonolo	Tonolo (2019)	Programa de Pós-Graduação em Sistemas de Energia da Universidade Tecnológica Federal do Paraná	Automação e Sistemas de Energia	Demonstra os principais fatores de perdas em SF bem como a sujidade que é um dos fatores a serem analisados no atual trabalho
Caracterização físico-química da sujidade depositada sobre módulos fotovoltaicos instalados em zonas climáticas de Minas Gerais	Suellen Caroline Silva Costa Amanda Mello Faria Veloso Abreu Marcelo Machado Viana Pedro Paiva Brito Cristiana Brasil Maia Antônia Sonia Alves Cardoso Diniz Lawrence Lee Kazmerski	Costa <i>et al.</i> (2016)	VI Congresso Brasileiro de Energia Solar – Belo Horizonte	Conversão térmica e fotovoltaica da energia solar	Utiliza dados do INMET para sua análise, assim como a pesquisa atual, além de estudar a morfologia das partículas de sujeira.
<i>Soiling losses for solar photovoltaic systems in California</i>	Felipe A. Mejia Jan Kleissl	Mejia e Kelissl (2013)	<i>Center for Renewable Resources and Integration, University of California, San Diego 9500 Gilman Dr., La Jolla, USA</i>	Engenharia mecânica e aeroespacial	O artigo apresenta uma análise de 186 sistemas fotovoltaicos estudados durante um período de estiagem ao longo de 2010. Mostra perdas médias de 0,051% ao dia. Chegando a mais de 0,1% em instalações que possuem pequenos ângulos de inclinação.

<i>Nome do trabalho</i>	<i>Autores</i>	<i>Citação</i>	<i>Local de publicação</i>	<i>Área de concentração</i>	<i>Importância para a atual pesquisa</i>
Determinação do tempo ótimo para limpeza de painéis fotovoltaicos para obtenção da melhor produtividade - Estudo de caso dos SFVCR's Implantados na UTFPR	João Nicolau Gaio Kleverson Moisés Apolinário de Campos	Gaio (2017)	Departamento Acadêmico de Eletrotécnica (DAELT) da Universidade Tecnológica Federal do Paraná (UTFPR).	Engenharia Elétrica	Demonstra o melhor tempo entre limpezas periódicas afim de obter um melhor desempenho em SFV's. Sua metodologia foi em parte utilizada na pesquisa atual.
Tratamento de dados meteorológicos e análise de desempenho do sistema fotovoltaico da EMC/UFG	Jéssica Alice Alves da Silva	Silva (2019)	Universidade Federal de Goiás Escola de Engenharia Elétrica, Mecânica e de Computação	Sistemas Eletroeletrônicos	Utiliza uma metodologia importante para o atual trabalho, ao passo que, elenca os principais dados meteorológicos que serão utilizados para a análise de desempenho dos módulos fotovoltaicos
O impacto no desempenho de sistemas fotovoltaicos causado pelo acúmulo de sujeira sobre módulos FV - Metodologia e avaliação através de curvas IxV em campo	Bernardo Meyer Hickel	Hickel (2017)	Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-graduação em Engenharia Civil	Engenharia Civil	Traz uma análise mais experimental sobre a influência da sujidade, observando parâmetros de tensão e corrente contínua.

<i>Nome do trabalho</i>	<i>Autores</i>	<i>Citação</i>	<i>Local de publicação</i>	<i>Área de concentração</i>	<i>Importância para a atual pesquisa</i>
Impactos da sujidade e efeitos da limpeza no desempenho de módulos fotovoltaicos	Danielly Norberto Araújo Suellen Caroline Silva Costa Ivonne Montero Dupont Paulo C. M. Carvalho	Araújo <i>et al.</i> (2020)	VIII Congresso Brasileiro de Energia Solar – Fortaleza	Energia solar	O trabalho apresenta diferenças de performance entre períodos secos e chuvosos de em média 17% devido a sujidade depositada no ar e nos módulos FV. Mostra também que mesmo em períodos chuvosos a irradiação solar atingindo níveis mais baixos do que no período de estiagem, o desempenho da planta FV é melhor.
Estimativa da taxa de sujidade em módulos fotovoltaicos utilizando dados de densidade gravimétrica	Túlio Pinheiro Duarte Suellen Caroline Silva Costa Antônia Sonia Alves Cardoso Diniz Lawrence Lee Kazmerski	Duarte <i>et al.</i> (2020)	VIII Congresso Brasileiro de Energia Solar – Fortaleza	Energia solar	O autor apresenta uma metodologia alternativa a esse trabalho para quantificar a sujidade, usando cupons. Ao final de um período de estiagem de 75 dias, seu trabalho demonstra uma diminuição de aproximadamente 14% na eficiência dos módulos FV. Sendo assim, é interessante pois mesmo com outras metodologias pode-se observar resultados semelhantes.
Sujidade depositada sobre módulos fotovoltaicos instalados em Goiânia: morfologia e composição química	Pedro Victor Valadares Romanholo Bernardo Pinheiro de Alvarenga Enes Gonçalves Marra Sérgio Pires Pimentel	Romanholo <i>et al.</i> (2019)	VII Congresso Brasileiro de Energia Solar – Gramado	Energia solar	Utiliza uma técnica de microscopia que por sua vez apresenta dados que mostram algumas partículas que não são facilmente eliminadas por chuva e vento, que por sua vez, mostra a importância das manutenções (limpezas) periódicas nos módulos FV.

A seguir serão apresentados conceitos que tem o propósito de auxiliar no acompanhamento da metodologia proposta.

### **2.3.1 Estudo estatístico**

Segundo Bayer *et al.* (2004), a estatística é uma ciência que estuda e pesquisa sobre: a) o levantamento de dados, com a máxima quantidade de informação possível para um dado custo; b) o processamento de dados para a quantificação da quantidade de incerteza existente na resposta para um determinado problema; c) tomada de decisões sob condições de incerteza, sob o menor risco possível.

Sendo assim, algumas análises são de extrema importância para entendimento dos dados. Valores como a média, mediana e o desvio padrão, mostram como os dados estão alocados e algumas de suas peculiaridades.

Como cita Kurokawa (2002), o histograma é uma das ferramentas estatísticas da qualidade. Ele é utilizado para representar graficamente uma grande quantidade de dados numéricos. Através da análise do histograma é possível interpretar estas informações de forma mais fácil e simples, do que acompanhando uma grande tabela ou um relatório com somente números e/ou valores.

O Histograma é um gráfico de colunas utilizado na estatística. É composto de vários retângulos adjacentes, representando a tabela de frequências com perda de informações (valores agrupados por classes) de um conjunto de valores. Na escala horizontal, marcam-se os intervalos de classes, e cada intervalo é a base de cada retângulo ou barra; na escala vertical, marcam-se as alturas dos retângulos ou barras, que são as respectivas frequências absolutas das classes conforme cita Lopes (1999).

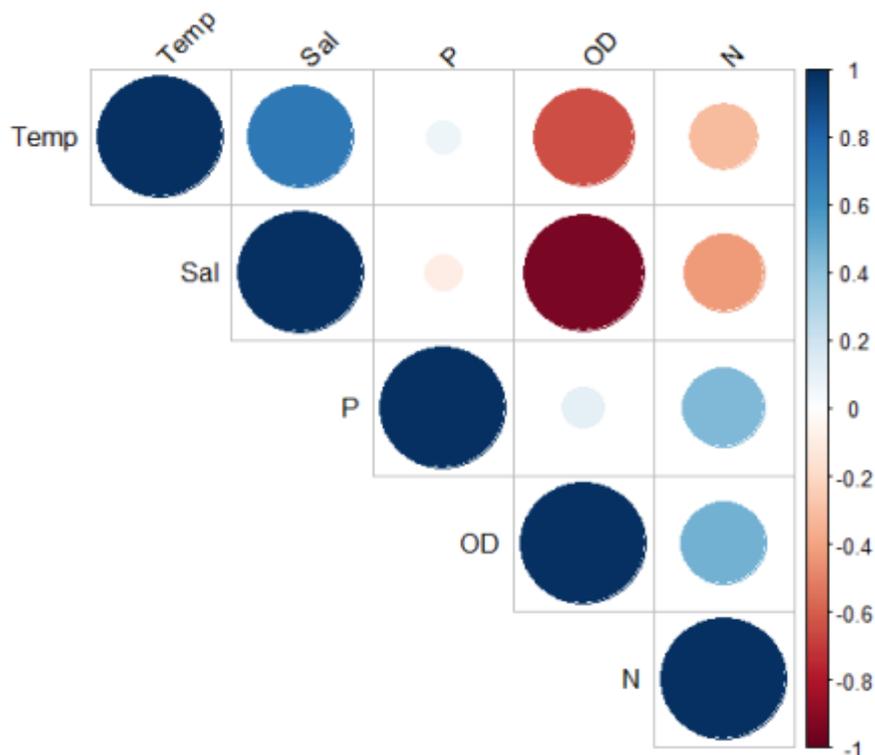
O primeiro passo para analisar a correlação entre duas variáveis quantitativas deve ser a visualização do diagrama de dispersão, a fim de identificar se existe uma variabilidade gradual entre os conjuntos de dados, se essa variação é monotônica (predominantemente ascendente ou descendente), se assume uma tendência proporcional (linear) e se a

distribuição subjacente dos dados é normal ou apresentam equidispersão. Diferentes combinações dessas premissas indicam diferentes técnicas de análise de correlação.

A análise da correlação linear dos dados é de extrema importância para o problema pois evita a multicolinearidade. Segundo Fávero (2017) a multicolinearidade surge como problema quando há correlações elevadas entre as variáveis explicativas, indicando relação linear entre elas, surgindo principalmente quando apresentam tendências semelhantes, por exemplo, uma variável que indica um dia chuvoso e uma outra variável que indica que o dia está nublado, estas terão uma correlação, apresentando assim multicolinearidade. Outro fator possivelmente determinante para multicolinearidade é análise de amostra com poucas observações.

O *Software RStudio* possui um pacote para a análise da correlação linear dos dados, chamado de *corrplot*, onde nele é mostrada a partir de figuras as correlações, como mostra um exemplo do próprio *software*. Conforme mostra a Figura 12 a variável Temp que é a variável de saída tem uma correlação alta e positiva com a variável Sal e uma correlação alta e negativa com a variável OD. No entanto há indícios de multicolinearidade entre as variáveis OD e Sal, ao passo que estão com uma correlação alta, indicando uma tendência semelhante entre elas.

Figura 12 – Exemplo de Pacote corrplot no *software RStudio*.



Fonte: R CORE TEAM (2019).

Quando se tem evidências de multicolinearidade, geralmente faz-se testes de explicação do modelo, eliminando a variável que possui menor correlação com a variável de saída. Se identificado que a variável é redundante, ela é eliminada do modelo de explicação do problema.

Por fim, a resposta é geralmente atrelada a um erro. Esse erro é apresentado de diferentes formas de acordo com o *software* trabalhado ou mesmo da resposta que o problema aborda. O erro absoluto médio é uma medida estatística que quantifica a média das diferenças absolutas entre os valores observados e os valores esperados em um conjunto de dados. É uma maneira de avaliar o quão próximo ou distante os valores observados estão dos valores esperados. Já o erro quadrático médio é uma medida estatística que quantifica a média dos quadrados das diferenças entre os valores observados e os valores esperados em um conjunto de dados. É amplamente utilizado em modelos de previsão ou regressão (BISHOP, 2006).

## 2.4 Sistemas computacionais

O ser humano utiliza o computador para análise e implementação de dados. No entanto, a pouco tempo, algumas ferramentas vêm tendo destaque para uso de mineração de dados e análise deles. As Redes Neurais Artificiais (RNAs) são uma ferramenta atrativa quando o assunto é a análise de entradas e saídas de informação e apresenta uma alternativa aos algoritmos estatísticos que são normalmente usados para resolução de problemas.

De uma maneira resumida, a rede neural simula o cérebro de um ser humano: recebe uma informação, a processa levando em consideração vários fatores e finalmente encaminha uma resposta (saída).

Para alcançarem bom desempenho, as redes neurais empregam uma interligação maciça de células computacionais simples, denominadas de “neurônios” ou unidades de processamento (Haykin, 2001). A arquitetura formada por esses neurônios é o que faz o sistema RNA ser satisfatório para a resolução do problema. Por isso, encontrar a melhor arquitetura é a maior dificuldade para o pesquisador que busca uma resposta mais assertiva para seu problema.

### 2.4.1 Redes MLP

Existem diversos tipos de RNAs, sendo que algumas atendem a problemas diversos e outros a problemas específicos, que seriam resolvidas por equações mais simples ou as mais complexas. Em 1958 Frank Rosenblatt propôs pela primeira vez uma topologia chamada MLP (*Multilayer Perceptron*), que consiste em camada (s) de entrada, camada (s) intermediária (s) e camada(s) de saída. As camadas intermediárias realizam combinações internas de acordo com os padrões de entrada, que são utilizadas na camada de saída, segundo Silva *et al.* (2010).

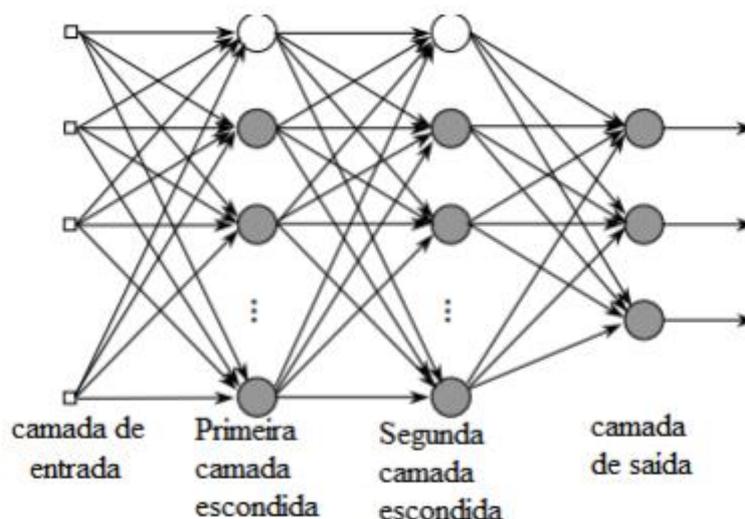
Para o processamento de cada neurônio de uma camada é definido uma combinação de processamentos, com seus pesos e limiares, realizados pelos neurônios da camada anterior ligados a ele. Os valores dos pesos e limiares são obtidos através de fórmula matemática que é chamada de cálculo do erro, onde o erro é obtido pela diferença entre a saída desejada e a saída corrente da rede, para redes de uma única camada. Já para redes de múltiplas camadas,

esse procedimento pode ser aplicado para a camada de saída, visto que não existem saídas desejadas definidas para as camadas intermediárias, tornando-se um problema de como calcular ou estimar o erro das camadas intermediárias, conforme Braga *et al.* (2014).

Atualmente as redes MLP são as mais utilizadas na resolução de problemas, pois elas conseguem identificar padrões e encontrar a melhor forma de chegar a uma saída compatível com o problema. Além disso, a capacidade de “aprender” através de exemplos dados pelo usuário e posteriormente generalizá-los é o que torna a rede RNA tão empregada ao redor do mundo.

Esse modelo também é chamado por muitos como aprendizado de máquina. A camada de entrada distribui as informações “entradas” para a(s) camada(s) intermediária(s) escondida(s) da rede. As camadas intermediárias possuem uma determinada quantidade de neurônios (determinada de acordo com a arquitetura do problema). Segundo Santos (2005), as camadas escondidas têm como função separar as camadas de entrada e de saída. Na camada de saída, a solução ou mesmo soluções do problema são obtidas, conforme mostra a Figura 13, que apresenta um sistema com 2 camadas ocultas e 3 saídas.

Figura 13 – Arquitetura MLP com duas camadas intermediárias.



Fonte: Fleck (2016).

Por via da Figura 13, pode-se notar que cada sinal de entrada da rede será propagado aos neurônios de saída da rede. Conforme explicam Silva *et al.* (2010), nota-se que as saídas dos neurônios da primeira camada escondida serão as entradas da segunda camada escondida e a saída da segunda camada escondida será a entrada dos neurônios pertencentes à camada de saída, tornando assim, a propagação dos sinais de entrada sempre em um único sentido, ou seja, da camada de entrada para a camada de saída da rede.

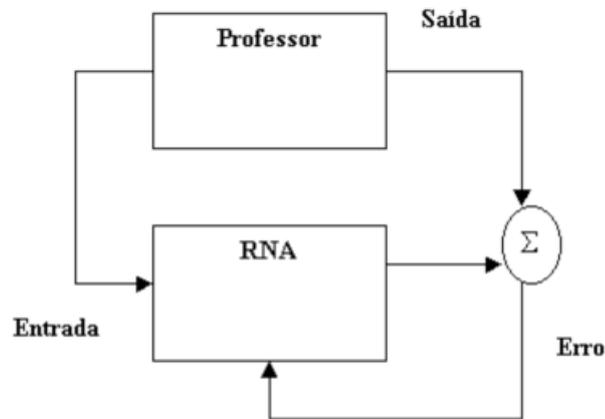
Silva *et al.* (2010) também elencam que a camada de entrada recebe os sinais e a partir desses sinais se inicia o funcionamento da rede. As camadas escondidas são responsáveis por extrair a maior parte das informações do seu comportamento e compilar por meio dos pesos sinápticos e dos limiares dos neurônios, tornando assim uma representação própria do sistema a ser solucionado. Por fim, a camada de saída recebe os sinais vindos da última camada escondida e produz um padrão de resposta para a rede.

Já Braga (2007) diz que, dado um número suficientemente grande de unidades intermediárias, é possível formar representações internas para qualquer conjunto de padrões de entrada. Existem vários métodos de determinação de camadas intermediárias e nós. Geralmente é consenso, em meio aos estudiosos da área, que uma rede com uma camada intermediária já é capaz de implementar funções contínuas, e de duas camadas, de qualquer função proposta.

#### **2.4.2 Aprendizado Supervisionado**

O aprendizado supervisionado é amplamente empregado em processos computacionais e treinamento de máquinas. São assim chamados porque as entradas e saídas desejadas são fornecidas pelo próprio usuário, a fim de que a máquina encontre os parâmetros de interligação (pesos), para cada um dos nós e sendo assim, reajuste a cada iteração, buscando uma menor porcentagem de erro, conforme a Figura 14.

Figura 14 – Aprendizado supervisionado.



Fonte: Braga (2007).

Já no aprendizado não supervisionado não há a existência de um “professor” ou supervisor, para esse algoritmo só os padrões de entrada são fornecidos para a máquina e posteriormente é gerada uma saída esperada.

### 2.4.3 *Backpropagation*

Um algoritmo geralmente utilizado na aprendizagem da máquina se dá através de um mecanismo chamado algoritmo BP (*Backpropagation*). Esse algoritmo trabalha em duas etapas: “ida” e “volta”. A ida é utilizada para definir os nós de saídas da rede, ou também chamado de resposta, e a volta utiliza essa resposta definida pela Inteligência Artificial (IA) e compara com a resposta já conhecida (dos exemplos) para o ajuste dos pesos, possibilitando assim resultados mais assertivos na modelagem do problema. Braga (2014) diz que a funcionalidade principal do algoritmo é estimar o erro das camadas intermediárias, por meio de uma estimativa do efeito que causam no erro da camada de saída. Logo, o erro de saída da rede é calculado e este é retroalimentado para as camadas intermediárias, proporcionando o ajuste dos pesos aos valores das conexões entre camadas.

Esse algoritmo tem como principal alvo evitar os chamados mínimos locais, que por sua vez são pontos onde a informação não ajuda a explicar o problema, e se alocada a solução nesses mínimos locais causa uma variação não desejada de resposta ou saída. Sendo assim, o *backpropagation* é um algoritmo muito importante para a obtenção de uma rede MLP que

explique o problema. Desde sua criação esse algoritmo sofre aperfeiçoamentos, e um dos mais importantes foi a inserção do chamado parâmetro *momentum*.

O termo *momentum* se refere a taxa de aprendizado da máquina que pode encontrar caminhos mais rápidos para chegar a uma mesma resposta, ou mesmo uma resposta mais satisfatória, sendo assim, refere-se a atualização dos pesos das conexões dos nós. A *Backpropagation* é uma técnica muito utilizada por diversos *softwares* que trabalham com redes MLP, por ser efetiva e relativamente simples para a resolução de problemas práticos, acelerando o aprendizado em regiões lineares e suprimindo o erro em mínimos locais.

Outro termo importante para a rede MLP é a taxa de aprendizagem, ou mesmo *learningRate*. É ela quem determina o quão rápido a rede vai “aprender” com os exemplos, se colocar um valor muito alto, pode ser que a rede não identifique pontos característicos dos problemas em questão, e no caso de valores muito pequenos, pode ser que a rede perca muito tempo de simulação, tentando aprender dados muito específicos que por sua vez, não são interessantes para a explicação do problema. Por isso, testes são feitos em redes RNA para encontrar a melhor taxa de aprendizagem para a rede obter o resultado com menor erro possível.

Além disso, o número de épocas de treinamento em uma rede neural se refere ao número de vezes que um conjunto de dados de treinamento é passado pela rede durante o processo de treinamento (NIELSEN, 2015). Por exemplo, se tiver um conjunto de dados de treinamento com 1000 amostras e treinar por 10 épocas isso significa que as 1000 amostras foram apresentadas a rede 10 vezes durante o treinamento. Portanto, após a busca pelo embasamento teórico ter sido concluída, a metodologia do trabalho está descrita no próximo capítulo.

### 3 METODOLOGIA

Esta pesquisa foi desenvolvida na área de Engenharia, com o apoio de tecnologia computacional na obtenção das informações, através de dados obtidos em campo e a sua manipulação através do uso de Inteligência Artificial. O tipo de pesquisa é, portanto, quantitativo, porque serão utilizados para a análise os dados numéricos obtidos do *software* de monitoramento de geração de energia solar, e demais parâmetros necessários.

Neste contexto, apresenta-se como uma abordagem multidisciplinar tendo em vista que traz informações das áreas de engenharias e computação e de métodos quantitativos, com levantamento de dados numéricos e posterior análises deles para a explicação de um problema. É importante destacar que essa pesquisa é do tipo aplicada, buscando a solução de um problema prático, e é descritiva pela observação do comportamento de sistemas fotovoltaicos.

Os dados possuem origem mista, com dados primários e secundários, utilizando-se tanto de dados do Instituto Nacional de Meteorologia (INMET), quanto de dados da pesquisa obtidas em campo, fornecidos por meio de *software* de monitoramento de geração solar. Esta pesquisa também tem caráter experimental, ao passo que faz-se necessária, além da coleta dos dados, a seleção deles, a fim de elucidar as variáveis necessárias para a explicação do problema, trabalhando, enfim, com base hipotético-dedutiva.

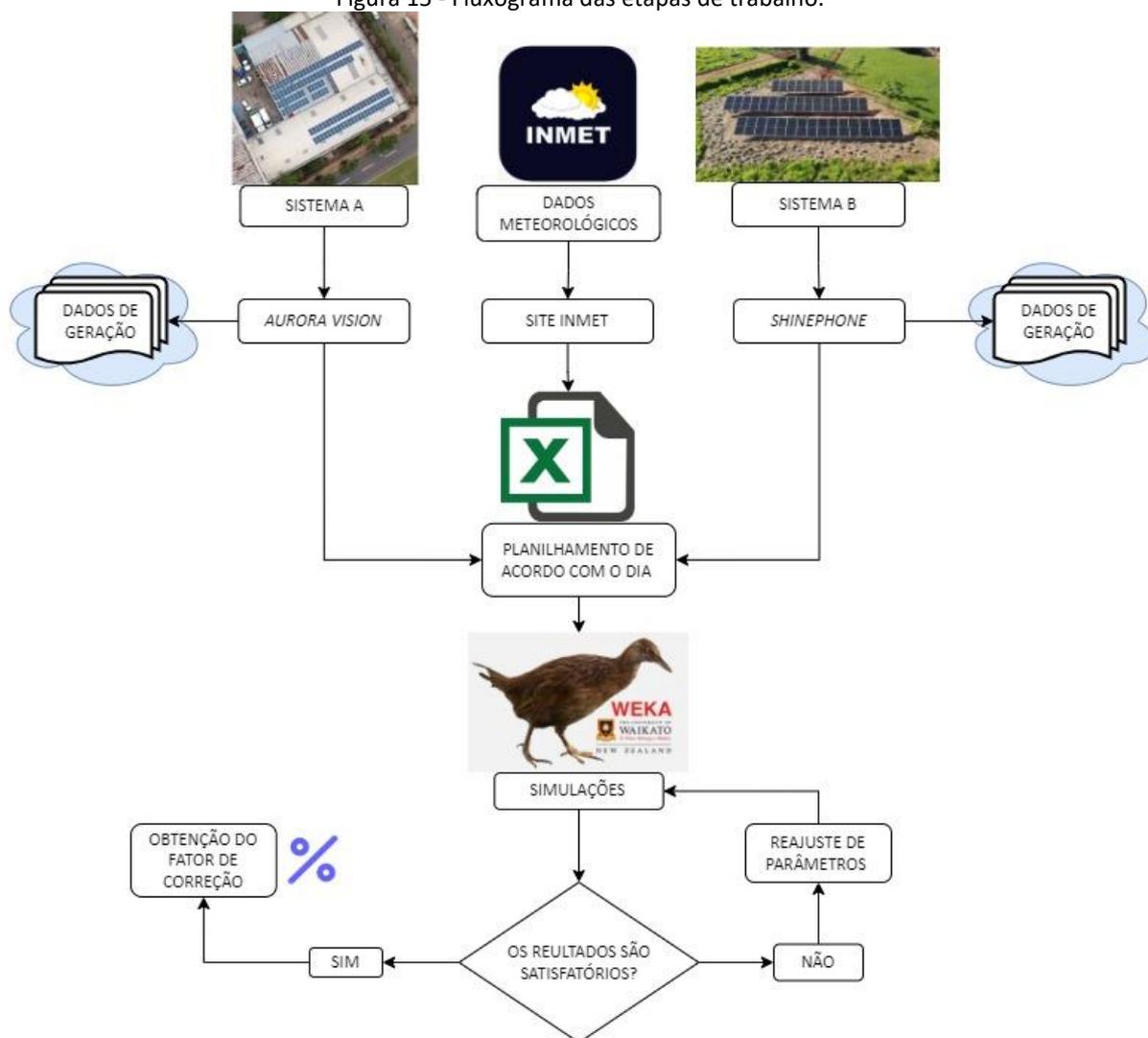
O método deste trabalho para se obter o fator de sobredimensionamento de projeto de sistema fotovoltaico foi dividido em quatro etapas:

- Etapa 1: determinação do local de projeto, onde há perdas de geração;
- Etapa 2: levantamento dos dados, tanto sobre os parâmetros solar, quanto das variáveis climáticas;
- Etapa 3: determinação de variáveis significativas para o estudo, com modelagem computacional no *software* WEKA;
- Etapa 4: finalização com a análise dos resultados obtidos.

A Figura 15 mostra um fluxograma de como foi realizado o trabalho, desde a análise dos dados até a obtenção do fator de sobredimensionamento. Neste fluxograma os dados

foram obtidos de um sistema fotovoltaico, denominado de Sistema A, e de uma outra usina fotovoltaica, denominado Sistema B, além de dados fornecidas pelo INMET. Os dados do sistema A e sistema B ficam armazenados em nuvem, e cada um deles é obtido por *softwares*. Esses dados foram formatados em planilhas, no aplicativo Excell, e foram inseridos no *software* WEKA (*Waikato Environment for Knowledge Analysis*), onde foi implementada a rede neural artificial MLP (*Multi Layer Perceptron*). Várias simulações foram realizadas, com mudanças de arquitetura da rede MLP e parâmetros, até ser encontrada a melhor configuração, com menor erro. O resultado apresenta o fator de sobredimensionamento.

Figura 15 - Fluxograma das etapas de trabalho.



Fonte: próprio autor.

As etapas são descritas com mais detalhes nas seções seguintes.

## **Etapa 1: Local de projeto**

Na geração de sistemas fotovoltaicos, após pesquisas em algumas bibliografias como Araújo *et al.* (2020), Mejia e Kelissl (2013) e Tonolo (2019) e baseado em experiência profissional, é possível afirmar que em alguns meses do ano há um decaimento na geração de energia, que até pode ser significativo, em relação ao “dimensionamento padrão previsto” em projetos.

Esse efeito, portanto, é notado há anos por alguns pesquisadores da área, com observação de decaimento mais intenso em regiões afetadas por períodos longos de estiagem. No entanto, ainda não há um consenso sobre como esse decaimento pode ser dimensionado em projetos, para essas regiões.

O Mato Grosso do Sul é um dos estados onde acontecem esses períodos longos sem chuvas, e por isso uma análise de dados de geração em dois locais diferentes foi realizada. A primeira instalação denominada de instalação A, tem um sistema fotovoltaico de 55,68 kWp instalado na cidade de Campo Grande – MS, conforme a Figura 16. Esta instalação foi escolhida por se encontrar em frente a um posto de coleta de dados meteorológicos do Instituto Nacional de Meteorologia (INMET), e sendo assim, possuir dados apurados de condições climáticas dos últimos anos. Com o estudo dessa instalação, busca-se retratar outros sistemas instalados em áreas urbanas.

Figura 16 – Imagem aérea da Instalação A.



Fonte: próprio autor.

O outro sistema fotovoltaico escolhido, para as análises, está instalado em uma área rural, no município de Bela Vista – MS. Tal sistema foi projetado para a potência de 18,45 kWp e será nomeado de Instalação B, conforme Figura 17. Esse sistema foi escolhido por também possuir uma estação do INMET, com dados históricos da região próxima à sua localização e por sua vez, poder retratar uma análise rural do problema em estudo.

Figura 17 – Imagem aérea da Instalação B.



Fonte: próprio autor.

## **Etapa 2: Levantamento e Coleta de dados.**

Dados de geração elétrica fotovoltaica foram coletados para os sistemas A e B, através de seus respectivos *softwares* de monitoramento. Para o Sistema A, como o inversor é da fabricante ABB, os dados do monitoramento ficam contidos na nuvem da fabricante, e podem ser acessados pelo *site* chamado *Aurora Vision*. O sistema tem sido monitorado desde a instalação e possui os dados que podem ser facilmente solicitados, e por sua vez, quando baixados já se encontram alocados em uma planilha para o estudo. Os possíveis dados coletados pelo *software Aurora Vision* são: tensão, corrente, potência de geração, potência instantânea e outros mais. A extensão do arquivo é *CSV* e o relatório pode ser obtido em planilha do *Excell*.

Já para o sistema B, o fabricante do inversor é a *Growatt* e assim como a ABB, armazena dados em nuvem e estes dados de geração do sistema podem ser facilmente solicitados e já se encontram planilhados, ficando à disposição do usuário para análises em seu *site*. Os dados possíveis de serem obtidos assim como no outro *software* são: tensão, corrente, potência de geração, potência instantânea e outros mais. A extensão do arquivo é *XLS* e o relatório pode ser obtido em planilha do *Excell*.

Também foram coletados dados meteorológicos do INMET que foram posteriormente planilhados pelo CEMTEC (Centro de Monitoramento do Tempo e do Clima de Mato Grosso do Sul) de ambas as localidades, Campo Grande – MS e Bela Vista- MS. Os dados meteorológicos ficam à disposição de qualquer pessoa que acesse o *site* do INMET. Os dados possíveis de serem obtidos são: rajada de vento, milímetros de chuva, temperatura instantânea, temperatura máxima, temperatura mínima, umidade instantânea, umidade máxima, umidade mínima, velocidade do vento, e outras mais. A extensão do arquivo é *XLSX* e o relatório pode ser obtido em planilha do *Excell*.

De posse desses dados, tanto os dados de geração de cada sistema como os dados fornecidos pelo INMET/CEMTEC, e estes alocados em uma mesma planilha de acordo com suas respectivas datas, foi possível analisar a geração real e a pretendida, com o cruzamento de dados climatológicos, que posteriormente passaram por simulações.

Sendo assim, algumas variáveis foram analisadas para possível explicação do problema. Essas variáveis foram inicialmente selecionadas, sendo estas consideradas importantes para encontrar o fator de sobredimensionamento. Diversos testes foram realizados com todas as variáveis disponíveis, tanto dos *softwares* dos equipamentos quanto dos dados climáticos. A Figura 18 contém as variáveis inicialmente selecionadas e suas fontes de informação.

Figura 18 – Variáveis pesquisadas e consideradas possivelmente relevantes para a solução do problema.

#	Variável	Nome	Fonte
1	Geração Real	GEREAL	PLATAFORMA DO INVERSOR
2	Geração estimada	ESTIMADO	PVwatts
3	Rajada de vento	RAJVENTO	INMET
4	Milímetros de chuva	CHUVA	INMET
5	Temperatura instantânea	TEMPINST	INMET
6	Temperatura máxima	TEMPMAX	INMET
7	Temperatura mínima	TEMPMIN	INMET
8	Umidade instantânea	UMIDINST	INMET
9	Umidade mínima	UMIDMIN	INMET
10	Umidade máxima	UMIDMAX	INMET
11	Velocidade do vento	VELVENTO	INMET

Fonte: próprio autor.

Para um melhor entendimento destes dados foi feito, inicialmente, um estudo estatístico com auxílio do *software RStudio*, em cada um ou conjuntos de variáveis, onde serão elencadas as seguintes informações:

- Histograma dos dados;
- Mediana;
- Média;
- Desvio padrão;
- Correlação;
- Equidispersão.

Dessa forma, foi possível verificar, por exemplo, o relacionamento entre as variáveis. Quando duas delas estão correlacionadas, significa que variam juntas. Uma correlação positiva significa que “pontuações” altas em uma estão associadas a pontuações altas em outra e as pontuações baixas em uma estão associadas a pontuação baixa em outra. Por outro

lado, uma correlação negativa significa que pontuações altas na primeira estão associadas a pontuações baixas na segunda e vice-versa (Schmuller, 2019).

Analisando os dados pela correlação, por exemplo, pode-se identificar que algumas variáveis têm o mesmo comportamento e pode-se considerar que apenas uma delas seria suficiente como entrada de dados, no sistema de aprendizado a ser implementado. Nota-se também, que algumas outras variáveis não consideradas relevantes até o momento, poderão, se necessárias, serem inseridas.

Posteriormente, os dados passaram por tratamento e novo planilhamento, já que o *software WEKA* faz processamento diferente do *software RStudio*. Sendo assim, várias simulações foram feitas, visando a melhor obtenção das respostas, para então serem simulados via Rede Neural Artificial do tipo MLP.

### **Etapa 3: Implementação da Rede Neural Artificial do tipo MLP**

Para a implementação da Rede Neural Artificial do tipo MLP foi utilizado o pacote de *software Weka (Waikato Environment for Knowledge Analysis)*, desenvolvido na Universidade de Waikato na Nova Zelândia, sendo esse de código aberto sob a licença pública *General Public License (GNU)*.

Esse *software* pode ser classificado como uma coleção de algoritmo de aprendizado de máquinas, que contém ferramentas para preparação, classificação, regressão e visualização de dados (WEKA, 2011). A escolha do Weka ocorreu devido à necessidade de se implementar uma Rede Neural Artificial do tipo MLP para a análise dos dados fornecidos e gerar informações estatísticas a partir dos resultados encontrados.

Para este trabalho foram construídas várias redes MLP, cujos parâmetros utilizados seguiram a heurística descrita a seguir e, quando necessário, é apresentada uma breve justificativa (BARBOSA, 2004).

- Número de camadas: 3 (totalmente conectadas na topologia original) observando que quando a função consiste em um número finito de pontos ou, quando é

utilizada uma função contínua, uma rede de três camadas é capaz de aprender esta função (BRAGA *et al.*, 2014).

- Número de neurônios da camada de entrada: O número de neurônios na camada de entrada, normalmente, corresponde ao número de variáveis de entrada do processo a ser modelado. Para este trabalho cada neurônio corresponde a uma variável que influencia na geração fotovoltaica e que poderia alterar o projeto da geração estimada (ESTIMADO). Atualmente, estão sendo consideradas 11 variáveis, então serão um neurônio para cada variável, ou seja, 11 neurônios na camada de entrada.
- Número de neurônios na camada de saída: serão feitos testes para observação da melhor aprendizagem, no entanto, como o alvo do trabalho será o fator de sobredimensionamento será buscado apenas 1 neurônio para a saída desejada.
- Número de padrões necessário no conjunto de treinamento: Este valor depende da quantidade de exemplos disponível. Neste trabalho, serão realizados testes com 66 % das informações para treinamento e os outros 33 % divididos entre validação e teste da rede como é usual para o software utilizado.
- Número de neurônios na camada intermediária: Na literatura encontram-se várias heurísticas para definir a quantidade e solucionar tal problema. Será utilizada inicialmente a heurística proposta por Barbosa (2004), que considera que o número de neurônios da camada intermediária depende do número de exemplos utilizado no treinamento. Dessa forma, para definir a quantidade de ligações sinápticas entre dois neurônios, encontra-se o Número Total de Pesos (NTP) que, caso possua milhares de exemplo no treinamento será definido pela Equação 1 para poucos exemplos, e Equação 2 , como será utilizado nesse trabalho.

$$NTP = \frac{NET}{2}, \text{ para poucos exemplos}$$

Equação 1

$$NTP = \frac{NET}{10}, \text{ para vários exemplos}$$

Equação 2

onde:

NTP: Número Total de Pesos

NET: Número de exemplos de treinamento (linhas de entrada)

- Algoritmo de treinamento: *backpropagation* com momento.
- Ordem de apresentação dos exemplos: por período de tempo.
- Taxa de aprendizagem: adaptadas para a melhor resolução com menor erro.
- *Momentum*: adaptadas para a melhor resolução com menor erro.
- Número de épocas: adaptadas para a melhor resolução com menor erro.

Posteriormente, foram considerados outros modelos de treinamento supervisionado, novas simulações foram feitas tanto por tentativa e erro, quanto pela heurística elencada por Hagan *et. al.* (2014) que evita o sobre ajuste, que por sua vez, evita as não generalizações.

O algoritmo de treinamento utilizado foi o *backpropagation* e a taxa de aprendizagem, *momentum* e número de épocas foram adaptadas para a melhor resolução com menor erro absoluto médio e quadrático médio.

Resumidamente, após vários testes e configurações foi implementada uma Rede Neural MLP, com menor erro, para definir qual o valor de sobredimensionamento deverá ser aplicado ao projeto, conforme condições meteorológicas do local a ser implantado o sistema de geração fotovoltaica.

#### **Etapa 4: Análise dos resultados**

Nessa etapa, com os resultados da melhor configuração encontrada para a rede MLP, foi feita a análise dos resultados. Com possibilidade de verificar se o uso da rede MLP foi adequado para esse problema, se o resultado encontrado em uma região rural foi muito diferente da região urbana. Também foi possível verificar se há muita diferença de geração de energia com projeto implementado com ou sem correção. Enfim, após os dados obtidos e processados, diversas informações foram geradas e que posteriormente poderão ser apresentadas aos projetistas e comunidade acadêmica.

## 4 RESULTADOS

Nesse capítulo serão apresentados os resultados obtidos no estudo.

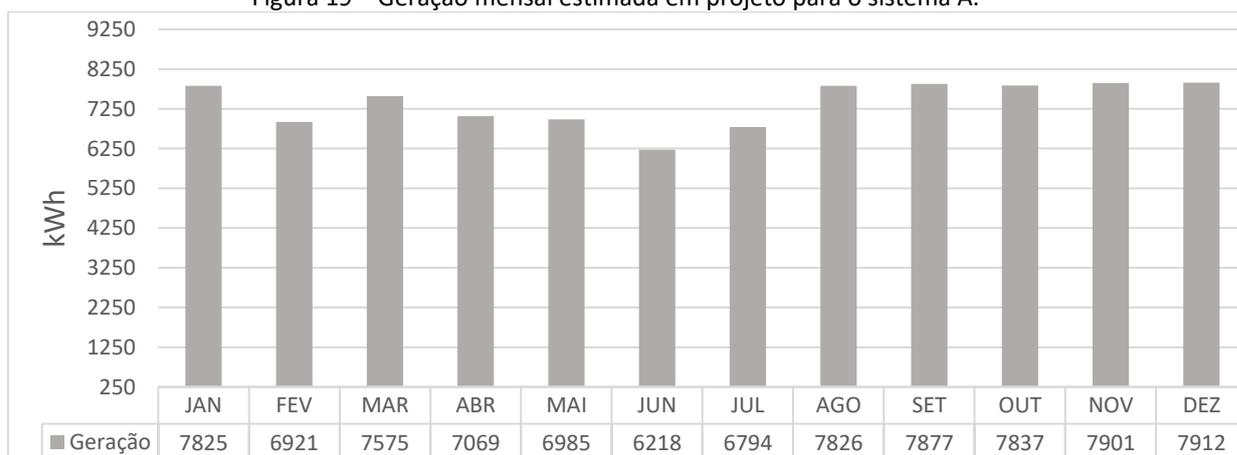
### 4.1 Dimensionamento do sistema

Tais instalações foram dimensionadas observando suas características de localidade, inclinação, azimute e outros fatores relevantes para a geração solar.

#### Sistema A

O sistema A foi instalado em fevereiro do ano de 2021. O consumo da instalação em questão possuía uma média anual em torno de 6.000 kWh/mês. Sendo assim, projetou-se a instalação de um sistema sobredimensionado em torno de 7.400 kWh/mês, atendendo às futuras necessidades de energia. Foi observado um Azimute de 340 graus Norte na instalação, e uma inclinação em média de 10 graus de telhado, bem como a latitude de -20,51 e longitude de -54,62. Para o sistema foram consideradas as perdas de 12% provenientes dos cabos, das conexões, do inversor, do transformador e da *stringbox*. As simulações foram realizadas no programa livre *PVwatts* do *National Renewable Energy Laboratory* (NREL) e a potência de 55,68 kWp foi determinada para atendimento ao cliente gerando uma média anual de 7.395 kWh/mês, conforme pode ser observado na Figura 19, que apresenta os dados obtidos do programa *PVwatts*.

Figura 19 – Geração mensal estimada em projeto para o sistema A.



Fonte: adaptado pelo autor do simulador *PVwatts National Renewable Energy Laboratory* (NREL).

Sendo assim, para o sistema foram dimensionados 128 painéis da marca *Longi* de 435W R4-72HPH-435M monocristalinos (totalizando 55,68 kWp), ligados em oito strings de

16 painéis cada e um inversor *on-grid* do fabricante ABB, série PVS-50-TL-SX - 380 V com 3 *maximum power point tracking*<sup>1</sup> (mppt), conforme Figura 20. O sistema fotovoltaico foi instalado, juntamente com um sistema de aterramento que não havia na edificação e foi providenciada a documentação junto a concessionária, no que diz respeito a projetos e execução.

Figura 20 – Sistema A - Inversor e *Stringbox*.



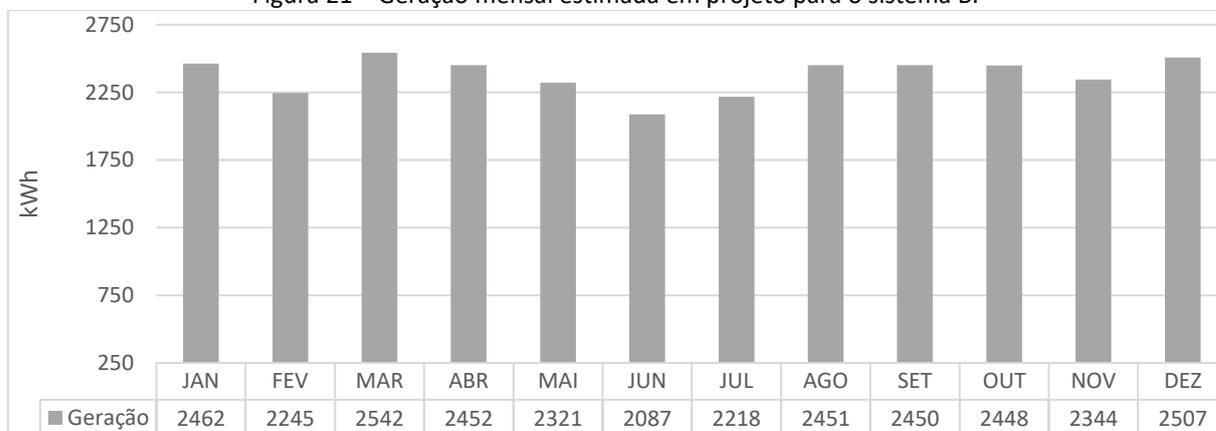
Fonte: próprio autor.

### Sistema B

O sistema B, foi instalado em janeiro do ano de 2022. O consumo da instalação em questão possuía uma média anual em torno de 2.200 kWh/mês. Sendo assim, projetou-se a instalação de um sistema em torno de 2.400 kWh/mês, atendendo às futuras necessidades de energia. Foi ajustado um Azimute de 0 graus Norte na instalação, visto que a instalação seria realizada em solo, e uma inclinação de 20 graus, bem como a latitude de -20,99 e longitude de -56,54. Para o sistema foram consideradas as perdas no sistema de 13%, provenientes dos cabos, das conexões, do inversor, do transformador e da *stringbox*. As simulações foram realizadas no programa livre *PVwatts* do *National Renewable Energy Laboratory* (NREL) e a potência de 18,45 kWp foi determinada para atendimento ao cliente, gerando uma média

anual de 2380 kWh/mês, conforme a Figura 21 que apresenta os dados obtidos do programa PVwatts.

Figura 21 – Geração mensal estimada em projeto para o sistema B.



Fonte: adaptado pelo autor do simulador PVwatts National Renewable Energy Laboratory (NREL).

Sendo assim, para o sistema foram dimensionados 41 painéis da marca CANADIAN 450W - CS3W 450MS monocristalinos (totalizando 18,45 kWp), ligados em cinco *strings*, duas de 7 painéis conectados a um inversor *on-grid* do fabricante Growatt, série MIN 5000TL-X 5 kW - 220 V com 2 *maximum power point tracking*<sup>1</sup> (mppt) e três de 9 painéis conectados a um inversor *on-grid* do fabricante Growatt, série MIN 10000TL-X 10 kW - 220 V com 3 *maximum power point tracking*<sup>1</sup> (mppt) conforme a Figura 22. O sistema fotovoltaico foi instalado, juntamente a um sistema de aterramento que não tinha na edificação e foi providenciada a documentação com a concessionária, no que diz respeito a projetos e execução.

Figura 22 – Sistema B - Inversor e Stringbox.



Fonte: próprio autor.

## 4.2 Visualização de dados

Para o melhor entendimento dos dados que seriam trabalhados, foram obtidos os gráficos de histogramas e de correlação, além dos cálculos do desvio padrão, da mediana, da média, do primeiro e terceiro quartil e, da correlação.

### 4.2.1 Filtragem de dados

Inicialmente foi feita uma filtragem dos dados, comum para dados computacionais. Essa filtragem tem o intuito de eliminar dados que não representem a realidade da amostra, como por exemplo, dias em que os sistemas não geraram adequadamente devido à falta de energia elétrica na rede, visto que, o sistema por ser *on-grid* depende da rede da concessionária para seu funcionamento.

No total foram retirados 7 dias no sistema A e 11 dias no sistema B, todos esses apresentavam “cortes” na geração, devido a falhas na rede elétrica da concessionária. Nas sessões seguintes, a análise estatística dos dados será apresentada.

### 4.2.2 Histograma

Para os sistemas A e B, foram feitos histogramas das suas 10 variáveis de entrada e de sua variável de saída (fator de sobredimensionamento). Os Histogramas, como em qualquer análise de dados, buscam o melhor entendimento do objeto de estudo. Além dos histogramas foi feita a análise dos principais indicadores de cada variável: média, mediana, mínimo, máximo, primeiro quartil, terceiro quartil e desvio padrão. Por fim, foi gerado um gráfico da correlação entre as variáveis onde cada variável é relacionada com outra, e mostra o quão correlacionada cada variável é, entre si e entre as outras demais.

#### Sistema A

Para o sistema A instalado na área urbana de Campo Grande MS, foram coletados dados de 681 dias, ao longo de dois anos (anos de 2021 e 2022). A Tabela 2 apresenta um resumo completo dos dados obtidos das 10 variáveis de entrada e da variável de saída.

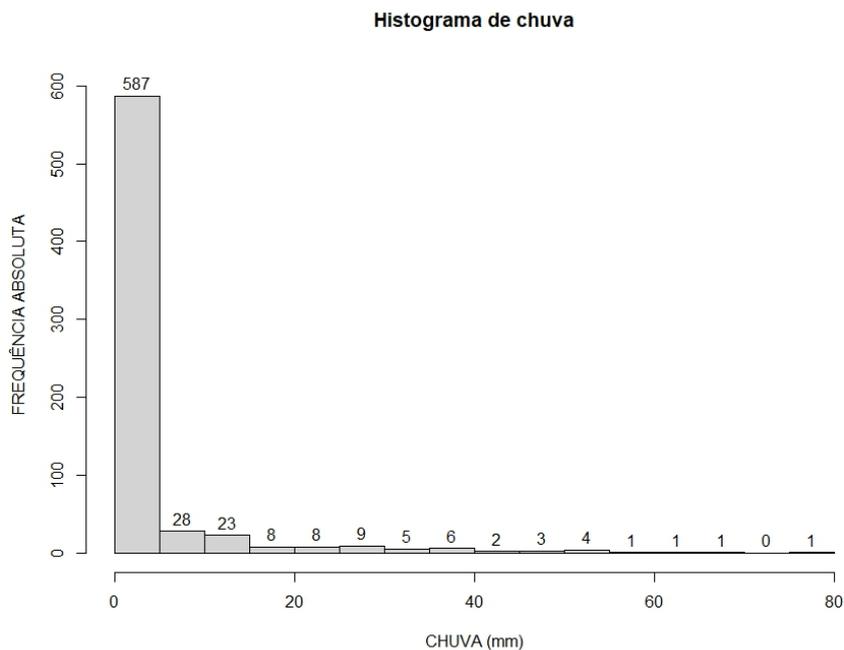
Também foram feitos histogramas das 11 variáveis que se encontram nas Figuras 23 a 34, onde ESTIMADO é a energia estimada de geração em kWh, RAJVENTO é a maior rajada de vento diária em km/h, CHUVA são os milímetros de chuva diários, TEMPINST, TMAX E TMIN, são as temperaturas instantâneas, máximas e mínimas respectivamente, médias diárias em graus Celsius, UMIDINST, UMIDMAX, UMIDMIN que são umidade instantânea, máxima e mínima respectivamente, VELVENTO que é a velocidade do vento em km/h e GEREAL que é a geração real em kWh.

Tabela 2 – Resumo de dados das variáveis do sistema A.

VARIÁVEL	ESTIMADO (kWh)	RAJVENTO (km/h)	CHUVA (mm)	TEMPINST (°C)	TMAX (°C)	TMIN (°C)
Mínimo	192,0	15,12	0	10,12	12,5	5,3
1º quartil	216,1	29,52	0	22,32	28,7	17,5
Mediana	238,4	35,64	0	24,87	31,0	20,0
Média	234,1	37,08	3,4	24,17	30,1	19,5
3º quartil	248,0	43,56	0,6	26,66	32,2	22,5
Máximo	268,6	102,24	75,2	31,10	39,3	26,3
Desvio padrão	22,31	10,59	9,73	3,48	3,64	4,09
VARIÁVEL	UMIDINST (%)	UMIDMAX (%)	UMIDMIN (%)	VELVENTO (km/h)	GEREAL (kWh)	
Mínimo	22	30	7	0,0	42,49	
1º quartil	53	70	30	11,9	209,20	
Mediana	65	80	44	16,2	241,77	
Média	62	78	44	16,3	240,27	
3º quartil	74	88	59	21,4	280,96	
Máximo	96	100	89	45,0	393,85	
Desvio padrão	15,10	14,25	17,23	7,20	64,12	

Observa-se na Figura 23 que dos 681 dias analisados, 587 ficam sem nenhuma chuva ou com menos de 5 mm, demonstrando um período de vários dias de estiagem. Também é possível verificar que o volume de chuva (mm de chuva) é baixo, sendo o maior valor de 80 mm com uma frequência 1.

Figura 23 – Histograma de chuva em (mm).



Conforme mostra a Figura 24 dos 681 dias analisados, a maior concentração das frequências fica para a geração de 220 a 250 kWh. A Figura 25, que mostra a geração real em kWh mostra a maior concentração de 200 a 300 kWh. Mostrando uma paridade nos dados de geração real e estimada.

Figura 24 - Geração estimada em (kWh/dia).

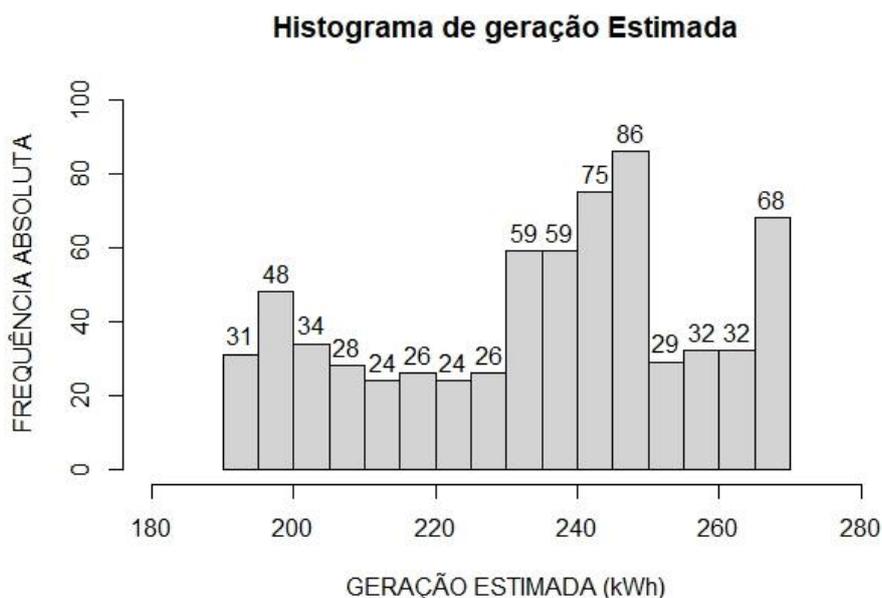
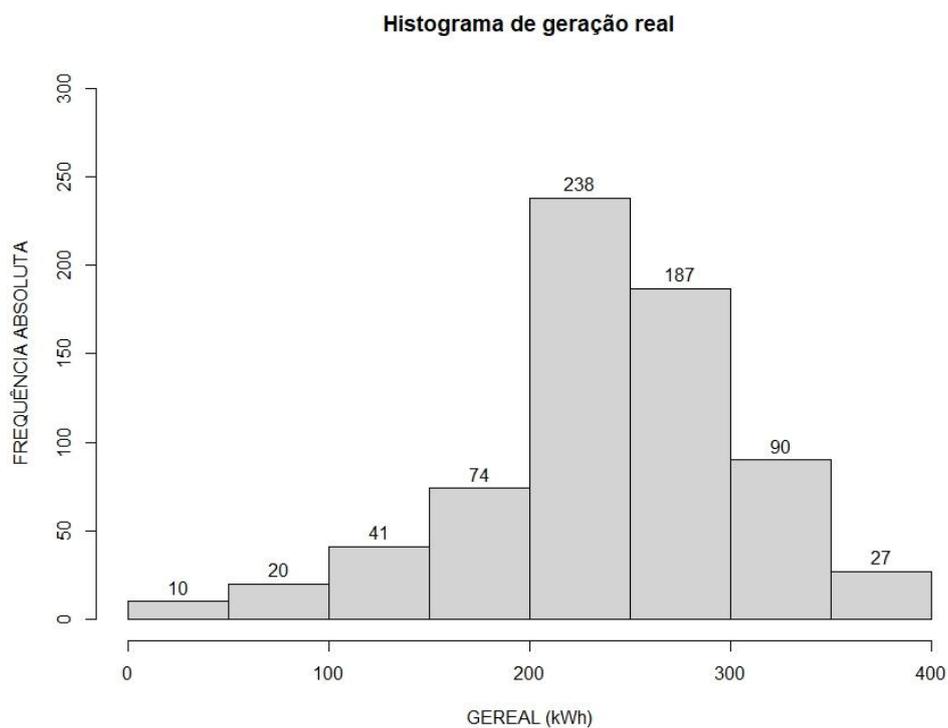
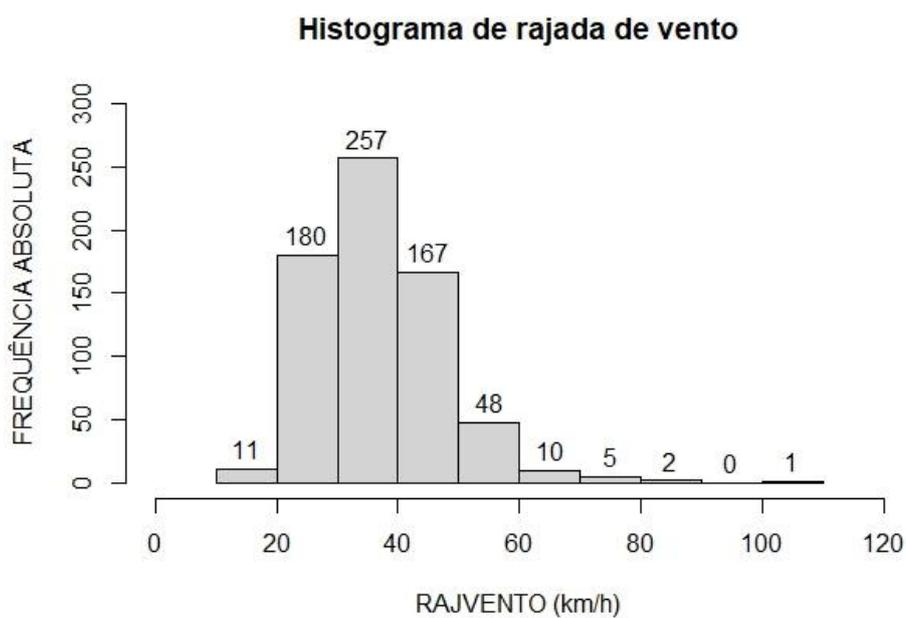


Figura 25 - Histograma da geração real em (kWh/dia).



A Figura 26 mostra uma maior concentração de rajadas de vento de 20 a 50 km/h sendo que em apenas 1 dia no período, foi registrado mais de 100 km/h.

Figura 26 - Histograma de rajada de vento (km/h).



As Figuras 27 a 29 mostram as temperaturas instantâneas, máximas e mínimas, que por sua vez têm as maiores frequências entre 24 e 28 graus Celsius, 28 e 34 graus Celsius e 18 e 24 graus Celsius, respectivamente.

Figura 27 - Histograma de temperatura Instantânea (°C).

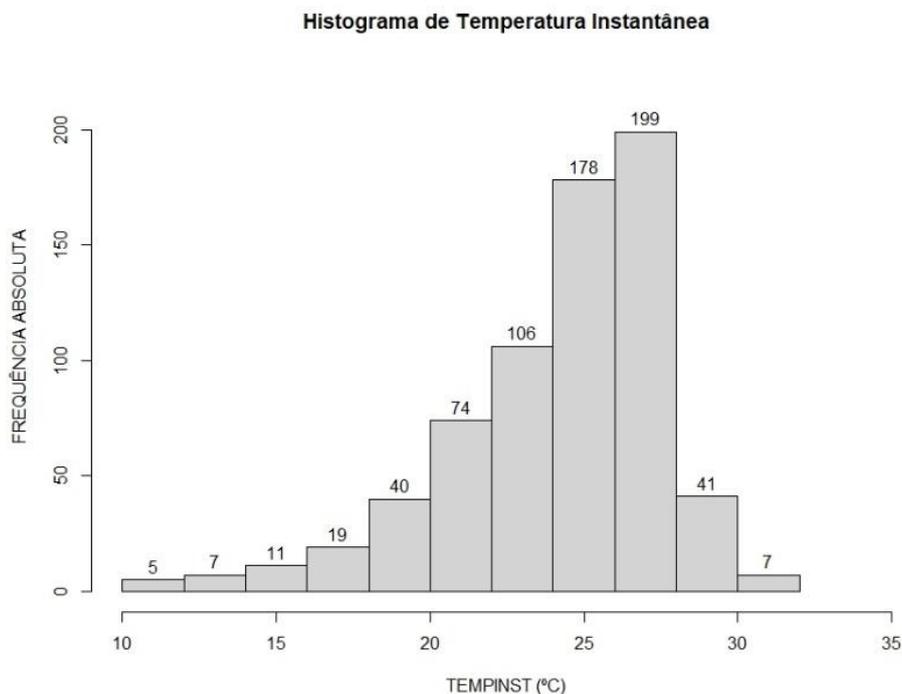


Figura 28 - Histograma de temperatura Máxima (°C).

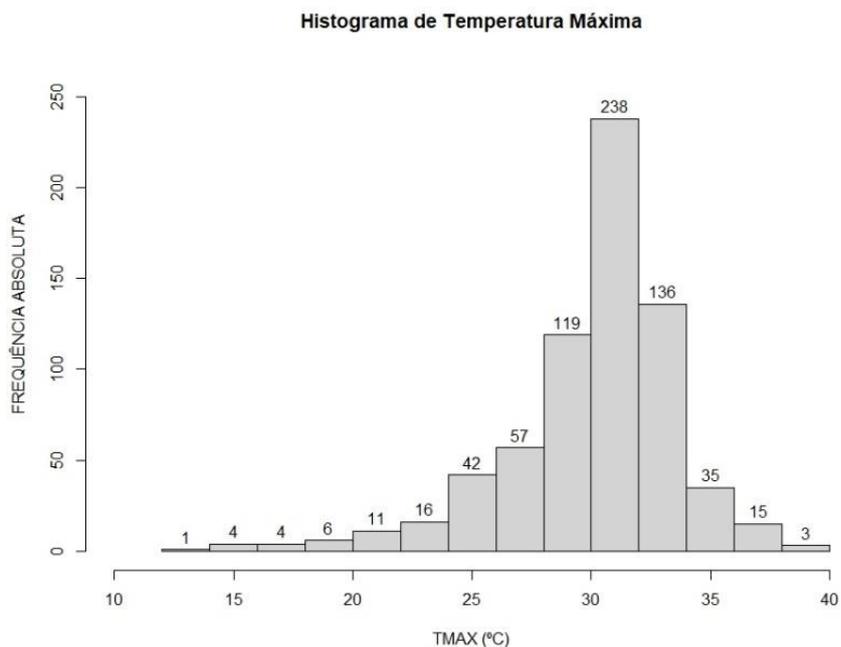
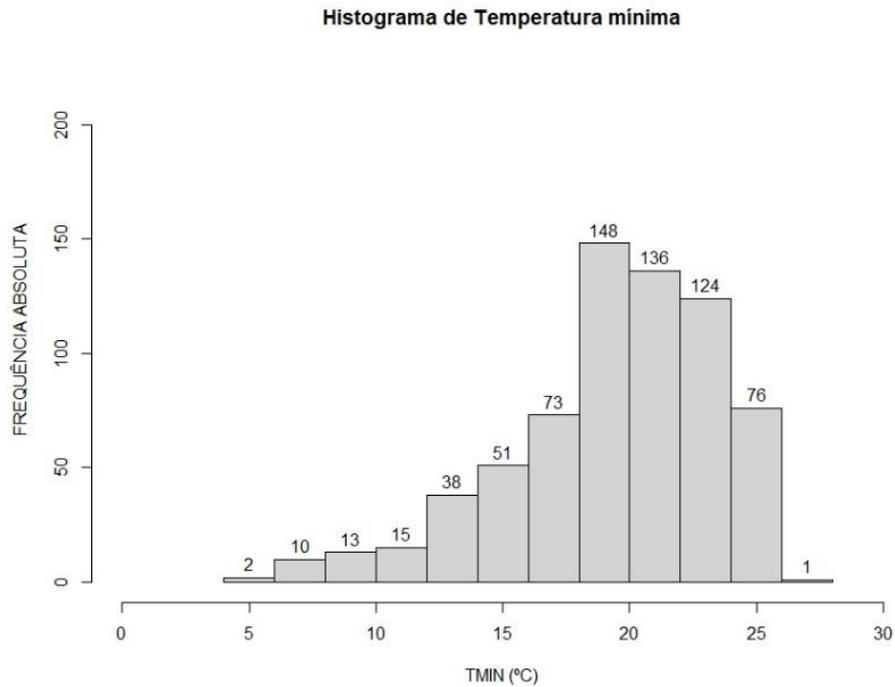


Figura 29 - Histograma de temperatura Mínima (°C).



As Figuras 30 a 32 mostram as umidades instantâneas, máximas e mínimas, que por sua vez têm as maiores frequências entre 55 e 80 por cento, 75 e 85 por cento e 20 e 65 por cento respectivamente.

Figura 30 - Histograma de umidade Instantânea (%).

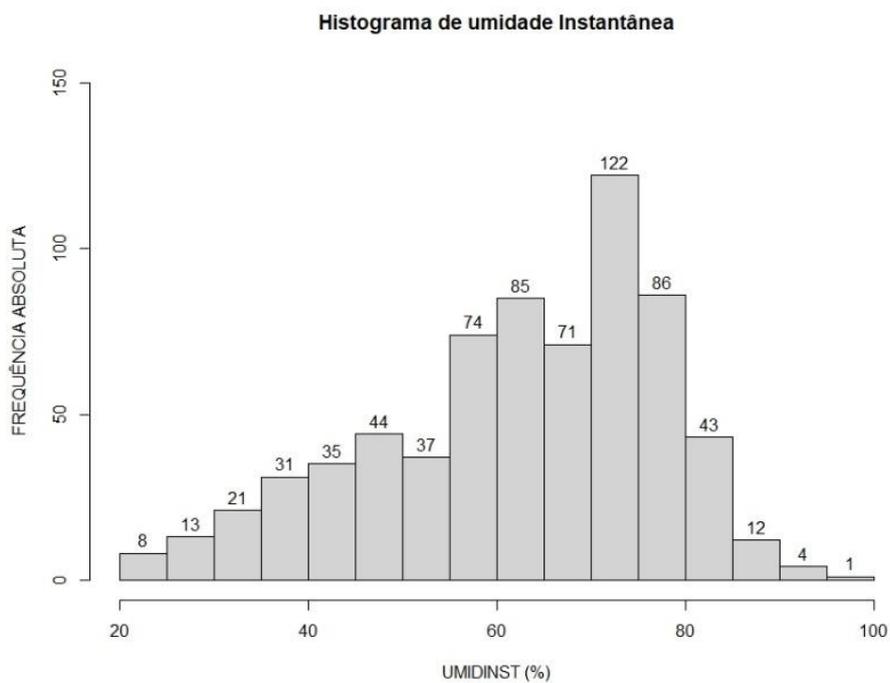


Figura 31 - Histograma de umidade Máxima (%).

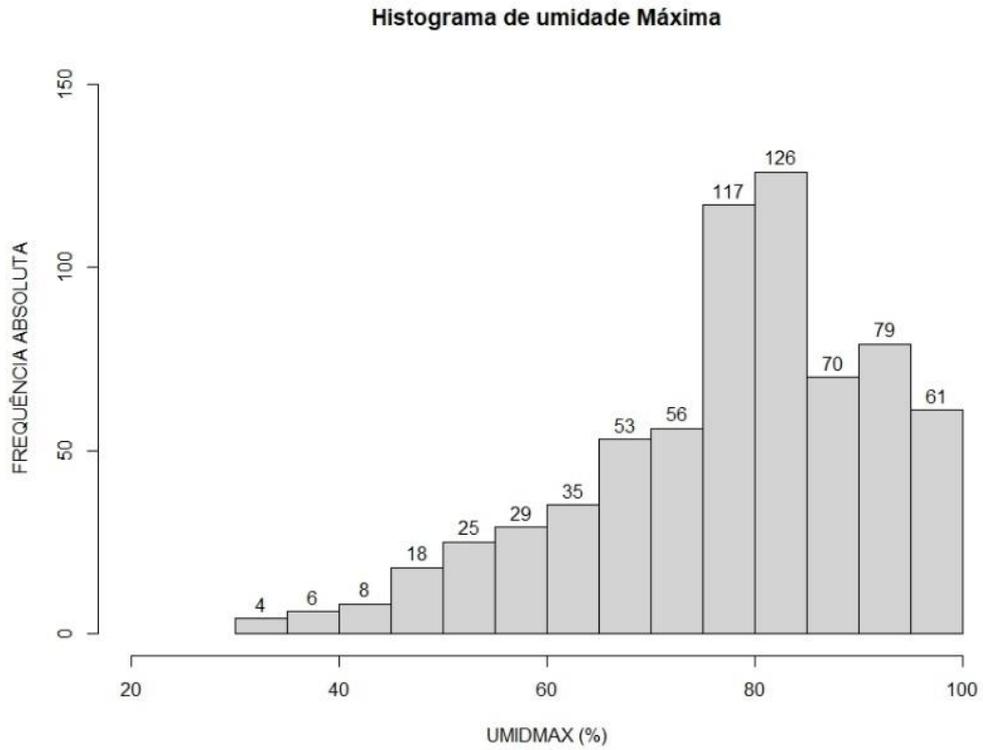
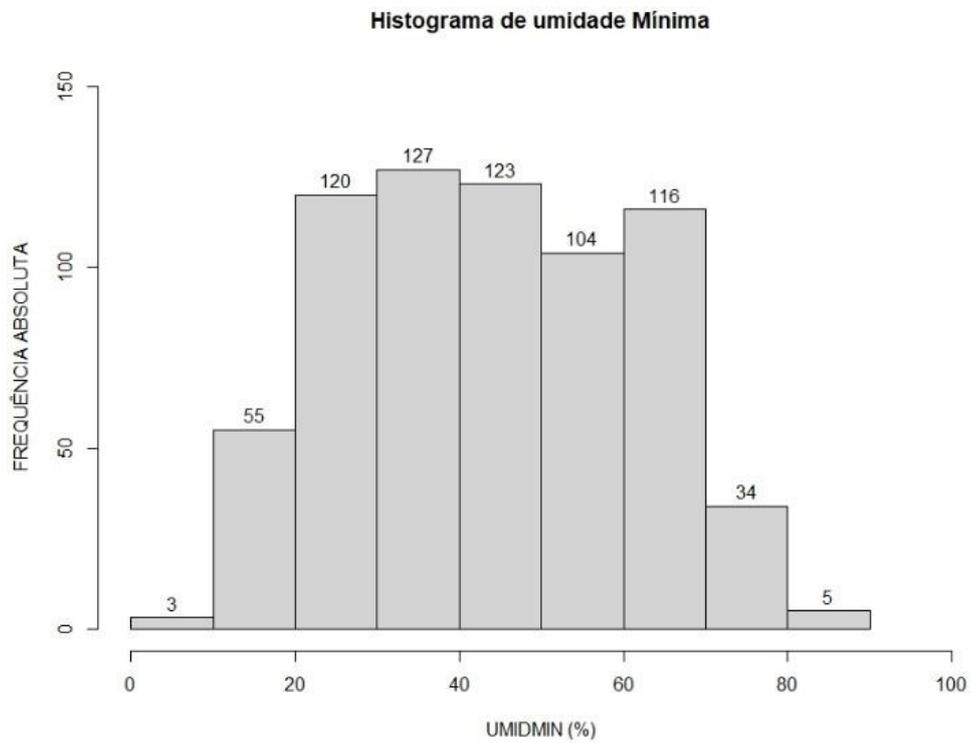
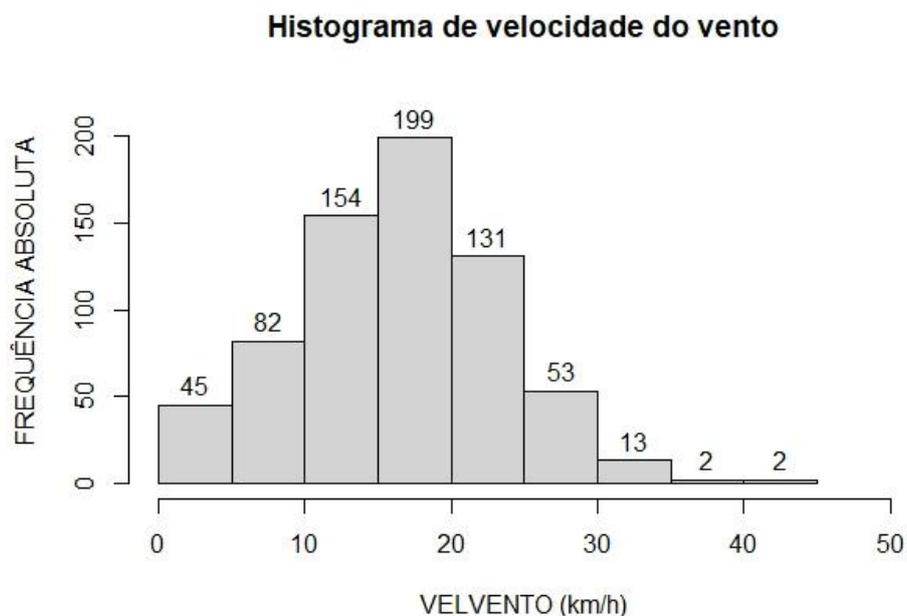


Figura 32 - Histograma de umidade mínima (%).



A Figura 33 mostra uma maior concentração de velocidade média do vento de 10 a 25 km/h sendo que em apenas dois dias no período, foi registrado mais de 40 km/h.

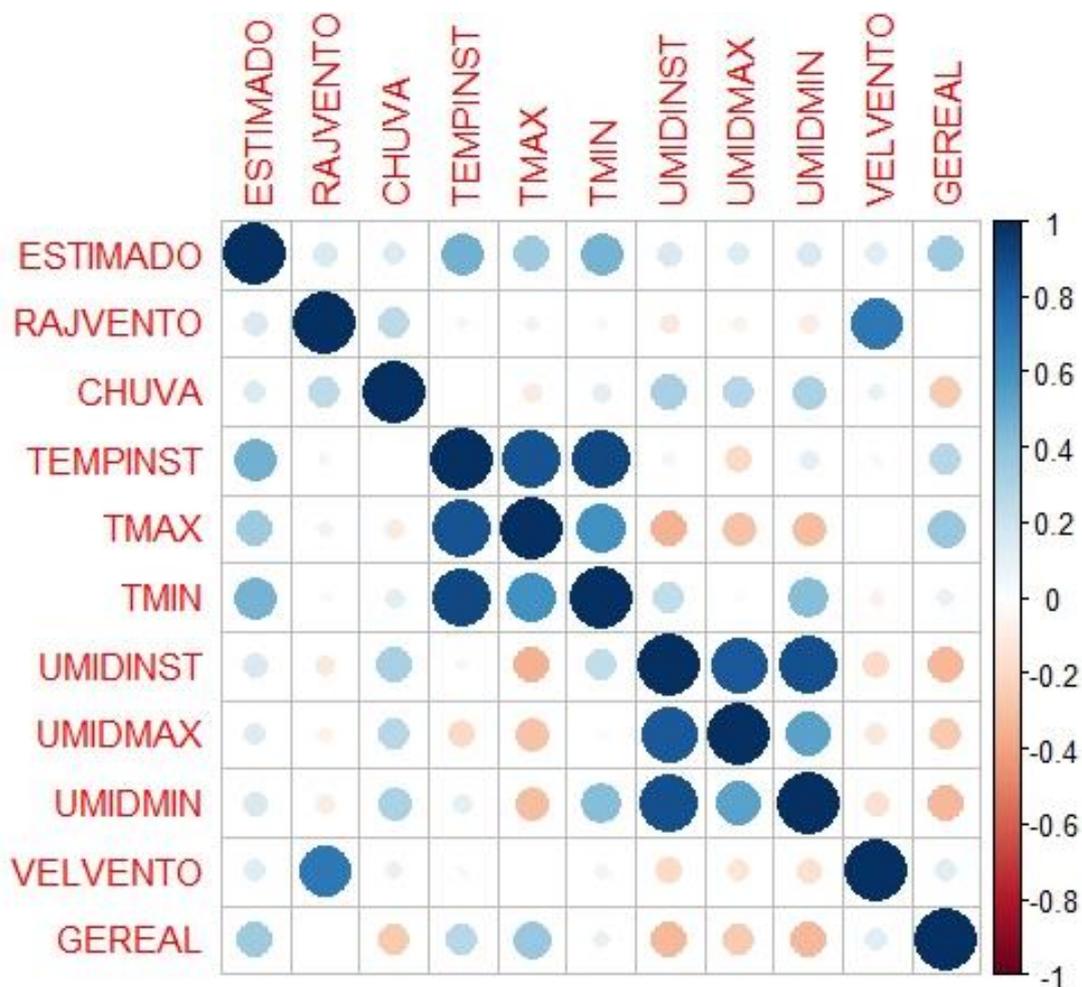
Figura 33 - Histograma de velocidade do vento (km/h).



Por fim para uma melhor análise dos dados e de como se correlacionam foi feito o gráfico de correlação entre as variáveis. Quanto maior a circunferência, maior é a correlação entre as variáveis, quanto mais azul (escuro) mais as variáveis são correlacionadas positivamente, e quanto maior a circunferência e mais vermelho (escuro), maior é a correlação negativa. Portanto, como mostra a Figura 34, para uma hipótese inicial, no sistema A, as variáveis ESTIMADO, CHUVA, TEMPINST, TMAX, UMIDINST, UMIDMAX e UMIDMIN, são importantes para a explicação da variável de saída GEREAL, porque não apresentaram alta correlação.

No entanto, posteriormente uma análise mais pontual foi feita visto que algumas dessas variáveis que possivelmente seriam importantes, têm também correlação alta entre si, o que significa que são redundantes, ou seja, que possivelmente carreguem informações similares para a explicação do problema. Essa verificação foi feita com a RNA implementada.

Figura 34 - Gráfico de correlação entre as variáveis do sistema A.



### Sistema B

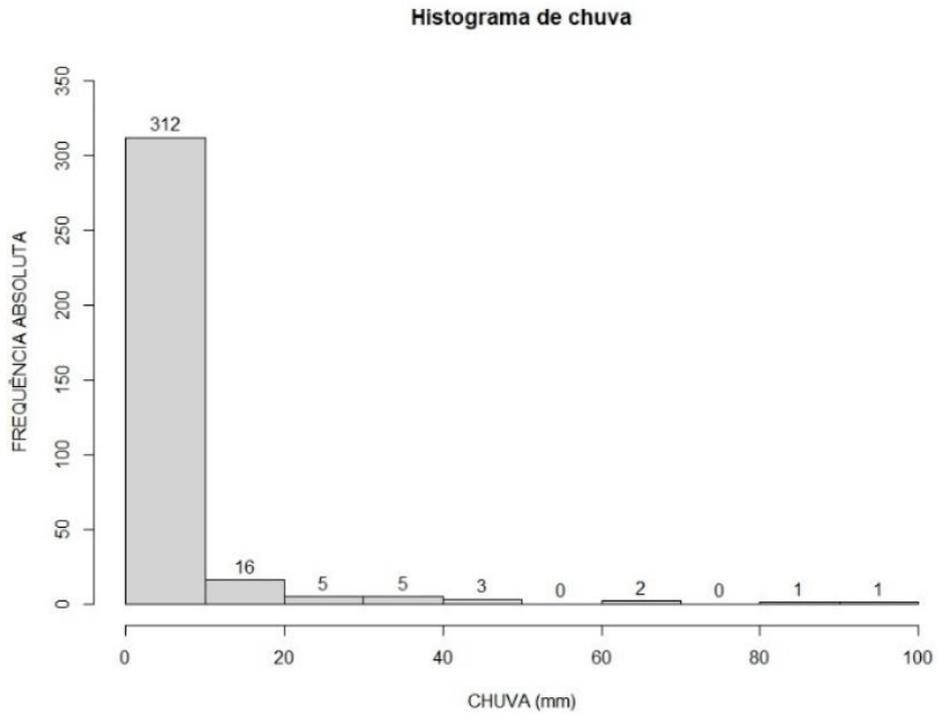
Para o sistema B instalado na área rural da cidade de Bela Vista/MS, foram coletados dados de 334 dias ao longo de um ano. A Tabela 3 mostra um resumo completo dos dados obtidos das 10 variáveis de entrada e da variável de saída. Também foram feitos histogramas das 11 variáveis que se encontram nas Figuras 35 a 46.

Tabela 3 - Resumo de dados das variáveis do sistema B.

VARIÁVEL ->	ESTIMADO (kWh)	RAJVENTO (km/h)	CHUVA (mm)	TEMPINST (°C)	TMAX (°C)	TMIN (°C)
Mínimo	77,4	12,60	0,0	12,25	17,2	4,7
1º quartil	90,6	22,68	0,0	21,27	28,4	15,5
Mediana	94,4	27,36	0,0	24,41	31,8	18,8
Média	92,2	28,91	3,1	24,02	30,8	18,48
3º quartil	95,7	33,12	0,2	27,27	34,0	22,3
Máximo	97,3	33,12	93,4	31,89	39,9	26,6
Desvio padrão	5,31	9,30	10,70	4,12	4,61	4,57
VARIÁVEL ->	UMIDINST (%)	UMIDMAX (%)	UMIDMIN (%)	VELVENTO (km/h)	GEREAL (kWh)	
Mínimo	35	26	6	3,9	16,70	
1º quartil	53	72	26	9,4	68,65	
Mediana	63	78	37	11,9	86,30	
Média	63	76	39	13,3	82,76	
3º quartil	72	83	49	15,5	103,80	
Máximo	89	94	84	52,2	124,00	
Desvio padrão	12,32	12,28	16,76	6,39	26,30	

Observa-se na Figura 35 que dos 334 dias analisados, 312 ficam sem nenhuma chuva ou com menos de 10mm, demonstrando um período de vários dias de estiagem. Também é possível verificar que o volume de chuva (mm de chuva) é baixo, sendo o maior valor entre 80 e 100 mm com uma frequência de dois dias apenas.

Figura 35 – Histograma de chuva em (mm).



Conforme mostra a Figura 36, dos 334 dias analisados, a maior concentração das frequências fica para a geração em torno de 95 kWh. A Figura 37, que mostra a geração real em kWh mostra a maior concentração de 80 a 120 kWh. Mostrando uma paridade nos dados de geração real e estimada.

Figura 36 - Geração estimada em (kWh/dia).

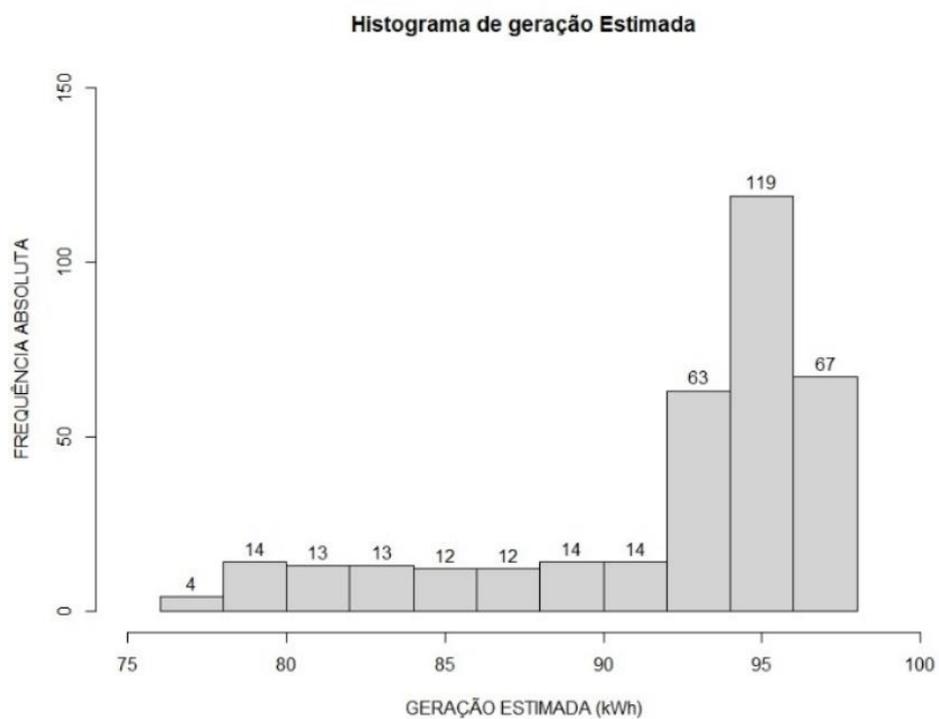
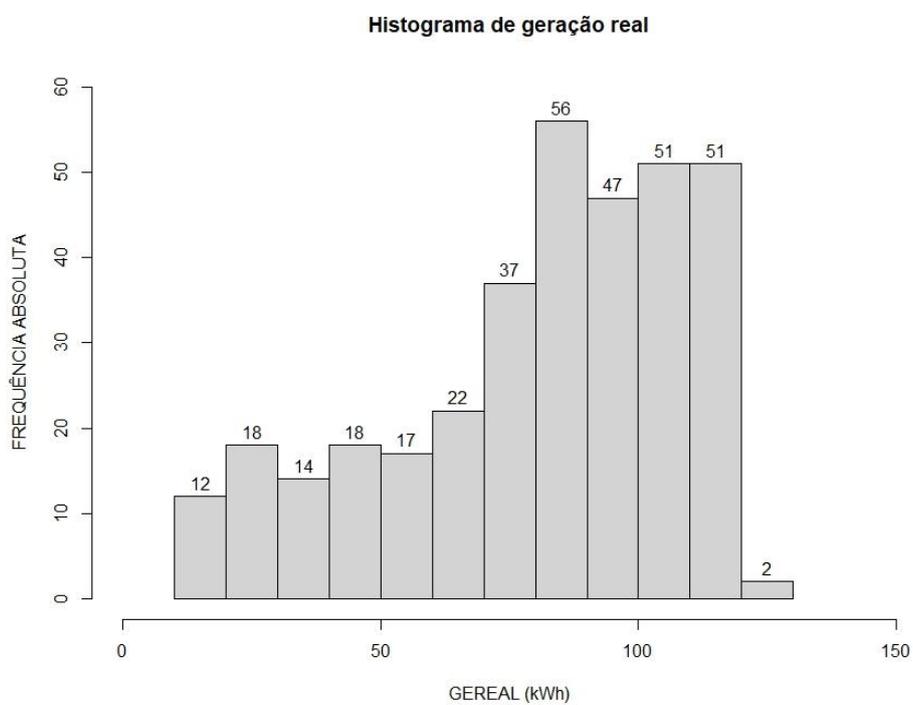
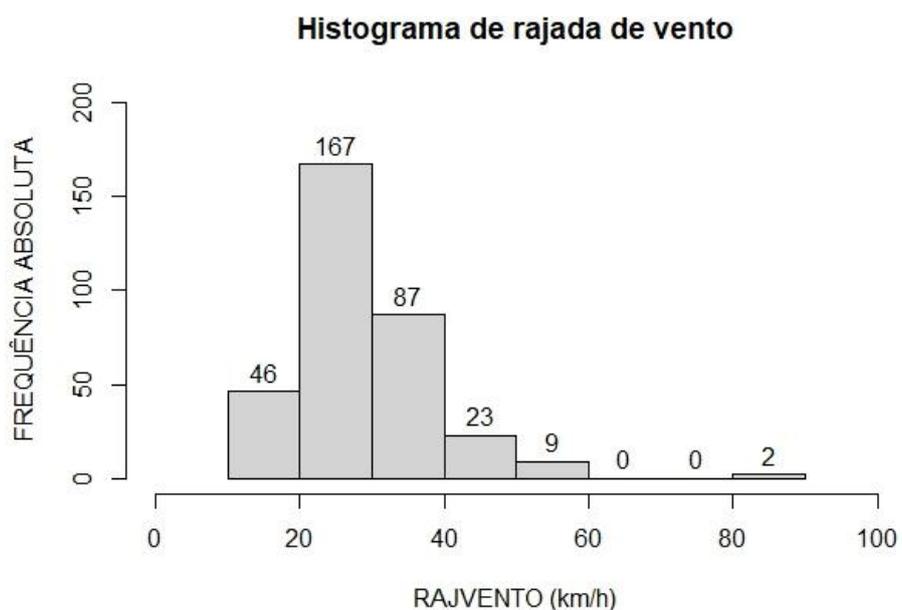


Figura 37 - Geração real em (kWh/dia).



A Figura 38 mostra uma maior concentração de rajadas de vento de 20 a 30 km/h sendo que em apenas dois dias no período, foi registrado mais de 80 km/h.

Figura 38 - Histograma de rajada de vento (km/h).



As Figuras 39 a 41 mostram as temperaturas instantâneas, máximas e mínimas, que por sua vez tem as maiores frequências entre 20 e 30 graus Celsius, 30 e 36 graus Celsius e 14 e 24 graus Celsius, respectivamente.

Figura 39 - Histograma de temperatura Instantânea (°C).

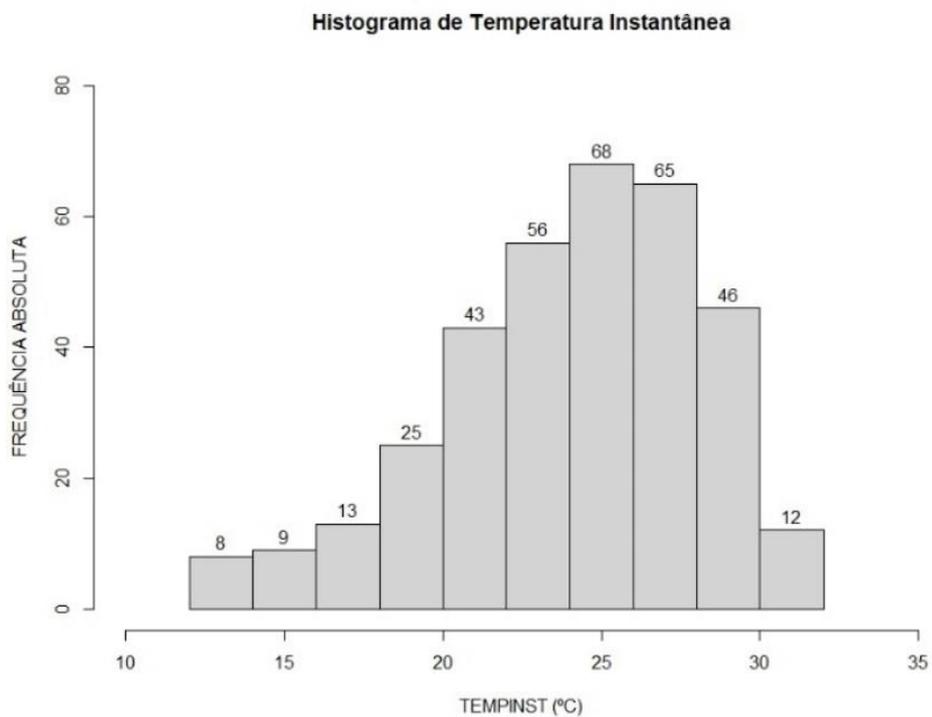


Figura 40 - Histograma de temperatura máxima (°C).

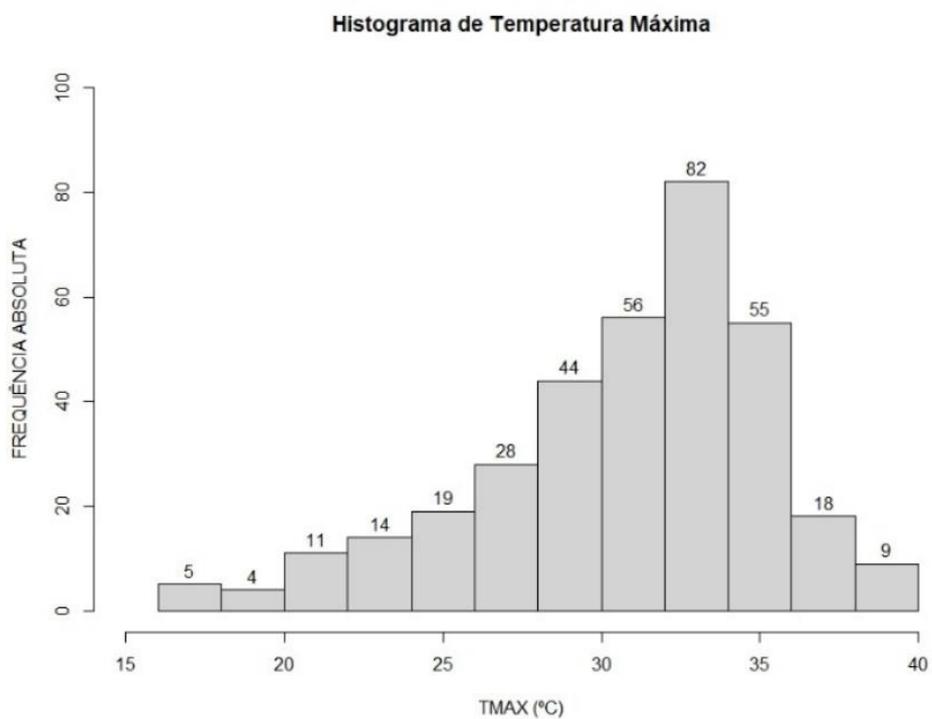
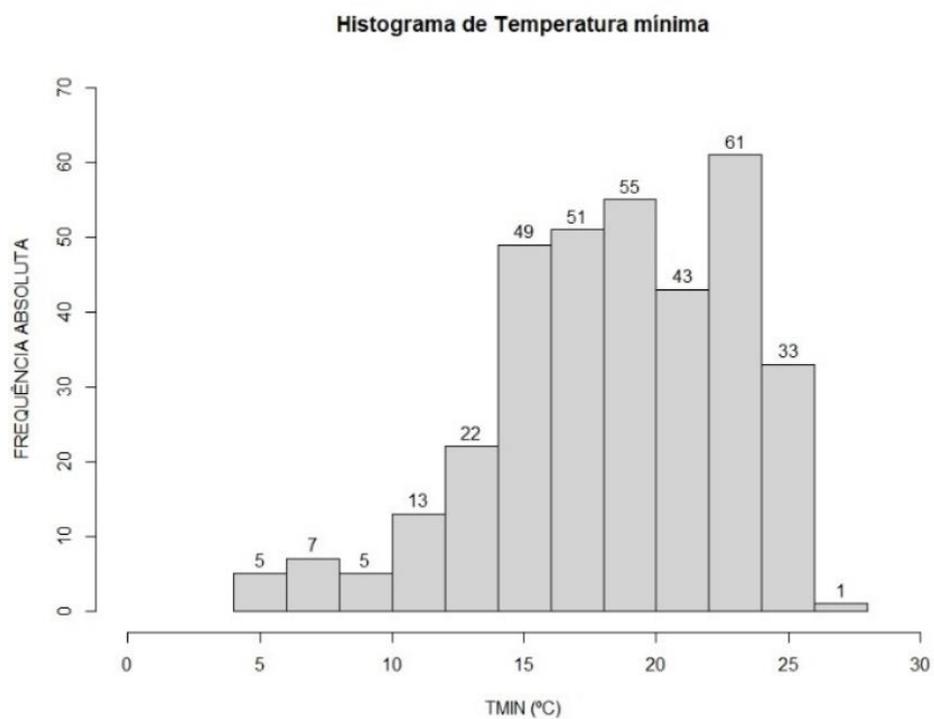


Figura 41 - Histograma de temperatura mínima (°C).



As Figuras 42 a 44 mostram as umidades instantâneas, máximas e mínimas, que por sua vez tem as maiores frequências entre 65 e 70 por cento, 70 e 90 por cento e 20 e 50 por cento, respectivamente.

Figura 42 - Histograma de umidade Instantânea (%).

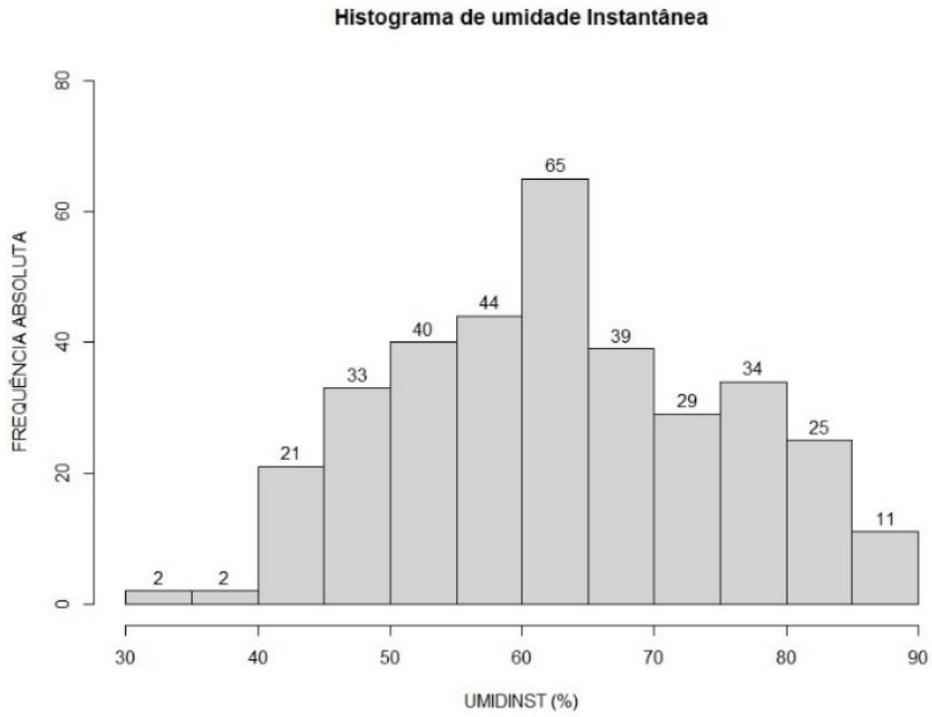


Figura 43 - Histograma de umidade máxima (%).

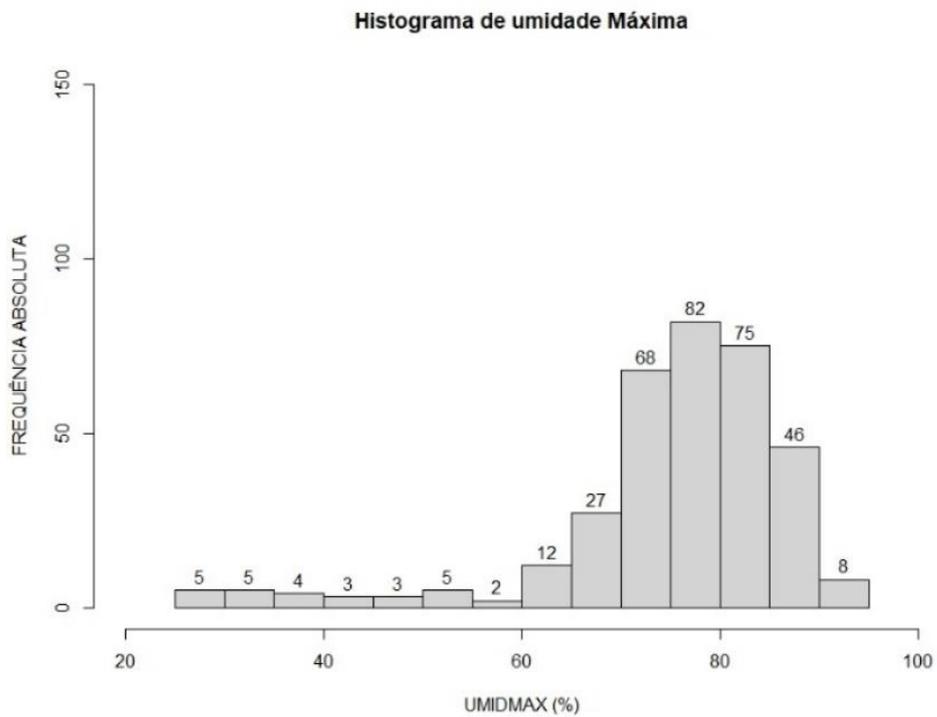
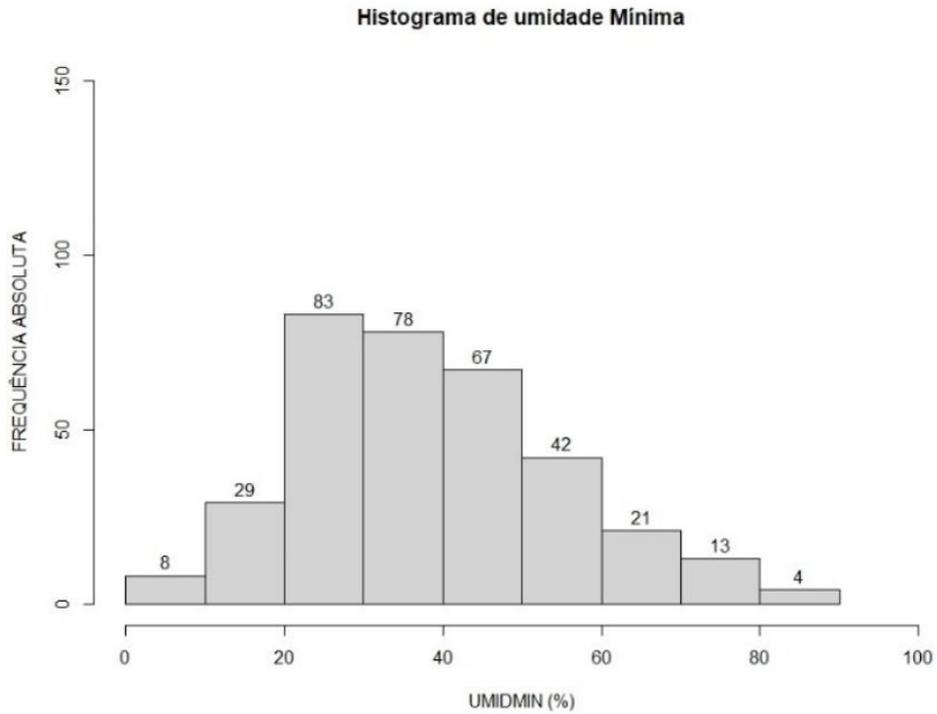
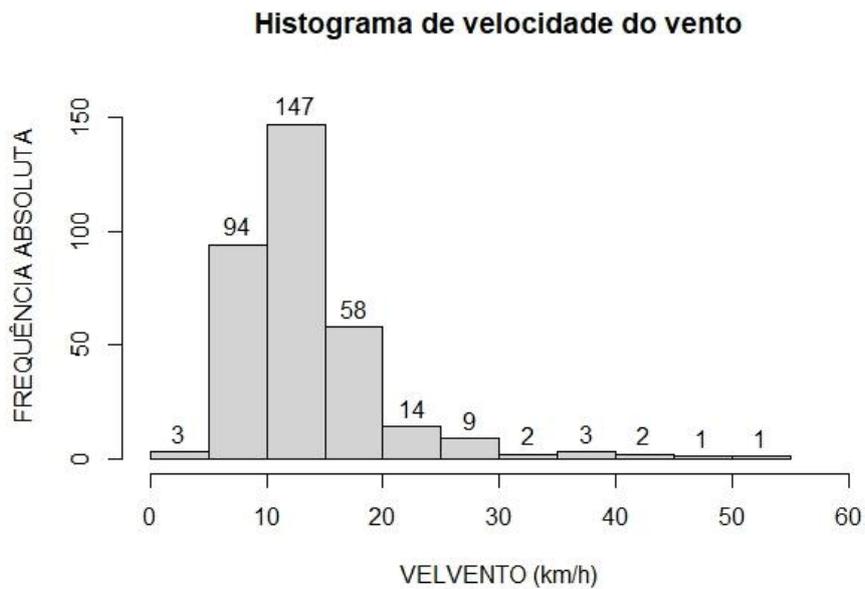


Figura 44 - Histograma de umidade mínima (%).



A Figura 45 mostra uma maior concentração de velocidade média do vento de 5 a 20 km/h sendo que em apenas um dia no período, foi registrado mais de 50 km/h.

Figura 45 - Histograma de velocidade do vento (km/h).

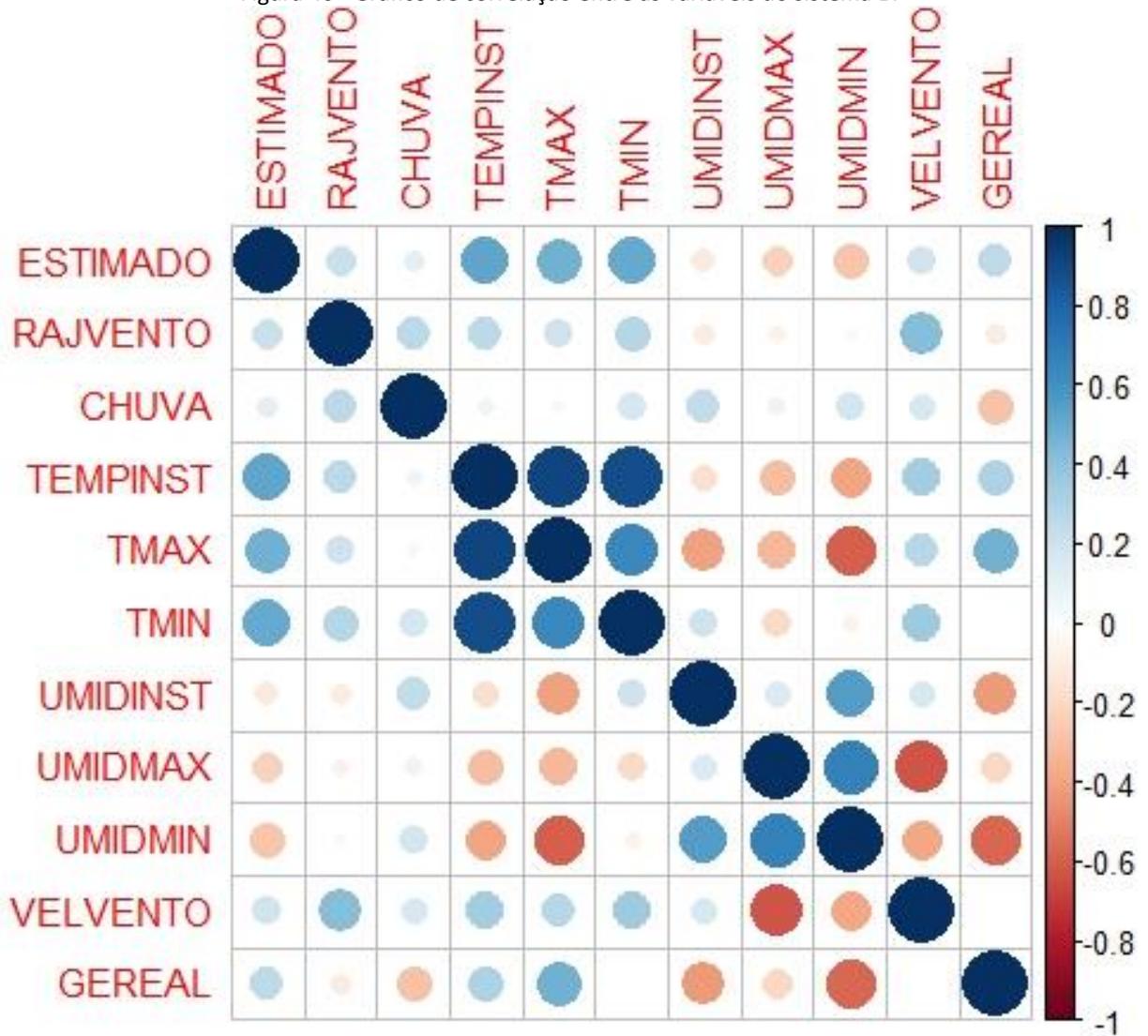


Assim como no sistema A, para uma melhor análise dos dados e de como se correlacionam, foi feito para o sistema B o gráfico de correlação entre as variáveis como mostra a Figura 46. Quanto maior a circunferência, maior é a correlação entre as variáveis, quanto mais azul (escuro) mais as variáveis são correlacionadas positivamente, e quanto maior a circunferência e mais vermelho (escuro), maior é a correlação negativa.

Desta forma, para uma hipótese inicial, no sistema B, as variáveis ESTIMADO, CHUVA, TEMPINST, TMAX, UMIDINST, UMIDMAX e UMIDMIN, são importantes para a explicação da variável de saída GEREAL.

No entanto, posteriormente, após uma análise mais pontual foi verificado que algumas dessas variáveis, que possivelmente seriam importantes, têm também correlação alta entre si, o que significa que são redundantes. Ou seja, possivelmente estes dados carreguem informações similares para a explicação do problema. Outro fator observável é que mesmo com dados completamente distintos de geração e de fatores meteorológicos, as variáveis que hipoteticamente mais importam para a explicação do problema são as mesmas.

Figura 46 - Gráfico de correlação entre as variáveis do sistema B.



### 4.3 Seleção dos atributos

Tendo em vista que algumas variáveis mostraram uma alta correlação quando simuladas no *software* RStudio, testes iniciais foram feitos no *software* Weka com o intuito de analisar o comportamento da RNA, com relação a estes dados. Os testes foram inicialmente realizados com as configurações padrão do *Weka* e, posteriormente, foram aplicadas as heurísticas destacadas na metodologia.

Inicialmente foram colocadas todas as variáveis disponíveis para o sistema A, que possuía 681 instâncias (dados para treinamento, teste e validação). Inicialmente foram feitos

testes gerais com os dados, que por sua vez destacaram 9 variáveis que possuíam alto índice de assertividade para explicação do problema, são elas:

- ESTIMADO;
- CHUVA;
- TEMPINST;
- TMÁX;
- TMÍN;
- UMIDINST;
- UMIDMIN;
- VELVENTO;
- RAJVENTO.

Posteriormente, também foram consideradas todas as variáveis disponíveis para o sistema B, que possuía 334 instâncias. Inicialmente, foram feitos testes gerais com os dados, que por sua vez, destacaram variáveis que possuíam alto índice de assertividade para explicação do problema, são elas:

- ESTIMADO;
- CHUVA;
- TEMPINST;
- TMÁX;
- TMÍN;
- UMIDINST;
- VELVENTO;
- RAJVENTO.

Visto que comparando-se os dois sistemas (A e B) verificou-se que as variáveis, destacadas pelo software *Weka*, e utilizadas para a solução do problema eram praticamente as mesmas. Desta forma, foram feitos testes iniciais para verificar como a rede MLP se comportava com as variáveis escolhidas para cada um dos sistemas.

Pode-se notar também que entre o sistema A e o sistema B, apenas a variável UMIDTMIN (umidade mínima) não foi um consenso. Por isso, foram realizados testes com a rede MLP, com e sem a variável.

Para verificar qual o melhor desempenho do processamento dois testes foram feitos com cada sistema (A e B), e para cada um deles foi testado com as variáveis coincidentes e as não coincidentes (último teste).

### SISTEMA A

A Tabela 4 apresenta os resultados das simulações iniciais para o sistema A e a rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 9 (teste 1) e 8 (teste 2);
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO; UMIDMIN;
- camada intermediária: 7 (teste 1) e 30 (teste 2) neurônios;
- valor de *Momentum*: 0,1;
- taxa de Aprendizagem: 0,4;
- número de épocas: 10.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem *split* 66% (451 para treinamento e 230 para teste).

Ainda na Tabela 4 são apresentados os resultados do erro absoluto médio, do erro quadrático médio, do erro relativo e do erro relativo quadrático.

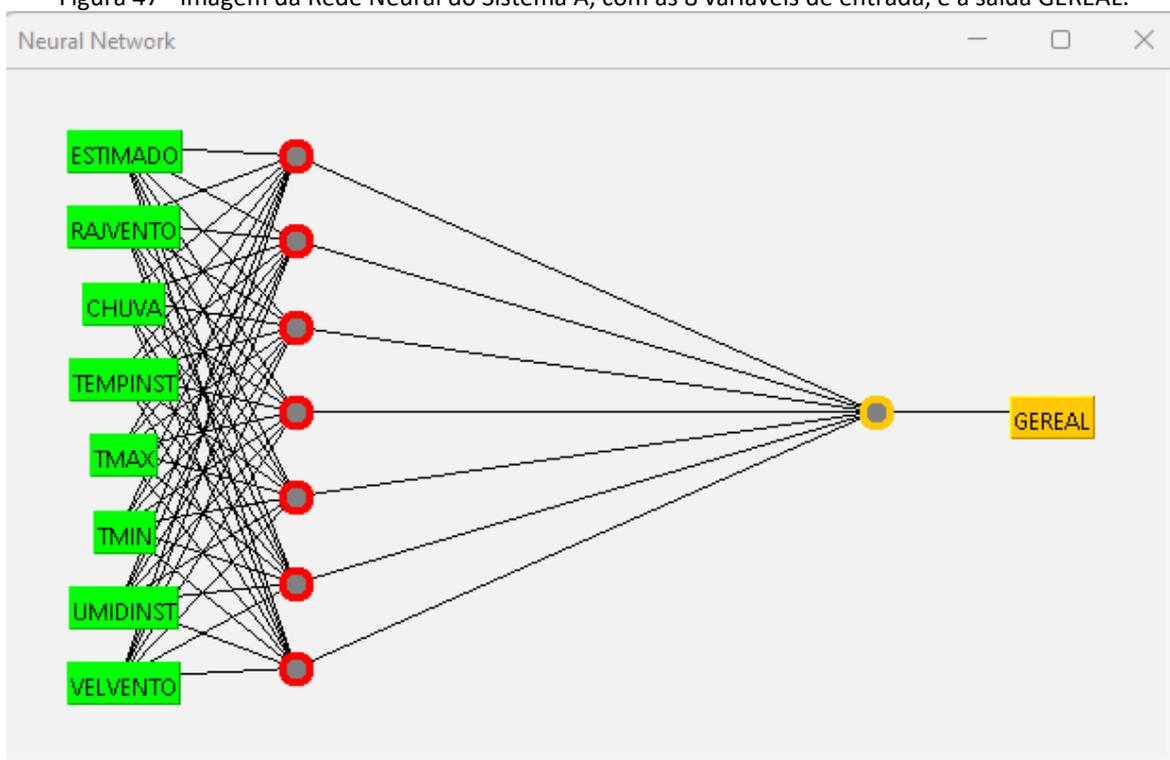
Tabela 4 – Comparativo entre uso com e sem a variável UMIDMÍN para resolução do problema no sistema A.

Resultados	Teste	Com UMIDMÍN	Sem UMIDMÍN
Coefficiente de correlação		0,9872	0,9872
Erro absoluto médio		3,2395	3,3582
Erro quadrático médio		4,2791	4,3944
Erro absoluto relativo		15,2617 %	15,8206 %
Erro quadrático relativo		16,6045 %	17,0518 %

Sendo assim, a partir do que mostra a Tabela 4 pode-se notar que a variável UMIDMIN quando retirada praticamente não modificou o erro relativo, portanto, devido ao gasto computacional das simulações, optou-se pela sua retirada. Para o sistema A foram escolhidas as 8 variáveis descritas a seguir, e apresentadas na Figura 47:

- ESTIMADO;
- CHUVA;
- TEMPINST;
- TMÁX;
- TMÍN;
- UMIDINST;
- VELVENTO;
- RAJVENTO.

Figura 47 - Imagem da Rede Neural do Sistema A, com as 8 variáveis de entrada, e a saída GEREAL.



Fonte: próprio autor, retirado do *software* WEKA.

## SISTEMA B

A Tabela 5 apresenta os resultados das simulações iniciais para o sistema B e a rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 9 (teste 1) e 8 (teste 2);
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO; UMIDMIN;
- camada intermediária: 16 (teste 1) e 19 neurônios (teste 2) (sistema A e B);
- valor de *Momentum*: 0,1;
- taxa de Aprendizagem: 0,4;
- número de épocas: 10.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem split 66% (220 para treinamento e 114 para teste).

Ainda na Tabela 5, são apresentados os resultados do erro absoluto médio, do erro quadrático médio, do erro relativo e do erro relativo quadrático.

Tabela 5 - Comparativo entre uso com, sem UMIDMÍN para resolução do problema no sistema B.

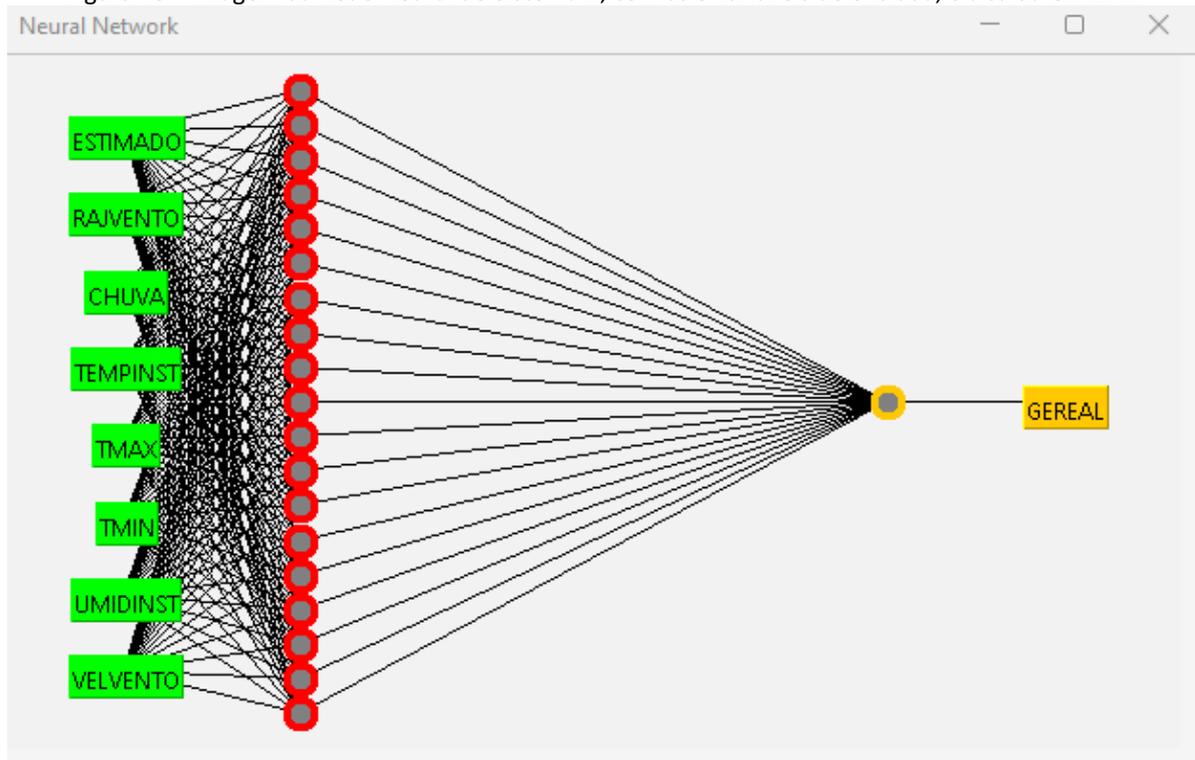
Resultados	Teste	Com UMIDMÍN	Sem UMIDMÍN
Coefficiente de correlação		0,9956	0,9953
Erro absoluto médio		1,2228	1,0048
Erro quadrático médio		1,4865	1,3459
Erro absoluto relativo		10,7463 %	8,8302 %
Erro quadrático relativo		11,3985 %	10,3279 %

Sendo assim, a partir do que mostra a Tabela 5 pode-se notar que a variável UMIDMÍN quando retirada, diminuiu o erro relativo. Portanto, para o sistema B foram escolhidas as 8 variáveis descritas a seguir e apresentadas na Figura 48:

- ESTIMADO;
- CHUVA;
- TEMPINST;

- TMÁX;
- TMÍN;
- UMIDINST;
- VELVENTO;
- RAJVENTO.

Figura 48 – Imagem da Rede Neural do Sistema B, com as 8 variáveis de entrada, e a saída GEREAL.



Fonte: próprio autor, retirado do *software* WEKA.

Portanto, como para o sistema A e B as mesmas variáveis foram selecionadas, observa-se uma concordância dos dados. No entanto, posteriormente em todos os testes foram novamente refeitas análises com todas as variáveis.

#### 4.4 Implementação da RNA do tipo MLP

Para uma melhor implementação da Rede Neural Artificial do tipo MLP, algumas estratégias foram adotadas quanto à disposição dos dados para simulação. Além disso, também foi implementado um teste inicial com uma equação simples para ver a assertividade

da rede e se essa se comportava da forma correta. Ou seja, com esse procedimento foi possível constatar se o uso do simulador WEKA estava correto.

### TESTE 0

Uma equação simples foi desenvolvida para o teste da rede e até mesmo do *software*. A Equação 3 representa a equação que foi colocada em uma planilha no *software* EXCELL com dados aleatórios de 5 variáveis  $x$ ,  $y$ ,  $z$ ,  $w$  e  $a$ , sendo que para cada instância foi calculado o valor de  $r$  (variável resposta). Cabe observar que na implementação deste teste foi utilizada a mesma heurística para a definição dos parâmetros e da arquitetura da rede que os demais testes que foram posteriormente implementados.

$$2x + 4y + 6z + 8w + 10a = r \quad \text{Equação 3}$$

Sendo assim, foram colocadas 199 instâncias para observar o comportamento da rede. A Tabela 6 apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 5;
- variáveis:  $x$ ,  $y$ ,  $z$ ,  $w$ ,  $r$ ;
- camada intermediária: 33 neurônios;
- valor de *Momentum*: 0,2;
- taxa de Aprendizagem: 0,3;
- número de épocas: 500, 1.000, 10.000, 50.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem split 66% (131 para treinamento e 68 para teste).

Tabela 6 – Resultados de erro para o TESTE 0 em diferentes épocas de treinamento.

Teste Resultados	Nº de épocas			
	500	1.000	10.000	50.000
Coefficiente de correlação	1	1	1	1
Erro absoluto médio	0,2189	0,1672	0,1159	0,0154
Erro quadrático médio	0,2595	0,2009	0,1294	0,0199
Erro absoluto relativo	0,6875 %	0,5238 %	0,363 %	0,0483 %
Erro quadrático relativo	0,6537 %	0,5062 %	0,326 %	0,0502 %

Portanto, como o TESTE 0 apresentou um erro muito pequeno (menos que 1%), como esperado na solução deste problema simples, confirmou-se o correto uso do simulador e da heurística utilizada na determinação dos parâmetros e da arquitetura da rede. Desta forma, realizou-se a sequência de testes na rede MLP dos sistemas que buscarão os pesos (conhecimento da rede) para a solução do problema de obtenção do fator de correção da geração para o projeto do sistema fotovoltaica, em locais de estiagem prolongada.

### TESTE 1

Com o *software* testado, iniciou-se o trabalho de simulação para os dois sistemas. Primeiramente ajustando o *momentum* e taxa de aprendizagem.

Desta foram, foram colocadas as 681 instâncias (exemplos de treinamento – dias de dados de geração fotovoltaica) do sistema A conforme a Figura 49, que mostra 10 linhas de dados do *Software Excell*, e as 334 instâncias (exemplos de treinamento – dias de dados de geração fotovoltaica) do sistema B conforme a Figura 50, que mostra 10 linhas de dados do *Software Excell*, onde DATA (coluna A) é a data do dado, ESTIMADO (coluna B) é a energia estimada de geração no projeto em kWh, RAJVENTO (coluna C) é a maior rajada de vento diária em km/h, CHUVA (coluna D) são os milímetros de chuva diários, TEMPINST (coluna E), TMAX (coluna F) E TMIN (coluna G), são as temperaturas instantâneas, máximas e mínimas respectivamente, médias diárias em graus Celsius, UMIDINST (coluna H) é a umidade instantânea, VELVENTO (coluna H) que é a velocidade do vento em km/h e GEREAL (coluna J) que é a geração real em kWh. A coluna A não é usada para simulação, apenas as entradas de B a I que são as entradas da rede e a coluna J que são as saídas da rede. A Tabela 7 apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO;
- camada intermediária: 7 e 3 neurônios (sistema A e B);
- valor de *Momentum*: 0,01;
- taxa de Aprendizagem: 0,2;
- número de épocas (3 testes, um com cada época): 5.000, 10.000 e 25.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem *split* 66% (451 para treinamento e 230 para teste no sistema A e 220 para treinamento e 114 para teste no sistema B).

Figura 49 - Exemplo de 10 linhas da planilha de dados contendo as 8 variáveis e a variável de saída GEREAL.

	A	B	C	D	E	F	G	H	I	J
1	DATA	ESTIMADO	RAJVENTO	CHUVA	TEMPINST	TMAX	TMIN	UMIDINST	VELVENTO	GEREAL
2	22/04/2021	244	29,52	2,6	22,99	28,8	19,1	79,36	11,88	240,99
3	23/04/2021	244,22	22,32	0	24,78	30,9	19,5	75,18	9,36	196,36
4	24/04/2021	244,44	44,28	11,4	24,37	29,8	19,9	77,63	11,52	255,8
5	25/04/2021	244,67	37,8	9,8	22,95	28,1	19,7	86,45	15,48	88,19
6	26/04/2021	244,89	33,84	41,8	21,54	24,4	19,5	87,38	17,64	289,49
7	27/04/2021	245,11	26,28	0,4	22,66	30,6	17,7	79,27	11,88	356,27
8	28/04/2021	245,33	27,72	0	25,2	31,4	19,4	59,04	14,76	362,88
9	29/04/2021	245,56	25,56	0	24,54	31,9	17,5	48,58	6,12	328,94
10	30/04/2021	245,78	20,88	0	24,94	32,2	17,4	49,33	10,44	306

Figura 50 - Exemplo de 10 linhas da planilha de dados contendo as 8 variáveis e a variável de saída GEREAL.

	A	B	C	D	E	F	G	H	I	J
1	DATA	ESTIMADO	RAJVENTO	CHUVA	TEMPINST	TMAX	TMIN	UMIDINST	VELVENTO	GEREAL
2	11/08/2021	94,3	38,88	0	31,13	39,2	22,7	41,33	15,12	95,3
3	12/08/2021	94,36	45	1,8	28,35	35,4	24,6	62,17	11,16	69,1
4	13/08/2021	94,42	30,6	0	26,98	32,9	23	71,25	10,08	85,9
5	14/08/2021	94,49	37,8	0	28,46	35,6	23,4	61,08	14,04	82,1
6	15/08/2021	94,55	31,32	0	30,26	37,1	24,4	51,04	15,48	115,5
7	16/08/2021	94,61	46,8	0	30,47	39	23,4	50,29	11,16	111,2
8	17/08/2021	94,68	29,52	0	30,13	38,3	25	57,5	7,92	110,2
9	18/08/2021	94,74	30,6	2	28,97	38	23,6	61,08	14,4	76,6
10	19/08/2021	94,8	29,16	0	29,88	37,3	23,7	55,58	14,04	110,2

Tabela 7 – Resultados de erro para o TESTE 1 em diferentes épocas de treinamento.

Resultados	5.000		10.000		25.000	
	Sistema A	Sistema B	Sistema A	Sistema B	Sistema A	Sistema B
Coeficiente de correlação	0,6233	0,682	0,5925	0,6833	0,5766	0,6878
Erro absoluto médio	35,1185	15,6426	37,2262	15,2703	38,8168	14,8879
Erro quadrático médio	47,0747	20,2329	48,9498	19,8973	50,7209	19,5294
Erro absoluto relativo	79,2534%	78,3129%	84,0098%	76,4489%	87,5995%	74,5344%
Erro quadrático relativo	78,6132%	80,7707%	81,7446%	79,4311%	84,7022%	77,9624%

Conforme mostrado na Tabela 7 os resultados da rede MLP, com os dados inseridos sem nenhum tratamento, apontam erros considerados elevados (na ordem de 80%). Isto ocorre porque os dados obtidos do INMET e dos projetos são apenas inseridos na rede MLP na data que obtidos. Por exemplo, pode-se citar dias em que mesmo com chuva, a geração de energia ainda é alta, como em dias de verão que tem sol e algumas pancadas de chuva (passageira). Além disso, também existem dias em que não chove, mas a geração é muito baixa devido ao alto índice de nuvens no céu. Tudo isso “confunde” o aprendizado da rede MLP, que não fornece um índice assertivo aceitável. Por isso, algumas propostas foram colocadas em prática nas próximas simulações.

## TESTE 2

A primeira proposta de como tornar a rede mais assertiva foi fazer uma média dos 7 dias posteriores à data da geração real para o parâmetro chuva, visto que, a chuva é um fator que interfere na geração solar, pois acontece uma “limpeza natural” nos módulos fotovoltaicos proveniente da precipitação. A Tabela 8 apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO;
- camada intermediária: 7 e 3 neurônios (sistema A e B);
- valor de *Momentum*: 0,01;

- taxa de Aprendizagem: 0,2;
- número de épocas: 5.000, 10.000 e 25.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem *split* 66% (451 para treinamento e 230 para teste no sistema A e 220 para treinamento e 114 para teste no sistema B).

Tabela 8 – Resultados de erro para o TESTE 2 em diferentes épocas de treinamento.

Resultados	5.000		10.000		25.000	
	Sistema A	Sistema B	Sistema A	Sistema B	Sistema A	Sistema B
Coeficiente de correlação	0,682	0,682	0,6833	0,6833	0,6878	0,6878
Erro absoluto médio	15,6426	15,6426	15,2703	15,2703	14,8879	14,8879
Erro quadrático médio	20,2329	20,2329	19,8973	19,8973	19,5294	19,5294
Erro absoluto relativo	78,3129%	78,3129%	76,4489%	76,4489%	74,5344%	74,5344%
Erro quadrático relativo	80,7707%	80,7707%	79,4311%	79,4311%	77,9624%	77,9624%

Novamente pode-se notar a partir da Tabela 8, que os resultados não foram assertivos como esperado (erro absoluto relativo maior que 70 %). Desta forma, outros testes foram realizados, com mudança no valor de inserção dos dados.

### TESTE 3

No TESTE 3 proposto, além da média dos 7 dias posteriores para a variável chuva, também foi feita a média dos 7 dias posteriores para a variável GEREAL (geração real), visto que a chuva impacta diretamente da geração real do sistema.

A Tabela 9 apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO;
- camada intermediária: 7 e 3 neurônios (sistema A e B);

- valor de *Momentum*: 0,01;
- taxa de Aprendizagem: 0,2;
- número de épocas: 5000, 10000 e 25000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem split 66% (451 para treinamento e 230 para teste no sistema A e 220 para treinamento e 114 para teste no sistema B).

Tabela 9 – Resultados de erro para o TESTE 3 em diferentes épocas de treinamento.

Resultados	5000		10000		25000	
	Sistema A	Sistema B	Sistema A	Sistema B	Sistema A	Sistema B
Coeficiente de correlação	0,7108	0,5568	0,7158	0,5456	0,677	0,537
Erro absoluto médio	19,8816	9,6006	19,5872	9,5226	20,4534	9,7281
Erro quadrático médio	25,1135	12,2801	24,9715	12,5544	26,8908	12,8203
Erro absoluto relativo	73,8569%	81,3891%	72,7631%	80,7276%	75,9812%	82,4696%
Erro quadrático relativo	71,6162%	86,1877%	71,2113%	88,1124%	77,9624%	89,9786%

Assim como no TESTE 2 pode-se notar a partir da Tabela 9, que os resultados não foram assertivos como esperado.

#### TESTE 4

Uma nova proposta com o TESTE 4 foi implementado, onde além da média dos 7 dias posteriores para a variável chuva e GEREAL (geração real), para todas as variáveis foi feito a média dos 7 dias posteriores.

A Tabela 10 apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO;
- camada intermediária: 7 e 3 neurônios (sistema A e B);

- valor de *Momentum*: 0,01;
- taxa de Aprendizagem: 0,2;
- número de épocas: 5.000, 10.000 e 25.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem split 66% (451 para treinamento e 230 para teste no sistema A e 220 para treinamento e 114 para teste no sistema B).

Tabela 10 – Resultados de erro para o TESTE 4 em diferentes épocas de treinamento.

Resultados	5000		10000		25000	
	Sistema A	Sistema B	Sistema A	Sistema B	Sistema A	Sistema B
Coefficiente de correlação	0,8703	0,8485	0,8659	0,854	0,8642	0,8622
Erro absoluto médio	13,332	5,8548	13,5453	5,7641	13,8035	5,4133
Erro quadrático médio	17,193	7,5217	17,4917	7,3933	17,6614	7,1991
Erro absoluto relativo	49,5263%	49,6337%	50,3185%	48,8653%	51,2778%	45,8907%
Erro quadrático relativo	49,0293%	52,791%	49,8811%	51,8895%	50,3649%	50,5267%

Pode-se notar a partir da Tabela 10 que os resultados melhoraram em relação aos anteriores, com um erro absoluto relativo em média de 50%, entretanto, esse resultado ainda não é satisfatório.

## TESTE 5

Um TESTE 5 e um TESTE 6 foram propostos, onde no TESTE 5 a rede MLP foi executada com inserção de dados com a média de todas as variáveis para os próximos 15 dias e o TESTE 6, a média das variáveis para os próximos 30 dias.

A Tabela 11 apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO;

- camada intermediária: 7 e 3 neurônios (sistema A e B);
- valor de *Momentum*: 0,2;
- taxa de Aprendizagem: 0,15;
- número de épocas: 5.000, 10.000 e 25.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem *split* 66% (451 para treinamento e 230 para teste no sistema A e 220 para treinamento e 114 para teste no sistema B).

Tabela 11 – Resultados de erro para o TESTE 5 em diferentes épocas de treinamento.

Resultados	5.000		10.000		25.000	
	Sistema A	Sistema B	Sistema A	Sistema B	Sistema A	Sistema B
Coeficiente de correlação	0,9539	0,9586	0,9558	0,9624	0,9573	0,9672
Erro absoluto médio	6,7255	2,9445	6,5433	2,8108	6,6939	2,5935
Erro quadrático médio	8,8042	3,6573	8,6259	3,4948	8,5457	3,2476
Erro absoluto relativo	28,513%	26,2534%	27,7406%	25,0611%	28,3793%	23,1242%
Erro quadrático relativo	29,8923%	28,5663%	29,2871%	27,2973%	29,0148%	25,3665%

Os resultados melhoraram significativamente como pode ser observado no TESTE 5, visto que os dados se encontram com um erro absoluto relativo em média de 26%, entretanto, sendo assim, o teste 6 foi feito, com o intuito de observar melhores resultados.

## TESTE 6

Como explicado anteriormente o TESTE 6 foi feito com a média das variáveis para os próximos 30 dias.

A Tabela 12 apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 8;

- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO;
- camada intermediária: 7 e 19 neurônios (sistema A e B);
- valor de *Momentum*: 0,1;
- taxa de Aprendizagem: 0,15;
- número de épocas: 5.000, 10.000 e 25.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem *split* 66% (451 para treinamento e 230 para teste no sistema A e 220 para treinamento e 114 para teste no sistema B).

Tabela 12 – Resultados de erro para o TESTE 6 em diferentes épocas de treinamento.

Resultados	5.000		10.000		25.000	
	Sistema A	Sistema B	Sistema A	Sistema B	Sistema A	Sistema B
Coeficiente de correlação	0,9993	0,9995	0,9995	0,9996	0,9995	0,9997
Erro absoluto médio	0,8817	0,2442	0,7812	0,1968	0,6729	0,1811
Erro quadrático médio	1,1221	0,3063	0,9959	0,2594	0,8619	0,2301
Erro absoluto relativo	4,0035%	3,3151%	3,5471%	2,6708%	3,0557%	2,4586%
Erro quadrático relativo	4,2571%	3,3808%	3,7781%	2,8633%	3,2699%	2,5399%

O TESTE 6 foi, entre todos realizados, o mais satisfatório com resultado de erro absoluto relativo de aproximadamente 3 %, o que é uma porcentagem com erro muito baixo, visto que são dados obtidos em campo, de dois locais distintos, demonstrando que a proposta inicial de se utilizar uma ferramenta de Inteligência artificial, mais especificamente com o uso da rede neural artificial do tipo MLP para simular tal situação é adequada, e que se pode obter modelagens factíveis para tal sistema.

Os pesos da rede neural artificial do tipo MLP, que estabelece o conhecimento obtido pela rede sobre este problema, poderão ser utilizados em trabalhos futuros, e encontram-se no Apêndice B para o sistema A e Apêndice C para o sistema B.

## TESTE 7

Visto a inquietude de obter resultados mais assertivos para ambos os sistemas, foi proposto um último teste (TESTE 7). Nesse teste assim como no TESTE 6, a rede MLP foi executada com inserção de dados com a média de todas as variáveis para os próximos 30 dias, no entanto, os dados do SISTEMA A e SISTEMA B foram inseridos juntos (misturados) a fim de obter uma resposta genérica do problema.

A Tabela 13 – Resultados de erro para o TESTE 7 em diferentes épocas de treinamento. apresenta os resultados da rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: ESTIMADO; CHUVA; TEMPINST; TMÁX; TMÍN; UMIDINST; VELVENTO; RAJVENTO;
- camada intermediária: 10 neurônios;
- valor de *Momentum*: 0,05;
- taxa de Aprendizagem: 0,4;
- número de épocas: 5.000, 10.000, 25.000 e 50.000;
- normalização dos atributos: verdadeiro;
- método de execução do teste: Percentagem *split* 66% (670 para treinamento e 345 para teste).

Tabela 13 – Resultados de erro para o TESTE 7 em diferentes épocas de treinamento.

Teste	Nº de épocas			
	5000	10.000	25.000	50.000
Resultados				
Coefficiente de correlação	0,9999	0,9999	0,9999	1
Erro absoluto médio	0,9333	0,7209	0,6411	0,5991
Erro quadrático médio	1,2141	0,9243	0,8282	0,7668
Erro absoluto relativo	1,3689 %	1,0574 %	0,9403 %	0,8786 %
Erro quadrático relativo	1,5969 %	1,2158 %	1,0893 %	1,0086 %

O TESTE 7, além de mostrar resultados muito bons com erro na casa de 1%, nos mostra que quanto mais exemplos forem inseridos no treinamento da rede MLP, melhores são os

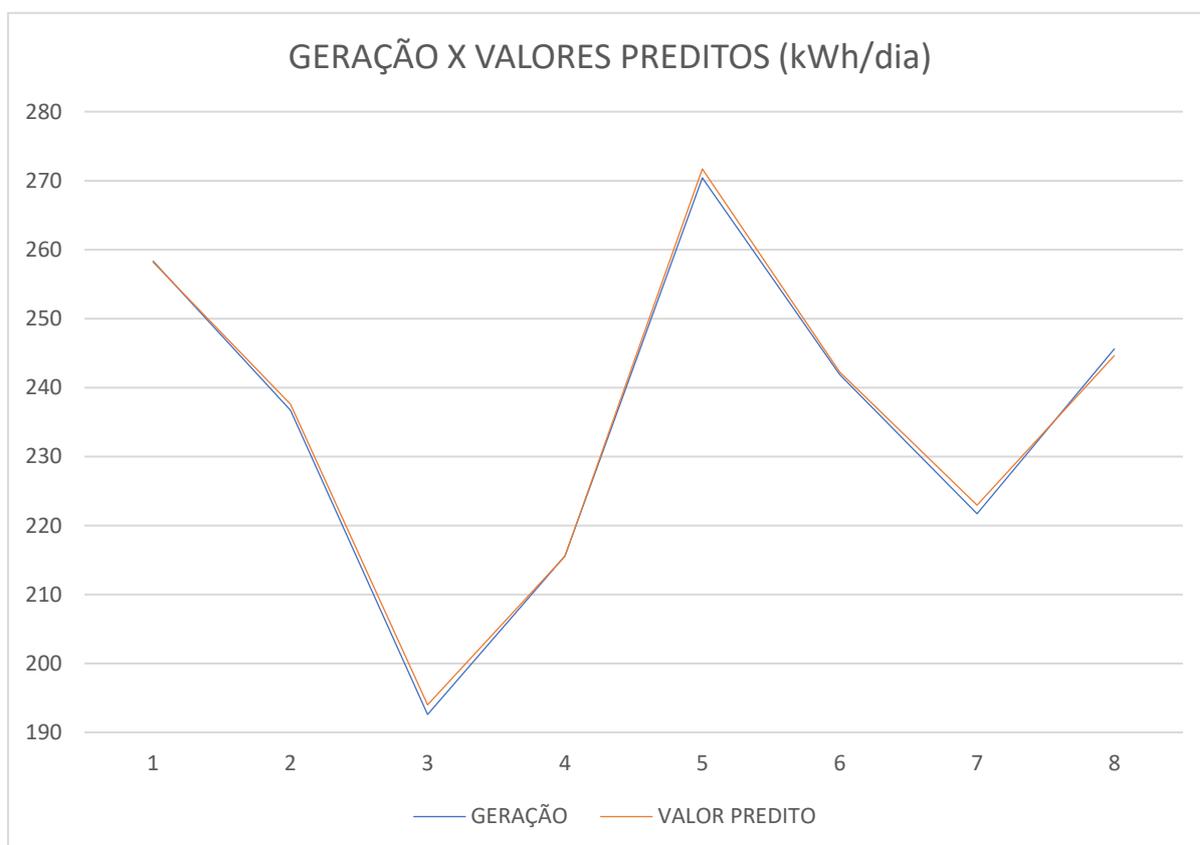
resultados obtidos. Além disso, apesar dos dados serem misturados, ela aprendeu para o caso urbano e rural.

Os pesos da rede neural MLP para uso em possíveis trabalhos futuros encontram-se no Apêndice A.

#### 4.5 Predição

Foram removidos 8 valores ao acaso da planilha do TESTE 7 e novamente simulados com as mesmas configurações do TESTE 7. Vale ressaltar que esses 8 valores removidos são referentes a 8 dias de produção fotovoltaica que não entraram para o treinamento da rede, ou seja, a rede nunca viu esses exemplos, não os conhece. Os valores reais de geração e preditos encontra-se na Figura 51.

Figura 51 – Gráfico de geração x valores preditos pelo *software WEKA*, em kWh/dia.



A partir do gráfico pode-se notar que a diferença entre a geração real e a geração predita pelo software é muito pequena, na ordem de 0,5 kWh/dia. Sendo assim, pode-se dizer que a RNA do tipo MLP aprendeu com os exemplos e, conforme simulação, conclui-se que atende a ideia inicial de que é uma ferramenta adequada para predição de valores de geração fotovoltaica.

#### 4.6 Discussão

Os resultados obtidos nas simulações demonstram que quanto maior o número de exemplos, melhores são os resultados obtidos. Desde o TESTE 0, o entendimento da rede foi de grande importância, pois garantiu que resultados promissores fossem alcançados e possibilitou que a implementação da rede MLP fosse melhorada, ou seja, foi diminuindo o erro.

Um erro final de aproximadamente 1 % como o do TESTE 7 foi melhor do que o previsto no início da pesquisa, mostrando que além de encontrar as principais variáveis que influenciam na geração fotovoltaica, o trabalho também poderá ser utilizado para futuras pesquisas e até mesmo no desenvolvimento de *softwares* de dimensionamento fotovoltaico, em que o pesquisador ou o engenheiro projetista da área poderá saber quais serão as perdas de energia fotovoltaica para sistemas onde não haverá manutenção periódica.

É importante observar também que o erro absoluto médio, ficou na ordem de 0,9 para o sistema A e 0,3 para o sistema B, que quando se divide pela média da geração estimada de cada sistema, obtém-se o valor de 0,38% e 0,32%, o que além de serem próximos, demonstra a concordância dos resultados em sistemas diferentes, com um erro muito pequeno na modelagem da rede MLP.

Outro resultado a ser destacado é a diferença entre a geração real e a geração estimada. Para o Sistema A, instalado na área urbana, pode-se observar uma redução entre 3% e 4% na geração real se comparada com a de projeto, visto que esse sistema não sofreu nenhum tipo de limpeza, apenas a autolimpeza proporcionada pela chuva. Já para o sistema

B, instalado em área rural, pode-se observar uma redução entre 9% e 10% na geração real se comparada com a de projeto, também, apenas com a autolimpeza dos módulos.

A situação piora em períodos de estiagem, onde o sistema A apresenta diferença na ordem de 6% e o sistema B de 12%, entre a geração real e a estimada. Sendo assim, é importante a observação por parte dos engenheiros e técnicos que em sistemas de difícil acesso, onde a manutenção é difícil e complicada de ser realizada e, muitas vezes, não ocorre, o sistema fotovoltaico deva ser majorado em projeto. Tal incremento de potência, se explica para atingir a quantidade de energia gerada pretendida.

Tais valores convergem com alguns outros estudos que foram citados nesse trabalho, a Tabela 14 traz um comparativo entre as perdas de geração elencada pelos autores referenciados nesse trabalho e o resultado obtido. Para tal, pode-se notar que alguns autores citam perdas de geração próximas aos dados obtidos pelo trabalho atual, outros já elencam perdas maiores ou menores. Essa divergência entre alguns dos dados pode ser ocasionada pois alguns trabalhos foram realizados em regiões diferentes do trabalho atual, onde os dados meteorológicos possivelmente também não são os mesmos, como precipitação e até incidência solar. Portanto, é importante que para cada simulação em outros lugares seja alocado os dados de geração e meteorológicos do local em questão.

Tabela 14 – Comparativo entre as perdas de geração para os trabalhos citados e o trabalho atual.

<i>Nome do trabalho</i>	<i>Autores</i>	<i>Citação</i>	<i>Perdas em sistemas devido a sujeidade</i>
Análise do fator de sobredimensionamento em projetos para sistemas fotovoltaicos com o uso de rede neural artificial.	Gabriel Edgar Hermann, Andrea Teresa Riccio Barbosa.	Hermann (2024)	Redução de 3 a 4 % em sistemas instalados em áreas urbanas e entre 9 a 10% em sistemas instalados em áreas rurais. Em períodos de estiagem, a diferença passa para 6% em sistemas urbanos e 12% em sistemas rurais.
Análise dos fatores de perdas nos sistemas fotovoltaicos conectados à rede elétrica em Curitiba.	Ana Júlia Nunes de Araújo, Narah luata Rank, Talita Bezerra de Araujo Bueno.	Araujo (2016)	Perdas na ordem de 10% para sistemas sem limpeza periódica.
Avaliação do potencial dos sistemas fotovoltaicos conectados à rede elétrica com geradores de diversas tecnologias.	Suellen Caroline Silva Costa.	Costa (2011)	Perdas em torno de 30% no rendimento de um sistema antes e depois de limpeza periódica.
Estimativa da taxa de sujeidade em módulos fotovoltaicos utilizando dados de densidade gravimétrica.	Túlio Pinheiro Duarte, Suellen Caroline Silva Costa, Antônia Sonia Alves Cardoso Diniz, Lawrence Lee Kazmerski	Duarte (2020)	Demonstra que a sujeidade reduz a geração de energia em até 14,02% após um período de 75 dias de exposição sem chuva.
Influência da sujeira na geração fotovoltaica.	Elismar Ramos Barbosa, Merlim dos Santos Ferreira de Faria, Fabio de Brito Gontijo.	Barbosa (2019)	Para o ensaio, o autor mostra uma diferença de 10% de geração em sistemas que possuíam limpeza regular e aqueles que não possuíam.
<i>Soiling losses for solar photovoltaic systems in California.</i>	Felipe A. Mejia Jan Kleissl	Mejia e Kelissl (2013)	Os autores mostram perdas anuais de 6,8% em sistemas sem limpeza devido a sujeidade, em relação aos sistemas com limpeza regular.
Análise dos fatores de perdas em sistemas fotovoltaicos da UTFPR campus Curitiba.	Édwin Augusto Tonolo	Tonolo (2019)	Demonstra melhoras de 6% para limpeza e 10% para autolimpeza na geração de painéis fotovoltaicos.

## 5 CONCLUSÃO

É importante que projetistas e técnicos que dimensionam sistemas fotovoltaicos saibam fazer o correto dimensionamento quanto à potência do sistema de geração de energia em locais de estiagem. Em localidades, principalmente onde a limpeza e manutenção é precária, faz-se necessário um aumento da potência instalada de geração para que em períodos de estiagem o sistema obtenha a geração pretendida, como é o caso, por exemplo, em algumas regiões do Pantanal brasileiro.

Este trabalho teve como objetivo encontrar o valor de incremento de potência de projeto para áreas que possuam longos períodos de estiagem, como é o caso do Mato Grosso do Sul em áreas rurais e urbanas. O trabalho contou com uma análise estatística e posteriormente os dados foram inseridos utilizando Redes Neurais Artificiais do tipo *MultiLayer Perceptron* com o auxílio de ferramenta de simulação, o *software* Weka. No decorrer da pesquisa, foi construída uma base de dados com cruzando de informações dos *softwares* de geração dos equipamentos fotovoltaicos instalados e do INMET que possuía os dados meteorológicos. A partir dessa base de dados e testes, houve uma seleção das variáveis que afetam a geração de energia fotovoltaica sendo estas:

- ESTIMADO;
- CHUVA;
- TEMPINST;
- TMÁX;
- TMÍN;
- UMIDINST;
- VELVENTO;
- RAJVENTO.

ESTIMADO é a energia estimada de geração em kWh, RAJVENTO é a maior rajada de vento diária em km/h, CHUVA são os milímetros de chuva diários, TEMPINST, TMAX E TMIN, são as temperaturas instantâneas, máximas e mínimas respectivamente, médias diárias em

graus Celsius, UMIDINST é a umidade instantânea e VELVENTO que é a velocidade do vento em km/h.

Uma das dificuldades foi a definição do conjunto de treinamento da RNA, o que ocasionou inicialmente erros elevados. Depois de várias análises pode-se identificar que a rede MLP não se encontrava assertiva no início dos testes porque o conjunto de dados do treinamento, da maneira como eles eram fornecidos causavam uma “confusão” no aprendizado da rede. Por exemplo, em um dia de sol, em um período de verão, no final da tarde acontece uma elevada precipitação de chuva, e mesmo assim os valores de geração são altos naquele dia, e há situações em que durante o dia ocorre uma geração baixa devido ao alto índice de nuvens no céu, e não houve precipitação alguma. Como a rede aprende com exemplos, e os dados tinham parâmetros diferentes levando a mesmos resultados de saída, ficou difícil da rede “entender” e “aprender” com estes dados, sem um tratamento prévio.

Após alguns testes, notou-se que com utilizando-se no conjunto de treinamento dados com a média dos próximos 7, 15 e 30 dias esse problema diminuiu consideravelmente obtendo um erro final na de aproximadamente 1%, como no TESTE 7.

Com relação ao modelo escolhido, a Rede Neural Artificial do tipo MLP, mostrou-se bastante capaz para a execução dessa tarefa, alcançando o objetivo esperado, visto que é um modelo robusto com a capacidade de se adaptar conforme os dados disponíveis para treinamento. Vale ressaltar que como já previsto, quanto mais dados de treinamento forem utilizados na rede, menor o erro da resposta obtida.

Sobre o trabalho desenvolvido, acredita-se que é bastante significativo, pois apresenta resultados importantes no mercado de trabalho na área de projeto de sistemas fotovoltaicos. Demonstra que para sistemas instalados em área urbana e sem manutenção periódica o incremento de projeto deve ser de aproximadamente 4% e para áreas rurais aproximadamente 10%, reduzindo os problemas com falta geração pretendida que são muito comuns no dia a dia.

O trabalho mostra que a rede MLP é uma ferramenta eficiente para o modelamento de perdas em sistemas fotovoltaicos. Além de contribuições acadêmicas, visto que apresenta

os pesos da rede MLP, que poderão ser utilizadas em trabalhos futuros, dimensionando outras redes e possibilitando desenvolver *softwares* que permitam o correto dimensionamento dos sistemas fotovoltaicos, além de fornecer os fatores meteorológicos que influenciam na geração fotovoltaica da região.

Ainda há muito a ser explorado nesse assunto, e espera-se que esta pesquisa incentive novas pesquisas, e seja uma fonte de conhecimento para o tema abordado

### **5.1 Trabalhos futuros**

Como dito anteriormente, quanto maior o número de exemplos de treinamento, menor os erros de saída da rede. Portanto, uma ideia de trabalho futuro seria a extrapolação de dados em âmbito nacional, cruzando dados de geração em diferentes regiões do país, fazendo assim comparações entre regiões diferentes e até biomas diferentes com o intuito de buscar resultados que possam ser usados em todas as regiões.

Outra proposta seria a criação de um *software* ou até uma aplicação *web* ou *mobile* que conseguisse cruzar os dados de geração dos equipamentos fotovoltaicos com os dados meteorológicos, sem a necessidade de planilhamento, visto que, essa é uma etapa que demanda muito tempo do pesquisador pois os dados são apresentados dia a dia. E depois de planilhados automaticamente, o usuário consiga no próprio *software* ou *app* fazer o correto dimensionamento do sistema fotovoltaico já levando em consideração as perdas de geração devido a sujidade, levando em consideração sistemas com e sem limpeza periódica, possibilitando que se faça o registro de *software*.

## REFERÊNCIAS

ANEEL - Agência Nacional de Energia Elétrica. **Energia Solar 3**.

AL-SABOUNCHI, A. M.; YALYALI, S. A.; AL-THANI, H. A. **Design and performance evaluation of a photovoltaic gri-connected system in hot weather conditions**. *Renewable Energy*, v. 53, p. 71 – 78, 2013.

ARAÚJO, Danielly Norberto. **Impactos da sujidade e efeitos da limpeza no desempenho de módulos fotovoltaicos**. p. 11, 2020.

ARAÚJO, Ana Júlia Nunes de; RANK, Narah luata; BUENO, Talita Bezerra de Araujo. **Análise dos fatores de perdas nos sistemas fotovoltaicos conectados à rede elétrica em Curitiba**. 2016. 85 f. Trabalho de Conclusão de Curso (Graduação em Engenharia Elétrica) - Universidade Tecnológica Federal do Paraná, Curitiba, 2016.

BARBOSA, A T. R. **Mecanismo de Adaptação Baseado em Redes Neurais Artificiais para Sistemas Hipermedia Adaptativos**. Tese de Doutorado — Programa de Pós-graduação em Engenharia Elétrica, UFSC, Florianópolis, 2004 SC.

BARBOSA, Elismar Ramos; FARIA, Merlim dos Santos Ferreira de; GONTIJO, Fabio de Brito. **Influência da sujeira na geração fotovoltaica**. *In: paulo jayme pereira abdala. Energia solar e eólica 2*. 1. ed. [S.l.]: Antonella Carvalho de Oliveira, 2019. p. 121–131. DOI 10.22533/at.ed.6741922019. Disponível em: <https://www.atenaeditora.com.br/wp-content/uploads/2019/01/E-book-Energia-Solar-e-Eolica-2.pdf>. Acesso em: 21 out. 2022.

BAYER, A.; BITTENCOURT, H.; ECHEVEST, S. A **Estatística e sua História**. XIII CIAEM-IACME, Recife, Brasil, 2011. XII Simpósio Sul Brasileiro de Ensino e Ciências, 2004, Canoas.

BISHOP, Christopher M. **Pattern Recognition and Machine Learning**. Nova Iorque: Springer, 2006.

BRAGA, Antonio de Pádua e LUDERMIR, Teresa Bernarda e CARVALHO, André Carlos Ponce de Leon Ferreira. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC 2007.

COSTA, Suellen Caroline Silva; ABREU, Amanda Mello Faria Veloso; VIANA, Marcelo Machado; BRITO, Pedro Paiva; MAIA, Cristiana Brasil; DINIZ, Antônia Sônia Alves Cardoso; KAZMERSKI, Lawrence Lee. **Caracterização físico-química da sujidade depositada sobre módulos fotovoltaicos instalados em zonas climáticas de minas gerais.** *In:* Jaqueline Rezende Oliveira. 2016. DOI: <https://doi.org/10.59627/cbens.2016.1461>. Acesso em: 21 out. 2022.

COSTA, Suellen Caroline Silva. **Avaliação do potencial dos sistemas fotovoltaicos conectados à rede elétrica com geradores de diversas tecnologias.** 2011. 82 f. Projeto (Graduação). Pontifícia Universidade Católica de Minas Gerais – Curso Engenharia de Energia, Belo Horizonte, Minas Gerais.

DANTAS, S. G.; POMPERMAYER, F. M. **Viabilidade econômica de sistemas fotovoltaicos no Brasil possíveis efeitos no setor elétrico.** mai. 2018 (Texto para Discussão nº. 2388). Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada. Disponível em: [https://www.ipea.gov.br/portal/index.php?option=com\\_content&view=article&id=33431:td-2388-viabilidade-economica-de-sistemas-fotovoltaicos-no-brasil-e-possiveis-efeitos-no-setor-eletrico&catid=411:2018&directory=1](https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=33431:td-2388-viabilidade-economica-de-sistemas-fotovoltaicos-no-brasil-e-possiveis-efeitos-no-setor-eletrico&catid=411:2018&directory=1). Acesso em: 25 mar. 2021.

DE OLIVEIRA, Fernando Schuck; PRIEB, César Wilhelm Massen; KRENZINGER, Arno. **Avaliação de um arranjo fotovoltaico após onze anos de operação.** p. 5, 2018.

DUARTE, Tulio; COSTA, Suellen Caroline Silva. **Estimativa da taxa de sujidade em módulos fotovoltaicos utilizando dados de densidade gravimétrica.** p. 7, 2020.

EL-NASHAR A. ***Effect of dust deposition on the performance of a solar desalination plant operating in an arid desert area.*** *Solar Energy* 2003.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados.** Rio de Janeiro: Elsevier, 2017.

FLECK, L; TAVARES, M. H. F.; EYNG, E; HELMANN, A. C.; ANDRADE, M. A. M.; **Redes neurais artificiais: princípios básicos.** Revista Eletrônica Científica Inovação e Tecnologia Universidade Tecnológica Federal do Paraná Câmpus Medianeira, Paraná, Brasil, v. 1, n. 13, p. 47-57, jan./jun. 2016.

GARG H. *Effect of Dirt on Transparent Covers in Flat-plate Solar Energy Collectors*. *Solar Energy* 1974;

GAIO, João Nicolau. **Determinação do tempo ótimo para limpeza de painéis fotovoltaicos para obtenção da melhor produtividade - estudo de caso implantados na utfpr**. p. 85, 2017.

GHAZI, S.; SAYIGH, A.; Ip., K. **Dust effect on flat surfaces - A review paper**. *Renewable and Sustainable Energy Reviews*. Issue 33, p.742- 751, 2014.

HAGAN M. T., H. B. Demuth, M. H. Beale, and O. De Jesus, **Neural Network Design**, ed. 2 (USA, Martin Hagan, 2014).

HAYKIN, S. *Redes Neurais- Princípios e Práticas*. BOOKMAN, São Paulo, 2ª ed. 2001. 900 p.

HICKEL, Bernardo Meyer. UNIVERSIDADE FEDERAL DE SANTA CATARINA PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL. p. 110, 2017.

KUROKAWA, Edson; BORNIA, Antônio Cezar. **Utilizando o histograma como uma ferramenta estatística de análise da produção de água tratada em Goiânia**. Congresso Interamericano de engenharia sanitária e ambiental. Cancún, México, 2002.

LOPES, Paulo Afonso. **Probabilidades & Estatística**. 1. ed. Rio de Janeiro: R&A, 1999, 174 p.

MEJIA, F.; KLEISSL, J.; BOSCH, J.L. **The Effect of Dust on Solar Photovoltaic Systems**. *Energy Procedia*, v. 49, p. 2370–2376, 2014. DOI 10.1016/j.egypro.2014.03.251.

MEJIA, Felipe A.; KLEISSL, Jan. **Soiling Losses for Solar Photovoltaic Systems in California**. *Solar Energy*, v. 95, p. 357–363, set. 2013.

MICHELS, Roger N.; SANTOS, José Airton A. dos; GNOATTO, Estor; FISCHBORN, Marcos; KAVANAGH, Edward. **Influência do acúmulo de poeira sobre a eficiência de painéis fotovoltaicos na região oeste paranaense**. *Revista Agrogeoambiental*, v. 1, n. 3, 1 dez. 2009. DOI 10.18406/2316-1817v1n32009221. Disponível em: <https://agrogeoambiental.ifsuldeminas.edu.br/index.php/Agrogeoambiental/article/view/221>. Acesso em: 21 out. 2022.

MONTO M, Rohit P. ***Impact of Dust on Solar Photovoltaic (PV) Performance: Research Status, Challenges and Recommendations.*** 2010;

NIELSEN, M. A. ***Neural networks and deep learning.*** Determination Press. (2015).

NDIAYE, A., CHARKI, A., KOBI, A., KÉBÉ, C. M. F., NDIAYE, P. A., SAMBOU, V., 2013. ***Degradations of silicon photovoltaic modules: A literature review.*** Solar Energy, vol. 96, pp. 140-151.

PINHO, J. T.; GALDINO, M. A. ***Manual de Engenharia para Sistemas Fotovoltaicos.*** Rio de Janeiro: CEPEL - CRESESB, 2014.

R CORE TEAM (2019). ***R: A language and environment for statistical computing.*** R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

ROMANHOLO, Pedro Victor Valadares; ALVARENGA, Bernardo Pinheiro de; MARRA, Enes Gonçalves; PIMENTEL, Sérgio Pires. ***Sujidade depositada sobre módulos fotovoltaicos instalados em goiânia: morfologia e composição química. In: paulo jayme pereira abdala. Energia solar e eólica 2. 1. ed. [S.l.]: Antonella Carvalho de Oliveira, 2019. p. 263–274. DOI 10.22533/at.ed.67419220120. Disponível em: <https://www.atenaeditora.com.br/wp-content/uploads/2019/01/E-book-Energia-Solar-e-Eolica-2.pdf>. Acesso em: 21 out. 2022.***

RUTHER, R. ***Edifícios solares fotovoltaicos: o potencial da geração solar fotovoltaica integrada a edificações urbanas e interligada à rede elétrica pública no brasil.*** Florianópolis: LABSOLAR, 2004.

SCHMULLER, Joseph. ***Análise estatística com R para leigos.*** Rio de Janeiro, Alta Books, 2019.

TONOLO, Édwin Augusto. ***Análise dos fatores de perdas nos sistemas fotovoltaicos da utfpr campus curitiba.*** p. 140, 2019.

TONOLO, Édwin Augusto; MARIANO, Juliana D'Angela; URBANETZ JUNIOR, Jair. ***Análise do efeito do acúmulo de sujeira nos sistemas fotovoltaicos da UTFPR: Câmpus Curitiba.*** VII Congresso Brasileiro de Energia Solar, Gramado, 17 a 20 de abril de 2018.

SANTOS, A. M.; SEIXAS, J. M.; PEREIRA, B. B.; MEDRONHO, R. A. Usando Redes Neurais Artificiais e Regressão Logística na predição da Hepatite A. Revista Brasileira de Epidemiologia, v.8, n.2, p. 117-126, 2005.

SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes neurais artificiais: para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010.

SILVA, Jéssica Alice Alves da. **Tratamento de dados meteorológicos e análise de desempenho do sistema fotovoltaico da emc/ufgp**. 190, [s.d.]. Dissertação (Mestrado) - Universidade Federal de Goiás, Escola de Engenharia Elétrica, Mecânica e de Computação (EMC), Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Goiânia, 2019.

VERA, Luís Horácio. **Comparação do desempenho de módulos fotovoltaicos após seis anos de operação**. p. 8, [s.d.], 2006.

## APÊNDICE A – TABELA COM OS PESOS DA REDE MLP COMBINADAS

Tabela 15 – Resultado dos neurônios intermediários para o neurônio de saída da MLP para os sistemas rural e urbano combinados.

Função Linear Node 0	
Inputs	Weights
Threshold	0.3160275128755254
Node 1	-0.9529560510442828
Node 2	-0.3276270212055282
Node 3	-1.0985491211911267
Node 4	-0.6041386014516622
Node 5	0.46130315187311793
Node 6	0.9327220758177361
Node 7	0.348052872266027
Node 8	0.980259261527884
Node 9	-1.0804744722467194
Node 10	-0.8607349784224919

Tabela 16 - Valor dos Pesos das entradas para os neurônios intermediários da MLP para os sistemas misturados.

Função Sigmoidal neurônio intermediario 1		Sigmoid Node 3		Sigmoid Node 5			
Inputs	Weights	Inputs	Weights	Inputs	Weights		
Threshold	-0.8692029779215964	Threshold	-5.924112170455725	Threshold	-1.3044034170130592		
Attrib ESTIMADO	-5.636788087414185	Attrib ESTIMADO	-3.214254631245239	Attrib ESTIMADO	-2.2072824506438624		
Attrib RAJVENTO	-3.9435387036468432	Attrib RAJVENTO	-0.08433234088587127	Attrib RAJVENTO	2.06619400505873		
Attrib CHUVA	-0.7503210262893966	Attrib CHUVA	-0.6044413559565973	Attrib CHUVA	1.6868341797242001		
Attrib TEMPINST	1.6142966883959389	Attrib TEMPINST	1.311748689562362	Attrib TEMPINST	0.39772619837428597		
Attrib TMAX	-1.5943339950978734	Attrib TMAX	1.1125264980784717	Attrib TMAX	-0.13123820150035898		
Attrib TMIN	4.065372198758839	Attrib TMIN	-2.774959170576792	Attrib TMIN	1.3610548088552863		
Attrib UMIDINST	1.1179545017126127	Attrib UMIDINST	-1.0686716227588824	Attrib UMIDINST	-0.9448843039061053		
Attrib VELVENTO	-1.1650549005257467	Attrib VELVENTO	2.0367523670590404	Attrib VELVENTO	-0.2518121369254776		
Sigmoid Node 2		Sigmoid Node 4		Sigmoid Node 6		Sigmoid Node 10	
Inputs	Weights	Inputs	Weights	Inputs	Weights	Inputs	Weights
Threshold	-4.6714738459237	Threshold	-1.197601980642725	Threshold	-6.018543552304461	Threshold	-5.615559788214252
Attrib ESTIMADO	4.7645524376430375	Attrib ESTIMADO	-2.83779115255706	Attrib ESTIMADO	9.311799985727436	Attrib ESTIMADO	3.3973470437035296
Attrib RAJVENTO	1.6614441969417528	Attrib RAJVENTO	-1.9173286335713375	Attrib RAJVENTO	-1.165221720818766	Attrib RAJVENTO	0.5849452096610686
Attrib CHUVA	0.2958388854671739	Attrib CHUVA	2.0887490778816504	Attrib CHUVA	-0.966211471121935	Attrib CHUVA	0.9181851043515009
Attrib TEMPINST	-1.8519989280731761	Attrib TEMPINST	-2.4685764803664902	Attrib TEMPINST	-0.41226417898107104	Attrib TEMPINST	-0.8829319817635736
Attrib TMAX	7.39563971403363	Attrib TMAX	-0.7336835713682356	Attrib TMAX	0.14927493156498403	Attrib TMAX	-0.3938019354098877
Attrib TMIN	-3.225658293548183	Attrib TMIN	4.769955994171136	Attrib TMIN	0.2872743505221054	Attrib TMIN	-1.3837093859454024
Attrib UMIDINST	-0.22834835499770803	Attrib UMIDINST	-4.625466706983059	Attrib UMIDINST	0.043971459862043	Attrib UMIDINST	4.060081209961106
Attrib VELVENTO	-4.142851199246917	Attrib VELVENTO	0.2375651498833286	Attrib VELVENTO	0.09686222871251808	Attrib VELVENTO	2.837361032280105
Sigmoid Node 7		Sigmoid Node 8		Sigmoid Node 9			
Inputs	Weights	Inputs	Weights	Inputs	Weights		
Threshold	-4.790296798801808	Threshold	-5.028722546060273	Threshold	-3.5709116696636576		
Attrib ESTIMADO	4.150690952473633	Attrib ESTIMADO	0.6984280268026369	Attrib ESTIMADO	-1.6847705777333921		
Attrib RAJVENTO	0.46298076866985344	Attrib RAJVENTO	1.5016288728225415	Attrib RAJVENTO	2.729730586857033		
Attrib CHUVA	1.9176801804145434	Attrib CHUVA	1.1924534111829401	Attrib CHUVA	-0.2736149835579675		
Attrib TEMPINST	-0.7696745323996457	Attrib TEMPINST	-1.187109796196774	Attrib TEMPINST	-0.8225085200484729		
Attrib TMAX	1.3736290467519001	Attrib TMAX	-1.4132317342118805	Attrib TMAX	-0.7890623130699935		
Attrib TMIN	-3.994345775504521	Attrib TMIN	-1.060237340577925	Attrib TMIN	0.1729522937910023		
Attrib UMIDINST	-0.7488801414888573	Attrib UMIDINST	2.589538031670805	Attrib UMIDINST	1.5512868443819379		
Attrib VELVENTO	0.5399174979195236	Attrib VELVENTO	1.876512892837863	Attrib VELVENTO	-1.2312611185596822		

## APÊNDICE B – TABELA COM OS PESOS DA REDE MLP PARA O SISTEMA A

Tabela 17 – Resultado dos neurônios intermediários para o neurônio de saída da MLP para o sistema A.

Linear Node 0	
Inputs	Weights
Threshold	-0.005118648232187894
Node 1	-1.2139737395683408
Node 2	-0.9280304821489479
Node 3	1.1589483394641928
Node 4	-1.1505388827223149
Node 5	-0.3361556267701103
Node 6	1.5539364423975008
Node 7	-0.8468035497865599

Tabela 18 - Valor dos Pesos das entradas para os neurônios intermediários da MLP para o sistema A.

Função Sigmoidal neurônio intermediario 1		Sigmoid Node 3		Sigmoid Node 5	
Inputs	Weights	Inputs	Weights	Inputs	Weights
Threshold	0.08738820195173472	Threshold	-3.902499619727223	Threshold	-4.279447071470832
Attrib ESTIMADO	-2.410131653096561	Attrib ESTIMADO	5.1024057343744555	Attrib ESTIMADO	-3.1984020869303356
Attrib RAJVENTO	0.5617792096779964	Attrib RAJVENTO	0.3292475498177065	Attrib RAJVENTO	2.7144525623451052
Attrib CHUVA	0.4902508585805507	Attrib CHUVA	-0.6213411233879957	Attrib CHUVA	-4.771106869085033
Attrib TEMPINST	-1.5456181226756236	Attrib TEMPINST	1.02060466055102	Attrib TEMPINST	1.3452512843170235
Attrib TMAX	1.7561233472551747	Attrib TMAX	-4.307308633253856	Attrib TMAX	1.3986254286603175
Attrib TMIN	0.002947979641925828	Attrib TMIN	2.9400995452660306	Attrib TMIN	3.8217442952500993
Attrib UMIDINST	-1.7420616473560175	Attrib UMIDINST	-0.5324506737222684	Attrib UMIDINST	1.4204565562145253
Attrib VELVENTO	-2.3851119124590685	Attrib VELVENTO	1.1047873731000033	Attrib VELVENTO	1.4438836700995659
Sigmoid Node 2		Sigmoid Node 4		Sigmoid Node 6	
Inputs	Weights	Inputs	Weights	Inputs	Weights
Threshold	-4.794739202105884	Threshold	-2.7328647482381765	Threshold	1.0581106251271062
Attrib ESTIMADO	-4.1866776294829595	Attrib ESTIMADO	1.951104731923377	Attrib ESTIMADO	0.31103365540088906
Attrib RAJVENTO	-0.17756434454305986	Attrib RAJVENTO	-1.7905317558789793	Attrib RAJVENTO	-0.8802426696781531
Attrib CHUVA	0.41969389330618767	Attrib CHUVA	0.1086785660122842	Attrib CHUVA	-0.39342954788700235
Attrib TEMPINST	1.7825884386558595	Attrib TEMPINST	-0.17793164717731513	Attrib TEMPINST	-1.134604556299032
Attrib TMAX	-5.1565659417179095	Attrib TMAX	0.873266419929761	Attrib TMAX	1.95307098493438
Attrib TMIN	5.55054013205206	Attrib TMIN	-0.39404129591679093	Attrib TMIN	-1.4790396166982933
Attrib UMIDINST	-1.8108479948315515	Attrib UMIDINST	1.4107224506091989	Attrib UMIDINST	-2.511827175804855
Attrib VELVENTO	-4.621247907285199	Attrib VELVENTO	2.6396123706974404	Attrib VELVENTO	-0.7387109185220349
Sigmoid Node 7					
Inputs	Weights				
Threshold	-0.6589305910328952				
Attrib ESTIMADO	0.4829930740955613				
Attrib RAJVENTO	-1.223251381262927				
Attrib CHUVA	-1.0796843185237954				
Attrib TEMPINST	-0.9573342623400617				
Attrib TMAX	0.5947038050437572				
Attrib TMIN	-1.1236039156460897				
Attrib UMIDINST	-0.8007011241868863				
Attrib VELVENTO	0.9826690688036965				

## APÊNDICE C – TABELA COM OS PESOS DA REDE MLP PARA O SISTEMA B

Tabela 19 – Resultado dos neurônios intermediários para o neurônio de saída da MLP para o sistema B.

Linear Node 0	
Inputs	Weights
Threshold	-0.34026857386271536
Node 1	1.0495723528860401
Node 2	-0.840588865560528
Node 3	0.13105863215102154
Node 4	0.7689006093456472
Node 5	-2.0856769782514943
Node 6	-0.1937087802019858
Node 7	1.2368665843062092
Node 8	0.20149741122342169
Node 9	0.21856112971604322
Node 10	0.17093734770501628
Node 11	0.439742279318185
Node 12	0.6545156524001612
Node 13	-1.9580754490424899
Node 14	0.31083142309038586
Node 15	0.3284544885215624
Node 16	1.1435856377504965
Node 17	0.19460236920595161
Node 18	0.8162053788776062
Node 19	-0.11587581523708128

Tabela 20 - Valor dos Pesos das entradas para os neurônios intermediários da MLP para o sistema B.

Função Sigmoidal neurônio intermediário 1		Sigmoid Node 4		Sigmoid Node 7	
Inputs	Weights	Inputs	Weights	Inputs	Weights
Threshold	-0.5086519452777053	Threshold	-1.8926504702087477	Threshold	-1.4289141472859799
Attrib ESTIMADO	0.20108922723962055	Attrib ESTIMADO	-0.9724252015024106	Attrib ESTIMADO	-0.7727155435394412
Attrib RAJVENTO	-0.12394959561390308	Attrib RAJVENTO	-1.344207322963372	Attrib RAJVENTO	-3.142617063044426
Attrib CHUVA	-2.1629652306932163	Attrib CHUVA	-0.032019063450726644	Attrib CHUVA	-1.0394922644019298
Attrib TEMPINST	-0.37412578648607553	Attrib TEMPINST	0.05177789513168444	Attrib TEMPINST	-1.593853643862218
Attrib TMAX	-0.09777463313432887	Attrib TMAX	0.4445211595579937	Attrib TMAX	-0.09806975699451811
Attrib TMIN	-1.188987206959083	Attrib TMIN	-0.9587881054664196	Attrib TMIN	-4.119552022955718
Attrib UMIDINST	-0.9279327758130054	Attrib UMIDINST	-0.31102740115680505	Attrib UMIDINST	1.321619583786147
Attrib VELVENTO	1.0295310112140037	Attrib VELVENTO	0.8143933338483397	Attrib VELVENTO	1.089054320998505
Sigmoid Node 2		Sigmoid Node 5		Sigmoid Node 8	
Inputs	Weights	Inputs	Weights	Inputs	Weights
Threshold	-1.679006376263593	Threshold	-0.4472696848871334	Threshold	-1.276548415061418
Attrib ESTIMADO	-0.04269057004120052	Attrib ESTIMADO	0.3798037208361593	Attrib ESTIMADO	-0.5876515748933749
Attrib RAJVENTO	0.6777395390681925	Attrib RAJVENTO	-1.3063758552915545	Attrib RAJVENTO	-0.23247373972794377
Attrib CHUVA	1.7639256540159387	Attrib CHUVA	0.19157992759508743	Attrib CHUVA	0.2714484428969703
Attrib TEMPINST	-0.6780776768582699	Attrib TEMPINST	-3.072220836728066	Attrib TEMPINST	-0.07820977699559552
Attrib TMAX	-0.323600286967583	Attrib TMAX	-1.8756407877256158	Attrib TMAX	-0.08547951796562665
Attrib TMIN	0.07508593641880097	Attrib TMIN	-1.4047073149590807	Attrib TMIN	0.09642859461831135
Attrib UMIDINST	-0.08978113898624732	Attrib UMIDINST	-1.086483712424311	Attrib UMIDINST	0.0840839495855165
Attrib VELVENTO	1.2675469975744884	Attrib VELVENTO	1.2752886249150455	Attrib VELVENTO	0.5354909731492086
Sigmoid Node 3		Sigmoid Node 6		Sigmoid Node 9	
Inputs	Weights	Inputs	Weights	Inputs	Weights
Threshold	-1.3080590730538768	Threshold	-1.3706320712377047	Threshold	-1.304327708243335
Attrib ESTIMADO	-0.5446238216327653	Attrib ESTIMADO	-0.16717019154042062	Attrib ESTIMADO	-0.5899696947099108
Attrib RAJVENTO	-0.0470399320894334	Attrib RAJVENTO	-0.05182796968073653	Attrib RAJVENTO	-0.2847815718861364
Attrib CHUVA	0.44684332964348655	Attrib CHUVA	0.8552454870603892	Attrib CHUVA	0.2695815200587968
Attrib TEMPINST	-0.3037142746022794	Attrib TEMPINST	-0.45267000472247815	Attrib TEMPINST	-0.009586478742840108
Attrib TMAX	-0.193543755466161	Attrib TMAX	-0.34310267372557623	Attrib TMAX	0.004471648664649293
Attrib TMIN	0.17596085071308465	Attrib TMIN	0.05348905784901749	Attrib TMIN	0.0053276584722151965
Attrib UMIDINST	0.055079918139172505	Attrib UMIDINST	0.267462936418352	Attrib UMIDINST	0.09729020448192628
Attrib VELVENTO	0.5096440139432641	Attrib VELVENTO	0.5300984742315166	Attrib VELVENTO	0.5078223495065467

Sigmoid Node 10		Sigmoid Node 13		Sigmoid Node 16			
Inputs	Weights	Inputs	Weights	Inputs	Weights		
Threshold	-1.3051702459502827	Threshold	-0.9257129698398936	Threshold	-1.6288757506240148		
Attrib ESTIMADO	-0.6028915682685974	Attrib ESTIMADO	-2.2374542256297816	Attrib ESTIMADO	-1.390246814753177		
Attrib RAJVENTO	-0.0809815006204381	Attrib RAJVENTO	0.4903073327756026	Attrib RAJVENTO	0.07210771711589727		
Attrib CHUVA	0.43119679441859493	Attrib CHUVA	-1.6314726010997131	Attrib CHUVA	1.3562632560411765		
Attrib TEMPINST	-0.29606552679557474	Attrib TEMPINST	0.07717101089641942	Attrib TEMPINST	-0.871963235597443		
Attrib TMAX	-0.17195166284652316	Attrib TMAX	0.041805784154274266	Attrib TMAX	-0.19536443917319107		
Attrib TMIN	0.20495041480790166	Attrib TMIN	-1.2541273996933584	Attrib TMIN	0.22773921906329603		
Attrib UMIDINST	0.007421325883242003	Attrib UMIDINST	0.7257752539771114	Attrib UMIDINST	-1.7052796445034104		
Attrib VELVENTO	0.4584519916442801	Attrib VELVENTO	1.5155811539749686	Attrib VELVENTO	0.17953311131508876		
Sigmoid Node 11		Sigmoid Node 14		Sigmoid Node 17		Sigmoid Node 19	
Inputs	Weights	Inputs	Weights	Inputs	Weights	Inputs	Weights
Threshold	-1.383232369025205	Threshold	-1.3184992621633609	Threshold	-1.285898077254715	Threshold	-1.3619640101305641
Attrib ESTIMADO	-0.6135274039131658	Attrib ESTIMADO	-0.5836552773434294	Attrib ESTIMADO	-0.6320493466155653	Attrib ESTIMADO	-0.42679985888223765
Attrib RAJVENTO	-0.7343420363346317	Attrib RAJVENTO	-0.4473553192683433	Attrib RAJVENTO	-0.09102173213463836	Attrib RAJVENTO	0.2687432100513545
Attrib CHUVA	-0.019661042338622467	Attrib CHUVA	0.10139826995782855	Attrib CHUVA	0.4469516257636377	Attrib CHUVA	0.7808566051763688
Attrib TEMPINST	0.20805148920226868	Attrib TEMPINST	0.1160045003704995	Attrib TEMPINST	-0.35695733099723476	Attrib TEMPINST	-0.576831714257748
Attrib TMAX	0.3031007492960237	Attrib TMAX	0.18199341810531053	Attrib TMAX	-0.16577655436908809	Attrib TMAX	-0.42848701277185974
Attrib TMIN	-0.28795484196429416	Attrib TMIN	-0.1332610257778881	Attrib TMIN	0.21723740925253185	Attrib TMIN	0.2962234394374399
Attrib UMIDINST	0.12670219786899123	Attrib UMIDINST	0.19391378883811375	Attrib UMIDINST	-0.09028641629698855	Attrib UMIDINST	0.09482074243675144
Attrib VELVENTO	0.6354819581357759	Attrib VELVENTO	0.5614847444543212	Attrib VELVENTO	0.47681187809859715	Attrib VELVENTO	0.821134461202551
Sigmoid Node 12		Sigmoid Node 15		Sigmoid Node 18			
Inputs	Weights	Inputs	Weights	Inputs	Weights		
Threshold	-3.156585717272144	Threshold	-1.3406108044060772	Threshold	-0.18614942306149476		
Attrib ESTIMADO	1.4799987109983055	Attrib ESTIMADO	-0.830447662869082	Attrib ESTIMADO	-0.14870238803609864		
Attrib RAJVENTO	3.1155071724221584	Attrib RAJVENTO	0.05520375849561949	Attrib RAJVENTO	-1.100851265617582		
Attrib CHUVA	0.7465183902531253	Attrib CHUVA	0.6273363050250155	Attrib CHUVA	-0.2512449760320275		
Attrib TEMPINST	0.0015029328639627603	Attrib TEMPINST	-0.594439176467149	Attrib TEMPINST	1.0473086688443396		
Attrib TMAX	-0.26511485252678624	Attrib TMAX	-0.2851767830292128	Attrib TMAX	0.8403275105563995		
Attrib TMIN	-2.1933348637305716	Attrib TMIN	0.23667076343350393	Attrib TMIN	0.5429338886573144		
Attrib UMIDINST	-1.6603970033453959	Attrib UMIDINST	-0.4860505537020769	Attrib UMIDINST	-0.31945564963501444		
Attrib VELVENTO	1.588904875949264	Attrib VELVENTO	0.41896018862204976	Attrib VELVENTO	1.9649538711420196		

