

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA**  
**ENGENHARIA ELÉTRICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**Denise Sanches de Melo**

**Aplicação de Redes Neurais LSTM em Medição e Verificação de Resultados de  
Projetos com Fontes Renováveis: Estudo de caso de projeto com sistema  
fotovoltaico**

Juiz de Fora  
2022

**Denise Sanches de Melo**

**Aplicação de Redes Neurais LSTM em Medição e Verificação de Resultados de  
Projetos com Fontes Renováveis: Estudo de caso de projeto com sistema  
fotovoltaico**

Dissertação apresentada a Pós –  
Graduação em Engenharia Elétrica da  
Universidade Federal de Juiz de Fora  
como requisito parcial à obtenção do título  
de Mestre em Engenharia Elétrica. Área  
de concentração: Sistema de Energia.

Orientador: Leonardo Willer de Oliveira, D.Sc.

Coorientador: Janaína Gonçalves de Oliveira, Ph.D.

Juiz de Fora

2022

Sanches de Melo, Denise .

Aplicação de Redes Neurais LSTM em Medição e Verificação de Resultados de Projetos com Fontes Renováveis: Estudo de caso de projeto com sistema fotovoltaico / Denise Sanches de Melo. -- 2022. 126 f. : il.

Orientador: Leonardo Willer de Oliveira

Coorientadora: Janaína Gonçalves de Oliveira

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Engenharia. Programa de Pós-Graduação em Engenharia Elétrica, 2022.

1. M&V. 2. Rede Neural LSTM. 3. Previsão de Irradiação Solar. 4. Energia Solar. I. Oliveira, Leonardo Willer de , orient. II. Oliveira, Janaína Gonçalves de , coorient. III. Título.

**Denise Sanches de Melo**

**Aplicação de Redes Neurais**

**LSTM em Medição e Verificação de Resultados de Projetos com Fontes Renováveis: estudo de caso com implementação de sistema fotovoltaico**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Sistemas de Energia Elétrica

Aprovada em 08 de fevereiro de 2022.

**BANCA EXAMINADORA**

**Prof. Dr. Leonardo Willer de Oliveira** - Orientador

Universidade Federal de Juiz de Fora

**Profa. Dra. Janaína Gonçalves de Oliveira**

Universidade Federal de Juiz de Fora

**Prof. Dr. Vitor Hugo Ferreira**

Universidade Federal Fluminense

**Prof. Dr. Ângelo Rocha de Oliveira**

Centro Federal de Educação Tecnológica de Minas Gerais

**Prof. Dr. Bruno Henriques Dias**

Universidade Federal de Juiz de Fora

Juiz de Fora, 27/12/2021.

---



Documento assinado eletronicamente por **Leonardo Willer de Oliveira, Professor(a)**, em 08/02/2022, às 14:02, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Angelo Rocha de Oliveira, Usuário Externo**, em 08/02/2022, às 14:15, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Bruno Henriques Dias, Professor(a)**, em 08/02/2022, às 15:01, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Vitor Hugo Ferreira, Usuário Externo**, em 09/02/2022, às 22:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Janaina Goncalves de Oliveira, Professor(a)**, em 17/02/2022, às 22:11, conforme horário oficial de Brasília, 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 Portal do SEI-Ufjf ([www2.ufjf.br/SEI](http://www2.ufjf.br/SEI)) através do ícone Conferência de Documentos, informando o código verificador **0627865** e o código CRC **811A9E1D**.

---

## AGRADECIMENTOS

Antes de tudo agradecer a Deus por sempre ter colocado pessoas maravilhosas na minha vida, que me ajudaram nessa caminhada.

Agradecer aos meus pais, Ana Sanches e José das Graças, pela educação, pelo suporte, pelo amor e pelo apoio em minhas decisões.

Às minhas irmãs, Cláudia e Danielle, pelo carinho e amizade de sempre.

Agradecer ao meu marido Anderson pelo apoio, incentivo e paciência, durante todo o mestrado.

Agradecer a toda minha família, aos meus amigos, professores, amigos das aulas do mestrado e as empresas que tive o prazer de trabalhar, além de muitos dos meus clientes, que sempre estiveram torcendo por mim, ajudando e impactando de alguma forma na minha vida acadêmica e profissional, positivamente.

Ao meu orientador Leonardo Willer de Oliveira, pela excelência na orientação e dedicação dispensada para a realização deste trabalho e por acreditar no meu trabalho, além da paciência quanto a compreensão do meu dia a dia, sendo dividido pelo trabalho e vida acadêmica. Agradeço também a minha coorientadora Janaína Gonçalves, que abriu as portas do mestrado para mim, me mostrando que era possível acreditar neste sonho.

Aos amigos da DEODE que direta ou indiretamente, contribuíram e me apoiaram no desenvolvimento deste trabalho. Agradeço ao Frederico Araújo, por ser compreensivo e apoiador nesta etapa do mestrado. Um agradecimento especial ao Ricardo Gomes que me ensinou muito sobre redes neurais e em muitos momentos me apoiou e me mostrou caminhos para concretizar esta dissertação, frente aos desafios da programação. A Lara Lacera, Fabiane Barbosa e Walquíria Silva, minhas amigas, agradeço imensamente pelo grande apoio, carinho e me ajuda nos momentos difíceis que vivenciei frente ao grande desafio, chamado mestrado.

A UFJF e ao LABSPOT (Laboratório de Sistemas de Potência da Faculdade de Engenharia Elétrica da Universidade Federal de Juiz de Fora), pela estrutura disponibilizada para a realização deste trabalho.

Finalmente, agradeço a CAPES, ao CNPQ e ao INERGE pelo apoio financeiro para este trabalho.

“A persistência é o caminho do êxito.” (Chaplin, 1997).

## RESUMO

As campanhas de medição e verificação (M&V) de resultados em projetos de eficiência energética desempenham um papel fundamental na avaliação dos reais benefícios oriundos das ações implementadas. Em projetos que contemplam a implementação de geração solar, umas das variáveis independentes consideradas no processo é a irradiação solar, de que depende da variação da energia provida pelo sistema. Neste intuito, uma metodologia com aplicação de redes neurais LSTM (*Long-Short Term Memory*) para a previsão da “irradiação solar” é proposta, com o objetivo de reduzir o tempo de acompanhamento dos resultados do projeto de doze para seis meses, tendo em vista as dificuldades enfrentadas em períodos longos de monitoramento. A avaliação do método aplicado ao problema na M&V para projetos com sistema fotovoltaico (SFV), apresentou resultados satisfatórios, identificando uma solução alternativa com a aplicação de redes neurais LSTM.

Palavras-chave: M&V, Rede Neural LSTM, Previsão de Irradiação Solar, Energia Solar.

## **ABSTRACT**

The measurement and verification (M&V) campaigns of results in energy efficiency projects play a fundamental role in evaluating the real benefits arising from the implemented actions. In projects that include the implementation of solar generation, one of the independent variables considered in the process is “solar irradiation”, on which the variation of energy provided by the system depends. With this in mind, a methodology with the application of LSTM neural networks (Long-Short Term Memory) for the prediction of "solar irradiation" is proposed, with the objective of reducing the time for monitoring the project results from twelve to six months, taking in given the difficulties faced in long periods of monitoring. The evaluation of the method applied to the problem in M&V for projects with a photovoltaic system (SFV), showed satisfactory results, identifying an alternative solution with the application of LSTM neural networks.

Keywords: M&V, LSTM Neural Network, Solar Radiation Forecast, Solar Energy.

## LISTA DE ILUSTRAÇÕES

Gráfico 4.1 – Treinamento x Validação .....	51
Gráfico 5.1 – Gráfico de dispersão Irradiação Solar x Geração .....	68
Gráfico 5.2 - Regressão linear (geração x irradiação).....	72
Gráfico 5.3 – Dados reais de “Geração e Irradiação” x Dados modelo proposto “Geração e Irradiação” .....	81
Gráfico 5.4 – Regressão linear (geração x irradiação).....	88
Gráfico 5.5 – Dados reais de “Geração e Irradiação” x Dados modelo proposto “Geração e Irradiação” .....	96
Figura 2.1 - Registro da geração e irradiação .....	19
Figura 2.2 – Componentes da irradiância solar.....	26
Figura 2.3 – Esquema de uma rede neural com duas camadas ocultas.....	30
Figura 2.4 – Uma unidade de célula LSTM .....	32
Figura 2.5 – Unidade de célula LSTM – Portão de esquecimento .....	32
Figura 2.6 – Unidade de célula LSTM – Portão de entrada .....	33
Figura 2.7 – Unidade de célula LSTM – Estado da célula.....	35
Figura 2.8 – Unidade de célula LSTM – Estado da saída .....	36
Figura 3.1 – Fluxo do processo de M&V no Método tradicional .....	40
Figura 4.1 – Diagrama do fluxo de entradas e saídas da rede.....	47
Figura 4.2 – Diagrama Esquemático do <i>Framework</i> proposto .....	48
Figura 4.3 – Fluxo do processo de M&V na Metodologia Proposta.....	61
Figura 5.1 – Características do SFV .....	66
Figura 5.2 – Resultado de desempenho – Linha de Base.....	66
Figura 5.3 – Resultado de irradiação e geração – Linha de Base.....	67
Figura 5.4 – Características do SFV. ....	83
Figura 5.5 – Resultado de desempenho – Linha de Base.....	83
Figura 5.6 – Resultado de irradiação e geração – Linha de Base.....	84

## LISTA DE TABELAS

Tabela 2.1– Visão geral das opções do PIMVP .....	22
Tabela 3.1 - Descrição dos passos da M&V no Método Tradicional .....	41
Tabela 4.1 - Descrição dos passos da rede LSTM.....	49
Tabela 4.2 - Descrição dos passos da M&V no método proposto .....	62
Tabela 5.1 – Dados de Irradiação Solar no Plano Horizontal (Ano padrão).....	65
Tabela 5.2 – Regressão linear (geração estimada PVSyst x irradiação solar).....	67
Tabela 5.3 – Registro da geração e irradiação solar .....	70
Tabela 5.4 – Dados da Estação Meteorológica .....	71
Tabela 5.5 – Comparativo entre dados meteorológicos entre cidades.....	72
Tabela 5.6 – Regressão linear (geração medida x irradiação solar INMET). .....	73
Tabela 5.7 – Determinação dos Resultados de Geração “Normalizada” .....	74
Tabela 5.8 – Dados de irradiação solar comparativos LSTM x INMET .....	76
Tabela 5.9 – Dados de irradiação solar comparativos MÉDIA x INMET .....	77
Tabela 5.10 – Dados de irradiação solar comparativos MÉDIA x INMET .....	77
Tabela 5.11 – Regressão linear (geração x irradiação solar).....	78
Tabela 5.12 – Método proposto com LSTM .....	79
Tabela 5.13 – Regressão linear (geração x irradiação solar).....	79
Tabela 5.14 – Modelagem da geração de energia “Normalizada” dados LSTM .....	80
Tabela 5.15 – Dados de Irradiação Solar no Plano Horizontal (Ano padrão).....	82
Tabela 5.16 – Regressão linear (geração estimada PVSyst x irradiação solar).....	85
Tabela 5.17– Registro da geração e irradiação.....	87
Tabela 5.18– Dados da Estação Meteorológica.....	87
Tabela 5.19 – Regressão linear (geração medida x irradiação solar INMET). .....	88
Tabela 5.20 – Determinação dos Resultados de Geração “Normalizada” .....	90
Tabela 5.21 – Dados de irradiação solar comparativos LSTM x INMET .....	91
Tabela 5.22 – Dados de irradiação solar comparativos LSTM x INMET com 4 anos	92
Tabela 5.22 – Regressão linear (geração x irradiação solar).....	93
Tabela 5.23 – Método proposto com LSTM .....	94
Tabela 5.24 – Regressão linear (geração x irradiação solar).....	94
Tabela 5.25 – Modelagem da geração de energia “Normalizada” dados LSTM .....	95

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>13</b>
1.1	OBJETIVOS .....	14
1.2	PUBLICAÇÕES E/OU SUBMISSÕES DECORRENTES DA PESQUISA ....	15
1.3	ESTRUTURA DO TRABALHO .....	15
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b> .....	<b>17</b>
2.1	BASES CONCEITUAIS .....	18
<b>2.1.1</b>	<b>Construção da Estratégia e Linha de Base de M&amp;V</b> .....	<b>18</b>
2.1.1.1	Estratégia de M&V .....	20
2.1.1.2	Plano de M&V .....	23
<b>2.1.2</b>	<b>Definição do Período de Determinação dos Resultados de M&amp;V</b> .....	<b>24</b>
<b>2.1.3</b>	<b>Variável independente “Irradiação Solar”</b> .....	<b>25</b>
<b>2.1.4</b>	<b><i>Deep Learning</i> e suas Aplicações em Previsões Temporais</b> .....	<b>28</b>
<b>2.1.5</b>	<b>Rede Neural LSTM</b> .....	<b>30</b>
<b>2.1.6</b>	<b>Previsão de Séries</b> .....	<b>37</b>
<b>3</b>	<b>DESCRIÇÃO DO PROBLEMA DE M&amp;V EM SFV</b> .....	<b>38</b>
3.1	INTRODUÇÃO .....	38
3.2	OBJETIVOS .....	38
3.3	M&V no Método Tradicional .....	39
<b>3.3.1</b>	<b>Fluxo do processo de M&amp;V no Método tradicional</b> .....	<b>39</b>
<b>3.3.2</b>	<b>Permanência das condições construtivas</b> .....	<b>41</b>
<b>3.3.3</b>	<b>Condições dos dados no monitoramento a distância</b> .....	<b>42</b>
<b>3.3.4</b>	<b>Atendimento às condições contratuais</b> .....	<b>43</b>
3.4	CONCLUSÕES PARCIAIS .....	43
<b>4</b>	<b>METODOLOGIA PROPOSTA</b> .....	<b>45</b>
4.1	INTRODUÇÃO .....	45
4.2	REDE NEURAL LSTM .....	46
<b>4.2.1</b>	<b>Tratamento dos Dados</b> .....	<b>46</b>
<b>4.2.2</b>	<b>Rede LSTM Proposta</b> .....	<b>47</b>
4.2.2.1	Treinamento da rede .....	49
4.2.2.2	Normalização .....	53
4.2.2.3	Aplicação da rede LSTM .....	53
<b>4.2.3</b>	<b>Métrica para a análise</b> .....	<b>54</b>
4.3	CÁLCULO DA GERAÇÃO NO PERÍODO DE PREVISÃO .....	55
4.4	NORMALIZAÇÃO DA GERAÇÃO .....	56

4.5	APLICAÇÃO DAS MÉTRICAS DE AVALIAÇÃO .....	57
4.6	FLUXOGRAMA DA METODOLOGIA .....	60
4.7	CONCLUSÕES PARCIAIS .....	62
<b>5</b>	<b>TESTES E RESULTADOS .....</b>	<b>64</b>
5.1	INTRODUÇÃO.....	64
5.2	SFV NO ESTADO DE TOCANTINS .....	64
<b>5.2.1</b>	<b>M&amp;V método tradicional.....</b>	<b>64</b>
5.2.1.1	Definição da Linha de Base .....	64
5.2.1.1.1	Definição das Metas e Condições Físicas do Projeto .....	65
5.2.1.1.2	Variável Independente .....	67
5.2.1.1.3	Fronteira de Medição e Efeitos Interativos com o Projeto.....	68
5.2.1.1.4	Opção do PIMVP .....	69
5.2.1.1.5	Fatores Estáticos .....	69
5.2.1.2	Definição do período de determinação da geração obtida .....	69
5.2.1.2.1	Dados dos resultados durante o período de monitoramento do projeto	70
5.2.1.2.2	Variável Independente .....	72
5.2.1.2.3	Geração de energia do SFV “Normalizada” .....	74
<b>5.2.2</b>	<b>Rede neural LSTM para o Projeto em Tocantins .....</b>	<b>75</b>
5.2.2.1	Resultados de irradiação solar projetados pela rede neural LSTM .....	75
5.2.2.2	Aplicação da Rede neural LSTM no método proposto para M&V .....	78
5.2.2.3	Método proposto para M&V com a normalização da geração solar .....	79
5.3	SFV NO ESTADO DE GOIÁS .....	81
<b>5.3.1</b>	<b>M&amp;V método tradicional.....</b>	<b>81</b>
5.3.1.1	Definição da Linha de Base .....	82
5.3.1.1.1	Definição das Metas e Condições Físicas do Projeto .....	82
5.3.1.1.2	Variável Independente .....	84
5.3.1.1.3	Fronteira de Medição e Efeitos Interativos com o Projeto.....	85
5.3.1.1.4	Opção do PIMVP .....	86
5.3.1.1.5	Fatores Estáticos .....	86
5.3.1.2	Período de Determinação da Economia e Condições para o ajuste de base	86
5.3.1.2.1	Dados dos resultados durante o período de monitoramento do projeto:	86
5.3.1.2.2	Variável Independente .....	88
5.3.1.2.3	Geração de energia do SFV “Normalizada” .....	89
<b>5.3.2</b>	<b>Rede neural LSTM para o Projeto em Goiânia .....</b>	<b>90</b>

5.3.2.1	Resultados de irradiação solar projetados pela rede neural LSTM .....	91
5.3.2.2	Aplicação da Rede neural LSTM no método tradicional de M&V .....	92
5.3.2.3	Método proposto para M&V com a normalização da geração solar .....	94
5.4	CONCLUSÕES PARCIAIS .....	96
<b>6</b>	<b>CONCLUSÕES</b> .....	<b>97</b>
6.1	CONSIDERAÇÕES FINAIS.....	97
6.2	TRABALHOS FUTUROS.....	98
	<b>REFERÊNCIAS</b> .....	<b>99</b>
	<b>APÊNDICE A – Distribuição <math>t</math> de Student e Estatística <math>t</math></b> .....	<b>103</b>
	<b>APÊNDICE B – Desenvolvimento do código e o código</b> .....	<b>105</b>

## 1 INTRODUÇÃO

A avaliação dos resultados energéticos de ações de eficiência energética (EE) passa, necessariamente, por medições de campo, mas não se restringem a elas, já que não se pode medir diretamente a EE. Faz-se necessário, portanto, um processo de análise que possa estimar o consumo da instalação antiga nas condições após a intervenção. Para tal, é necessário que se façam medições antes da ação de EE. Para o caso de projetos com sistema de energia solar, a medição a ser feita é da geração obtida. Caso seja identificada a existência de variáveis que determinam ou influenciam o consumo de energia, deve ser estabelecido um modelo matemático que associe tais variáveis a esse consumo de energia (ANEEL, 2020).

A atividade de avaliação dos resultados energéticos de projetos com recursos oriundos do PROPEE (Procedimentos do Programa de Eficiência Energética) da ANEEL (Agência Nacional de Energia Elétrica) é baseada no Protocolo Internacional para Medição e Verificação de Performance (PIMVP), que descreve as melhores práticas atualmente disponíveis para medir e verificar os resultados de projetos de EE. O PIMVP não determina quais medições devem ser feitas, quais variáveis considerar, qual modelo determinar, dada a diversidade de situações que se apresentam na prática (ANEEL, 2020).

Mas determina apenas alguns requisitos básicos a observar, os cuidados que se deve ter, critérios para selecionar as variáveis e opções disponíveis para avaliar a EE, até porque medições mais prolongadas, com um número maior de variáveis e com maior precisão, podem diminuir a incerteza, porém aumentam os custos. Assim, o equilíbrio entre precisão e custo é fundamental para uma boa prática de M&V (ANEEL, 2020).

A energia solar vem sendo a fonte que apresenta o maior incremento de capacidade instalada anualmente no mundo, podendo ser explicado dados aos preços decrescentes verificados nos últimos anos, à robustez tecnológica tendo em vista projetos com mais de 30 anos em funcionamento, o vasto potencial técnico existente e a não emissão de gases de efeito estufa durante a operação dos parques. Assim, dada a redução de custos considerada, a fonte solar fotovoltaica se apresenta como alternativa competitiva no fornecimento de energia, podendo contribuir com os compromissos nacionais de redução de gases de efeito estufa, atrelados aos benefícios oriundos dos projetos de eficiência energética (PNE, 2020).

Quando se tem projetos com implementação de sistemas fotovoltaicos, uma das principais variáveis independentes é a irradiação solar, sendo fundamental o estabelecimento desta variável para a indicação da variação da energia. Além disso, a M&V deve ter acompanhamento de doze meses para a contabilização dos resultados de geração, ou na disponibilidade dos dados locais, trabalhar com um período menor de acompanhamento (ANEEL, 2020).

Diante deste cenário, alguns trabalhos foram encontrados no meio acadêmico, com a aplicação de *Deep Learning* (DL) em sistemas de energia usando *Long-Short Term Memory* (LSTM). Em Muhammad et. al. (2019), apresenta-se um estudo de caso sobre a previsão de irradiância solar necessária para a geração fotovoltaica usando redes neurais LSTM. E em Malakar et. al. (2021), é detalhada a previsão de curto prazo da irradiância no plano horizontal global, com um modelo com rede neural recorrente e LSTM, respectivamente, em termos de precisão de previsão.

Com inspiração nos assuntos abordados até o momento, o presente trabalho apresenta uma proposta de metodologia com aplicação de redes neurais LSTM na previsão da variável independente do processo de M&V de projetos com SFV, a “irradiação solar”, com foco na redução do tempo de acompanhamento dos resultados, passando de doze meses para seis meses, em casos reais localizados no estado de Goiás e Tocantins. No estudo, dados de medição da geração solar real de doze meses e dados medidos obtidos do banco de dados histórico do INMET (Instituto Nacional de Metrologia, Normalização e Qualidade Industrial) são aplicados.

## 1.1 OBJETIVOS

O presente trabalho tem como objetivo desenvolver uma metodologia para M&V em projetos com SFV com recursos do PEE da ANEEL, utilizando redes neurais LSTM para prever a variável independente “irradiação”, visando reduzir o tempo de acompanhamento das medições de 12 para 6 meses. Esta metodologia alia a aplicação das redes neurais LSTM, para tratar os dados da irradiação local subsequentes a implementação dos projetos, prevendo seis meses a frente, mês a

mês, possibilitando a projeção da geração solar dos seis meses que deixaram de ser acompanhados. Para tanto, os objetivos específicos são listados a seguir:

- Tratar os dados horários de irradiação solar da cidade onde o projeto foi implementado;
- Tratar os dados de geração do projeto, correlacionando aos dados de irradiação solar;
- Implementar esta nova metodologia em dois projetos, um na cidade de Goiânia – GO e um na cidade de Araguatins - TO
- Comparar os resultados utilizando seis meses com os dados reais de geração de um ano, com o intuito de analisar os resultados e validar o modelo.

## 1.2 PUBLICAÇÕES E/OU SUBMISSÕES DECORRENTES DA PESQUISA

A pesquisa desenvolvida no âmbito da presente dissertação resultou em uma publicação em anais de evento técnico-científico nacional na área de conhecimento, cujos dados são fornecidos a seguir:

“Aplicação de Redes Neurais Artificiais para Previsão de Velocidade de Vento Considerando Variáveis Correlacionadas”, MELO, Denise S.; OLIVEIRA, Leonardo W.; OLIVEIRA, Janaína G.. Simpósio Brasileiro de Sistemas Elétricos-SBSE, v. 1, n. 1, 2020.

## 1.3 ESTRUTURA DO TRABALHO

Além deste capítulo introdutório, a presente dissertação contém outros cinco capítulos e dois apêndices. O capítulo II apresenta uma revisão bibliográfica e o capítulo III descreve o problema da M&V em SFV pelos métodos tradicionais.

O capítulo IV apresenta a metodologia proposta para M&V em projetos com SFV, utilizando a rede neural LSTM para a previsão da variável independente, a irradiação solar.

No capítulo V, são apresentados e discutidos os resultados obtidos mediante a aplicação da metodologia proposta.

No capítulo VI, são apresentadas as principais conclusões do trabalho e sugestões para desenvolvimentos futuros.

## 2 REVISÃO BIBLIOGRÁFICA

Conforme determina a legislação específica, em particular a Lei nº 9.991, de 24 de julho de 2000, as empresas concessionárias ou permissionárias de distribuição de energia elétrica, devem aplicar um percentual mínimo da receita operacional líquida (ROL) em PEE e Pesquisa e Desenvolvimento (P&D), segundo regulamentos da Agência Nacional de Energia Elétrica – ANEEL, sendo destes recursos que sujam os projetos de EE direcionados a todos os tipos de consumidores pertencentes a área de concessão de cada concessionária de energia (ANEEL, 2020).

As campanhas de M&V de projetos de EE oriundos do PEE, no caso específico de geração a partir de Fonte Incentivada<sup>1</sup>, desempenham um papel fundamental na avaliação dos reais benefícios energéticos oriundos das ações implementadas, os quais são o foco da avaliação dos projetos por parte da ANEEL (ANEEL, 2020).

Dada a importância do problema, em se aferir os resultados de projetos desta natureza, a ANEEL, através do PROPEE (ANEEL, 2020), estabelece em seu módulo 8 de Medição e Verificação de Resultados, que as atividades de avaliação dos resultados energéticos de projetos deverão ser baseadas no Protocolo Internacional para Medição e Verificação de Performance (PIMVP - EVO, 2012). Esse protocolo descreve as melhores práticas atualmente disponíveis para medir e verificar os resultados de projetos de EE.

O trabalho de Gallagher et. al. (2018) propõe uma metodologia de M&V com adequação de aprendizado de máquina para minimizar a incerteza na medição e verificação de economia de energia de projetos em EE, colocando que algoritmos de *Machine Learning* (ML) oferecem uma abordagem alternativa para treinar esses modelos com técnicas baseadas em dados. Além disso, tratando-se de aplicação de aprendizado de máquinas em sistema de energia, Muhammad et. al. (2019) apresenta o estado da arte da aplicação de DL em sistemas de energia, e traz um estudo de caso sobre a previsão de irradiância, usando LSTM.

---

<sup>1</sup> Entende-se como geração a partir de Fonte Incentivada a central geradora de energia elétrica definida na Resolução Normativa nº 482, de 17 de abril de 2012 (ANEEL, 2012), e suas alterações.

Nesta seção, o estado da arte é apresentado de forma organizada por níveis de metodologias, envolvendo a abordagem para a medição e verificação dos resultados com dois níveis, sendo: a construção da linha de base nos moldes tradicionais e a construção do período de determinação da economia segundo (PIMVP - EVO, 2012). Por fim, apresenta-se uma abordagem voltada para o desenvolvimento proposto neste trabalho, com DL e suas aplicações em previsões temporais e redes neurais LSTM.

## 2.1 BASES CONCEITUAIS

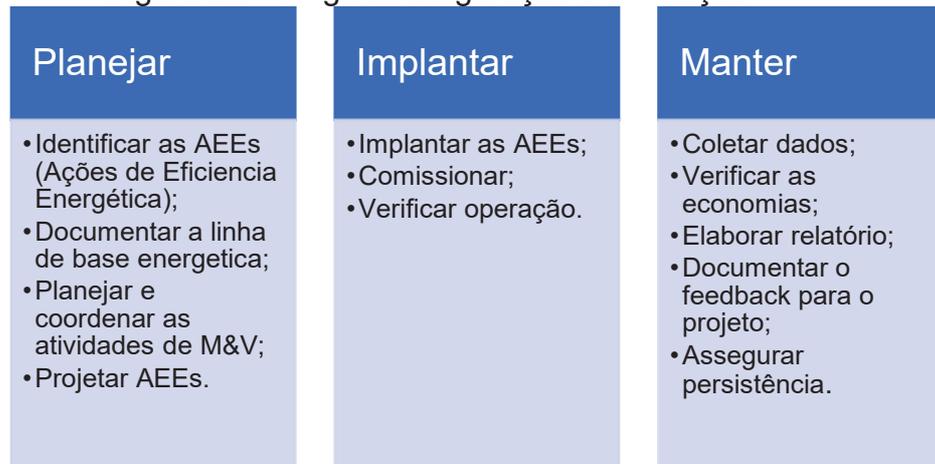
### 2.1.1 Construção da Estratégia e Linha de Base de M&V

A M&V é o processo de utilização de medições em projetos de EE para determinar, de modo seguro, a economia real gerada dentro de uma instalação individual por um projeto. A economia não pode ser medida diretamente, uma vez que representa a ausência de consumo de energia. Nesse caso, a economia é determinada pela comparação do consumo medido antes e depois da implementação de um projeto, com ajustes adequados, tendo em conta alterações nas condições. As atividades de M&V consistem em algumas ou todas as seguintes ações (EVO, 2012):

- Instalação, calibração e manutenção de medidores;
- Coleta e tratamento de dados;
- Desenvolvimento de um método de cálculo e estimativas aceitáveis;
- Cálculos com os dados medidos;
- Relatórios, garantia de qualidade e verificação de relatórios por terceiros.

O processo de M&V tem total sinergia com as atividades de um projeto conforme Figura 2.1, o que permite que as responsabilidades e papéis sejam distribuídas para os próprios integrantes da equipe, de acordo com cada especialidade de profissional presente no projeto (EVO, 2012).

Figura 2.1 - Registro da geração e irradiação



Fonte: EVO (2012), adaptada pelo autor (2021).

Os resultados da M&V devem ser tão precisos quanto possa ser justificado pelo valor do projeto. Os custos de M&V devem ser normalmente ‘pequenos’ em relação ao valor da economia em avaliação.

As despesas oriundas de M&V em projetos devem ser consistentes com as implicações financeiras de se sub ou superestimar o desempenho de um projeto. A precisão da metodologia de M&V e seu custo devem ser avaliados como parte do desenvolvimento do projeto. O compromisso entre custo e precisão deve ser acompanhado por um maior conservadorismo quando houver um aumento de valores estimados e julgamentos. A consideração de todos os fatores que afetam a precisão de forma razoável é um princípio orientador do PIMVP (EVO, 2012).

Segundo o PIMVP (EVO, 2012), em um relatório de economia, o nível aceitável de incerteza está relacionado com o custo de reduzir esta incerteza a um nível adequado para o valor previsto da economia, sendo que estes custos médios típicos anuais de M&V devem ser inferiores a 10% da economia média anual a ser avaliada.

O PROPEE, em seu módulo 8 (ANEEL, 2020), entende que o PEE sendo um instrumento de política pública de EE e cumprindo uma função no planejamento energético, deve garantir sua meta com uma determinada confiabilidade, compatível com o sistema elétrico em que se insere. A base de geração do sistema elétrico brasileiro é de origem hídrica, dependente do regime de chuvas. A previsão hidrológica é probabilística, e a garantia do suprimento é atendida nos estudos com confiabilidade mínima de 95% (Tolmasquim, 2011, p. 87). Assim, o PROPEE estabelece que deve ser atingido uma precisão de 10% com 95% de confiabilidade

nos projetos de EE, onde entende-se que assim seria garantido que os investimentos para a sociedade sejam vantajosos, já que a RCB<sup>2</sup> (Relação Custo-Benefício) máxima dos projetos está fixada em geral em 0,80.

O PIMVP (EVO, 2012) admite dois modos de se medir a economia de energia: pela energia evitada, quando se consideram as condições do período de determinação da economia, e pela economia normalizada, quando estas condições são fixas, conforme um padrão estabelecido. Dentro dos projetos do PEE, a recomendação é que esses projetos sigam a segunda opção (economia normalizada), onde alguns padrões (clima, por exemplo) podem ser definidos através de estudos. Para os estudos de caso da presente dissertação, sendo um projeto com implementação de SFV, a consideração de padrão segue para a irradiação solar mensal e anual, variável independente que se correlaciona com a geração provida da usina (ANEEL, 2020).

A M&V apresenta duas fases durante o projeto, sendo elas:

- a) **avaliação ex ante:** com resultados esperados estimados, na fase de definição, em procedimento baseado em análises de campo, dados típicos, experiências anteriores e cálculos de engenharia.
- b) **avaliação ex post:** com valores mensurados, consideradas a economia de energia e a redução de demanda na ponta avaliadas por ações de Medição e Verificação, a partir de medições feitas nas fases de Execução (período da linha de base) e Verificação (período de determinação da economia), e análise para determinação da EE.

#### 2.1.1.1 Estratégia de M&V

Antes de chegar ao período *ex post*, na fase de diagnóstico energético<sup>3</sup> conhecido como *ex ante* ou “Estratégia de M&V”, define-se as bases para as atividades de M&V. Premissas importantes que constam nessa estratégia:

---

<sup>2</sup> A Relação Custo-Benefício (RCB) é o principal critério para avaliação da viabilidade econômica de um projeto do PEE. O benefício considerado é a valoração da energia economizada e da redução da demanda na ponta durante a vida útil do projeto para o sistema elétrico (ANEEL, 2020).

- Variáveis independentes: indicação de quais variáveis (clima, produção e ocupação) explicam a variação da energia e como poderão ser medidas (local, equipamentos, períodos de medição - linha de base e de determinação da economia).
- Fronteira de medição: determina o limite, dentro da instalação, onde serão observados os efeitos da ação de EE, isolado por medidores, e eventuais efeitos interativos com o resto da instalação. (EVO, 2017).
- Opção do PIMVP: opção A, B, C ou D do PIMVP que será usada para medir a economia de energia. A Tabela 2.1 apresenta um resumo sobre estas opções.
- Modelagem: Modelos matemáticos são usados em M&V para preparar o termo referente aos ajustes de rotina nas várias versões da equação de cálculo da economia de energia. A modelagem da energia da linha de base e do período de determinação tem como objetivo buscar uma relação matemática entre as variáveis dependentes e independentes. A variável dependente, a energia, é modelada como sendo regida por uma ou mais variáveis independentes (também conhecidas como “variáveis explicativas”), quando tem-se esta variável. Este tipo de modelagem é chamado de análise de regressão.
- Cálculo das economias: definir como será calculada a economia de energia e a redução de demanda na ponta.
- Amostragem: para o cálculo de amostragem (ANEEL, 2014), calcula-se duas estimativas iniciais do tamanho da amostra: a primeira, segundo a NBR 5426 (Normas, 1985) com o regime de inspeção severa, nível I; e a segunda calcula-se também a amostra necessária supondo o coeficiente de variância<sup>4</sup> informado e uma precisão desejada de 10% a 95% de confiabilidade.

---

<sup>3</sup> Elaboração de todo o projeto a ser implementado, com a contabilização dos benefícios, custos e resultados pretendidos do projeto de eficiência energética.

<sup>4</sup> É o coeficiente de variância, definido como o desvio padrão das leituras dividido pela média. Até que a média e desvio padrão da população possa ser estimados pelas amostras, usar os valores de projetos semelhantes anteriores, ou 0,5 se não houver dados (EVO, 2017).

Tabela 2.1– Visão geral das opções do PIMVP

Opção do PIMVP	Definição	Cálculo da Economia	Aplicação Típica
<b>A.</b> Isolação da AEE: medição dos parâmetros chave	<p>» A economia é determinada pela medição no campo dos parâmetros chave, que definem o consumo de energia ou demanda dos sistemas afetados pela AEE ou o sucesso do projeto.</p> <p>» A frequência de medição varia desde medições de curto prazo até contínua, dependendo da variação esperada do parâmetro medido e da duração do período de determinação da economia. Os parâmetros não selecionados para medição no campo são valores estimados. As estimativas podem ser baseadas em dados históricos, especificação do fabricante ou juízo de engenharia.</p> <p>» A fonte da justificativa da estimativa é requerida. O erro plausível derivado da estimativa deve ser avaliado.</p>	<p>» Cálculo de engenharia da energia da linha de base e do período de determinação da economia baseado em: medições de curto prazo a contínua dos parâmetros e valores estimados</p> <p>» Ajustes de rotina e não de rotina como requerido. Parâmetros chave medidos durante ambos os períodos da linha de base e determinação da economia.</p>	<p>» AEE de iluminação onde a potência é o parâmetro chave medido e as horas de operação estimadas com base na programação da instalação e hábito dos ocupantes.</p>
<b>B.</b> Isolação da AEE: medição de todos os parâmetros	<p>» A economia é determinada pela medição no campo do consumo de energia e demanda e/ou variáveis independentes ou proxies dos sistemas afetados pela AEE.</p> <p>» A frequência de medição varia desde medições de curto prazo até contínua, dependendo da variação esperada do parâmetro medido e da duração do período de determinação da economia.</p>	<p>» Medições de curto prazo a contínua da energia dos períodos de linha de base e determinação, ou cálculos de engenharia usando medições de <i>proxies</i><sup>5</sup> do consumo da energia e demanda.</p> <p>» Ajustes de rotina e não de rotina como requerido.</p>	<p>» Aplicação de acionador de velocidade variável em motor para ajustar a vazão de bomba. Medir a potência com um wattímetro instalado na alimentação elétrica do motor, com leituras a cada minuto. Medir por uma semana na linha de base para verificar a potência constante. Medir durante o período de determinação o consumo e a demanda.</p>
<b>C.</b> Toda a instalação	<p>» A economia é determinada pela medição do consumo de energia e demanda de toda a instalação pelo medidor da concessionária.</p> <p>» Medições contínuas do consumo de energia e demanda de toda a instalação são feitas durante o período de determinação da economia.</p>	<p>» Análise dos dados do medidor da concessionária nos períodos de linha de base e determinação da economia.</p> <p>» Ajustes de rotina como requerido, usando técnicas como simples comparação ou análise de regressão.</p> <p>» Ajustes não de rotina como requerido.</p>	<p>» Programas de gestão energética multifacetados afetando muitos sistemas energéticos em uma instalação. Medir o consumo de energia e demanda com os medidores das concessionárias de gás e energia elétrica por doze meses na linha de base e durante o período de determinação da economia.</p>
<b>D.</b> Simulação calibrada	<p>» A economia é determinada através da simulação do consumo de energia e demanda de toda a instalação ou parte dela.</p> <p>» Rotinas de simulação demonstram que o desempenho energético da instalação é adequadamente modelado.</p> <p>» Esta opção requer habilidade considerável para calibrar a simulação.</p>	<p>» Simulação do consumo de energia e demanda calibrada com dados horários ou mensais da conta de energia. Medidores em usos finais e medições de desempenho podem ser usados para refinar o modelo.</p>	<p>» Programas de gestão energética multifacetados afetando muitos sistemas energéticos em uma instalação onde não havia medidores no período da linha de base.</p> <p>» As medições do consumo de energia e demanda, após a instalação de medidores de gás e eletricidade, são usadas para calibrar a simulação.</p> <p>» A energia da linha de base, determinada pela simulação calibrada, é comparada com a simulação do período de determinação da economia.</p>

Fonte: PIMVP Conceitos Básicos – EVO (2016), adaptada pelo autor (2021).

<sup>5</sup> A determinação do consumo de energia pode ser feita por medição direta do fluxo de energia, ou por medição direta de representantes (“proxies”) de consumo que fornecem indicação direta do consumo de energia (EVO, p. 12. 2017).

### 2.1.1.2 Plano de M&V

Seguindo para o período *ex post*, antes da implementação do projeto e com as definições finais do que será o escopo, inicia-se a elaboração do Plano de Medição e Verificação, conforme etapas a seguir (ANEEL, 2014):

- **Opções de PIMVP:** Conforme já definido na “Estratégia de M&V” com as condições finais do contrato, é ratificada a forma de medição da energia, bem como as variáveis independentes condicionadas anteriormente, de acordo com a Opção do PIMVP.
- **Verificação operacional:** Consiste na verificação de que as AEEs foram executadas e operaram de forma adequada ao planejado. São verificações em campo das condições das AEEs.
- **Variáveis independentes:** Define as variáveis independentes que serão usadas (medidas ou estimadas) para explicar a variação da energia. Os parâmetros que tiverem um efeito significativo na energia do período da linha de base são incluídos nos ajustes de rotina para determinar a economia de energia.
- **Fronteira de Medição:** Descreve a localização dos medidores para medir a energia e as variáveis independentes, que definem a fronteira de medição a ser usada e isolam os fluxos de energia que afetam o desempenho da AEE.
- **Efeitos interativos<sup>6</sup>:** Define os efeitos da AEE fora da fronteira de medição e como serão considerados (ignorados ou estimados).
- **Fatores Estáticos:** Define os fatores estáticos que podem mudar o padrão de uso da energia e que deverão ser acompanhados nos estudos de longo prazo.

---

<sup>6</sup> Quaisquer efeitos energéticos que ocorram para além da fronteira de medição imaginária serão denominados ‘efeitos interativos’. É preciso encontrar uma forma de estimar a magnitude destes efeitos interativos, a fim de determinar a economia. Alternativamente, tais efeitos podem ser ignorados, desde que o Plano de M&V inclua a discussão de cada efeito e sua provável magnitude (PIMVP - EVO, 2012).

- **Período, energia e condições da linha de base:** Definição de todos os parâmetros utilizados na determinação da linha de base para apuração das economias.
- **Período, energia e condições de determinação:** Indica o período em que serão feitas as medições para determinação das economias (energia e variáveis independentes).
- **Base de ajuste:** Indica o conjunto de condições (ou seja, conjunto de variáveis independentes) ao qual os valores de energia medidos serão ajustados, se houver ajustes a serem considerados.
- **Especificações dos medidores:** Especificar em cada ponto de medição o tipo de medidor, fabricante, modelo, precisão, rotina de calibração, método de leitura e controle de qualidade, verificação operacional, forma de tratamento dos dados perdidos.
- **Precisão esperada:** Avaliar a precisão esperada associada à medição, amostragem e modelagem. Em princípio, cada etapa do processo de obtenção de dados (amostragem, modelagem e medição) deverá perseguir uma meta de 10% de precisão a 95% de confiabilidade.
- **Orçamento M&V:** Definir os recursos e custos associados para a determinação da economia. Em projetos de PEE em Chamadas Públicas de Projetos (CPP), as concessionárias limitam a 5% do valor dos projetos em média, e ainda, às vezes, colocam um valor máximo que dificulta estabelecer os modelos necessários em alguns projetos.

### 2.1.2 Definição do Período de Determinação dos Resultados de M&V

O período de determinação da economia é um intervalo selecionado para avaliar e quantificar o desempenho energético após a instalação das ações de EE (EVO, 2014).

O PIMVP recomenda que a “duração do período de determinação da economia deve ser definido com a devida consideração pela duração da ação de EE e pela probabilidade de degradação da economia originalmente obtida ao longo do tempo” (EVO, 2012). Os dados que farão parte do relatório final do projeto, retratando o “período de determinação da economia”, são:

- Datas do período de medição;
- Dados de energia e demanda;
- Valor das variáveis independentes;
- Valor do coeficiente de determinação ( $R^2$ ), quando houver modelo do período de referência;
- Para amostragens: tamanho da amostra, precisão ( $e$ ) e coeficiente de variação ( $cv$ ) obtidos;
- Precisão dos instrumentos utilizados;
- Descrição e justificativa de quaisquer correções feitas aos dados observados;
- Valores estimados acordados;
- Valores da energia e demanda utilizados (ponto de vista do sistema elétrico e do consumidor);
- Desvio eventual das condições apresentadas no Plano de M&V. Apresentar cálculos de engenharia que fizeram o ajuste às novas condições;
- Economia calculada em unidades de energia e monetárias (ponto de vista do sistema elétrico e do consumidor);
- Desvio observado em relação à avaliação ex ante.

### 2.1.3 Variável independente “Irradiação Solar”

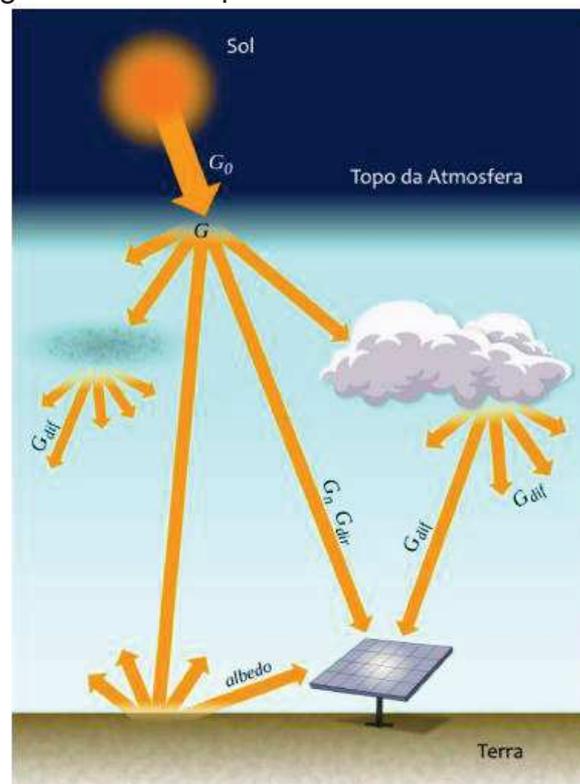
A irradiância solar ( $W/m^2$ ) que incide em uma superfície é composta por suas componentes direta e difusa, onde a direta apresenta direção de incidência na linha imaginária entre a superfície e o sol e a componente difusa, engloba a radiação proveniente de todas as demais direções que são decorrentes dos processos de espalhamento pelos gases e particulados presentes na atmosfera, onde tem-se (Pereira et. al. 2017):

- Irradiância extraterrestre ( $G_0$ ): é a taxa de energia incidente por unidade de área em um plano horizontal imaginário situado no topo da atmosfera. É também conhecido como irradiância no topo da atmosfera ou  $G_{TOA}$ .

- Irradiância direta normal ( $G_n$ ): também conhecida como DNI, é a taxa de energia por unidade de área proveniente diretamente do sol que incide perpendicularmente à superfície.
- Irradiância difusa horizontal ( $G_{dif}$ ): é a taxa de energia incidente sobre uma superfície horizontal por unidade de área, decorrente do espalhamento do feixe solar direto pelos constituintes atmosféricos (moléculas, material particulado, nuvens e etc.).
- Irradiância direta horizontal ( $G_{dir}$ ): é a taxa de energia por unidade de área do feixe solar direto numa superfície horizontal. Pode ser determinada como o produto entre a irradiância direta normal (DNI) e o cosseno do ângulo zenital solar. O ângulo zenital é o número de graus de latitude que separa a localização do indivíduo da declinação do sol.
- Irradiância global horizontal ( $G$ ): é a taxa de energia total por unidade de área incidente numa superfície horizontal.

A Figura 2.2 ilustra a irradiância assim que atinge a atmosfera e seu espalhamento.

Figura 2.2 – Componentes da irradiância solar



Fonte: Pereira et. al. 2017

A irradiância global é dada pela soma da composição destas variáveis e apresenta-se na Equação 2.1 (Pereira et. al. 2017):

$$G = G_{dif} + G_{dir} \text{ ou } G = G_{dif} + G_n \cdot \cos \theta_z \quad \text{Equação 2.1}$$

A integral da irradiância no tempo é definida como irradiação solar ( $\text{Wh/m}^2$ ) ou energia radiante incidente acumulada em um intervalo de tempo. A irradiação solar nada mais é do que uma determinada quantidade de radiação solar por unidade de área. (Pereira et. al. 2017).

Tanto na linha de base quanto no período de determinação dos processos de M&V a irradiância solar ( $\text{Wh/m}^2$ ) é utilizada como variável independente.

Nas simulações da linha de base do SFV através do PVsyst, dos *cases* apresentados no capítulo 5 no período do plano de M&V, utilizou-se os dados do CRESESB (Centro de Referência para as Energias Solar e Eólica Sérgio de S. Brito), das estações locais mais próximas dos projetos.

O CRESESB tem como missão promover o desenvolvimento das energias solar e eólica através da difusão de conhecimentos, da ampliação do diálogo entre as entidades envolvidas e do estímulo à implementação de estudos e projetos. Com isso disponibiliza o programa *SunData* em seu site, que se destina ao cálculo da irradiação solar diária e média mensal em qualquer ponto do território nacional. Os dados disponíveis foram produzidos a partir de um total de 17 anos de imagens de satélite e com informações de mais de 72.000 pontos em todo o território brasileiro, onde estes dados são indicativos e possuem as limitações dos modelos utilizados, sendo que a recomendação do CRESESB é utilizar os dados do INMET, para avaliações mais precisas com medição da irradiação no local de interesse (Pereira et. al. 2017). O modelo utilizado em (Pereira et. al. 2017) para obter os dados de irradiância é o BRASIL-SR, que é um modelo físico para obtenção de estimativas da radiação solar incidente na superfície que combina a aproximação de “dois-fluxos” (Meador et. al. 1980) na solução da equação de transferência radiativa com o uso de parâmetros determinados de forma estatística a partir de imagens de satélite (Pereira et. al. 2017). A obtenção de uma estimativa da irradiação solar incidente na superfície pode ser dividida em três etapas neste no modelo BRASIL-SR:

- 1 tratamento dos dados meteorológicos e imagens de satélite;
- 2 aplicação do método de transferência radiativa de “dois-fluxos”;

- 3 cálculo das componentes da irradiação solar incidente na superfície em Wh/m<sup>2</sup>: irradiação global horizontal ( $H$ ), irradiação direta normal ( $H_n$ ), irradiação difusa horizontal ( $H_{dif}$ ) e irradiação global no plano inclinado ( $H_i$ ).

No período de determinação dos resultados dos projetos com SFV, considerando a “Opção B” apresentado no subcapítulo 2.1.1, é necessário dados de medição real para a irradiação solar global no plano horizontal. Assim, a indicação para a obtenção desta variável independente é a utilização da base de dados no INMET, que possui um banco de dados histórico desde quando as estações meteorológicas existentes em cada região foram fundadas com dados horários, mensal e anual.

#### **2.1.4 Deep Learning e suas Aplicações em Previsões Temporais**

O DL permite que modelos computacionais compostos de várias camadas de processamento aprendam representações de dados com vários níveis de abstração. Esses métodos melhoraram drasticamente o estado da arte em reconhecimento de fala, reconhecimento de objeto visual, detecção de objeto e muitos outros domínios. O DL descobre a estrutura intrincada em grandes conjuntos de dados, usando o algoritmo de retropropagação para indicar como uma máquina deve alterar seus parâmetros internos, que são usados para calcular a representação em cada camada a partir da representação na camada anterior (Lecun et. al. 2015).

Redes como as *Convolutional Neural Network* (CNN) trouxeram avanços no processamento de imagens, vídeo, fala e áudio, enquanto as redes *Recurrent Neural Network* (RNN) trouxeram avanços para processamento de dados sequenciais, como texto e fala (Lecun et. al. 2015). Além disso, com as redes RNN e sua extensão com a LSTM, foi possível a aplicação em mercados de sistemas de energia, principalmente na previsão de séries temporais, como previsão de carga, resposta à demanda e de geração de energia renovável, devido à sua capacidade de resolver problemas de dados temporais (Muhammad et. al. 2019).

A eficácia do DL deve-se à disponibilidade de enorme quantidade de dados e computação de alto desempenho. Como em outros campos, também está

recebendo atenção em sistemas de energia. DL é um subcampo de aprendizado de máquina, feito a partir de *Artificial Neural Network* (ANN) (Muhammad et. al. 2019).

Com o avanço dos sistemas de energia e a alta penetração das energias renováveis, o resultado são operações de forma mais complexas, onde as técnicas convencionais de análise e controle dependem principalmente da modelagem física e numérica. Com isso, o DL é uma solução adequada para o desenvolvimento futuro de ferramentas e aplicações para o sistema de energia (Muhammad et. al. 2019).

Com a incerteza na natureza do cliente, a previsão de carga é um desafio no contexto do mercado de energia. O modelo baseado em LSTM é usado para prever cargas de uma hora e um minuto de passo em Marino et. al. (2016). Prever e encontrar flexibilidade energética no lado da demanda é importante para aplicar a resposta à demanda. Em Do Nascimento (2016), o DL é aplicado para identificar a flexibilidade das cargas e para fornecer referências para a demanda e resposta. Além disso, em Tornai et. al. (2017), a RNN é usada para classificar os consumidores e superar métodos tradicionais, atingindo precisão de quase 100%.

Entrando em aplicações para soluções com geração de energia renováveis, como solar e eólica, as redes neurais já são utilizadas com grande frequência, como em Melo et. al. (2020), em que se aplica redes neurais artificiais para previsão de velocidade do vento considerando variáveis correlacionadas, como irradiação solar e temperatura. Também em Ouammi et. al. (2012), um modelo de ANN é usado para prever a irradiação solar anual e mensal em Marrocos; os dados são inferidos usando um algoritmo ANN para estabelecer uma relação direta / reversa de correspondência entre longitude, latitude, altitude e irradiação solar.

Os métodos DL vêm ganhando mais espaço aplicados à previsão de series temporais de variáveis ambientais como irradiação, impulsionados devido à disponibilidade da enorme base de dados dos departamentos meteorológicos, como pode ser visto em Muhammad et. al. (2019), que apresenta o estado da arte da aplicação de DL em sistemas de energia. O DL consegue trabalhar com uma grande quantidade de dados, seu método de aprendizagem o torna mais assertivo e traz mais velocidade às análises.

Em Gökgöz et. al. (2018), com o propósito de explorar a eficácia da rede neural LSTM, faz-se uma aplicação em estimativa de energia renovável. O uso de DL encontra-se hoje em diversos artigos quando se trata de previsão de séries temporais, com em Brahma et. al. (2020), onde seu objetivo é desenvolver um

modelo de previsão baseado em metodologias de DL através de dados de vários locais, para prever a energia solar diária em dois locais da Índia. Para tanto, utiliza dados diários de irradiação solar obtidos da NASA (*National Aeronautics and Space Administration*) ao longo de 36 anos (1983–2019).

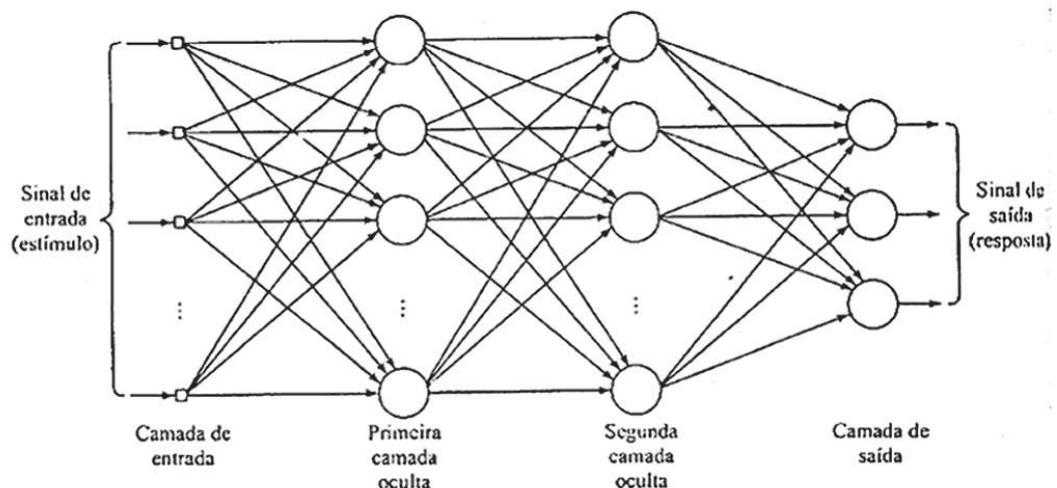
### 2.1.5 Rede Neural LSTM

As redes de memória de longo prazo, ou LSTM, são um tipo especial de RNN, capaz de aprender dependências de longo prazo (Olah, 2015). Elas foram introduzidas por Hochreiter & Schmidhuber (Hochreiter,1997), refinadas e popularizadas em muitos estudos.

As RNN são algoritmos recentes para dados sequenciais que, como são capazes de recordação de entradas, devido a uma memória interna, tornam-se apropriados para problemas que envolvem dados sequenciais. Há problemas de *vanish/exploding gradient* com as RNN que são resolvidos pela forma estendida de RNN chamada LSTM (Muhammad et. al. 2019). Os LSTM são explicitamente projetados para evitar o problema de dependência de longo prazo. Lembrar informações por longos períodos é praticamente o seu comportamento padrão (Olah, 2015).

Uma Rede de Memória Longa de Curto Prazo é composta por unidades LSTM, conforme Figura 2.3.

Figura 2.3 – Esquema de uma rede neural com duas camadas ocultas



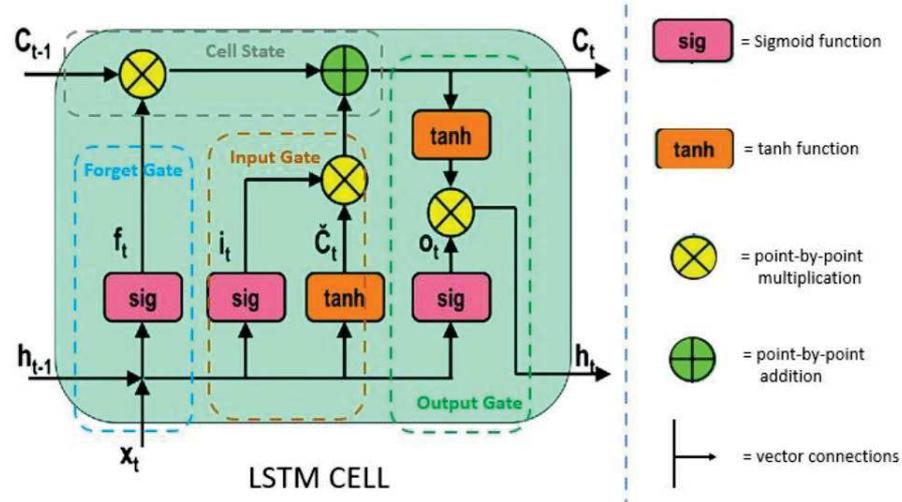
Fonte: VASSALLI (2018).

A ideia principal é que uma Rede Neural Recorrente pode ser alimentada reversamente com dados relevantes ao modelo, no caso séries históricas de irradiação solar, para realizar previsões. Em cada neurônio é atribuído um peso para as entradas recebidas e deles subtraídos um valor conhecido como bias e submetidos a uma função de ativação, como a função sigmoide, limitada entre  $[0,1]$ , para representar a ativação do neurônio. São esses pesos que representam o treinamento recebido e, portanto, constituem a memória da rede. Podemos entender que cada neurônio reconhece determinado tipo de situação e para essa situação de entrada, esse neurônio está mais ativado, portanto para um problema complexo, são utilizadas grandes quantidades de neurônios e camadas (Vassalli, 2018).

Uma unidade LSTM é composta de um neurônio e portões de entrada, esquecimento e saída, apresentado na Figura 2.4. O neurônio é responsável pela memória da rede, representada pela ativação baseada na soma ponderada e os portões são uma forma de, opcionalmente, permitir a informação. A rede LSTM tem seus componentes (Vassalli, 2018), conforme a seguir:

- Estado oculto: usado para determinar o que esquecer, entrar e sair no passo seguinte;
- Estado de entrada: combinação do estado oculto e a entrada atual;
- Estado interno: valores com função de memória;
- Portão de entrada: decide se a entrada alcança o estado interno;
- Portão de esquecimento: decide se o estado interno desconsidera o anterior;
- Portão de saída: decide se o estado interno é passado à saída e ao estado oculto no passo seguinte.

Figura 2.4 – Uma unidade de célula LSTM

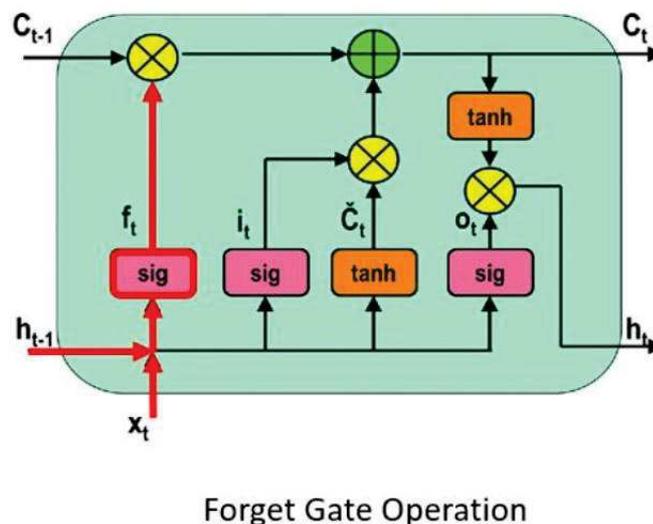


Fonte: Singhal (2020).

O portão de esquecimento decide quais informações precisam de atenção e quais podem ser ignoradas, a Figura 2.5 ilustra a operação na célula para esta etapa. As informações da entrada atual  $X_{(t)}$  e do estado oculto  $h_{(t-1)}$  são passadas pela função sigmóide. A função sigmoid gera valores entre 0 e 1. Ele conclui se a parte da saída antiga é necessária (dando a saída mais próxima de 1). Este valor de  $f_{(t)}$  será usado posteriormente pela célula para multiplicação ponto a ponto (Singhal, 2020).

A Equação 2.2 representa a operação que é feita nesta etapa.

Figura 2.5 – Unidade de célula LSTM – Portão de esquecimento



Fonte: Singhal (2020).

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \quad \text{Equação 2.2}$$

Onde:

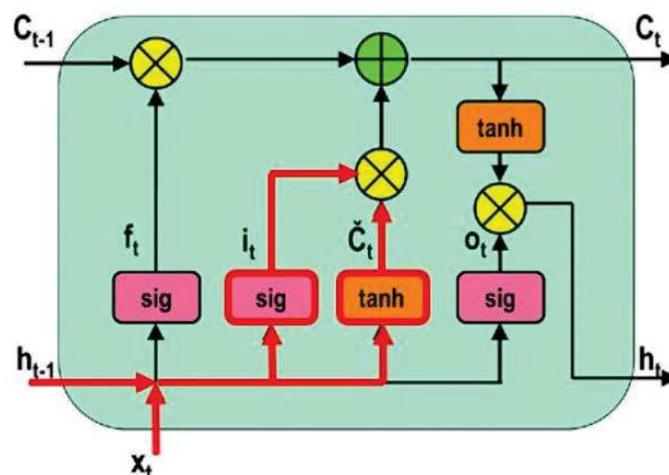
- $t$  = passo de tempo
- $f_t$  = portão de esquecimento em  $t$
- $x_t$  = entrada
- $h_{t-1}$  = estado oculto anterior
- $W_t$  = matriz de peso entre o portão de esquecimento e o portão de entrada
- $b_t$  = conexão bias em  $t$

A porta de entrada executa as seguintes operações para atualizar o status da célula, conforme ilustrado na Figura 2.6. Primeiro, o estado atual  $X_{(t)}$  e o estado anteriormente oculto  $h_{(t-1)}$  são passados para a segunda função sigmóide. Os valores são transformados entre 0 (importante) e 1 (não importante).

Em seguida, as mesmas informações do estado oculto e do estado atual serão passadas pela função  $\tanh$ . Para regular a rede, o operador  $\tanh$  criará um vetor ( $\tilde{C}_t$ ) com todos os valores possíveis entre -1 e 1. Os valores de saída gerados pelas funções de ativação estão prontos para multiplicação ponto a ponto (Singhal, 2020).

As equações 2.3 e 2.4 retratam o primeiro passo e o segundo passo, respectivamente.

Figura 2.6 – Unidade de célula LSTM – Portão de entrada



Fonte: Singhal (2020).

$$i_t = \sigma (W_i \cdot [h_{t-1}, x_t] + b_i) \quad \text{Equação 2.3}$$

$$\tilde{C}_t = \tan h (W_c \cdot [h_{t-1}, x_t] + b_c) \quad \text{Equação 2.4}$$

Onde:

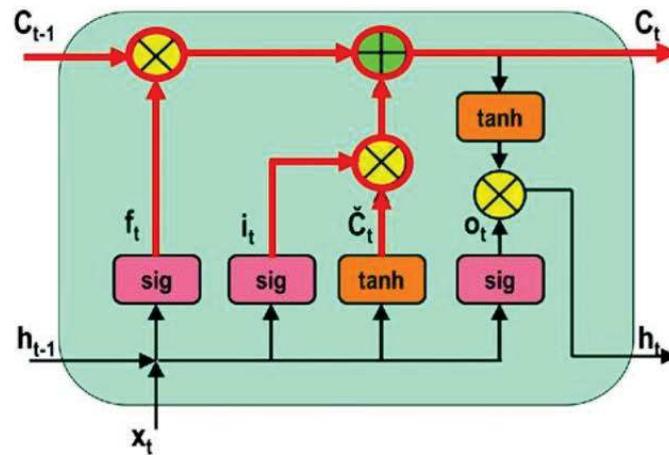
- $t$  = passo de tempo
- $i_t$  = portão de entrada em  $t$
- $\tilde{C}_t$  = valor gerado pela  $\tan h$
- $W_i$  = matriz de peso do operador sigmóide entre a porta de entrada e a porta de saída
- $W_c$  = matriz de peso do operador  $\tan h$  entre as informações do estado da célula e a saída da rede saída
- $b_t$  = vetor *bias* em  $t$
- $b_c$  = vetor *bias* em  $t$ , w. r.  $tW_c$

A rede tem informações suficientes do portão de esquecimento e do portão de entrada. O próximo passo é decidir e armazenar as informações do novo estado no estado da célula. A Figura 2.7 ilustra esta operação.

O estado da célula anterior  $C_{(t-1)}$  é multiplicado pelo vetor de esquecimento  $f_{(t)}$ . Se o resultado for 0, os valores serão descartados no estado da célula. Em seguida, a rede pega o valor de saída do vetor de entrada  $i_{(t)}$  e realiza a adição ponto a ponto, que atualiza o estado da célula dando à rede um novo estado de célula  $C_{(t)}$  (Singhal, 2020).

A Equação 2.5 mostra a operação relatada do estado da célula.

Figura 2.7 – Unidade de célula LSTM – Estado da célula



Cell State Operation

Fonte: Singhal (2020).

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad \text{Equação 2.5}$$

Onde:

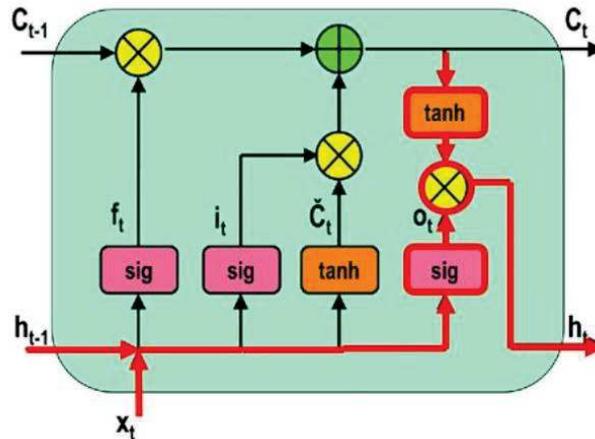
- $t$  = passo de tempo
- $C_t$  = informações do estado da célula
- $C_{t-1}$  = passo de tempo anterior

A porta de saída determina o valor do próximo estado oculto. Este estado contém informações sobre entradas anteriores.

Primeiro, os valores do estado atual e do estado oculto anterior são passados para a terceira função sigmóide. Em seguida, o novo estado da célula gerado a partir do estado da célula é passado pela função  $\tanh$ . Ambas as saídas são multiplicadas ponto a ponto. Com base no valor final, a rede decide quais informações o estado oculto deve transportar. Este estado oculto é usado para previsão. Finalmente, o novo estado da célula e o novo estado oculto são transportados para a próxima etapa de tempo (Singhal, 2020).

As equações 2.6 e 2.7 retratam o primeiro passo e o segundo passo, respectivamente.

Figura 2.8 – Unidade de célula LSTM – Estado da saída



Output Gate Operation

Fonte: Singhal (2020).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad \text{Equação 2.6}$$

$$h_t = o_t \times \tan h(C_t) \quad \text{Equação 2.7}$$

Onde:

- $O_t$  = portão de saída em  $t$
- $W_o$  = matriz de peso da porta de saída
- $h_t$  = LSTM saída
- $b_o$  = vetor *bias* em  $t$ , w. r.  $tW_o$

A metodologia proposta na presente dissertação utiliza o algoritmo LSTM para previsão da irradiação solar com base em dados históricos extraídos do INMET.

Ao contrário das redes neurais *feedforward* padrão, o LSTM tem conexões de *feedback*, além disso elas resolvem o problema do gradiente introduzindo mais portas. A célula lembra valores em intervalos de tempo arbitrários e três portas regulam o fluxo de informações para dentro e para fora da célula. O portão de entrada determina se a entrada será salva na célula de memória. O portão de saída

determina se a célula de memória atual será a saída. O portão de esquecimento determina se o conteúdo atual da memória será esquecido (Gökgöz, 2018).

### 2.1.6 Previsão de Séries

As séries temporais são dados sequenciais medidos em determinados intervalos em relação a qualquer processo. Os intervalos usados nas séries temporais podem ser de tamanhos diferentes, desde que sejam igualmente amostrados. Geralmente são medidos em intervalos de hora, dia, mês e ano. A população anual, o número diário de passageiros no metrô, a irradiação solar e a taxa de câmbio horária são exemplos de séries temporais. Na série temporal, os registros devem ser ordenados cronologicamente. Cada registro pode conter informações sobre um ou mais recursos. O termo univariada é usada para dados de séries temporais contendo informações únicas, e multivariada é usada para dados contendo mais de um tipo de informação. Neste contexto, a previsão de séries temporais pode ser definida como a previsão de dados futuros usando séries temporais de dados passado. Os dados da série temporal são usados para criar modelos por diferentes métodos. O processo de criação de modelos pela formulação dos dados é chamado de análise de série temporal. A previsão é realizada por meio desses modelos (Sorkun et. al. 2020).

Podemos dividir os modelos de previsão de séries temporais em duas categorias, lineares e não lineares. Modelos lineares produzem previsões tomando a composição linear de observações anteriores. Exemplos de modelagem linear são regressão e modelo auto-regressivo integrado de médias móveis (*autoregressive integrated moving average* ou ARIMA, na sigla em inglês). Modelos não lineares são usados quando os dados são multivariados ou contêm relacionamentos complexos em que os modelos lineares são insuficientes para se encaixar. Exemplos de modelos não lineares são RNA (ou ANN, *Artificial Neural Networks*), SVM (Support Vector Machines - Máquina de vetores de suporte), DNN (*Deep Neural Networks*) e árvores de decisão. A seguir, para previsão da irradiação solar, sendo uma previsão de séries temporais não lineares, o modelo LSTM é adotado (Sorkun et. al. 2020).

### 3 DESCRIÇÃO DO PROBLEMA DE M&V EM SFV

#### 3.1 INTRODUÇÃO

Conforme as diretrizes dos projetos de EE da ANEEL (ANEEL, 2020), todo projeto que são providos dos recursos oriundos do PEE, devem ter M&V para a aferição e avaliação dos resultados e benefícios energéticos proporcionados pelos projetos e comprovação dos recursos pleiteados para a implementação.

O PEE permite a inclusão nos projetos com fonte incentivada, desde que todas as ações de EE nos usos finais das unidades consumidoras já tenham sido contempladas:

Só poderão ser realizados investimentos em geração de energia a partir de fontes incentivadas com recursos do PEE se as ações de eficiência energética economicamente viáveis apuradas em diagnóstico energético nas instalações do consumidor beneficiado, de acordo com o estabelecido no Módulo 7 – Cálculo de Viabilidade, forem ou já tiverem sido implementadas (ANEEL, 2020, seção 6, p. 3).

Então, para projetos com fonte incentivada, com sistemas de geração fotovoltaica instalados deverão ser providos de medidores que registrem a energia gerada e a demanda provida no horário de ponta e fora de ponta. Podendo ser os registros providos diretamente do inversor de frequência ou medidores fixos instalados nas saídas dos transformadores. Outra condição estabelecida é que as medições para apuração da energia e demandas geradas deverão ser feitas por um ano. Caso haja dados locais sobre a disponibilidade da fonte utilizada, este tempo poderá ser reduzido (ANEEL, 2020).

Nesta seção, será apresentada a realidade do processo tradicional de M&V, atrelados ao: fluxo de realização da M&V, custo, risco e resultados esperados e obtidos para o projeto.

#### 3.2 OBJETIVOS

Apresentar a forma tradicional de se fazer M&V em projetos com instalação de geração fotovoltaica no formato do PEE e das concessionárias de energia,

trazendo todas as condições que norteiam o processo ao longo dos doze meses de acompanhamento.

### 3.3 M&V no Método Tradicional

A forma de fazer a M&V em projetos com recursos do PEE com “fonte incentivada”, no caso energia solar, vem se tornando um grande desafio para quem realiza a implementação deste porte de projeto.

Os recursos aportados e permitidos pelo PEE para este tipo de projeto, compreendem somente as ações atreladas diretamente ao benefício gerado. Sendo para o SFV, alguns custos normais existentes ao longo do ciclo de vida do projeto, não serem permitidos, tais como: inserção de custos para manutenção do sistema, pacotes de internet durante o período de M&V para o monitoramento a distância, a inserção de custos para obras civis que não representem ganhos energéticos (como a construção de uma garagem coberta para colocação de módulos fotovoltaicos) e aquisição de imóveis (como um terreno para instalação da geração por fonte incentivada) (ANEEL, 2018).

Neste sentido, surgem os desafios porque em doze meses muitas coisas podem acontecer, principalmente quando a responsabilidade de manter o sistema operando é do beneficiário, a unidade consumidora contemplada no PEE. Este subcapítulo retrata o processo de M&V e as dificuldades encontradas, trazendo a realidade associada a condições construtivas e sua permanência, monitoramento a distância e atendimento as condições contratuais como limites de custos para M&V e atingimento das metas estabelecidas nos projetos para os benefícios projetados.

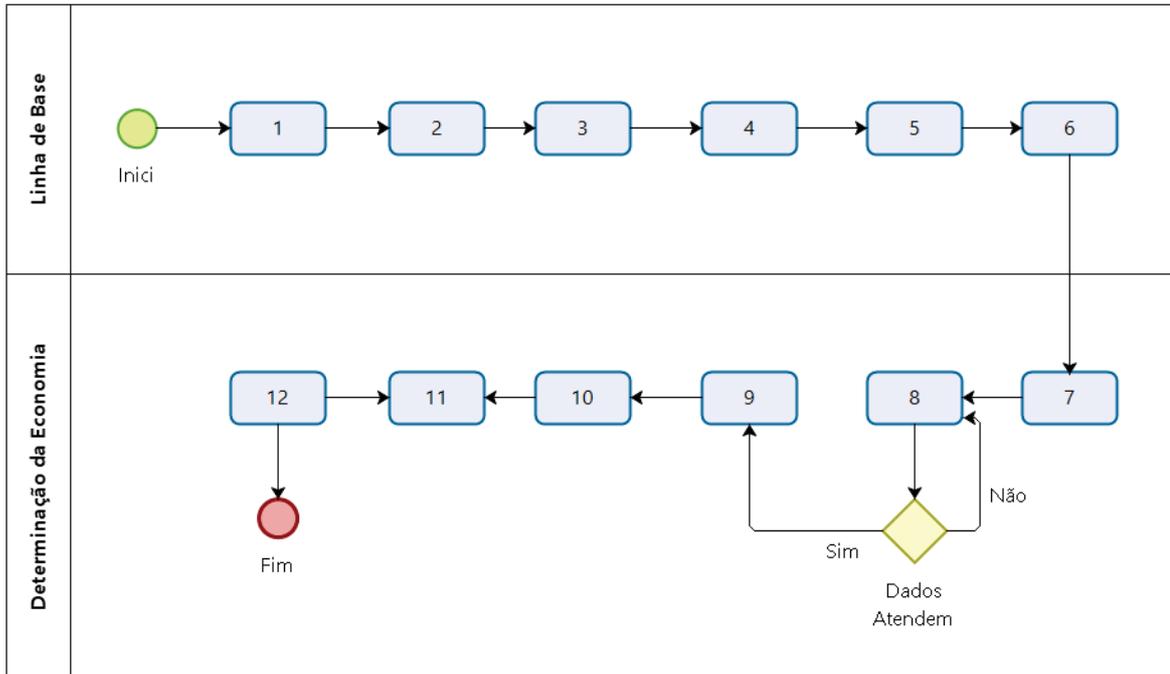
#### 3.3.1 Fluxo do processo de M&V no Método tradicional

O processo de M&V segue as condições estabelecidas no subcapítulo 2.1.1 com a construção da estratégia de M&V no período de diagnóstico energético, na assinatura do contrato antes da implementação do projeto. A definição da linha de base e a definição do período de determinação da economia são definidas após a assinatura do contrato, conforme subcapítulo 2.1.2.

O fluxo do processo de M&V do método tradicional é definido na Figura 3.1 com as fases desde a linha de base até o período de determinação da economia,

que no caso de projeto com SFV a economia torna-se a geração de energia em kWh obtida anual do sistema.

Figura 3.1 – Fluxo do processo de M&V no Método tradicional



Fonte: Criado pelo autor (2021).

O processo da Figura 3.1 encerra-se com a computação da geração anual “normalizada” compatível a prevista na linha de base, atingindo as metas estabelecidas em contrato. A Tabela 3.1 traz a descrição dos passos de cada etapa do processo da M&V no método tradicional.

Tabela 3.1 - Descrição dos passos da M&amp;V no Método Tradicional

Passo	Descrição do Passo	Etapa do processo
1	Obtenção dos dados de irradiação no CRESESB e condições físicas do projeto	Linha de Base
2	Simulação da geração solar com projeção mensal e anual	Linha de Base
3	Definição da variável independente através da regressão linear (geração simulada x irradiação solar "CRESESB")	Linha de Base
4	Definição fronteira de medição e efeitos interativos com o projeto	Linha de Base
5	Definição da opção de PIMVP	Linha de Base
6	Definições dos fatores estáticos	Linha de Base
7	Implementação do projeto	Determinação da Economia
8	Obtenção dos dados de irradiação solar com base no INMET e geração de energia do SFV com monitoramento de 12 meses	Determinação da Economia
9	Validação das incertezas das amostras das medições da geração de energia do SFV	Determinação da Economia
10	Validação da regressão linear (geração medida x irradiação solar medida "INMET")	Determinação da Economia
11	Normalização da geração de energia anual com dados "padrão" de irradiação solar (CRESESB) e juntamente com a equação da regressão linear obtida no passo 10	Determinação da Economia
12	Obtenção final da geração normalizada mensal e anual do SFV e comparação com a linha de base	Determinação da Economia

Fonte: Criado pelo autor (2021).

### 3.3.2 Permanência das condições construtivas

Após o *startup* do projeto com a vistoria da concessionária, um treinamento é realizado junto ao beneficiário pelo projeto com todas as condições operativas, manutenções periódicas, o que pode ser feito e o que não pode ser feito e suas responsabilidades mediante ao projeto recebido.

Condições adversas podem acontecer durante o percurso associadas as condições operativas e construtivas do sistema, impactando no tempo de acompanhamento da M&V, tais como:

- O sistema ser desligado por pessoal associado a instituição, para uma manutenção interna, e não ser ligado imediatamente e nem ser comunicado a empresa responsável pela M&V para validação da informação;
- Algum dispositivo de proteção queimar e o sistema ficar inoperante. Este impacto faz com que a empresa responsável pelo M&V tenha que se deslocar até o local ou ter parcerias perto da instituição para colocar o sistema para

operar novamente. Muitas vezes os projetos são em estados diferentes, o que ocasiona muitos dias sem medição;

- Condições das manutenções preventivas previstas para o sistema, que fica de responsabilidade da unidade consumidora, conforme instruções em treinamento e manuais de utilização. Estes beneficiários não seguem as instruções previstas e periódicas, o que pode impactar em diversas falhas, como a falta de reaperto dos terminais e conexões do sistema, impactando em possíveis sobreaquecimentos que pode impactar em desligamento do sistema;
- Outra condição, é na questão da limpeza dos módulos, podendo ocasionar perda de geração percentual ao que foi projetado.

Todos os itens citados podem trazer aumento em tempo de sistema desligado, ocasionando mais dias de acompanhamento na M&V e impactando diretamente nos custos de M&V, que não são repassados dentro do contrato com às concessionárias, e no prazo.

### **3.3.3 Condições dos dados no monitoramento a distância**

Como no PEE, as condições de desembolso atrelados ao projeto remetem aos itens que de fato proporcionam impacto no benefício, manter um pacote de internet funcional durante todo o tempo necessário para completar os 12 meses de monitoramento do projeto passa a ser de responsabilidade do beneficiário. Neste ponto, iniciam-se os desafios.

Segundo o PROPEE (ANEEL, 2018), permite-se a instalação de medidores registradores de energia gerada com instalação fixa ou extrair os dados dos inversores, quando possível, e que os mesmos tenham um sistema que permita o monitoramento remoto através de uma plataforma.

Para diminuir os custos com este acompanhamento, o sistema é proposto com um monitoramento online mensal, com relatórios feitos pela empresa responsável pela M&V, e ao fim de 12 meses completos é gerado um relatório final de M&V consolidado.

Com a falta do monitoramento remoto, por causa de problemas oriundos da transmissão dos dados, o acesso ao local deve ser realizado, impactando em tempo

de sistema sem dados e sem condições de identificação sobre possibilidades do real problema ocasionado na geração solar.

### **3.3.4 Atendimento às condições contratuais**

Na elaboração da proposta para o projeto, condições são estabelecidas para os custos de M&V, cada concessionária estabelece um critério sempre quantificados dentro do percentual máximo a 10% do valor global do projeto, percentual este sugerido no PMIVP, conforme subcapítulo 2.1.1. Existem algumas variações nos editais das concessionárias nas CPP, algumas estabelecem 10% do valor dos projetos, outras até 5% do valor do projeto, mas limitados a um valor.

Outras condições que envolvem a M&V nestes projetos é a rigorosidade de alguns contratos frente aos resultados de redução de demanda na ponta e/ou da energia economizada, no caso do SFV a energia gerada, medido após a realização da implementação do projeto. Após a etapa do Plano de M&V com a implementação do projeto, os resultados traçados na linha de base devem atingir pelo menos 95% (noventa e cinco por cento) dos valores previstos no carregamento inicial do projeto, se não acontecer, multas são estabelecidas de pelo menos 10% do valor contratual a empresa responsável pela elaboração da M&V.

Com isso, a M&V acaba tendo valores financeiros pré-fixados, o que impede o detalhamento de um escopo que de fato inclua os possíveis riscos de um processo de M&V para projetos com SFV, oriundos de recursos do PEE. E as condições de cláusulas contratuais, podem trazer uma sensação de limitação aos responsáveis da M&V capacitados em criar alternativas de engenharia que equilibrem boas práticas de M&V agregados a soluções técnicas, com anterioridades e segurança ao contratante que no caso, são as concessionárias de energia que são auditadas pela ANEEL dentro do PEE.

## **3.4 CONCLUSÕES PARCIAIS**

Este capítulo teve como objetivo trazer as condições reais da realização de M&V em projetos com implementação de SFV, oriundos dos recursos do PEE. O propósito é discutir sobre as diversas dificuldades existentes atreladas as reais

necessidades do processo de M&V nestes projetos. Além disso, justificar o porquê do desenvolvimento deste trabalho nesta dissertação.

A dissertação traz uma proposta de M&V onde o tempo de monitoramento pode ser reduzido, oriundo da forma que a variável independente pode ser trabalhada junto a aplicação de redes neurais LSTM.

## 4 METODOLOGIA PROPOSTA

### 4.1 INTRODUÇÃO

Sabendo que a principal variável independente em um processo de M&V para a geração fotovoltaica é a irradiação solar e que a mesma é uma série temporal, este trabalho traz como aplicação o uso de redes neurais LSTM para a previsão dessa irradiação no plano horizontal, durante meses de um ano.

Com a possibilidade de prever os meses de forma assertiva com base nos dados reais históricos medidos pelas estações solarimétricas pertencentes ao INMET, espera-se poder impactar não somente na forma de realização da M&V em projetos com geração fotovoltaica nas considerações tradicionais, mas impulsionar uma nova forma de se fazer M&V, com a redução do tempo de acompanhamento das medições de doze para seis meses, o que pode resultar em diversos benefícios tangíveis ao processo, conforme apresentado no subcapítulo 3.3.

A rede LSTM é aplicada para a previsão de seis meses da irradiação solar de um ano, subsequente a instalação do SFV e seu acompanhamento por 6 meses de medições reais. No caso, são projetados os próximos seis meses de geração do sistema através da regressão linear obtida com os dados de acompanhamento (geração e irradiação medidos) após a instalação, e a irradiação dos próximos seis meses extraída da rede LSTM, sendo estes dados imputados como variável independente.

A solução proposta é aplicada no período de determinação da economia (geração solar do sistema), conforme Figura 4.1. Os resultados obtidos com a geração no período de determinação da economia são comparados com os dados projetados no período da linha de base, onde os dados podem ser simulados através de programas computacionais.

Assim, este capítulo 4 apresenta a metodologia proposta para a projeção da geração solar em processos de M&V a partir da previsão da variável independente “irradiação solar”, utilizando rede neural LSTM. Para o desenvolvimento do código, foi utilizado a linguagem de programação Python.

## 4.2 REDE NEURAL LSTM

### 4.2.1 Tratamento dos Dados

Os dados utilizados foram extraídos do banco de dados histórico do INMET, em formato “csv.” de cada estação solarimétrica.

Os dados importados do INMET estão divididos em diversos arquivos separados por cidade e ano. Com isso, a unificação dos arquivos em uma única base de dados foi necessária, para a realização do tratamento e a normalização. Após o mapeamento dos arquivos, os dados são importados para um único <sup>7</sup>*dataset*.

Os dados importados do banco histórico do INMET contêm algumas falhas que no processo são corrigidas, como: valores nulos, dias não existentes, entre outros *outliers* a serem resolvidos. Assim, foi realizado um procedimento de “processamentos nos dados”, identificado na Figura 4.1 em tratamento dos dados, a fim de deixá-los aptos para o treinamento. Entre estes procedimentos, foram realizados:

- Remoção ou substituição de valores nulos;
- Utilizado somente o período útil de insolação diária das 07:00 até as 19 horas;
- Conversão dos dados de radiação solar em kJ/m<sup>2</sup> para irradiação solar kWh/m<sup>2</sup>;
- Os dados obtidos são dados horários, desta maneira como deseja-se realizar a previsão de dias, realizou-se um *resample*<sup>8</sup> no *dataset* e executou-se uma soma diária dos valores horários, assim o resultado é um *dataset* com a soma de irradiação solar de cada dia dos meses.

Os dados de irradiação do INMET são extraídos em radiação global kJ/m<sup>2</sup>. Para a utilização da irradiação no capítulo 5 foi necessário transformar na mesma unidade da energia gerada pelo SFV (kWh), onde multiplicou-se os valores de kJ/m<sup>2</sup> por 0,000278 (Sganzerla, 2018).

---

<sup>7</sup> Conjunto de dados.

<sup>8</sup> Método de conveniência para conversão de frequência e reamostragem de séries temporais (Pandas, 2021)

### 4.2.2 Rede LSTM Proposta

A arquitetura da LSTM é composta por  $ndm$  ( $n^\circ$  de entradas) – 50 ( $n^\circ$  neurônios na primeira camada oculta) - 30 ( $n^\circ$  de neurônios na segunda camada oculta) -  $ndm$  ( $n^\circ$  de saídas), onde  $ndm$  é o número de dias do mês correspondente à rede neural.

O número de camadas ocultas foi definido através de testes na construção da rede, assim como na definição do número de neurônios nas mesmas. Observou-se que não havia variações nos resultados encontrados e que com base em pesquisas, em diversos materiais acadêmicos, as aplicações seguiam com condições próximas na construção dos modelos.

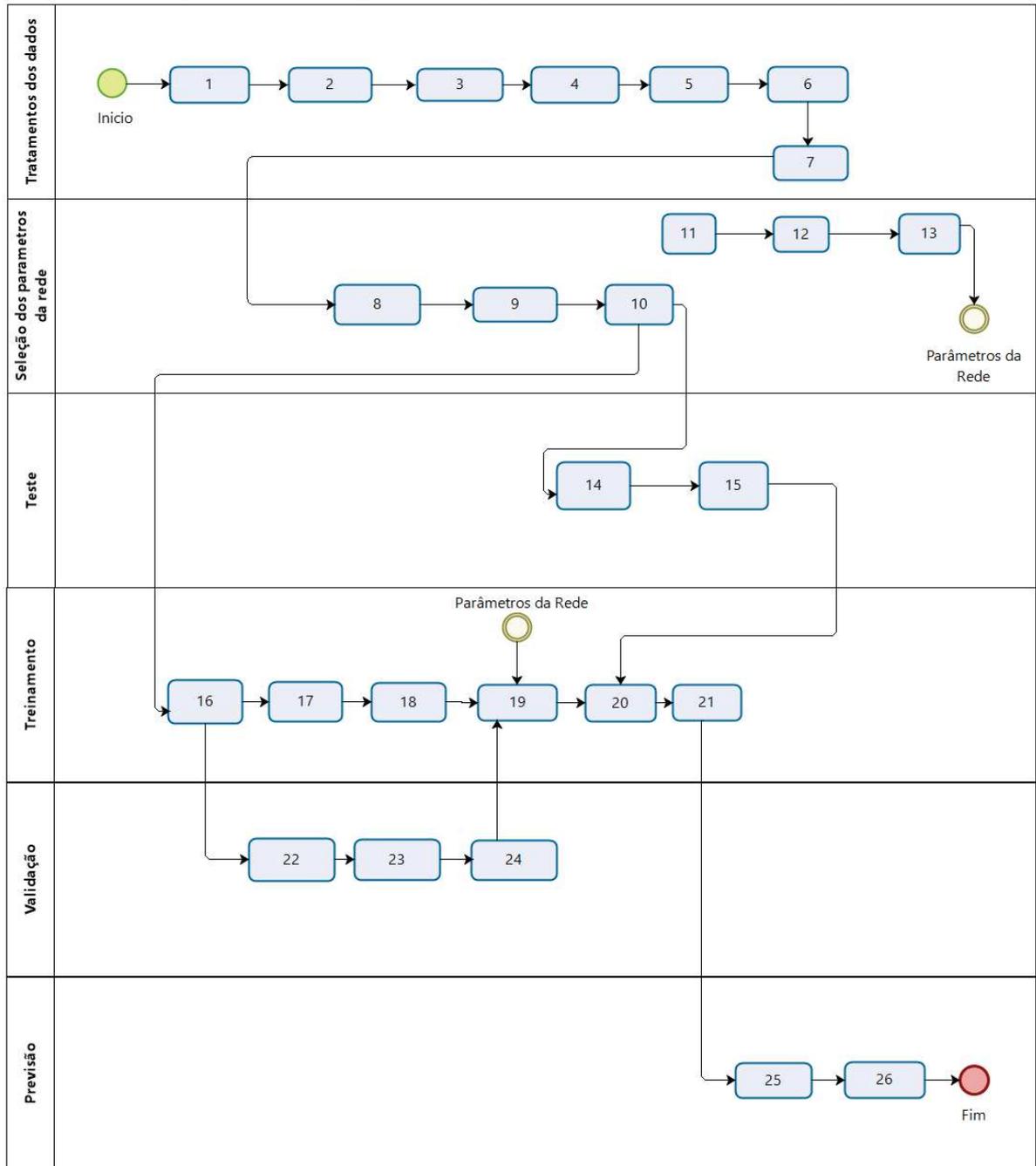
A Figura 4.1 apresenta o diagrama simplificado do fluxo de entradas e saídas da rede LSTM.

Figura 4.1 – Diagrama do fluxo de entradas e saídas da rede



Fonte: Criado pelo autor (2021).

O esboço da estrutura proposta com a rede LSTM é mostrado na Figura 4.2, com as etapas de forma macro, quanto ao processo de previsão de irradiação solar. Na tabela 4.1 segue a descrição dos passos da rede LSTM, conforme o diagrama da Figura 4.2.

Figura 4.2 – Diagrama Esquemático do *Framework* proposto

Fonte: Criado pelo autor (2021).

Tabela 4.1 - Descrição dos passos da rede LSTM

Passo	Descrição do Passo	Etapa do processo
1	Dados csv. INMET ano "Estação Solarimetrica"	Tratamento de Dados
2	Unificação dos arquivos formando um único "dataset"	Tratamento de Dados
3	Remoção ou substituição de valores nulos	Tratamento de Dados
4	Período útil de insolação diária das 07:00 até as 19 horas	Tratamento de Dados
5	Conversão dos dados de irradiação (kJ/m <sup>2</sup> ) para irradiação (kWh/m <sup>2</sup> )	Tratamento de Dados
6	Período útil (meses completos)	Tratamento de Dados
7	Transformação dados horários para soma diária	Tratamento de Dados
8	Seleção da estação solarimetrica	S. dos parâm. da rede
9	Seleção do ano e mês a serem previstos	S. dos parâm. da rede
10	Normalização dos dados	S. dos parâm. da rede
11	Definir número de neurônios na primeira camada oculta	S. dos parâm. da rede
12	Definir número de neurônios na 2 <sup>a</sup> camada oculta	S. dos parâm. da rede
13	Definir o número de épocas	S. dos parâm. da rede
14	Separação dos dados para teste	Teste
15	<i>ndm</i> a ser previsto	Teste
16	Separação dos dados para treinamento	Treinamento
17	Expansão dos dados	Treinamento
18	80% para treinamento	Treinamento
19	Realização do Treinamento	Treinamento
20	Testando a rede com os dados de teste	Treinamento
21	Normalização reversa dos resultados da previsão	Treinamento
22	Separação dos dados para validação	Validação
23	Expansão dos dados	Validação
24	20% dos dados separados para a validação	Validação
25	Análise do erro RMSE	Previsão
26	Análise do erro entre os dados da previsão e dados medidos na usina	Previsão

Fonte: Criado pelo autor (2021).

#### 4.2.2.1 Treinamento da rede

Com relação ao treinamento da rede neural, uma das alternativas para resolver o problema da parada do treinamento é a técnica da validação cruzada. Ao

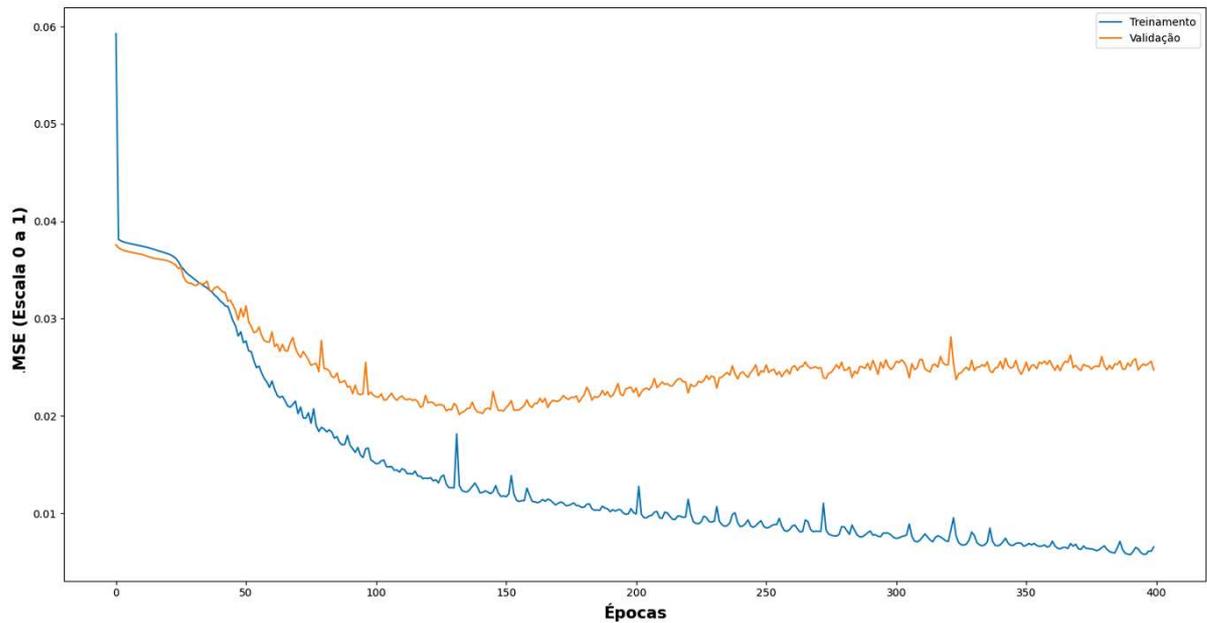
invés de definir o número exato de iterações de ajuste de pesos no treinamento, divide-se o conjunto de dados em 3 subconjuntos: treinamento, validação e testes.

Com isso, a cada iteração a rede é treinada, já com os pesos ajustados, e ela é testada com o subconjunto de validação, assim o erro da predição é calculado ao final de cada iteração. A motivação é ajustar os pesos com os dados do subconjunto de treinamento e calcular o erro com os dados do subconjunto de validação, fornecendo assim, dados diferentes à rede. Desta forma é possível avaliar o erro do treinamento da rede a cada iteração. No caso analisado, o erro da validação cruzada inicialmente é alto, e vai se reduzindo à medida que o treinamento avança até encontrar um determinado número de iterações que diverge do erro de treinamento.

Enquanto o erro da validação está reduzindo na mesma dinâmica que o de treinamento, podemos presumir que a rede está conseguindo generalizar o resultado, porém a partir de 50 iterações, pode-se perceber que o erro de validação não acompanha mais a dinâmica do erro de treinamento, e ao contrário o erro de treinamento reduz drasticamente, nesse caso a rede pode estar tendo sobreajuste (*overfitting*), perdendo a capacidade de generalizar. Assim, o treinamento pode ser finalizado no ponto de divergência das curvas.

Não foi adotado um critério específico para determinar o número de iterações máximas para a parada do treinamento da rede. O procedimento executado para determinar a parada do treino, foi de modo empírico com tentativa e erro. O processo foi conduzido através do *input* de vários números de épocas diferentes, aumentando gradativamente e observando o impacto no erro. Assim, no final das tentativas foi constatado que a partir de 50 épocas, inicia-se o *overfitting*, conforme apresentado no Gráfico 4.1 com o erro da validação e o erro do treinamento.

Gráfico 4.1 – Treinamento x Validação



Fonte: Criado pelo autor (2021).

Para um melhor entendimento quanto a aplicação, um mês dentro dos anos foi selecionado. Tomando como base um mês de janeiro, com dados de sequências de 2002 a 2018 com 31 dias e 16 anos, tem-se um total de 496 dias. Então, o *dataset* total inclui:

- **Treinamento e validação:** Para o treinamento, utilizou-se os dados dos anos 2002 a 2018; assim tem-se um vetor de tamanho 496 dias. Do valor total, foram utilizados 80% da amostra para treinamento e 20% para a validação dos resultados.
- **Teste:** Para o teste, foram utilizados os dados de 2019 como entrada para prever a saída em 2020, e os dados reais de 2020 são utilizados como comparação com o erro. Os dados de teste foram separados do *dataset* original, acompanhando o exemplo utilizado do mês de janeiro, a entrada ficaria os 31 dias de janeiro de 2019 e como saída os 31 dias de janeiro de 2020. O conjunto de dados utilizado no teste não tem contato com o treinamento.

Para a rede LSTM, tem-se a entrada de dados com a seguinte característica do vetor: [*samples*, *timesteps*, *features*], em que:

- “*samples*” - quantidade de amostras a serem utilizadas;
- “*timesteps*” - quantos dias passados a rede lembrará durante o treinamento;
- “*features*” – é a dimensão da saída do sistema, no caso a saída será um vetor de irradiação com 31 dias previstos.

Seguindo essa premissa, para o mês de janeiro do exemplo, conforme exemplo anterior, tem-se: [496, 31, 1].

Onde:

- “*samples*”: 496 é número de amostras
- “*timesteps*” 496 vetores, cada um com 31 posições
- “*features*” significa o número de variáveis que se quer projetar: 1

Uma estratégia para aumentar o número de amostras foi aplicada impactando positivamente na eficiência do treinamento, onde um novo *dataset* de treinamento foi criado. Esta estratégia é aplicada quando se deseja aumentar a amostragem de um vetor deslocando-o em  $N$  dias (*timesteps*) da previsão em uma janela de dados.

Considerando como exemplo um vetor de entrada contendo 31 dias de cada mês de janeiro durante 16 anos. Totaliza-se 496 dias de dados para treinamento. Como saída utiliza-se 31 dias do último ano a ser previsto.

Aplicando uma janela deslizante de reamostragem de 31 dias, com incremento de 1 dia de deslocamento a cada passo, no qual percorre todo o conjunto de entrada (496 dias) e um novo vetor é criado, ao qual foi expandido para 15.376 dias. O mesmo processo é aplicado à saída. Para ilustrar a aplicação desse processo, assume-se  $X$  como a entrada e  $Y$  como saída:

$$X = [[1,2,3,4,5,\dots,31], [2,3,4,5,6,\dots,1], [3,4,5,6,7,\dots,2], [4,5,6,7,8,\dots,3], \dots, [1,2,3,4,5,\dots,31]]$$

1° a 31° janeiro/2002	2° janeiro/2002 a 1° janeiro/2003	3° janeiro/2003 a 2° janeiro/2004	4° janeiro/2004 a 3° janeiro/2005	1° janeiro/2018 a 31° janeiro/2018
--------------------------	--------------------------------------	--------------------------------------	---	---------------------------------------

$$Y = [[1,2,3,4,5,\dots,31], [2,3,4,5,6,\dots,1], [3,4,5,6,7,\dots,2], \dots, [1,2,3,4,5,\dots,31]]$$

1° a 31° janeiro/2003	2° janeiro/2003 a 1° janeiro/2004	3° janeiro/2004 a 2° janeiro/2005	1° janeiro/2018 a 31° janeiro/2018
--------------------------	--------------------------------------	--------------------------------------	---------------------------------------

#### 4.2.2.2 Normalização

Como outras redes neurais, as LSTM tendem a se comportar melhor quando os dados estão dentro de uma escala para serem submetidos à função de ativação da rede.

A normalização é uma técnica geralmente como parte da preparação de dados para aprendizado de máquina, com o objetivo de alterar os valores das colunas numéricas no conjunto de dados para usar uma escala comum, sem distorcer as diferenças nos intervalos de valores ou perder informações. Os valores dos coeficientes de escala (mínimo e máximo) devem ser calculados no conjunto de dados de treinamento e aplicados para dimensionar o conjunto de dados de teste e as previsões. Isso evita contaminar o experimento com conhecimento do conjunto de dados de teste (De Melo, 2018).

As redes neurais não obtêm uma boa performance quando os dados possuem uma grande diferença em seu range, por isso foi necessário normalizar os dados antes do treinamento, onde utilizou-se a classe “MinMaxScale” da biblioteca “skitlearn”, que normaliza os dados de treinamento, teste e validação no intervalo entre [0 a 1]. A Equação 4.1 apresenta o método de normalização utilizado (Carlos, 2020).

$$z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad \text{Equação 4.1}$$

Onde:

- $x$  é um valor dentro de um vetor menos o valor mínimo deste mesmo vetor.
- O resultado desta subtração é então dividido pelo valor máximo menos o valor mínimo do vetor.
- Como resultado, todos os valores do vetor estarão compreendidos entre 0 e 1.

#### 4.2.2.3 Aplicação da rede LSTM

Na aplicação do modelo, uma rede LSTM é definida para cada mês desejado na previsão.

Foi utilizada uma rede neural LSTM com 50 neurônios na primeira camada oculta e 30 neurônios na segunda camada. A camada de saída prevê um vetor com *ndm*, ou seja, uma posição para cada dia do mês correspondente. Quanto ao tamanho do lote, ou *batch\_size*<sup>9</sup>, foi considerado 1, para que as amostras sejam analisadas de uma a uma, sendo atualizadas após cada passo de treinamento, que é o padrão da biblioteca usada.

O otimizador *Adam* foi utilizado por ser rápido e eficiente. O otimizador *Adam* utiliza um método de otimização que calcula taxas de aprendizado adaptativas para cada parâmetro e mantém uma média exponencial decrescente dos gradientes anteriores (Ribeiro et. al. 2020). O total de épocas de treinamento utilizadas foi de 50 épocas.

A métrica para treinamento utilizada é o erro médio quadrático (RMSE), conforme subcapítulo 4.3.4. Para plotar os resultados das previsões, foi usado o *MinMaxScale* para retornar os dados a escala original, realizando a normalização inversa.

Com a previsão da irradiação solar extraída da rede LSTM, é feito o comparativo da previsão com os dados previstos pelo CRESESB e pelos dados medidos reais correspondentes ao mês de geração do SFV. O desenvolvimento do código e o código em si constam no Apêndice B.

### 4.2.3 Métrica para a análise

Nos modelos criados pelos algoritmos de DL é necessário fazer avaliações, onde as ferramentas estatísticas auxiliam bastante neste processo. Para obter o desempenho da rede neural LSTM em estudo, foi utilizado a ferramenta estatística, o RMSE (Raiz do Erro Médio Quadrático).

O RMSE é a medida que calcula "a raiz quadrática média" dos erros entre valores observados (reais) e previsões (hipóteses). Esta métrica penaliza muito mais os erros maiores. Desta forma, uma técnica que apresente ótimos resultados na

---

<sup>9</sup> Tamanho do lote (*Batch Size*) é um termo usado em aprendizado de máquina e refere-se ao número de exemplos de treinamento usados em uma iteração (Book, 2021).

maioria das previsões, porém tenha erros elevados em uma previsão específica, irá fornecer um alto RMSE (Mentaschi et al. 2013). A Equação 4.2 apresenta o cálculo do RMSE.

$$RMSE = \frac{1}{N} \sum_{k=i}^N \sqrt{\left| \frac{a_k - y_k}{a_k} \right|^2} \quad \text{Equação 4.2}$$

Sendo  $a_k$  e  $y_k$  os dados de entrada e previstos, respectivamente. Pode-se dizer que quanto menores forem os valores do RMSE, melhor é o desempenho da rede neural (Olatomiwa et al. 2015). O RMSE expressa o erro médio do modelo preditivo, em relação aos dados originais (treino e/ou teste).

#### 4.3 CÁLCULO DA GERAÇÃO NO PERÍODO DE PREVISÃO

O objetivo da metodologia proposta com a utilização da rede LSTM é reduzir o tempo de acompanhamento do monitoramento do SFV, utilizando as previsões de irradiação solar da LSTM dos próximos seis meses, subsequentes ao monitoramento realizado de seis meses após a implementação do projeto.

Para a modelagem proposta, são utilizados seis meses de dados de geração extraídos dos inversores da usina e dados de irradiação solar medidos pela estação meteorológicas pertencentes ao INMET, presente no local do projeto.

De posse dos dados, uma análise de regressão segundo (EVO, 2014) é aplicada, onde modelo tenta explicar a variação da energia resultante das alterações nas variáveis independentes individuais. Assim, a Equação 4.3 expressa a geração solar a partir da variação da variável independente irradiação solar, considerando os dados de monitoramento de seis meses de operação da usina.

$$\text{Geração} \left( \frac{kWh}{\text{mês}} \right) = b + a \times \text{IRRADIAÇÃO} \quad \text{Equação 4.3}$$

Onde:

- $a$  coeficiente angular da regressão (para a variável independente);
- $b$  coeficiente linear da regressão;
- $IRRADIAÇÃO$  é a variável independente obtida dos seis meses de medição do banco de dados do INMET.

Com a Equação 4.3 é possível projetar os outros seis meses de geração solar, considerando os dados de irradiação mensal extraídos da rede LSTM como variável independente. Assim, mês a mês de geração solar é projetado fazendo a complementação dos seis meses que faltavam para o fechamento dos dozes meses de geração do SFV.

#### 4.4 NORMALIZAÇÃO DA GERAÇÃO

Para aferir a geração solar obtida com o projeto, no plano de medição e verificação é estabelecido como será a opção de M&V, segundo PIMVP, que será praticada no período de determinação da economia, no caso de geração anual.

Para projetos de SFV, usualmente trabalha-se com a opção B, conforme Tabela 2.1.

Além disso, modelos matemáticos são usados em M&V para definir como será calculada a economia de energia, no caso geração de energia obtida. Admitindo um dos modelos permitidos pelo PIMVP, o modo de se medir a energia gerada é pela “economia normalizada” ou “geração normalizada”, quando estas condições são fixas, de um padrão estabelecido.

Recomenda-se que seja utilizado esta opção (economia normalizada), onde alguns padrões podem ser definidos através de estudos. Para sistemas de fontes incentivadas que possuem como variáveis independentes como a irradiação solar, é usada a média mínima dos últimos 10 anos destes fatores para definição das condições padrão em um ano típico, sendo recomendado utilizar os dados do INMET da estação mais próxima da unidade consumidora onde o sistema será instalado.

Para a obtenção da geração mensal e anual “normalizada”, o primeiro passo foi gerar outra Equação 4.3 com os dados mensais de geração e irradiação obtidos no subcapítulo 4.3. O segundo passo foi projetar mais seis meses de irradiação solar com a rede LSTM, complementares aos seis meses projetados, completando um ano.

Com a Equação 4.4 e a irradiação mensal prevista pela rede LSTM, projeta-se um ano com condições fixas, de um padrão estabelecido.

$$\text{Geração normalizada} \left( \frac{kWh}{mês} \right) = b + a \times \text{IRRADIAÇÃO} (cf) \quad \text{Equação 4.4}$$

Onde:

- $a$  coeficiente angular da regressão;
- $b$  coeficiente linear da regressão;
- $IRRADIAÇÃO (cf)$  é a variável independente obtida dos doze meses projetados pela rede LSTM.

#### 4.5 APLICAÇÃO DAS MÉTRICAS DE AVALIAÇÃO

A primeira métrica de avaliação aplicada na análise de M&V no período de determinação da economia (geração solar obtida), é o processo de amostragem, conforme subcapítulo 2.1.1. A meta para a validação da amostragem selecionada, é que a amostragem fique dentro dos 10% de precisão. Caso contrário, amplia-se a amostra. A Equação 4.5 apresenta o cálculo da precisão obtida  $e$ , de acordo com o tamanho da amostra medida, o  $cv$  (coeficiente de variação será igual o desvio padrão  $S$  dividido pela média  $\bar{X}$ ) obtido, para uma confiabilidade de 95% ( $z = 1,96$ ).

$$e = \frac{z \times cv}{\sqrt{n_0}} \quad \text{Equação 4.5}$$

Onde:

- $e$  precisão obtida
- $z$  valor padrão da distribuição normal (confiabilidade de 95%) = 1,96
- $cv$  coeficiente de variação das medidas
- $n_0$  tamanho da amostra

$$\bar{X} = \frac{\sum X_i}{n} \quad \text{Equação 4.6}$$

Onde:

- $\bar{X}$  média das amostras
- $X_i$  amostra  $i$

- $n$  número de amostras

$$S = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}} \quad \text{Equação 4.7}$$

Onde:

- $S$  desvio padrão das amostras

Na condição de projetos com variável independente, mesmo se a primeira condição de precisão das amostras não for atendida, no modelo de regressão com as informações de erro padrão e números de amostras, aplica-se as mesmas fórmulas e condições acima para validação das incertezas.

Na sequência, parte-se para a validação do modelo de regressão na condição de existência da variável independente. Para avaliar quão bem um determinado modelo de regressão explica a relação entre o uso da energia e as variáveis independentes, três testes são mais usados:

- Coeficiente de determinação ( $R^2$ ): O primeiro passo que deve ser dado para avaliar a precisão de um modelo é examinar coeficiente de determinação  $R^2$ , que é uma medida do quanto da variação da energia é explicado pelo modelo. Matematicamente o  $R^2$  é explicado pela Equação 4.9.

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \quad \text{Equação 4.8}$$

Onde:

- $\hat{Y}_i$  valor da energia previsto pelo modelo em um ponto dado pela medição do valor da variável independente (ou seja, obtido pela inserção dos valores de  $X$  no modelo de regressão)
- $\bar{Y}$  média dos  $n$  valores medidos da energia
- $Y_i$  valor real medido da energia

A faixa de variação possível para o  $R^2$  é de. Quanto maior for o coeficiente de determinação, melhor a descrição do modelo da relação entre as variáveis dependentes e independentes. Embora não haja um padrão universal para um valor mínimo de aceitação, 0,75 é usualmente considerado um indicador mínimo razoável de uma boa relação causal entre a energia e as variáveis independentes.

- Erro padrão da estimativa ( $EP_{\hat{Y}}$ ): Quando um modelo é usado para prever um valor de energia ( $Y$ ) para um conjunto de variáveis independentes, a precisão da predição é medida pelo erro padrão da estimativa ( $EP_{\hat{Y}}$ ). A Equação 4.10 apresenta o cálculo:

$$\hat{Y} \pm t \times EP_{\hat{Y}} \quad \text{Equação 4.9}$$

Onde:

- $\hat{Y}$  valor da energia previsto pelo modelo em um ponto dado pela medição do valor da variável independente (ou seja, obtido pela inserção dos valores de  $X$  no modelo de regressão)
- $t$  é o valor obtido da estatística  $t$  de Student (Apêndice A)
- $EP_{\hat{Y}}$  é o erro padrão da estimativa (predição), calculado como:

$$\hat{Y} \pm t \times EP_{\hat{Y}} = \sqrt{\frac{\sum(\hat{Y}_i - Y_i)^2}{n-p-1}} \quad \text{Equação 4.10}$$

Onde:

- $p$  é o número de variáveis independentes na equação de regressão.
- Esta estatística é usualmente referida como erro médio quadrático ( $EMQ$ ). Dividindo-se o  $EMQ$  pela média do consumo de energia, tem-se o coeficiente de variação do  $EMQ$ , ou  $CV(EMQ)$ .

$$CV(EMQ) = \frac{EP_{\hat{Y}}}{\bar{Y}} \quad \text{Equação 4.11}$$

- Estatística *t*: A estatística *t* é um teste estatístico para determinar se uma estimativa tem significado estatístico. Segundo (EVO, 2014), uma regra empírica estabelece que se o valor absoluto da estatística *t* for maior que 2, significa que o coeficiente estimado é significativo em relação ao erro padrão. Entretanto, uma estatística *t* perto de 2 significa que a precisão do valor do coeficiente está em torno de  $\pm 100\%$ , o que pode gerar desconfiança no valor calculado. No APÊNDICE A, constam as considerações completas da Estatística *t*.

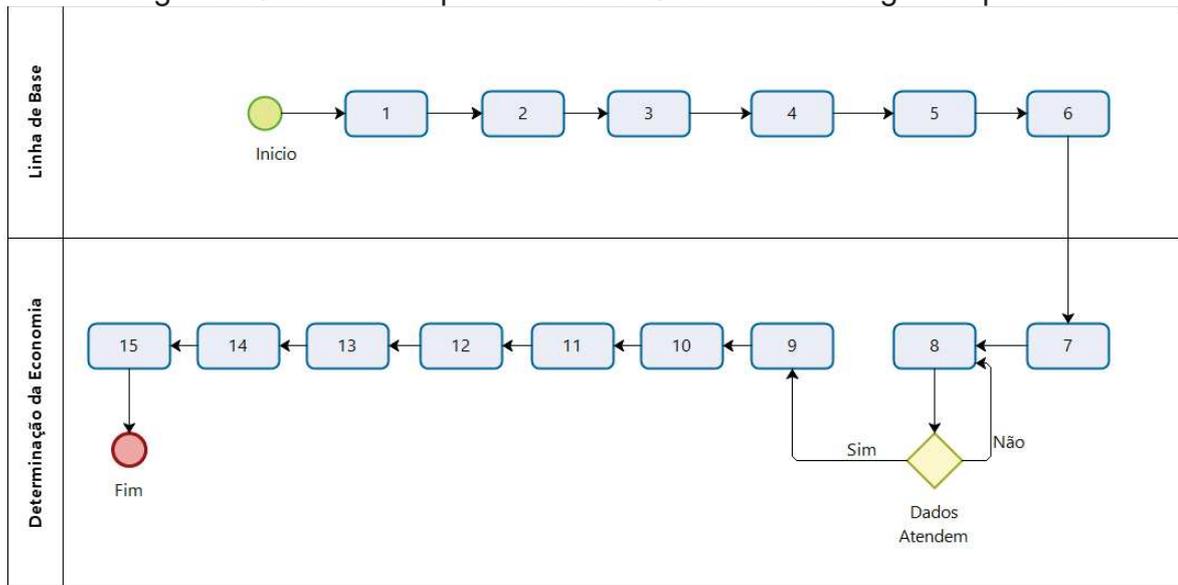
Para os projetos de EE da ANEEL, as considerações seguem o mesmo conceito com validação da regressão em  $R^2$  maior que 0,75, *CV (EMQ)* menor que 5% e estatística *t* ser maior que 2. Se dois deles atenderem aos critérios, há uma sugestão de adotar-se a regressão, que deve ser confirmada. Caso valide, segue-se com a regressão, caso contrário não são consideradas variáveis independentes.

Para os cases apresentados no capítulo 5, somente com seis meses de medições de acompanhamento do SFV operando, foi possível encontrar a precisão dentro dos 10% com 95% de confiabilidade, e demais indicadores estatísticos para a validação da regressão. Assim, no método proposto, por garantia para projetos futuros, o acompanhamento dos 6 meses torna-se necessário para evitar o não atingimento dos critérios necessários para validação do modelo.

#### 4.6 FLUXOGRAMA DA METODOLOGIA

A Figura 4.2 apresenta o fluxo do processo de M&V na Metodologia Proposta com as fases desde a linha de base até o período de determinação da economia, que no caso sendo um projeto com SFV, a economia torna-se a geração de energia em kWh obtida anual do sistema. Neste fluxo nos passos de 8 a 12 representados na Figura 4.2 retratam a parte do processo com a utilização da rede LSTM.

Figura 4.3 – Fluxo do processo de M&V na Metodologia Proposta



Fonte: Criado pelo autor (2021).

O processo da Figura 4.2 encerra-se com a computação da geração anual “normalizada” compatível a prevista na linha de base, atingindo as metas estabelecidas em contrato. A Tabela 4.2 traz a descrição dos passos de cada etapa do processo da M&V na Metodologia Proposta.

Tabela 4.2 - Descrição dos passos da M&amp;V no método proposto

<b>Passo</b>	<b>Descrição do Passo</b>	<b>Etapa do processo</b>
1	Obtenção dos dados de irradiação no CRESESB e condições físicas do projeto	Linha de Base
2	Simulação da geração solar com projeção mensal e anual	Linha de Base
3	Definição da variável independente através da regressão linear (geração simulada x irradiação solar "CRESESB")	Linha de Base
4	Definição da fronteira de medição e efeitos interativos com o projeto	Linha de Base
5	Definição da opção de PIMVP	Linha de Base
6	Definições dos fatores estáticos	Linha de Base
7	Implementação do projeto	Determinação da Economia
8	Obtenção dos dados de irradiação solar com base no INMET e geração de energia do SFV com monitoramento de seis meses	Determinação da Economia
9	Validação das incertezas das amostras das medições da geração de energia do SFV com seis meses	Determinação da Economia
10	Validação da regressão linear (geração medida x irradiação solar medida "INMET")	Determinação da Economia
11	Aplicação da rede LSTM para prever os outros seis meses complementares de irradiação solar, a partir dos dados históricos do "INMET")	Determinação da Economia
12	Aplicar a previsão da rede LSTM na equação de regressão linear do passo 10 e obter a projeção da geração solar dos seis meses complementares aos outros 6 meses de monitoramento real	Determinação da Economia
13	Validação da regressão linear com os 12 meses medidos e projetados (geração medida e projetada x irradiação solar medida "INMET" medida e a projetada)	Determinação da Economia
14	Normalização da geração de energia anual com dados "padrão" de irradiação solar (rede LSTM com a projeção dos seis meses após o monitoramento e seis meses antes do monitoramento) e juntamente com a equação da regressão linear obtida no passo 13	Determinação da Economia
15	Obtenção final da geração normalizada mensal e anual do SFV e comparação com a linha de base	Determinação da Economia

Fonte: Criado pelo autor (2021).

#### 4.7 CONCLUSÕES PARCIAIS

O capítulo 4 teve como objetivo apresentar a metodologia proposta, com a utilização da rede neural LSTM, as justificativas sobre o porquê da utilização deste método para a previsão da irradiação solar, a arquitetura para a solução do problema, como foi realizado o tratamento dos dados e o desenvolvimento do código, juntamente com o método de análise e validação da rede proposta. A rede

neural LSTM é bem utilizada para aplicação em previsão de séries temporais como a irradiação solar, trazendo bons resultados ao estudo em questão, conforme segue no capítulo 5.

## 5 TESTES E RESULTADOS

### 5.1 INTRODUÇÃO

Após a definição da metodologia no capítulo 4, com o propósito de fazer uma verificação real utilizando o método proposto, dois ‘cases’ de projetos com SFV nos moldes do PROPEE foram utilizados. Primeiramente, são apresentados os resultados do projeto com o método tradicional e, na sequência, apresenta-se a aplicação da metodologia com o uso da rede neural LSTM.

Os cases de projetos com SFV envolvem um no estado de Tocantins e outro localizado no estado de Goiás. Esses cases são abordados de duas formas: i) apresentação da linha de base dos projetos e o período de determinação da geração solar através do método tradicional, com acompanhamento por 12 meses; ii) apresentação do período de determinação da geração solar com acompanhamento de seis meses, projetando os outros seis meses subsequentes com a aplicação da rede LSTM para a variável independente irradiação solar.

### 5.2 SFV NO ESTADO DE TOCANTINS

O projeto está localizado na cidade de Araguatins, no estado de Tocantins, sendo a unidade consumidora um prédio público de ensino superior e técnico, com regime de funcionamento com três turnos de aulas. A instalação do sistema aconteceu no telhado da instituição no ano de 2019.

#### 5.2.1 M&V método tradicional

No projeto, antes da implementação do SFV e após a assinatura do contrato, é definida a “Linha de Base”, consolidando o plano de M&V. Com a implementação do projeto, inicia-se o acompanhamento e monitoramento da geração e irradiação solar por um período de doze meses, necessários para o fechamento da “determinação da economia”, no caso, da geração solar obtida, consolidando o relatório final de M&V.

##### 5.2.1.1 Definição da Linha de Base

No período de definição da linha de base, para a simulação da geração solar mensal e anual, foi utilizado o software PVsyst, através dos dados extraídos do projeto e suas condições físicas, além dos dados de irradiação solar do CRESESB. Os dados do CRESESB da cidade de Araguatins - TO, envolvendo a localização geográfica da cidade e a irradiação solar diária média, encontram-se na Tabela 5.1, obtidos de [[http://www.cresesb.cepel.br/index.php#localidade\\_50558](http://www.cresesb.cepel.br/index.php#localidade_50558)].

Tabela 5.1 – Dados de Irradiação Solar no Plano Horizontal (Ano padrão)

Irradiação solar diária média [kWh/m <sup>2</sup> .dia]															
Latitude [°]	Longitude [°]	Distância [km]	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez	Média
5,701	48,049	9,2	4,6	4,7	4,8	4,9	4,8	5,1	5,3	5,8	5,5	5	4,8	4,6	4,99

Fonte: Elaborado pelo autor (2021).

A irradiação solar da Tabela 5.1 encontra-se em um formato de contabilização média diária mensal. Para a visão da irradiação mensal total multiplica-se a média diária pelo número de dias de cada mês, conforme Figura 5.3 na coluna “*GlobHor*”.

#### 5.2.1.1.1 Definição das Metas e Condições Físicas do Projeto

O SFV tem potência nominal instalada de 31,5 kWp. As Figuras 5.1 a 5.4 apresentam os resultados extraídos do software PVsyst, trazendo desde as características do sistema até os resultados esperados com a implementação do projeto.

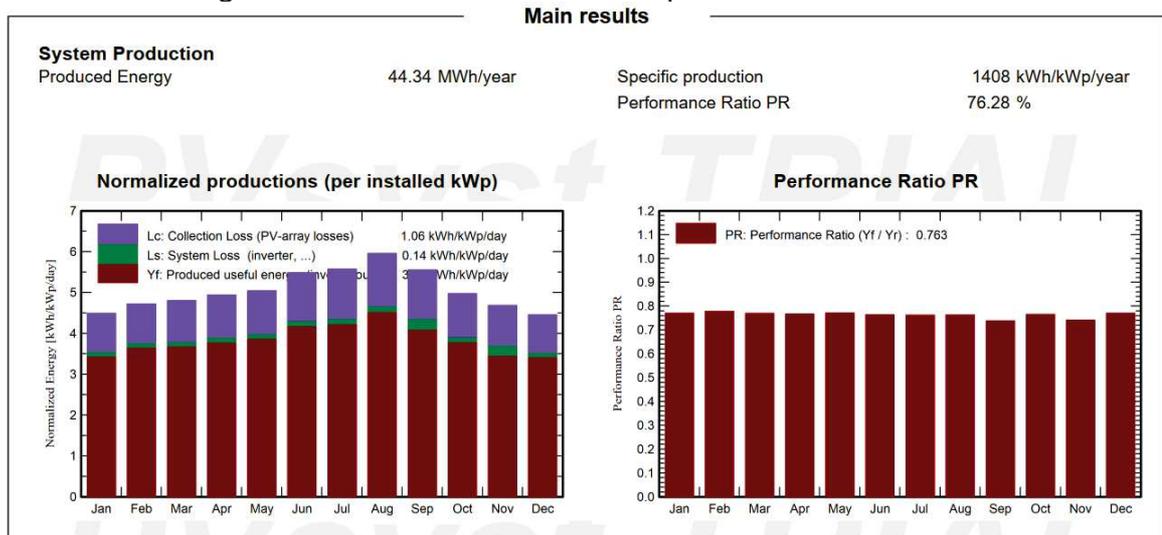
Figura 5.1 – Características do SFV

PV Array Characteristics			
<b>PV module</b>		<b>Inverter</b>	
Manufacturer	Generic	Manufacturer	Generic
Model	TSM-DEG14-(II)-350	Model	TRIO-27.6-TL-OUTD-400 (27,6 kWac max)
(Original PVsyst database)		(Original PVsyst database)	
Unit Nom. Power	350 Wp	Unit Nom. Power	27.6 kWac
Number of PV modules	90 units	Number of inverters	1 units
Nominal (STC)	31.5 kWp	Total power	27.6 kWac
Modules	6 Strings x 15 In series	Operating voltage	200-950 V
<b>At operating cond. (50°C)</b>		Pnom ratio (DC:AC)	1.14
Pmpp	28.47 kWp		
U mpp	518 V		
I mpp	55 A		
<b>Total PV power</b>		<b>Total inverter power</b>	
Nominal (STC)	32 kWp	Total power	28 kWac
Total	90 modules	Nb. of inverters	1 Unit
Module area	178 m <sup>2</sup>	Pnom ratio	1.14
Cell area	157 m <sup>2</sup>		

Fonte: Elaborado pelo autor no PVSyst (2021).

As Figuras 5.2 e 5.3 apresentam os resultados de desempenho, ou *Performance Ratio*<sup>10</sup>(PR) para o SFV, juntamente com os resultados esperados de geração mensal e anual para o projeto. Assim, o SFV foi projetado para uma PR de 76,28%, uma irradiação solar global no plano horizontal por ano de 1822,1 kWh/m<sup>2</sup> e geração solar anual calculada de 44,339 MWh.

Figura 5.2 – Resultado de desempenho – Linha de Base



Fonte: Elaborado pelo autor no PVSyst (2021).

<sup>10</sup> PR é a relação entre a estimativa de produção de energia, considerando todas as perdas do sistema fotovoltaico, e a produção esperada sem cômputo dessas perdas, para um determinado período de relatório, com base na classificação da placa de identificação do sistema (Normas, 1998).

Figura 5.3 – Resultado de irradiação e geração – Linha de Base

Balances and main results								
	GlobHor	DiffHor	T_Amb	GlobInc	GlobEff	EArray	E_Grid	PR
	kWh/m <sup>2</sup>	kWh/m <sup>2</sup>	°C	kWh/m <sup>2</sup>	kWh/m <sup>2</sup>	MWh	MWh	ratio
January	143.2	89.30	26.00	139.0	131.0	3.476	3.372	0.770
February	134.4	71.30	26.00	132.1	125.3	3.334	3.236	0.778
March	148.8	92.40	26.50	148.8	140.5	3.717	3.607	0.770
April	145.5	75.60	27.00	148.2	140.7	3.693	3.582	0.767
May	149.1	64.70	26.48	156.2	148.0	3.910	3.794	0.771
June	154.2	39.00	27.50	164.5	156.3	4.078	3.959	0.764
July	162.4	37.90	28.00	172.8	164.3	4.266	4.143	0.761
August	177.6	55.70	28.00	184.6	176.3	4.567	4.433	0.762
September	164.4	60.60	28.50	166.6	159.1	4.130	3.875	0.738
October	155.3	98.40	27.50	154.1	145.8	3.827	3.712	0.765
November	143.7	92.90	27.00	140.4	132.4	3.512	3.277	0.741
December	143.5	79.10	26.50	138.0	130.0	3.453	3.349	0.770
Year	1822.1	856.90	27.09	1845.3	1749.6	45.962	44.339	0.763

**Legends**

GlobHor	Global horizontal irradiation	EArray	Effective energy at the output of the array
DiffHor	Horizontal diffuse irradiation	E_Grid	Energy injected into grid
T_Amb	Ambient Temperature	PR	Performance Ratio
GlobInc	Global incident in coll. plane		
GlobEff	Effective Global, corr. for IAM and shadings		

Fonte: Elaborado pelo autor no PVSyst (2021).

#### 5.2.1.1.2 Variável Independente

No período de “Linha de Base”, foi identificada a irradiação solar como variável independente, conforme detalhado no subcapítulo 4.3. Essa indicação de variável independente foi definida aplicando-se a análise de regressão a partir dos dados de geração simulados (PVSyst) e de irradiação solar do CRESESB.

Seguindo os passos para a validação da regressão apresentados no subcapítulo 4.5, a Tabela 5.2 apresenta os resultados obtidos para essa validação, em que  $R^2$  é maior do que 0,75,  $CV(EMQ)$  é menor do que 5% e a estatística  $t$  é maior que 2.

Tabela 5.2 – Regressão linear (geração estimada PVSyst x irradiação solar)

<b>Coefficiente angular da regressão (a)</b>	29,23 m <sup>2</sup> / mês
<b>Coefficiente linear da regressão (b)</b>	-742,85 kWh/mês
<b>Estatística t</b>	8,33
<b>Coefficiente de determinação (R<sup>2</sup>)</b>	87%
<b>Coefficiente de Variação do Erro Médio Quadrático (CV (EMQ))</b>	3,7%
<b>Incertezas</b>	7,27%

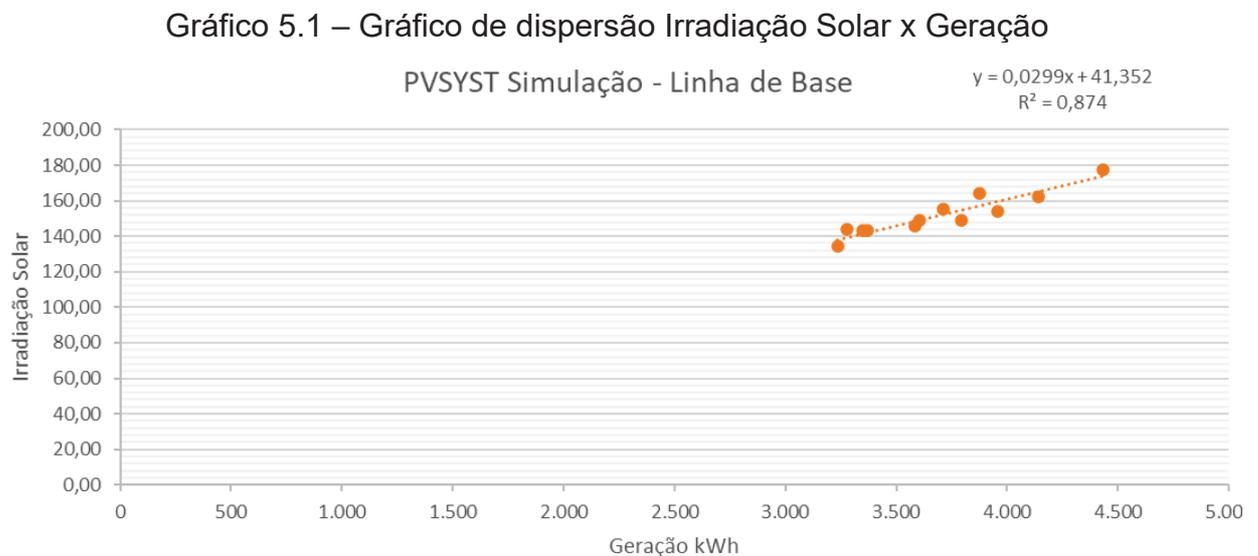
Fonte: Elaborado pelo autor (2021).

Quanto as incertezas, ou seja, a determinação dos níveis desejados de precisão e confiança para a estimativa a ser determinada, a confiança refere-se à probabilidade da estimativa ficar dentro da faixa de precisão, conforme apresentado no subcapítulo 2.1.1. Então, a incerteza encontrada e mostrada na Tabela 5.2 foi de 7,27% atingindo a meta de 10% de precisão a 95% de confiabilidade, quer-se dizer que a faixa definida por  $\pm 10\%$  em torno da média conterá o valor verdadeiro (que não se conhece) com a probabilidade de 95%, ou seja, em 95% dos casos.

A Equação 5.1, obtida através dos coeficientes  $a$  e  $b$  apresentados na tabela 5.2, indica que a variação da energia gerada é causada por mudanças na irradiação solar.

$$\text{Geração} \left( \frac{\text{kWh}}{\text{mês}} \right) = -742,85 + 29,23 \times \text{IRRADIAÇÃO} \quad \text{Equação 5.1}$$

O Gráfico 5.1, obtida através dos coeficientes  $a$  e  $b$  apresentados na tabela 5.2, indica que a variação da energia gerada é causada por mudanças na irradiação solar. A região de interesse da reta é quanto  $x \neq 0$ , onde somente existe geração se tem irradiação solar, onde  $y$  sendo a geração nunca será igual a  $b$ , no caso de  $x \neq 0$ .



Fonte: Elaborado pelo autor (2021).

### 5.2.1.1.3 Fronteira de Medição e Efeitos Interativos com o Projeto

Como fronteira de medição para o período da linha de base, considerou-se a entrada de energia da concessionária (para este prédio), levando-se em consideração as características dos locais da instalação, assim como a irradiação solar mensal. Para o período de determinação da economia, foi considerada a utilização de medições extraídas dos inversores do sistema.

Os inversores são do fabricante ABB (Asea Brown Boveri) e apresentam uma plataforma de gerenciamento de planta (Aurora Vision), que é o sistema de monitoramento da ABB, permitindo que se obtenha uma visão em tempo real de quanta energia que foi gerada.

Os possíveis efeitos interativos com o sistema de geração fotovoltaica que poderiam surgir ao longo do projeto, foram ignorados para a avaliação dos benefícios. Todavia, os efeitos gerados com o projeto foram monitorados após a implementação do projeto.

#### 5.2.1.1.4 Opção do PIMVP

A determinação das economias foi feita pelo prazo de cálculo do benefício do projeto, até completar um ano adotando a opção B – Medição isolada de todos os parâmetros (Tabela 2.1).

#### 5.2.1.1.5 Fatores Estáticos

Possíveis efeitos estáticos foram avaliados e monitorados durante as medições, e desde o plano de M&V foram estabelecidos como possibilidade de impacto, tais como: efeito de sombreamento nos painéis causados após a implementação do projeto como árvores ou construções não existentes e ou ampliações; mudanças significativas no perfil de utilização da instalação do projeto. Outro ponto importante, que foi observado é que como o sistema de geração de energia fotovoltaico é sincronizado com a rede da concessionária, a mesma deveria manter o fornecimento de energia, pois, caso contrário, não haveria geração. Este fato ocorre devido aos inversores do SFV operarem somente com o sistema energizado.

#### 5.2.1.2 Definição do período de determinação da geração obtida

O período de determinação da geração iniciou-se logo após a instalação do SFV, com um tempo total de um ano de monitoramento, com memória de massa da energia injetada na rede, garantindo a verificação dos dados de um ciclo de operação entendido como completo (ANEEL, 2020).

O benefício final para o projeto, a geração de energia, teve como base um ano padrão, no caso utilizou-se os dados de irradiação solar do CRESESB em relação à variável independente, considerando a geração normalizada conforme subcapítulo 4.4.

#### 5.2.1.2.1 Dados dos resultados durante o período de monitoramento do projeto

Os dados de geração extraídos do SFV, bem como de irradiação solar do INMET correspondentes ao período de geração real, seguem na Tabela 5.3. Os resultados obtidos frente ao projeto, em comparação aos dados estimados na linha de base no subcapítulo 5.2.1.1 na Figura 5.3, foram para a geração solar de 9,17%, e para a irradiação solar de 2,46%.

Tabela 5.3 – Registro da geração e irradiação solar

Ano	Mês	Dados Medidos	
		kWh	Irradiação Solar (kWh/m <sup>2</sup> . mês)
2020	fev	3.675,01	140,63
2020	mar	3.903,67	145,50
2020	abr	3.504,27	133,74
2020	mai	4.190,25	149,52
2020	jun	4.515,73	147,90
2020	jul	4.801,86	162,49
2020	ago	4.567,08	167,74
2020	set	4.414,59	161,16
2020	out	4.168,88	154,18
2020	nov	3.485,99	130,26
2020	dez	3.731,36	145,09
2021	jan	3.858,46	140,11
<b>Total</b>		<b>48.817,15</b>	<b>1.778,33</b>

Fonte: Elaborado pelo autor (2021).

Para a irradiação solar medida, a base de dados utilizada foi extraída da base de dados do INMET da estação A225, conforme Tabela 5.4.

Tabela 5.4 – Dados da Estação Meteorológica

REGIAO:	NE
UF:	MA
ESTACAO:	IMPERATRIZ
CODIGO (WMO):	A225
LATITUDE:	-5,555723
LONGITUDE:	-47,459794
ALTITUDE:	117,96
DATA DE FUNDACAO:	01/03/2008

Fonte: Elaborado pelo autor (2021).

A estação Imperatriz foi utilizada, pois a estação de Araguatins A021 apresentou, pelo INMET, falta de registro de dados ao longo de vários anos, inclusive no ano de análise do projeto, o que dificulta a aplicação da rede neural LSTM. Sendo assim, foram analisadas estações próximas, com o intuito de identificar qual teria mais semelhança com a de Araguatins, chegando à estação de Imperatriz.

Weber (2011) traz alternativas para a falta registros disponíveis de irradiação solar para uma determinada área de interesse, através do uso de estimativas. Existe uma grande variedade de métodos e para o presente trabalho, adotou-se o “Uso direto de dados de estações vizinhas”, levando-se em conta que condições climáticas, de distâncias entre as cidades e dados históricos apresentam uma certa similaridade. Tanto a cidade de Imperatriz (95 m altitude), quanto Araguatins (103 m altitude), tem clima tropical, com distância entre elas, em linha reta, de 72,51 km, condução 101 km (estrada) e tempo de deslocamento estimado 1 hora e 13 minutos.

Para dar uma ideia da proximidade entre os dados de irradiação solar destas cidades, os dados do CRESESB foram utilizados, onde observa-se diferenças que não ultrapassam 1% conforme Tabela 5.5, sendo que para o estudo em questão torna-se aceitável. Os dados do CRESESB, conforme subcapítulo 2.1.3, foram produzidos a partir de um total de 17 anos de imagens de satélite, assim tem-se um ano padrão com base em meses históricos.

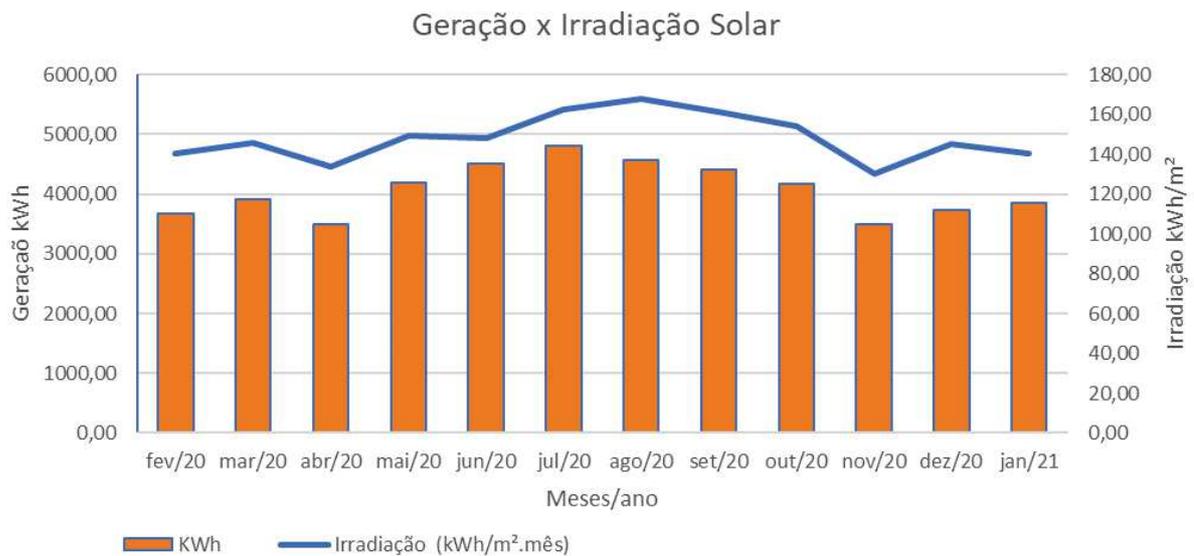
Tabela 5.5 – Comparativo entre dados meteorológicos entre cidades

Meses	Irradiação solar diária média [kWh/m <sup>2</sup> .dia]	
	Araguatins	Imperatriz
jan	4,67	4,64
fev	4,73	4,71
mar	4,89	4,8
abr	4,87	4,84
mai	4,81	4,79
jun	5,08	5,14
jul	5,22	5,26
ago	5,77	5,71
set	5,5	5,59
out	5,14	5,19
nov	4,86	4,91
dez	4,69	4,75
<b>Média</b>	<b>5,02</b>	<b>5,03</b>

Fonte: Elaborado pelo autor (2021).

O Gráfico 5.2 é representativo dos dados da Tabela 5.3, com a irradiação e geração de acordo com os meses e anos.

Gráfico 5.2 - Regressão linear (geração x irradiação)



Fonte: Elaborado pelo autor (2021).

#### 5.2.1.2.2 Variável Independente

Na construção da linha de base para o projeto é apresentada a possibilidade de uma variável independente, com a validação da regressão linear entre a geração simulada e a irradiação solar obtida pelo CRESESB, conforme subcapítulo 5.2.1.1.2. No período de determinação dos resultados do projeto, de posse dos dados reais de medição presentes na tabela 5.5, realiza-se a validação da regressão linear para a consolidação das condições da correlação entre os dados de geração medidos e de irradiação solar registrados pela base do INMET.

Seguindo os passos para a validação da regressão apresentados no subcapítulo 4.5, a Tabela 5.6 apresenta os resultados obtidos para a validação, em que  $R^2$  é maior do que 75% com 80,7%, o  $CV(EMQ)$  é menor do que 5% com 4,98% e a estatística  $t$  é maior que 2 com 6,46.

Tabela 5.6 – Regressão linear (geração medida x irradiação solar INMET).

<b>Coeficiente angular da regressão (a)</b>	34,27 m <sup>2</sup> / mês
<b>Coeficiente linear da regressão (b)</b>	-1010,25 kWh/mês
<b>Estatística t</b>	6,46
<b>Coeficiente de determinação (R<sup>2</sup>)</b>	80,7%
<b>Coeficiente de Variação do Erro Médio Quadrático (CV (EMQ))</b>	4,98%
<b>Incertezas</b>	10%

Fonte: Elaborado pelo autor (2021).

Com os resultados da Tabela 5.6, é possível construir a Equação 5.2, onde em *IRRADIAÇÃO* pode ser imputado valores de condições padrões de irradiação solar mensal, tendo como intuito trazer as condições de geração de energia no somatório de um ano, o mais perto de uma condição aproximada a uma média dos outros ano, e não somente do ano em que foi feito o monitoramento da geração após a conexão do SFV.

$$Geração \left( \frac{kWh}{mês} \right) = -1010,25 + 34,272 \times IRRADIAÇÃO \quad \text{Equação 5.2}$$

Com a Equação 5.2 é possível partir para a próxima e última etapa de aferição final dos resultados para o projeto no modelo tradicional apresentado no subcapítulo 5.2.1.2.3, trazendo a geração de cada mês compatível as variações da irradiação solar dos meses de um ano padrão, onde utiliza-se nesta análise os

dados extraídos do CRESESB, da mesma forma que foi utilizado na linha de base para construir as simulações, colocando tanto o período de linha de base quanto o período de determinação da geração final do SFV nas mesmas condições.

#### 5.2.1.2.3 Geração de energia do SFV “Normalizada”

Recomenda-se que seja utilizado a opção de “economia normalizada”, quando padrões podem ser definidos através de estudos, conforme apresentado no subcapítulo 4.4. Para sistemas de fontes incentivadas que possuam como variável independente a irradiação solar, é usada a média mínima dos últimos 10 anos destes fatores para definição das condições padrão em um ano típico. O ano padrão para as condições de irradiação solar foi baseado nos dados extraído do CRESESB.

A Tabela 5.7 mostra a análise final dos resultados obtidos com o projeto, onde a energia gerada anual normalizada foi de 50.317,08 MWh/ano. Comparando a geração anual normalizada com a geração real obtida na Tabela 5.3, a diferença encontrada foi de 3,07%. Na comparação com as condições da linha de base da Figura 5.3 a diferença encontrada na geração anual é de 11,88%.

Tabela 5.7 – Determinação dos Resultados de Geração “Normalizada”

Meses/Ano	Geração Solar (kWh) Normalizada
fev/20	3.595,40
mar/20	4.088,86
abr/20	3.975,77
mai/20	4.099,14
jun/20	4.273,91
jul/20	4.554,90
ago/20	5.075,78
set/20	4.623,44
out/20	4.311,60
nov/20	3.914,09
dez/20	3.907,24
jan/21	3.896,96
<b>Soma</b>	<b>50.317,09</b>

Fonte: Elaborado pelo autor (2021).

## **5.2.2 Rede neural LSTM para o Projeto em Tocantins**

No subcapítulo 4.6, um fluxograma do processo de M&V aplicando as previsões da irradiação solar via rede neural LSTM é apresentado, a partir do passo 11. Através dos seis meses de monitoramento da geração do SFV e irradiação solar medida pela INMET (fev/2020 a julho/2020), foi possível projetar os demais seis meses de geração solar (ago/2020 a jan/2021) do SFV através da equação de regressão linear obtida com os dados reais do monitoramento, e dos dados de irradiação solar obtidos da rede neural LSTM coincidindo com este mesmo período, conforme subcapítulo 5.2.2.2.

A decisão da utilização dos 6 meses de medição dos SFV e do INMET foi tomada a partir dos resultados encontrados da precisão e dos indicadores estáticos para a validação da regressão, conforme subcapítulo 4.3. Para a obtenção da equação de regressão linear para aplicação da metodologia proposta, menos meses foram testados, mas somente com 6 meses foi possível fazer a validação do modelo.

Para a verificação da assertividade dos resultados obtidos com a rede neural LSTM em comparação aos dados reais medidos, também foi projetado os meses do período de monitoramento, completando um ano de irradiação solar, ou seja, foi gerado uma rede neural LSTM para cada mês desde fev/2020 a jan/2021, apresentados no subcapítulo 5.2.2.1. Além disso, este um ano produzido pela rede LSTM de irradiação solar foi utilizado na aplicação da normalização da geração solar anual final para o projeto, considerando que este ano seria um ano padrão por ter sido criado com base em dados históricos extraídos do INMET, mostrado no subcapítulo 5.2.2.3.

### **5.2.2.1 Resultados de irradiação solar projetados pela rede neural LSTM**

Um ano completo foi projetado pela rede LSTM, a fim de comparar os resultados de irradiação solar mensal e anual com os dados medidos extraído do INMET do mesmo período, assim como apresentado na Tabela 5.8. Os dados utilizados do INMET da estação de Imperatriz – TO, iniciam-se de 1/1/2009 e terminam em 31/1/2020.

Tabela 5.8 – Dados de irradiação solar comparativos LSTM x INMET

Mês	RSME (kWh/m <sup>2</sup> )	Irradiação Solar INMET (kWh/m <sup>2</sup> )	LSTM (kWh/m <sup>2</sup> )	Diferença (INMET x LSTM)
fev/20	1,12	140,63	141,8	-0,83%
mar/20	1,12	145,5	145,81	-0,21%
abr/20	1,15	133,74	143,6	-7,37%
mai/20	0,84	149,52	150,25	-0,49%
jun/20	0,41	147,9	150,86	-2,00%
jul/20	0,45	162,49	158,37	2,54%
ago/20	0,63	167,74	161,46	3,74%
set/20	0,7	161,16	161,87	-0,44%
out/20	0,65	154,18	154,51	-0,21%
nov/20	1,45	130,26	139,25	-6,90%
dez/20	1,12	145,09	147,93	-1,96%
jan/21	1,09	140,11	140,36	-0,18%
	0,89	1.778,32	1.796,07	-1,00%

Fonte: Elaborado pelo autor (2021).

Os dados de irradiação solar projetados pela rede LSTM no somatório anual em comparação aos dados do INMET, apresentaram uma diferença de 1%. Os resultados do RMSE aplicado na análise, que comparam os valores medidos pelo INMET (medidos) com as previsões da irradiação solar da LSTM, foram, em média, de 0,89 kWh/m<sup>2</sup>.

Para fins de análise comparativa quanto ao método utilizado com redes neurais LSTM, também foi utilizado a média com base no mesmo banco de dados histórico de cada mês e ano do INMET utilizados nas redes LSTM. A Tabela 5.9 mostra os resultados.

Tabela 5.9 – Dados de irradiação solar comparativos MÉDIA x INMET

Mês	Irradiação Média 2009 a 2019 (kWh/m <sup>2</sup> )	Irradiação Solar INMET (kWh/m <sup>2</sup> )	Diferença (Média x INMET)
fev/20	130,024	140,63	-8,16%
mar/20	143,881	145,5	-1,12%
abr/20	141,530	133,74	5,50%
mai/20	142,971	149,52	-4,58%
jun/20	147,103	147,9	-0,54%
jul/20	154,158	162,49	-5,40%
ago/20	159,462	167,74	-5,19%
set/20	153,736	161,16	-4,83%
out/20	151,395	154,18	-1,84%
nov/20	137,843	130,26	5,50%
dez/20	139,142	145,09	-4,27%
jan/21	141,063	140,11	0,68%
	<b>1742,31</b>	<b>1.778,32</b>	<b>-2,07%</b>

Fonte: Elaborado pelo autor (2021).

A média dos dados de irradiação solar obtidos de 2009 a 2019, apresentou uma diferença de 2,07% em relação aos dados de irradiação solar medidos pelo INMET para o período de aferição da geração do SFV. Percebe-se que comparando mês a mês, a rede LSTM apresenta melhores resultados quando comparados aos resultados da média.

Entende-se que como se tem a média de um mesmo mês, durante vários anos conforme Tabela 5.10, com o mês de fevereiro de 2009 a 2019 como exemplo, a tendência é que o resultado seja próximo ao ano de referência da geração. Percebe-se que a média da forma que foi utilizada, similar a LSTM, também pode ser usada para a aplicação assim com os resultados obtidos do CRESESB. A rede LSTM, em comparação as duas aplicações, mostrou-se com um percentual menor quando comparados ao global de radiação solar do ano de referência da geração.

Tabela 5.10 – Dados de irradiação solar comparativos MÉDIA x INMET

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Média
<b>Fev</b>	141	126	124	123	135	121	125	164	119	117	135	130

Fonte: Elaborado pelo autor (2021).

Na condição de utilização de um banco de dados de pelo menos 10 anos, observa-se que a média apresenta resultados mais próximos ao ano de referência da geração, mas para condições 3 a 5 anos a diferença não se torna atrativa para a análise.

#### 5.2.2.2 Aplicação da Rede neural LSTM no método proposto para M&V

Para a modelagem da solução de M&V proposta, são utilizados seis meses prévios de dados de geração extraídos dos inversores da usina e de 6 de dados medidos pelo INMET da estação Imperatriz. A Tabela 5.101 apresenta o modelo de regressão linear utilizado com os dados, para os seis meses prévios de acompanhamento, tanto de geração quanto de irradiação solar medida no mesmo período.

Tabela 5.11 – Regressão linear (geração x irradiação solar)

<b>Coeficiente angular da regressão (a)</b>	48,11 m <sup>2</sup> / mês
<b>Coeficiente linear da regressão (b)</b>	-2.955,94 kWh/mês
<b>Estatística t</b>	5,01
<b>Coeficiente de determinação (R<sup>2</sup>)</b>	86,27%
<b>Coeficiente de Variação do Erro Médio Quadrático (CV (EMQ))</b>	4,93%
<b>Incertezas</b>	10%

Fonte: Elaborado pelo autor (2021).

A partir dos coeficientes ‘a’ e ‘b’ da Tabela 5.10, a Equação 5.3 é utilizada para gerar os seis meses de geração restantes para completar o ano, considerando os dados gerados pelas redes neurais LSTM de cada mês, apresentados na Tabela 5.11.

$$\text{Geração} \left( \frac{kWh}{\text{mês}} \right) = -2.955,94 + 48,11 \times \text{IRRADIAÇÃO} \quad \text{Equação 5.3}$$

A Tabela 5.12 apresenta a complementação da geração do SFV de ago/2020 a jan/2021, projetados a partir da Equação 5.3 e com os dados de irradiação solar previstos pelas redes neurais LSTM de cada mês correspondente, somados aos meses monitorados da geração de a irradiação, completando um ano de geração anual global para o projeto. A diferença final no somatório anual da geração

considerando os seis meses projetados, em relação à geração anual real, é de 3,27%.

Tabela 5.12 – Método proposto com LSTM

Meses	6 meses da irradiação solar INMET (kWh/m <sup>2</sup> ) + 6 meses da irradiação solar LSTM (kWh/m <sup>2</sup> )	6 meses da energia gerada (kWh) + 6 meses da energia gerada (rede LSTM) (kWh)
fev/20	140,63	3.675,01
mar/20	145,50	3.903,67
abr/20	133,74	3.504,27
mai/20	149,52	4.190,25
jun/20	147,90	4.515,73
jul/20	162,49	4.801,86
<b>ago/20</b>	<b>161,46</b>	<b>4.811,95</b>
<b>set/20</b>	<b>161,87</b>	<b>4.831,68</b>
<b>out/20</b>	<b>154,51</b>	<b>4.477,59</b>
<b>nov/20</b>	<b>139,25</b>	<b>3.743,42</b>
<b>dez/20</b>	<b>147,93</b>	<b>4.161,02</b>
<b>jan/21</b>	<b>140,36</b>	<b>3.796,83</b>
<b>Soma</b>	<b>1.785,16</b>	<b>50.413,28</b>

Fonte: Elaborado pelo autor (2021).

### 5.2.2.3 Método proposto para M&V com a normalização da geração solar

Com o mesmo propósito do subcapítulo 5.2.1.2.3, para finalizar o cálculo da geração solar anual final, utilizou-se os dados de irradiação solar obtidos da rede LSTM para um ano, conforme Tabela 5.8, para a variável independente. Na validação da regressão linear, a partir dos dados da Tabela 5.13, os resultados obtidos foram para  $R^2$  de 93,2%, para o  $CV(EMQ)$  de 3,1% e a para a estatística  $t$  de 11,74, além disso, a incerteza do modelo de regressão ficou em 7%.

Tabela 5.13 – Regressão linear (geração x irradiação solar)

<b>Coeficiente angular da regressão (a)</b>	48,11 m <sup>2</sup> / mês
<b>Coeficiente linear da regressão (b)</b>	-2.955,94 kWh/mês
<b>Estatística t</b>	11,74
<b>Coeficiente de determinação (<math>R^2</math>)</b>	93,2%
<b>Coeficiente de variação do EMQ (<math>CV(EMQ)</math>)</b>	3,1%
<b>Incertezas</b>	7%

Fonte: Elaborado pelo autor (2021).

Os coeficientes angular e linear da Tabela 5.13 são os mesmos valores da Equação 5.3, onde a mesma equação será utilizada para normalizar a geração da tabela 5.14.

Tabela 5.14 – Modelagem da geração de energia “Normalizada” dados LSTM

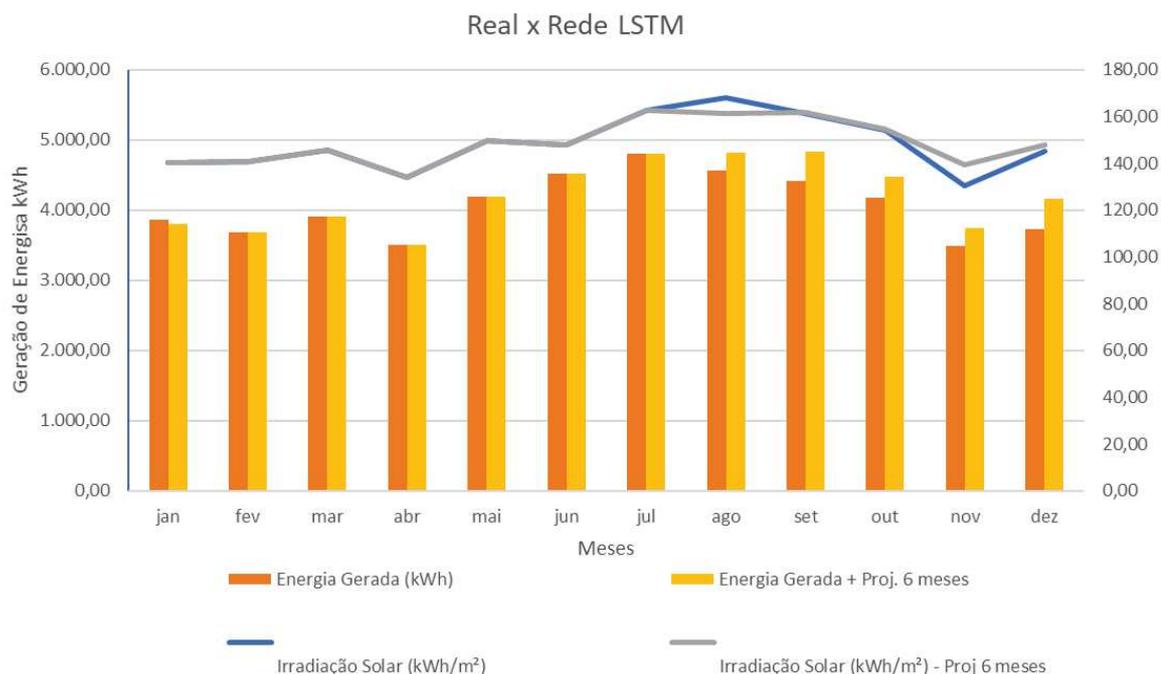
<b>Mês</b>	<b>Geração Solar (kWh) Normalizada</b>
Jan	3.796,83
Fev	3.866,11
Mar	4.059,03
abr	3.952,70
mai	4.272,64
jun	4.301,99
jul	4.663,29
ago	4.811,95
set	4.831,68
out	4.477,59
nov	3.743,42
dez	4.161,02
<b>Total</b>	<b>50.938,25</b>

Fonte: Elaborado pelo autor (2021).

Os resultados obtidos com geração anual normalizada para um ano, considerando a irradiação solar obtidas pela rede LSTM como uma condição padrão para um ano, em comparação com a geração obtida na Tabela 5.7, observa-se uma diferença de 1,23%.

O Gráfico 5.3 apresenta o comparativo entre geração e irradiação solar real versus a geração e irradiação solar extraída da metodologia proposta.

Gráfico 5.3 – Dados reais de “Geração e Irradiação” x Dados modelo proposto “Geração e Irradiação”



Fonte: Elaborado pelo autor (2021).

### 5.3 SFV NO ESTADO DE GOIÁS

O projeto está localizado na cidade de Goiânia, no estado de Goiás, sendo a unidade consumidora um prédio público de ensino superior, com regime de funcionamento com três turnos de aulas. A instalação do sistema aconteceu no telhado da instituição no ano de 2019.

A mesma sequência aplicada na análise do projeto em Araguaína – TO no subcapítulo 5.1 para M&V, foi aplicada para este projeto.

#### 5.3.1 M&V método tradicional

O projeto inicia-se com a definida a “Linha de Base”, consolidando o plano de M&V apresentado no subcapítulo 5.3.1.1. E com a implementação do projeto, inicia-se o acompanhamento e monitoramento da geração e irradiação solar por um período de doze meses, necessários para o fechamento da “determinação da economia”, no caso, da geração solar obtida.

### 5.3.1.1 Definição da Linha de Base

No período de definição da linha de base, para a simulação da geração solar mensal e anual, foi utilizado o software PVsyst, através dos dados extraídos do projeto e suas condições físicas, além dos dados de irradiação solar do CRESESB. Os dados do CRESESB da cidade de Goiânia - GO, envolvendo a localização geográfica da cidade e a irradiação solar diária média, encontram-se na Tabela 5.15, obtidos de [[http://www.cresesb.cepel.br/index.php#localidade\\_17365](http://www.cresesb.cepel.br/index.php#localidade_17365)].

Tabela 5.15 – Dados de Irradiação Solar no Plano Horizontal (Ano padrão)

Irradiação solar diária média [kWh/m <sup>2</sup> .dia]															
Latitude [°]	Longitude [°]	Distância [km]	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez	Média
16,701°S	49,249°O	2,4	5,5	5,5	5,1	5	4,8	4,6	4,8	5,7	5,6	5,6	5,4	5,5	5,25

Fonte: Elaborado pelo autor (2021).

A irradiação solar da Tabela 5.15 encontra-se em um formato de contabilização média diária mensal. Para a visão da irradiação mensal total multiplica-se a média diária pelo número de dias de cada mês, conforme Figura 5.6. na coluna “*GlobHor*”.

#### 5.3.1.1.1 Definição das Metas e Condições Físicas do Projeto

O SFV tem potência nominal instalada de 39,6 kWp. As Figuras 5.4 a 5.6 apresentam os resultados extraídos do software PVsyst, trazendo desde as características do sistema até os resultados esperados com a implementação do projeto.

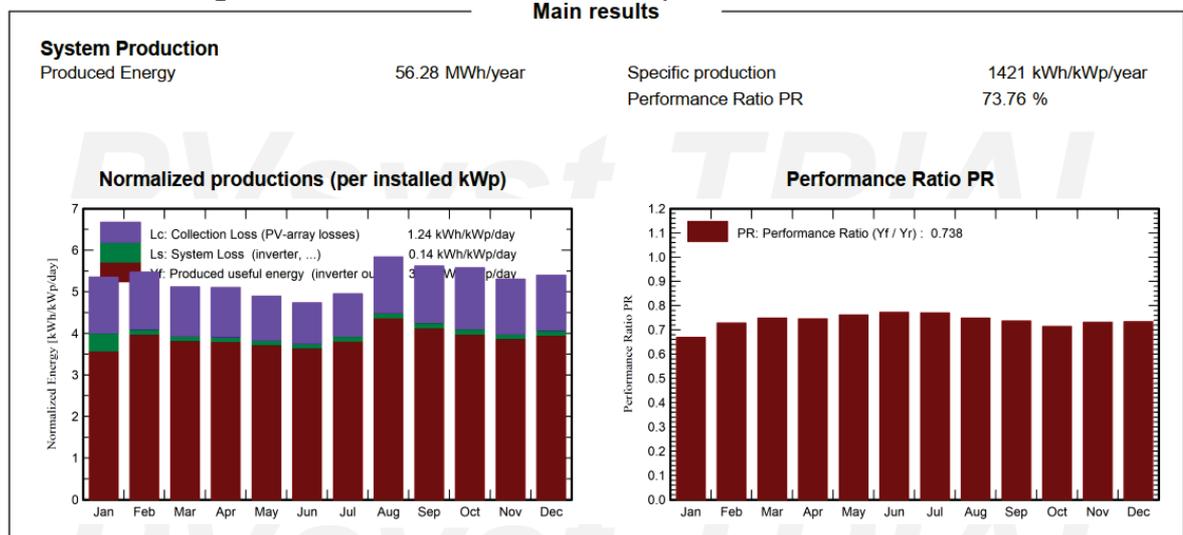
Figura 5.4 – Características do SFV.

PV Array Characteristics			
<b>PV module</b>		<b>Inverter</b>	
Manufacturer	Generic	Manufacturer	Generic
Model	GCL-P6/72H-330	Model	SUN2000-33KTL-A
(Original PVsyst database)		(Original PVsyst database)	
Unit Nom. Power	330 Wp	Unit Nom. Power	30.0 kWac
Number of PV modules	120 units	Number of inverters	4 * MPPT 25% 1 units
Nominal (STC)	39.6 kWp	Total power	30.0 kWac
Modules	6 Strings x 20 In series	Operating voltage	200-1000 V
<b>At operating cond. (50°C)</b>		Max. power (=>40°C)	33.0 kWac
Pmpp	35.8 kWp	Pnom ratio (DC:AC)	1.32
U mpp	678 V		
I mpp	53 A		
<b>Total PV power</b>		<b>Total inverter power</b>	
Nominal (STC)	40 kWp	Total power	30 kWac
Total	120 modules	Nb. of inverters	1 Unit
Module area	238 m <sup>2</sup>	Pnom ratio	1.32
Cell area	218 m <sup>2</sup>		

Fonte: Elaborado pelo autor no PVSyst (2021).

As Figuras 5.5 e 5.6 apresentam os resultados de desempenho para o SFV, juntamente com os resultados esperados de geração mensal e anual para o projeto. Assim, o SFV foi projetado para uma PR de 73,76%, uma irradiação solar global no plano horizontal por ano de 1914,8 kWh/m<sup>2</sup> e geração solar anual calculada de 56,28 MWh.

Figura 5.5 – Resultado de desempenho – Linha de Base



Fonte: Elaborado pelo autor no PVSyst (2021).

Figura 5.6 – Resultado de irradiação e geração – Linha de Base

Balances and main results								
	GlobHor	DiffHor	T_Amb	GlobInc	GlobEff	EArray	E_Grid	PR
	kWh/m <sup>2</sup>	kWh/m <sup>2</sup>	°C	kWh/m <sup>2</sup>	kWh/m <sup>2</sup>	MWh	MWh	ratio
January	169.0	82.03	25.15	166.0	154.4	4.923	4.399	0.669
February	154.6	74.84	24.57	153.3	143.1	4.554	4.419	0.728
March	158.7	79.67	24.43	158.5	147.7	4.840	4.698	0.748
April	150.6	55.59	24.12	153.1	143.1	4.659	4.521	0.746
May	147.3	48.53	23.23	151.7	141.8	4.712	4.573	0.761
June	136.8	40.24	21.61	141.9	132.2	4.470	4.340	0.772
July	148.2	42.39	20.88	153.5	143.2	4.826	4.686	0.771
August	176.1	45.44	23.90	180.9	169.3	5.524	5.366	0.749
September	167.1	71.03	27.37	168.5	157.5	5.059	4.913	0.736
October	173.9	70.31	26.75	173.1	161.0	5.045	4.895	0.714
November	161.7	83.11	25.49	159.1	146.9	4.744	4.607	0.731
December	170.8	86.06	24.80	167.4	154.9	5.007	4.861	0.733
Year	1914.8	779.21	24.35	1926.9	1795.0	58.363	56.280	0.738

**Legends**

GlobHor	Global horizontal irradiation	EArray	Effective energy at the output of the array
DiffHor	Horizontal diffuse irradiation	E_Grid	Energy injected into grid
T_Amb	Ambient Temperature	PR	Performance Ratio
GlobInc	Global incident in coll. plane		
GlobEff	Effective Global, corr. for IAM and shadings		

Fonte: Elaborado pelo autor no PVSyst (2021).

### 5.3.1.1.2 Variável Independente

No período de “Linha de Base”, foi identificada a irradiação solar como variável independente, conforme detalhado no subcapítulo 4.3. Essa indicação de variável independente foi definida aplicando-se a análise de regressão a partir dos dados de geração simulados (PVSyst) e de irradiação solar do CRESESB.

Seguindo os passos para a validação da regressão apresentados no subcapítulo 4.5, a Tabela 5.16 apresenta os resultados obtidos, onde  $R^2$  é menor do que 0,75, o  $CV(EMQ)$  é menor do que 5% e a estatística  $t$  é maior que 2. Mesmo que um dos resultados não esteja atingindo um valor esperado, no caso  $R^2$ , como os demais atingiram, a indicação é seguir com a sugestão de validação da regressão para o “período de determinação da economia”.

Tabela 5.16 – Regressão linear (geração estimada PVSyst x irradiação solar)

<b>Coeficiente angular da regressão (a)</b>	16,33 m <sup>2</sup> / mês
<b>Coeficiente linear da regressão (b)</b>	2083,72 kWh/mês
<b>Estatística t</b>	3,08
<b>Coeficiente de determinação (R<sup>2</sup>)</b>	49%
<b>Coeficiente de Variação do Erro Médio Quadrático (CV (EMQ))</b>	4,60%
<b>Incertezas</b>	9,02%

Fonte: Elaborado pelo autor (2021).

Quanto as incertezas, ou seja, a determinação dos níveis desejados de precisão e confiança para a estimativa a ser determinada. Então, a incerteza encontrada e mostrada na Tabela 5.16 foi de 9,02% atingindo a meta de 10% de precisão a 95% de confiabilidade, quer-se dizer que a faixa definida por  $\pm 10\%$  em torno da média conterá o valor verdadeiro (que não se conhece) com a probabilidade de 95%, ou seja, em 95% dos casos.

A Equação 5.4, obtida através dos coeficientes  $a$  e  $b$  apresentados na tabela 5.14, indica que a variação da energia gerada é causada por mudanças na irradiação solar.

$$\text{Geração} \left( \frac{kWh}{\text{mês}} \right) = 2083,72 + 16,33 \times \text{IRRADIAÇÃO} \quad \text{Equação 5.4}$$

#### 5.3.1.1.3 Fronteira de Medição e Efeitos Iterativos com o Projeto

Como fronteira de medição para os períodos da linha de base, considerou-se a entrada de energia da concessionária para este sistema (prédio), levando em consideração as características dos locais de instalação, assim como a irradiação solar mês. Para o período de determinação da economia, além do monitoramento através dos medidores da concessionária de energia destes prédios, foi considerado a utilização de medições extraídas dos inversores do sistema e as medições do sistema de monitoramento instalados juntos aos inversores de frequência (com memória de massa e monitorados).

Sobre os medidores de energia instalados juntos aos inversores do SFV, são os equipamentos CCK4400ME (CCK Automação), com 4 quadrantes e com memória de massa.

Os possíveis efeitos interativos com o sistema de geração fotovoltaica que poderiam surgir ao longo do projeto, foram ignorados para a avaliação dos benefícios. Todavia, os efeitos gerados com o projeto foram monitorados após a implementação do projeto.

#### 5.3.1.1.4 Opção do PIMVP

A determinação das economias foi feita pelo prazo de cálculo do benefício do projeto, até completar um ano adotando a opção B – Medição isolada de todos os parâmetros (Tabela 2.1).

#### 5.3.1.1.5 Fatores Estáticos

Possíveis efeitos estáticos foram avaliados e monitorados durante as medições, e desde o plano foram estabelecidos como possibilidade de impacto, tais como: efeito de sombreamento nos painéis causados após a implementação do projeto como árvores ou construções não existentes e ou ampliações; mudanças significativas no perfil de utilização da instalação do projeto. Outro ponto importante, que foi observado é que como o sistema de geração de energia fotovoltaico é sincronizado com a rede da concessionária, a mesma deveria manter o fornecimento de energia, pois, caso contrário, não haveria geração. Este fato ocorre devido aos inversores do SFV operarem somente com o sistema energizado.

#### 5.3.1.2 Período de Determinação da Economia e Condições para o ajuste de base

O período de determinação da geração iniciou-se logo após a instalação do SFV, de setembro de 2020 a agosto de 2021, com um tempo total de um ano de monitoramento, com memória de massa da energia injetada na rede.

O benefício final para o projeto, a geração de energia, teve como base um ano padrão, no caso foi utilizado os dados de irradiação solar do CRESESB em relação à variável independente.

##### **5.3.1.2.1** Dados dos resultados durante o período de monitoramento do projeto:

Os dados de geração extraídos do SFV, bem como de irradiação solar do INMET correspondentes ao período de geração real, seguem na Tabela 5.17. Os resultados obtidos frente ao projeto, em comparação aos dados estimados na linha de base no subcapítulo 5.3.1.1 na Figura 5.6, foram para a geração solar de 10,21%, e para a irradiação solar de 4,39%.

Tabela 5.17– Registro da geração e irradiação

Ano	Mês	Dados Medidos	
		kWh	Irradiação Solar (kWh/m <sup>2</sup> . mês)
2020	set	4.594,82	168,14
2020	out	4.039,13	156,21
2020	nov	4.761,57	176,44
2020	dez	4.648,02	170,43
2021	jan	4.714,59	166,07
2021	fev	3.859,75	134,56
2021	mar	4.400,10	144,80
2021	abr	4.013,19	150,42
2021	mai	3.918,85	138,88
2021	jun	3.590,09	124,20
2021	jul	3.924,16	141,98
2021	ago	4.069,44	158,72
<b>Total</b>		<b>50.533,71</b>	<b>1.830,84</b>

Fonte: Elaborado pelo autor (2021).

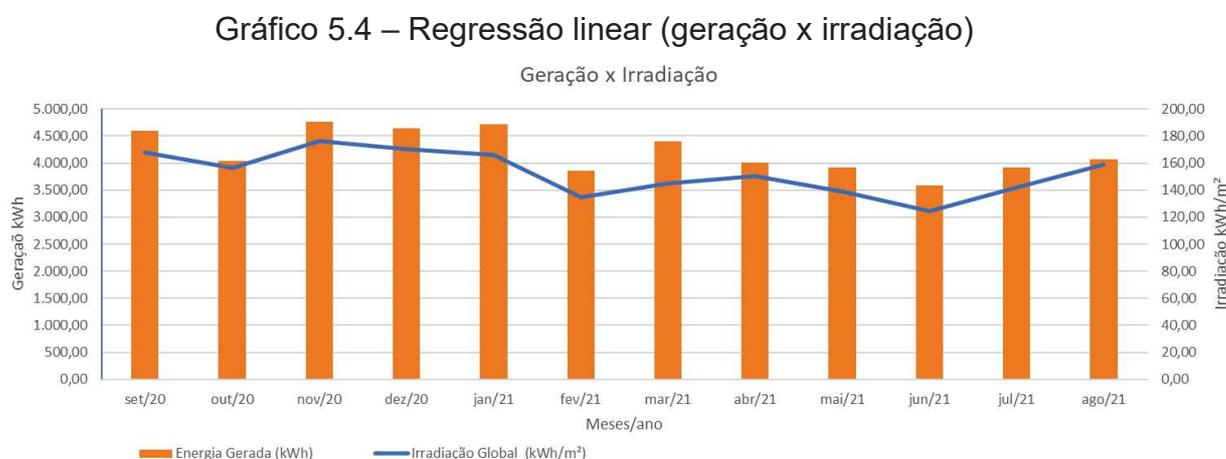
Para a irradiação solar medida, a base de dados utilizada foi extraída da base de dados do INMET da estação A225, conforme Tabela 5.18.

Tabela 5.18– Dados da Estação Meteorológica

REGIÃO:	CO
UF:	GO
ESTAÇÃO:	GOIANIA
CODIGO (WMO):	A002
LATITUDE:	-16,64277777
LONGITUDE:	-49,21999999
ALTITUDE:	770
DATA DE FUNDAÇÃO (YYYY-MM-DD):	29/05/2001

Fonte: Elaborado pelo autor (2021).

O Gráfico 5.4 é o representativo dos dados da Tabela 5.17, com a irradiação e geração de acordo com os meses e anos.



Fonte: Elaborado pelo autor (2021).

### 5.3.1.2.2 Variável Independente

Na construção da linha de base o projeto apresenta a possibilidade de uma variável independente, com a validação da regressão linear entre a geração simulada e a irradiação solar obtida pelo CRESESB, conforme subcapítulo 5.3.1.1.2. No período de determinação dos resultados do projeto, de posse dos dados reais de medição presentes na tabela 5.16, realiza-se a validação da regressão linear para a consolidação das condições da correlação entre os dados de geração medidos e de irradiação solar registrados pela base do INMET.

A Tabela 5.19 apresenta os resultados obtidos para a validação, em que  $R^2$  é maior do que 75% com 80,7%, o  $CV(EMQ)$  é menor do que 5% com 4,98% e a estatística  $t$  é maior que 2 com 6,46.

Tabela 5.19 – Regressão linear (geração medida x irradiação solar INMET).

<b>Coefficiente angular da regressão (<math>a</math>)</b>	21,75 m <sup>2</sup> / mês
<b>Coefficiente linear da regressão (<math>b</math>)</b>	892,93 kWh/mês
<b>Estatística <math>t</math></b>	6,27
<b>Coefficiente de determinação (<math>R^2</math>)</b>	80%
<b>Coefficiente de Variação do Erro Médio Quadrático (CV (EMQ))</b>	4,41%
<b>Incertezas</b>	9,82%

Fonte: Elaborado pelo autor (2021).

Com os resultados da Tabela 5.19, é possível construir a Equação 5.5, onde em *IRRADIAÇÃO* pode ser imputado valores de condições padrões de irradiação solar mensal, tendo como intuito trazer as condições de geração de energia no somatório de um ano o mais perto de uma condição aproximado a uma média dos outros ano, e não somente do ano em que foi feito o monitoramento da geração após a conexão do SFV.

$$Geração \left( \frac{kWh}{mês} \right) = 892,93 + 21,75 \times IRRADIAÇÃO \quad \text{Equação 5.5}$$

Com a Equação 5.5 é possível partir para a próxima e última etapa de aferição final dos resultados para o projetos no modelo tradicional apresentado no subcapitulo 5.3.1.2.3, trazendo a geração de cada mês compatível as variações da irradiação solar dos meses de um ano padrão, onde utiliza-se nesta análise os dados extraídos do CRESESB, da mesma forma que foi utilizado na linha de base para construir as simulações, colocando tanto o período de linha de base quanto o período de determinação da geração final do SFV nas mesma condições.

#### 5.3.1.2.3 Geração de energia do SFV “Normalizada”

Conforme aplicado no caso anterior no subcapitulo 5.2.1.2.3, o ano padrão para as condições de irradiação solar foi baseado nos dados extraído no CRESESB para esta análise.

A Tabela 5.20 mostra a análise final dos resultados obtidos com o projeto, onde a energia gerada anual normalizada foi de 52.359,68 MWh/ano. Comparando a geração anual normalizada com a geração real obtida na Tabela 5.17, a diferença encontrada foi de 3,49%. Na comparação com as condições da linha de base da Figura 5.6 a diferença encontrada na geração anual é de 6,97%.

Tabela 5.20 – Determinação dos Resultados de Geração “Normalizada”

Meses/Ano	Geração Solar (kWh) Normalizada
set/20	4.568,47
out/20	4.255,29
nov/20	4.344,46
dez/20	4.168,29
jan/21	4.096,52
fev/21	3.868,16
mar/21	4.116,10
abr/21	4.722,89
mai/21	4.527,15
jun/21	4.675,04
jul/21	4.409,70
ago/21	4.607,62
<b>Soma</b>	<b>52.359,68</b>

Fonte: Elaborado pelo autor (2021).

### 5.3.2 Rede neural LSTM para o Projeto em Goiânia

O mesmo procedimento adotado no subcapítulo 5.2.2. foi aplicado para este caso de projeto em Goiânia. Através dos seis meses de monitoramento da geração do SFV e irradiação solar medida pela INMET (set/2020 a fev/2021), foi possível projetar os demais seis meses de geração solar (mar/2021 a ago/2021) do SFV através da equação de regressão linear obtida com os dados reais do monitoramento, e dos dados de irradiação solar obtidos da rede neural LSTM coincidindo com este mesmo período, conforme subcapítulo 5.3.2.2.

A decisão da utilização dos 6 meses de medição dos SFV e do INMET foi tomada a partir dos resultados encontrados da precisão e dos indicadores estáticos para a validação da regressão, conforme subcapítulo 4.3.

Para a verificação da assertividade dos resultados obtidos com a rede neural LSTM em comparação aos dados reais medidos, também foi projetado os meses do período de monitoramento, completando um ano de irradiação solar, ou seja, foi gerado uma rede neural LSTM para cada mês desde mar/2021 a ago/2021, apresentados no subcapítulo 5.3.2.1. Além disso, este um ano produzido pela rede LSTM de irradiação solar foi utilizado na aplicação da normalização da geração solar anual final para o projeto, considerando que este ano seria um ano padrão por ter

sido criado com base em dados históricos extraídos do INMET, mostrado no subcapítulo 5.3.2.3.

#### 5.3.2.1 Resultados de irradiação solar projetados pela rede neural LSTM

Um ano completo foi projetado pela rede LSTM, a fim de comparar os resultados de irradiação solar mensal e anual com os dados extraídos do INMET no mesmo período, assim como apresentado na Tabela 5.21. Os dados utilizados do INMET da estação de Goiânia iniciam-se de 2001 e terminam em 2021.

Tabela 5.21 – Dados de irradiação solar comparativos LSTM x INMET

Mês	RSME (kWh/m <sup>2</sup> )	Irradiação Solar		Diferença (INMET x LSTM)
		INMET (kWh/m <sup>2</sup> )	LSTM (kWh/m <sup>2</sup> )	
set/20	1,50	168,14	168,12	0,01%
out/20	1,50	156,21	157,06	-0,54%
nov/20	1,60	176,44	168,13	4,71%
dez/20	1,58	170,43	163,66	3,97%
jan/21	1,29	166,07	167,21	-0,69%
fev/21	1,53	134,56	155,83	-15,81%
mar/21	1,18	144,80	152,42	-5,26%
abr/21	1,21	150,42	145,82	3,06%
mai/21	1,03	138,88	138,87	0,01%
jun/21	0,29	124,20	134,23	-8,08%
jul/21	0,41	141,98	140,99	0,70%
ago/21	0,39	158,72	141,85	10,63%
	<b>1,13</b>	<b>1.830,84</b>	<b>1.834,19</b>	<b>-0,18%</b>

Fonte: Elaborado pelo autor (2021).

Os dados de irradiação solar projetados pela rede LSTM no somatório anual em comparação aos dados do INMET, apresentaram uma diferença de 0,18%. Os resultados do RMSE aplicado na análise, que comparam os valores medidos pelo INMET (medidos) com as previsões da irradiação solar da LSTM, foram, em média, de 1,13 kWh/m<sup>2</sup>.

Para fins de análise comparativa quanto ao número de dados que seriam utilizados da base histórica de cada mês e ano do INMET, outros intervalos de tempo foram imputados no modelo, mostrados na Tabela 5.22, compreendidos de 2016 a 2019 para treinamento, para prever de set/2020 a ago/2021.

Tabela 5.22 – Dados de irradiação solar comparativos LSTM x INMET com 4 anos

Mês	RSME (kWh/m <sup>2</sup> )	Irradiação Solar		Diferença (INMET x LSTM)
		INMET (kWh/m <sup>2</sup> )	LSTM (kWh/m <sup>2</sup> )	
set/20	1,25	168,14	168,92	0,47%
out/20	1,61	156,21	169,02	8,20%
nov/20	1,53	176,44	149,62	-15,20%
dez/20	1,38	170,43	158,32	-7,10%
jan/21	1,22	166,07	179,28	7,95%
fev/21	1,62	134,56	156,73	16,48%
mar/21	1,06	144,80	157,26	8,60%
abr/21	1,22	150,42	143,45	-4,63%
mai/21	1,05	138,88	141,61	1,97%
jun/21	0,25	124,20	135,23	8,88%
jul/21	0,42	141,98	146,92	3,48%
ago/21	0,62	158,72	154,64	-2,57%
	<b>1,10</b>	<b>1.830,84</b>	<b>1.861,00</b>	<b>-1,65%</b>

Na primeira análise, na Tabela 5.21, observa-se que a diferença encontrada é menor que diferença encontrada na Tabela 5.22 considerando 3 anos de dados como base para a previsão 2020 a 2021.

Em análises para dimensionamento de SFV, base de dados para previsões são utilizados como o caso do CRESESB que trabalha com 17 anos de bases históricas (PEREIRA, 2017).

Para os projetos de eficiência energética do PROPEE, concessionárias estipulam premissas similares para a normalização dos dados e também previsões nos dimensionamentos dos sistemas, onde para fontes incentivadas que possuem como variáveis independentes fatores climáticos (temperatura, irradiação, velocidade de ventos, entre outras), deverão ser usados a média dos últimos 10 anos destes fatores para definição das condições padrão em um ano típico da estação mais próxima da unidade consumidora onde o sistema será instalado (CELESC, 2021).

No projeto apresentado neste subcapítulo utilizou-se os resultados na M&V da Tabela 5.21, seguindo as premissas que são direcionadas pelas concessionárias de energia.

### 5.3.2.2 Aplicação da Rede neural LSTM no método tradicional de M&V

Para a modelagem da solução de M&V proposta, são utilizados seis meses prévios de dados de geração extraídos dos inversores da usina e de 6 de dados

medidos pelo INMET da estação Goiânia. A Tabela 5.22 apresenta o modelo de regressão linear utilizado com os dados, para os seis meses prévios de acompanhamento, tanto de geração quanto de irradiação solar medida no mesmo período.

Tabela 5.23 – Regressão linear (geração x irradiação solar)

<b>Coeficiente angular da regressão (a)</b>	23,84 m <sup>2</sup> / mês
<b>Coeficiente linear da regressão (b)</b>	574,802 kWh/mês
<b>Estatística t</b>	4,887
<b>Coeficiente de determinação (R<sup>2</sup>)</b>	85,65%
<b>Coeficiente de Variação do Erro Médio Quadrático (CV (EMQ))</b>	3,68%
<b>Incertezas</b>	9,5%

Fonte: Elaborado pelo autor (2021).

A partir dos coeficientes ‘a’ e ‘b’ da Tabela 5.22, a Equação 5.6 é utilizada para gerar os seis meses de geração restantes para completar o ano, considerando os dados gerados pelas redes neurais LSTM de cada mês, apresentados na Tabela 5.10.

$$Geração \left( \frac{kWh}{mês} \right) = 574,8 + 23,84 \times IRRADIAÇÃO \quad \text{Equação 5.6}$$

A Tabela 5.23 apresenta a complementação da geração do SFV de mar/2021 a ago/2021, projetados a partir da Equação 5.6 e com os dados de irradiação solar previstos pelas redes neurais LSTM de cada mês correspondente, somados aos meses monitorados da geração e da irradiação, completando um ano de geração anual global para o projeto. A diferença final no somatório anual da geração considerando os seis meses projetados, em relação à geração anual real, é de 0,20%.

Tabela 5.24 – Método proposto com LSTM

Meses	6 meses da irradiação solar INMET real (kWh/m <sup>2</sup> ) + 6 meses da irradiação solar LSTM (kWh/m <sup>2</sup> )	6 meses da energia gerada real (kWh) + 6 meses da energia gerada com a irradiação solar gerada pela rede LSTM (kWh)
set/20	168,12	4.594,82
out/20	157,06	4.039,13
nov/20	168,13	4.761,57
dez/20	163,66	4.648,02
jan/21	167,21	4.714,59
fev/21	155,83	3.859,75
<b>mar/21</b>	<b>152,42</b>	<b>4.208,54</b>
<b>abr/21</b>	<b>145,82</b>	<b>4.051,20</b>
<b>mai/21</b>	<b>138,87</b>	<b>3.885,51</b>
<b>jun/21</b>	<b>134,23</b>	<b>3.774,89</b>
<b>jul/21</b>	<b>140,99</b>	<b>3.936,05</b>
<b>ago/21</b>	<b>141,85</b>	<b>3.956,55</b>
<b>Soma</b>	<b>1.834,19</b>	<b>50.430,62</b>

Fonte: Elaborado pelo autor (2021).

### 5.3.2.3 Método proposto para M&V com a normalização da geração solar

Na validação da regressão linear, a partir dos dados da Tabela 5.24, os resultados obtidos foram para  $R^2$  de 92,93%, para o  $CV(EMQ)$  de 2,46% e a para a estatística  $t$  de 11,47, além disso a incerteza do modelo de regressão ficou em 5,48% abaixo de 10%.

Tabela 5.25 – Regressão linear (geração x irradiação solar)

<b>Coefficiente angular da regressão (a)</b>	23,84 m <sup>2</sup> / mês
<b>Coefficiente linear da regressão (b)</b>	574,802 kWh/mês
<b>Estatística t</b>	11,47
<b>Coefficiente de determinação (<math>R^2</math>)</b>	92,93%
<b>Coefficiente de variação do EMQ (<math>CV(EMQ)</math>)</b>	2,46%
<b>Incertezas</b>	5,5%

Fonte: Elaborado pelo autor (2021).

Os coeficientes angular e linear da Tabela 5.24 são os mesmos valores da Equação 5.6, onde a mesma equação será utilizada para normalizar a geração da tabela 5.25.

Tabela 5.26 – Modelagem da geração de energia “Normalizada” dados LSTM

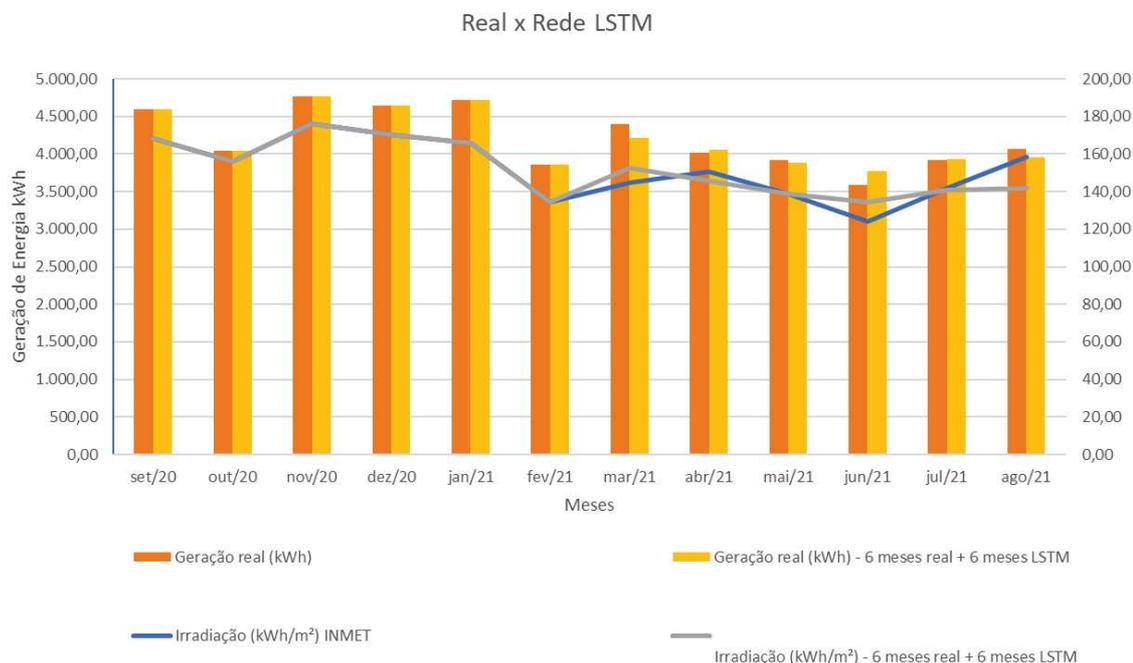
<b>Mês</b>	<b>Geração Solar (kWh) Normalizada</b>
set	4.582,84
out	4.319,16
nov	4.583,08
dez	4.476,51
jan	4.561,14
fev	4.289,84
mar	4.208,54
abr	4.051,20
mai	3.885,51
jun	3.774,89
jul	3.936,05
ago	3.956,55
<b>Total</b>	<b>50.625,31</b>

Fonte: Elaborado pelo autor (2021).

Os resultados obtidos com geração anual normalizada para um ano, considerando a irradiação solar obtidas pela rede LSTM como uma condição padrão para um ano, em comparação com a geração obtida na Tabela 5.18, observa-se uma diferença de 3,4%.

O Gráfico 5.4 apresenta o comparativo entre geração e irradiação solar real versus a geração e irradiação solar extraída da metodologia proposta.

Gráfico 5.5 – Dados reais de “Geração e Irradiação” x Dados modelo proposto “Geração e Irradiação”



Fonte: Elaborado pelo autor (2021).

#### 5.4 CONCLUSÕES PARCIAIS

O capítulo 5 teve como objetivo apresentar os testes e resultados para a aplicação de redes neurais LSTM para a previsão da variável independente irradiação solar em processos de realização da M&V em projetos de EE com sistemas fotovoltaicos segundo PROPEE da ANEEL.

O estudo mostrou resultados satisfatórios tanto para o trabalho com os doze meses de geração “normalizada” utilizando a rede neural LSTM, quanto para a redução do tempo de acompanhamento do projeto em seis meses trazendo maior assertividade quando comparados as estimativas extraídas do CRESESB.

Outra consideração importante, é quanto aos dados utilizados na previsão com a rede LSTM, extraídos do INMET medidos de hora em hora considerando desde o período anterior e durante a medição da geração, de forma atualizada na comparação ao CRESESB.

Além disso, para o desenvolvido do código da rede neural LSTM foi utilizada a linguagem de programação no Python, sendo este um programa gratuito e intuitivo, o que facilita a aplicação.

## 6 CONCLUSÕES

### 6.1 CONSIDERAÇÕES FINAIS

A presente dissertação de mestrado apresentou uma proposta de metodologia com aplicação de redes neurais LSTM para a previsão da variável independente, sendo a mesma usada para explicar a variação da energia gerada pelo SFV, com o intuito de dar mais velocidade ao processo de análise e assertividade quando comparado aos dados reais extraídos de um projeto real acompanhados por um período de 12 meses.

Portanto, a partir das análises realizadas no acompanhamento do projeto por doze meses, e na utilização dos dados extraídos das redes neurais LSTM substituindo os dados do CRESESB “estimados” no modelo de contabilização de geração anual “normalizada”, observou-se o potencial de aplicação de uma nova forma de modelagem da geração do SFV anual com redução do tempo de acompanhamento dos resultados de projetos deste porte. Com isso, foi proposta a redução do tempo de acompanhamento dos resultados no período de determinação da economia de doze meses para seis meses, utilizando regressão linear para gerar a equação da geração mensal do sistema com: dados de medição da geração real de seis meses; dados de irradiação solar projetados pela rede neural LSTM, e dados medidos obtidos do banco de dados histórico do INMET correspondente a cada base solarimétrica utilizada.

A avaliação do método aplicado ao problema de M&V para projetos com SFV segundo a ANEEL, apresentou resultados satisfatórios, identificando uma solução alternativa, trazendo a aplicação de redes neurais LSTM para a previsão da variável independente irradiação solar, impactando de forma assertiva na modelagem da geração anual final quando se comparado a dados reais medidos de um ano de dois casos reais utilizados na análise. Portanto, a modelagem proposta demonstrou eficácia na identificação de um plano alternativo a M&V. Destaca-se que a principal contribuição foi aliar aplicação de DL em processos de M&V de projetos de EE, usando redes neurais LSTM para a previsão de variáveis independentes.

Finalmente, pode-se listar como principais contribuições do trabalho:

- Aplicar redes neurais LSTM para previsão de series temporais, de forma mais assertiva, presente em processos de M&V, para a previsão da variável independente com a irradiação solar;

- Devido a assertividade no processo de previsão da irradiação solar com as redes neurais LSTM, dar credibilidade na exploração de novas formas de se fazer M&V em projetos com sistemas fotovoltaicos, com menos tempo de acompanhamento;
- Com a aplicação de redes neurais LSTM em projetos reais com processos de M&V, dar velocidade na análise de dados históricos e projeções, impactando em resultados seguros aos projetos frente a linha de base estabelecida inicialmente;
- Com as redes neurais LSTM, foi possível explorar a parte do PROPEE para M&V em projetos de EE com geração fotovoltaica, onde diz que para casos que haja dados locais sobre a disponibilidade da fonte utilizada o tempo de acompanhamento dos projetos poderá ser reduzido.

## 6.2 TRABALHOS FUTUROS

Seguindo a linha de pesquisa desenvolvida nesta dissertação, e tendo em vista os resultados obtidos, algumas sugestões são promissoras para trabalhos futuros:

- Incluir na metodologia proposta mais variáveis ambientais como variáveis independentes (temperatura ambiente, temperatura do modulo, umidade), com a utilização da rede neural LSTM;
- Trabalhar com redes neurais na aplicação de M&V em projetos de EE, sem aplicação de fonte incentivada, para prever mais outros tipos de variáveis independentes, tais como: taxa de ocupação, temperatura, vazão, pressão e produtividade;
- Trabalhar com redes neurais em mapeamento de perfis de cargas elétricas, processos, para que com os resultados, quanto ao comportamento, facilitar a aplicação de automação para controle de funcionamento, impactando resultados atrativos em projetos de EE;
- Trabalhar com todas as estações solarimétricas do INMET (diversas entradas) na rede neural, no mesmo contexto do trabalho apresentado.

## REFERÊNCIAS

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – **Guia de M&V**. Brasília – DF: ANEEL, 2014.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - Procedimentos do Programa de Eficiência Energética – PROPEE. 10 Módulos. Brasília – DF: ANEEL, 2020.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. Resolução Normativa nº 482, de 17 de abril de 2012. Estabelece as condições gerais para o acesso de microgeração e minigeração distribuída aos sistemas de distribuição de energia elétrica, o sistema de compensação de energia elétrica, e dá outras providências. Brasília-DF: ANEEL, 2012.

Book, Deep Learning. O Efeito do Batch Size no Treinamento de Redes Neurais Artificiais. Capítulo 37. Site: <https://www.deeplearningbook.com.br/o-efeito-do-batch-size-no-treinamento-de-redes-neurais-artificiais/>. Acessado em: 21 julho 2021.

BRAHMA, Banalaxmi; WADHVANI, Rajesh. Solar Irradiance Forecasting Based on Deep Learning Methodologies and Multi-Site Data. **Symmetry**, v. 12, n. 11, p. 1830, 2020.

CARLOS, Robert. Normalizando dados com R e Python: Transformando dados com Min-Max Normalization, 2020 Medium. Disponível em: <<https://datalivre.medium.com/normalizando-dados-com-r-e-python-fc048dd78ede#:~:text=A%20normaliza%C3%A7%C3%A3o%20de%20dados%20Min,0%20para%20o%20maior%20valor>> Acesso em: 01 set. 2021.

CELESC – Centrais Elétricas de Santa Catarina S. A.. Chamada Pública PEE Celesc nº 1/2021. Site: <http://site.celesc.com.br/peecelesc/index.php/noticias-site/274-chamada-publica-2021-pee-celesc>. Acessado em: 18 abril 2022.

CHAPLIN, Charles. Vida e Pensamentos, Martin Claret, p. 118,1997.

Coutinho, Thiago. Conheça 11 bibliotecas Python aplicadas na Ciência de Dados. Voitto. Site: <https://www.voitto.com.br/blog/artigo/principais-bibliotecas-python>. Acessado em: 25 junho 2020.

DA SILVA, T. Vieira et al. Performance analysis of neural network training algorithms and support vector machine for power generation forecast of photovoltaic panel. IEEE Latin America Transactions, v. 15, n. 6, p. 1091-1100, 2017.

DE MELO, Gedson Santos. Aplicação de Aprendizado de Máquina para Previsão de Fluxo de Caixa em ATMs. Universidade Federal de Pernambuco Centro de Informática, 2018.

DO NASCIMENTO, Pedro Paulo Marques. Applications of deep learning techniques on NILM. **Diss. Universidade Federal do Rio de Janeiro**, 2016.

EVO - EFFICIENCY VALUATION ORGANIZATION. Conceitos Básicos - Protocolo Internacional de Medição e Verificação de Performance - EVO 10000 - 1:2016 (Br). EVO, setembro 2017.

EVO – EFFICIENCY VALUATION ORGANIZATION. Protocolo Internacional de Medição e Verificação de Performance – Conceitos e Opções para a Determinação de Economias de Energia e de Água - vol. 1 - EVO 10000 – 1:2012 (Br). Sofia: EVO, 2012.

EVO – EFFICIENCY VALUATION ORGANIZATION. Protocolo Internacional de Medição e Verificação de Performance – Estatística e Incerteza para o PMIMVP - EVO 10100 - 1:2014 (Br). EVO, junho de 2014.

EVO - EFFICIENCY VALUATION ORGANIZATION. Renewables Applications Guide – International Performance Measurement and Verification Protocol - EVO 10200 - 1:2017. EVO, março 2017.

GALLAGHER, Colm V. et al. The suitability of machine learning to minimise uncertainty in the measurement and verification of energy savings. *Energy and Buildings*, v. 158, p. 647-655, 2018.

GERS, Felix A.; SCHMIDHUBER, Jürgen; CUMMINS, Fred. Learning to forget: Continual prediction with LSTM. *Neural computation*, v. 12, n. 10, p. 2451-2471, 2000.

GÖKGÖZ, Fazıl; FILIZ, Fahrettin. Deep Learning for Renewable Power Forecasting: An Approach Using LSTM Neural Networks. *International Journal of Energy and Power Engineering*, v. 12, n. 6, p. 416-420, 2018.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Memória longa de curto prazo. *Computação neural*, v. 9, n. 8, pág. 1735-1780, 1997.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. *nature*, v. 521, n. 7553, p. 436-444, 2015.

MALAKAR, Sourav et al. Designing a long short-term network for short-term forecasting of global horizontal irradiance. *SN Applied Sciences*, v. 3, n. 4, p. 1-15, 2021.

MARINO, Daniel L.; AMARASINGHE, Kasun; MANIC, Milos. Building energy load forecasting using deep neural networks. In: *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2016. p. 7046-7051.

MEADOR, W. E.; WEAVER, W. R. Two-stream approximations to radiative transfer in planetary atmospheres: A unified description of existing methods and a new improvement. *Journal of Atmospheric Sciences*, v. 37, n. 3, p. 630-643, 1980.

MUHAMMAD, Aslam et al. Deep learning application in power system with a case study on solar irradiation forecasting. In: *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIC)*. IEEE, 2019. p. 275-279.

NORMAS, A. B. N. T. Plano de Amostragem e Procedimentos na Inspeção por atributos-03.011-NBR 5426-Jan/1985 Portaria n ° 1004 de 11.12. 98-Regulamento Técnico Atribuição de Função de Aditivos. Aditivos e seus Limites Máximo de uso para a Categoria.

NORMAS, I.E.C. Photovoltaic System Performance Monitoring - Guidelines For Measurement, Data Exchange And Analysis. IEC 61724 Ed. 1.0 b:1998.

OLAH, Christopher. Understanding lstm networks. 2015.

OLATOMIWA, Lanre et al. A support vector machine–firefly algorithm-based model for global solar radiation prediction. *Solar Energy*, v. 115, p. 632-644, 2015.

OUAMMI, Ahmed et al. Artificial neural network analysis of Moroccan solar potential. **Renewable and Sustainable Energy Reviews**, v. 16, n. 7, p. 4876-4889, 2012.

PANDAS, Guia de usuário “Pandas.dataframe.resample”, 2008-2021 Equipe de desenvolvimento do pandas. Disponível em: <<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.resample.html>>. Acesso em: 01 set. 2021.

PEREIRA, Enio Bueno et al. Atlas brasileiro de energia solar. **São José dos campos: Inpe**, v. 1, 2017.

PNE - Plano Nacional de Energia 2050 / Ministério de Minas e Energia. Empresa de Pesquisa Energética. Brasília: MME/EPE, 2020.

RIBEIRO, Athyrson Machado et al. Um Estudo Comparativo Entre Cinco Métodos de Otimização Aplicados Em Uma RNC Voltada ao Diagnóstico do Glaucoma. *Revista de Sistemas e Computação-RSC*, v. 10, n. 1, 2020.

Singhal, Gaurav. Introdução às Unidades LSTM em RNN. Pluralsight, 9 de sete,brp de 2020. Site: <https://www.pluralsight.com/guides/introduction-to-lstm-units-in-rnn>. Acessado em: 03 de abril 2022.

SGANZERLA, Lucas Marino Bianchessi. Análise dos índices de mérito de sistemas fotovoltaicos conectados à rede em Curitiba. 2018.

SORKUN, Murat Cihan; İNCEL, Özlem Durmaz; PAOLI, Christophe. Time series forecasting on multivariate solar radiation data using deep learning (LSTM). **Turkish Journal of Electrical Engineering & Computer Sciences**, v. 28, n. 1, p. 211-223, 2020.

TOLMASQUIM, M. T. Novo Modelo do Setor Elétrico Brasileiro. Rio de Janeiro: Synergia; EPE: Brasília, 2011.

TORNAL, Kalman et al. Recurrent neural network based user classification for smart grids. In: **2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)**. IEEE, 2017. p. 1-5.

VASSALLI, Lucas Coutinho. Aplicação de Redes Neurais LSTM para a Previsão de Curto Prazo de Vazão do Rio Paraíba do Sul. Universidade Federal de Juiz de Fora, 2018, Curso de Graduação de Engenharia da Computação.

WANG, Xin et al. Predicting polarities of tweets by composing word embeddings with long short-term memory. In: **Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)**. 2015. p. 1343-1353.

WEBER, Eliseu Jose. Estimativa e mapeamento da irradiação solar incidente em superfícies com topografia heterogênea na zona de produção vitivinícola Serra Gaúcha. 2011.

## APÊNDICE A – Distribuição *t* de Student e Estatística *t*

- **Distribuição *t* de Student:**

Tabela 0 – Distribuição *t* de Student

Graus de liberdade GL	Nível de confiança				Graus de liberdade GL	Nível de confiança			
	95%	90%	80%	50%		95%	90%	80%	50%
1	12,71	6,31	3,08	1,00	16	2,12	1,75	1,34	0,69
2	4,30	2,92	1,89	0,82	17	2,11	1,74	1,33	0,69
3	3,18	2,35	1,64	0,76	18	2,10	1,73	1,33	0,69
4	2,78	2,13	1,53	0,74	19	2,09	1,73	1,33	0,69
5	2,57	2,02	1,48	0,73	21	2,08	1,72	1,32	0,69
6	2,45	1,94	1,44	0,72	23	2,07	1,71	1,32	0,69
7	2,36	1,89	1,41	0,71	25	2,06	1,71	1,32	0,68
8	2,31	1,86	1,40	0,71	27	2,05	1,70	1,31	0,68
Graus de liberdade GL	Nível de confiança				Graus de liberdade GL	Nível de confiança			
	95%	90%	80%	50%		95%	90%	80%	50%
9	2,26	1,83	1,38	0,70	31	2,04	1,70	1,31	0,68
10	2,23	1,81	1,37	0,70	35	2,03	1,69	1,31	0,68
11	2,20	1,80	1,36	0,70	41	2,02	1,68	1,30	0,68
12	2,18	1,78	1,36	0,70	49	2,01	1,68	1,30	0,68
13	2,16	1,77	1,35	0,69	60	2,00	1,67	1,30	0,68
14	2,14	1,76	1,35	0,69	120	1,98	1,66	1,29	0,68
15	2,13	1,75	1,34	0,69	□	1,96	1,64	1,28	0,67

Fonte: PIMVP Estatística e Incerteza para o PMIMVP - EVO (2014).

Nota: Calcular os GL da forma abaixo:

- $GL = n - 1$  (para uma distribuição amostral)
- $GL = n - p - 1$  (para um modelo de regressão)

Onde:

- $n$  = tamanho da amostra
- $p$  = número de variáveis independentes no modelo de regressão

- **Estatística *t*:**

Os coeficientes do modelo de regressão são estimativas estatísticas da verdadeira relação entre cada variável  $X$  e  $Y$ , são sujeitos a variação. A precisão da estimativa é medida pelo erro padrão do coeficiente e do valor respectivo da

estatística  $t$ . A estatística  $t$  é um teste estatístico para determinar se uma estimativa tem significado estatístico. Uma vez que o valor é estimado usando-se o teste, pode ser comparado com os valores críticos  $t$  da Tabela 1. A equação 1 abaixo se aplica para o caso de uma variável independente (EVO, 2014).

$$EP_b = \sqrt{\frac{\sum(Y_i - \hat{Y})^2 / (n-2)}{\sum(X_i - \bar{X})^2}} \quad \text{Equação 1}$$

Para os casos com mais de uma variável independente, a equação fornece uma aproximação razoável quando as variáveis independentes o são realmente (ou seja, não são correlacionadas). A faixa dentro da qual o valor verdadeiro do coeficiente deve estar é encontrada pela Equação 2.

$$b \pm t \times EP_b \quad \text{Equação 2}$$

O erro padrão do coeficiente  $b$  leva também ao cálculo da estatística  $t$ . Este teste, em última instância, determina se o coeficiente calculado tem significado estatístico. A estatística  $t$  é calculada usando a seguinte equação 3.

$$\text{Estatística } t = \frac{b}{SE_b} \quad \text{Equação 3}$$

Após a estimativa da estatística  $t$  pela equação acima, o valor pode ser comparado aos valores críticos da Tabela 1. Se o valor absoluto da estatística  $t$  calculada exceder o respectivo valor da Tabela 1, pode-se concluir que a estimativa do coeficiente é estatisticamente válida.

Uma regra empírica estabelece que se o valor absoluto da estatística  $t$  for maior que 2 significa que o coeficiente estimado é significativo em relação ao erro padrão, ou seja, há uma relação válida entre  $Y$  e o  $X$  do respectivo coeficiente; ou ainda, pode-se concluir que o  $b$  estimado não é zero. Entretanto, uma estatística  $t$  perto de 2 significa que a precisão do valor do coeficiente está em torno de  $\pm 100\%$ , o que não é muito para se confiar no valor de  $b$  calculado. Para se obter uma melhor precisão, digamos  $\pm 10\%$ , os valores da estatística  $t$  devem ficar em torno de 20, ou o erro padrão de  $b$  não ser maior que 10% do seu valor.

## APÊNDICE B – Desenvolvimento do código e o código

### Desenvolvimento do código

O código foi desenvolvido em *python*, uma das linguagens mais populares quando fala-se de processamento de dados. Existem bibliotecas dedicadas ao processamento de dados que abstraem a necessidade de desenvolvimento de código manual para as técnicas de *machine learning* conhecidas pela literatura. Essas bibliotecas são *open-source* e estão em constante validação por milhares de pessoas pelo mundo.

Algumas dessas bibliotecas é o “Pandas”, que foi criada para a manipulação e análise de dados. Permitindo a criação de *datasets* fazendo o papel de um “banco de dados” local para os dados, e permite várias ferramentas de análise, filtros, pesquisas, eliminação e substituição de valores, etc (Coutinho, 2020).

A biblioteca responsável pelas técnicas de *machine learning* é a *scikit-learn*, onde contém o conjunto de diferentes técnicas para modelos supervisionados, não supervisionadas, validação dos modelos, treinamento, processamento e transformação dos dados, entre outras (Coutinho, 2020).

Para os modelos de redes neurais de Deep Learning foi utilizada uma biblioteca dedicada chamada “Keras”, que é fácil de ser usada, extensível e modular, trazendo diversas funções que constroem partes fundamentais das redes neurais, como camadas, otimizadores, objetivos, entre outras. Esta biblioteca trabalha em cima do framework “TensorFlow”. O objetivo dessa biblioteca é abstrair a quantidade do código necessário para implementar o projeto, concentrando apenas nas configurações dos parâmetros necessários. O framework “TensorFlow” é um sistema para criação e treinamento de redes neurais para detectar e decifrar padrões e correlações (Coutinho, 2020).

### O Código

```
from pathlib import Path
import pathlib
from datetime import datetime
import PySimpleGUI as sg
```

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
import numpy as np
import pandas as pd
from keras.models import Sequential
from keras.layers import LSTM, Dense
import matplotlib.pyplot as plt
from keras.models import load_model
import math
from sklearn.metrics import mean_squared_error
from tabulate import tabulate
import operator
from matplotlib.ticker import NullFormatter # useful for `logit` scale
import matplotlib.pyplot as plt
import numpy as np
from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg
import PySimpleGUI as sg
import matplotlib
from io import BytesIO
from PIL import Image, ImageDraw
matplotlib.use('TkAgg')
diretorio = Path(__file__).parent.absolute()
lista_cidades_dados = ['CUIABA', 'GOIANIA',
                       'IMPERATRIZ', 'ARAGUATINS', 'PALMAS']
cidade = "GOIANIA" # A cidade de analise
ano = "2020"
mes_atual = 1 # Mes de analise
mes_atual_nome = "Janeiro"
repeticoes = 1 # Número de treinamentos
erro_maximo = 0 # Erro maximo desejado em relação a medição real
NUM_N_P = 50 # Primeira camada
NUM_N_S = 30 # Segunda camada
EPOCHS = 50 # Total de epocas camada
df = pd.DataFrame(columns=['A']) # Cria um dataset sem dados
```

```

df_por_dia_aux = pd.DataFrame(columns=['A']) # Cria um dataset sem dados
# Dias dos meses do ano
meses = [31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31]
meses_dias = meses[mes_atual-1]
MAX_ROWS, MAX_COLS, COL_HEADINGS = 1, 12, (
    'JAN', 'FEV', 'MAR', 'ABR', 'MAI', 'JUN', 'JUL', 'AGO', 'SET', 'OUT', 'NOV',
'DEZ')
COL_HEADINGS_PV = ('PVJAN', 'PVFEV', 'PVMAR', 'PVABR', 'PVMAI',
'PVJUN',
    'PVJUL', 'PVAGO', 'PVSET', 'PVOUT', 'PVNOV', 'PVDEZ')
COL_HEADINGS_MV = ('MVJAN', 'MVFEV', 'MVMAR', 'MVABR', 'MVMAI',
'MVJUN',
    'MVJUL', 'MVAGO', 'MVSET', 'MVOUT', 'MVNOV', 'MVDEZ')
# Valores de 12 meses de simulados para no pvsyst
pvsyst = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
# Valores de 12 meses de medição real medidos
real_calculado = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
# Atribui os valores das variveis de acordo com o mes escolhido para analise
previsto_pvsyst = pvsyst[mes_atual-1]
medido_real_calculado = real_calculado[mes_atual-1]
# O objetivo é prever os proximo mês (proximos_dias) a partir dos dias do
mesmo mês no ano anterior.
# A rede será treinada utilizando os dados do ultimos 10 anos do mês
escolhido.
look_back = meses[mes_atual-1]
forward_days = meses[mes_atual-1]
array_test = [] # Array para o teste
array_train = [] # Array para o treino
scl = MinMaxScaler() # Escala de 0 a 1
# armazena os dados da melhor rede
melhor_rede_modelo = None
melhor_rede = {}
# Lista de Redes Treinadas
lista_redes = [[0, 0, 0, 0, 0, 0, 0]]

```

```
lista_redes_melhor = [[0, 0, 0, 0, 0, 0, 0]]
def icon(check):
    box = (29, 29)
    background = (255, 255, 255, 0)
    rectangle = (3, 3, 20, 20)
    line = ((6, 14), (12, 20), (20, 6))
    im = Image.new('RGBA', box, background)
    draw = ImageDraw.Draw(im, 'RGBA')
    draw.rectangle(rectangle, outline='black', width=3)
    if check == 1:
        draw.line(line, fill='green', width=3, joint='curve')
    elif check == 2:
        draw.line(line, fill='green', width=3, joint='curve')
    with BytesIO() as output:
        im.save(output, format="PNG")
        png = output.getvalue()
    return png
check = [icon(0), icon(1), icon(2)]
headings = ['Ano', ""]
data_years = [
    ['2001', "],
    ['2002', "],
    ['2003', "],
    ['2004', "],
    ['2005', "],
    ['2006', "],
    ['2007', "],
    ['2008', "],
    ['2009', "],
    ['2010', "],
    ['2011', "],
    ['2012', "],
    ['2013', "],
    ['2014', "],
```

```

    ['2015', ""],
    ['2016', ""],
    ['2017', ""],
    ['2018', ""],
    ['2019', ""],
    ['2020', ""],
    ['2021', ""],
]
treedata = sg.TreeData()
for year, second in data_years:
    treedata.Insert("", year, year, values=[second],
                    icon=check[0])
# Funções
# Atribui o valor do mês atual
def atualisaLista(tree, years):
    tree.metadata = []
    for current_year, b in data_years:
        tree.metadata.append(str(current_year))

    for current_year in years:
        if str(current_year) in tree.metadata:
            tree.update(key=str(current_year), icon=check[1])
        else:
            tree.metadata.remove(str(current_year))
            tree.update(key=str(current_year), icon=check[0])
def defineMes(c):
    global mes_atual, previsto_pvsyst, medido_real_calculado,
mes_atual_nome
    mes_atual = int(idMes(c))
    meses_dias = meses[mes_atual-1]
    previsto_pvsyst = pvsyst[mes_atual-1]
    medido_real_calculado = real_calculado[mes_atual-1]
    mes_atual_nome = nomeMes(mes_atual)
    print("[LOG] Mês: ", mes_atual_nome, " selecionado. ", meses_dias, " dias")

```

```
print("[LOG] Valor Pvsyst ", previsto_pvsyst)
print("[LOG] Valor Medido ", medido_real_calculado)
print("")

# Retorna a data e hora atual
def hora():
    now = datetime.now()
    dt_string = now.strftime("%d/%m/%Y %H:%M:%S")
    return dt_string

def nomeMes(argument):
    switcher = {
        1: "Janeiro",
        2: "Fevereiro",
        3: "Março",
        4: "Abril",
        5: "Maio",
        6: "Junho",
        7: "Julho",
        8: "Agosto",
        9: "Setembro",
        10: "Outubro",
        11: "Novembro",
        12: "Dezembro"
    }
    return switcher.get(argument, "Mês inválido")

def idMes(argument):
    switcher = {
        "Janeiro": 1,
        "Fevereiro": 2,
        "Março": 3,
        "Abril": 4,
        "Maio": 5,
        "Junho": 6,
        "Julho": 7,
```

```
    "Agosto": 8,  
    "Setembro": 9,  
    "Outubro": 10,  
    "Novembro": 11,  
    "Dezembro": 12  
}  
return switcher.get(argument, 0)  
def limpaMelhorRede():  
    global melhor_rede  
    melhor_rede = {  
        "id": 0,  
        "mes_nome": mes_atual_nome,  
        "mes_atual": mes_atual,  
        "rsme": 0,  
        "valor_previsto": 0,  
        "valor_pvsyst": 0,  
        "valor_medido": 0,  
        "erro_pvsyst_medido": 0,  
        "erro_pvsyst": 0,  
        "erro_medido": 0,  
        "loss_treinamento": [],  
        "loss_validacao": [],  
        "X_test": [],  
        "y_test": [],  
        "X_test": [],  
        "y_test": [],  
        "teste_x": [],  
        "teste_y": [],  
        "treino_x": [],  
        "treino_y": [],  
        "dados_x": [],  
        "dados_y": []  
    }  
limpaMelhorRede()
```

```

def limpaListaRede():
    global lista_redes
    lista_redes = [[0, 0, 0, 0, 0, 0, 0]]
    for x in range(repeticoes-1):
        lista_redes.append([[0, 0, 0, 0, 0, 0, 0]])
def limpaListaRedeMelhor():
    global lista_redes_melhor
    lista_redes_melhor = [[0, 0, 0, 0, 0, 0, 0]]
def defineCidade(c):
    global cidade
    cidade = c
    print("[LOG]", hora(), " Cidade ", c, " selecionada")
    print("")
def filtraAnos(c, tree):
    global df, df_por_dia_aux, window
    df_por_dia_aux_ano = df_por_dia_aux
    l = tree.metadata
    l = np.array(l).astype(int)
    df_por_dia_aux_ano =
df_por_dia_aux_ano[df_por_dia_aux_ano['ANO'].isin(l)]
    preparaDados(df_por_dia_aux_ano)
    calculaDatas(df)
    print("[LOG]", hora(), " Ano ", c, "alterado")
    print("")
def carregaValoresReais():
    global pvsyst, diretorio
    pvsyst_valores = pd.read_excel(
        str(diretorio)+"/valores/valores.xlsx", index_col=0)
    pvsyst_valores.reset_index(inplace=True)
    carregados = pvsyst_valores.loc[(pvsyst_valores['CIDADE'] == cidade) & (
        pvsyst_valores['ANO'] == int(ano))]
    if(len(carregados) > 0):
        table =
        [[window[("pvr"+str(row),
"pvc"+str(col))].update(carregados.values[0][2+col]) for col in range(MAX_COLS)]

```

```

        for row in range(MAX_ROWS)]
    for col in range(MAX_COLS):
        pvsyst[col] = carregados.values[0][2+col]
    else:
        table = [[window[("pvr"+str(row), "pvc"+str(col))].update(1) for col in
range(MAX_COLS)]
                for row in range(MAX_ROWS)]
    for col in range(MAX_COLS):
        pvsyst[col] = 1
    print("[LOG] Valores de comparação não existe para o ano especificado")
def calulaDatas(df_data):
    global window
    if(len(df_data) > 0):
        d_inicial = datetime.strptime(str(df_data.index.min()), '%Y-%m-%d')
        d_final = datetime.strptime(str(df_data.index.max()), '%Y-%m-%d')
        window['-TXT-INICIAL-'].update(d_inicial.strftime('%d/%m/%y'))
        window['-TXT-FINAL-'].update(d_final.strftime('%d/%m/%y'))

    else:
        window['-TXT-INICIAL-'].update("--/--/----")
        window['-TXT-FINAL-'].update("--/--/----")
        window['-TAMANHO-DADOS-'].update(str(len(df_data))+ " dias")
def preparaDados(df_por_dia):
    global df, df_por_dia_aux, window
    cols_plot = ['RADIACAO GLOBAL (KJ/m²)']
    df_por_dia_group = df_por_dia.groupby(
        by=df_por_dia['DATA'].dt.date)[cols_plot].sum()
    df = df_por_dia_group
    d_inicial_filtro = '2001-01-01'
    d_final_filtro = str(ano)+'-12-31'
    d_inicial_filtro = pd.to_datetime(d_inicial_filtro).date()
    d_final_filtro = pd.to_datetime(d_final_filtro).date()
    df = df.loc[d_inicial_filtro:d_final_filtro]
    calulaDatas(df)

```

```

print("[LOG] Dados carregados")
def carregaDados(tree):
    global df, df_por_dia_aux, window
    carregaValoresReais()
    df_dados = pd.read_pickle(
        str(diretorio)+"/dados_tratados/dados_completos_"+str(cidade.lower()))
    df_dados = df_dados.reset_index()
    df_dados['DATA'] = df_dados['index'].astype(str).str.replace("\+00:00", "")
    df_por_dia = df_dados.loc[df_dados['CIDADE'] == cidade]
    df_por_dia['DATA'] = pd.to_datetime(df_por_dia['DATA'])
    df_por_dia = df_por_dia.loc[df_por_dia['ANO'] <= int(ano)]
    df_dia = df_por_dia.loc[df_por_dia['ANO'] == int(ano)]
    radicao = df_dia['RADIACAO GLOBAL (KJ/m²)'] * 1000/3600
    df_dia['RADIACAO GLOBAL (KJ/m²)'] = radicao
    df_dia = df_dia.reset_index()
    df_dia = df_dia.set_index(
        pd.DatetimeIndex(df_dia['DATA']))
    df_dia = df_dia.groupby(pd.Grouper(freq='M')).sum()
    df_dia = df_dia[df_dia['RADIACAO GLOBAL (KJ/m²)'] > 0]
    calculados_completo = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
    for col in range(len(df_dia)):
        t2 = df_dia.reset_index()
        itemAtual = t2["DATA"][col]
        t2 = t2[t2["DATA"] == itemAtual].index[0]
        if(t2 == col):
            calculados_completo[col] = round(
                df_dia["RADIACAO GLOBAL (KJ/m²)"][col]/1000, 2)
    if(len(df_dia) > 0):
        table = [[window[("mvr"+str(row),
"mvc"+str(col))].update(calculados_completo[col]) for col in range(MAX_COLS)]
                for row in range(MAX_ROWS)]
        for col in range(MAX_COLS):
            real_calculado[col] = calculados_completo[col]
    else:

```

```

        table = [[window[("mvr"+str(row), "mvc"+str(col))].update(1) for col in
range(MAX_COLS)]
                for row in range(MAX_ROWS)]
    for col in range(MAX_COLS):
        real_calculado[col] = 1
        print("[LOG] Valores de comparação não pode ser calculados")
    df_por_dia = df_por_dia.loc[df_por_dia['MES'] == mes_atual]
    df_por_dia['RADIACAO GLOBAL (KJ/m²)'] = df_por_dia['RADIACAO
GLOBAL (KJ/m²)'] * 1000/3600
    anos = df_por_dia['ANO'].unique()
    actualisaLista(tree, anos)
    df_por_dia_aux = df_por_dia
    preparaDados(df_por_dia)
    # print(df.info())
    # print("")
def defineAno(c):
    global ano
    ano = c
    print("[LOG] Ano: ", ano)
    print("")

def defineNP(c):
    global NUM_N_P
    NUM_N_P = c
    print("[LOG] Neuronios P. ", NUM_N_P)
    print("")

def defineNS(c):
    global NUM_N_S
    NUM_N_S = c
    print("[LOG] Neuronios S. ", NUM_N_S)
    print("")

def defineEP(c):
    global EPOCHS

```

```

EPOCHS = c
print("[LOG] Epocas. ", EPOCHS)
print("")
def defineRepeticoes(c):
    global repeticoes
    repeticoes = c
    print("[LOG] Repetições. ", repeticoes)
    print("")
def defineErro(c):
    global erro_maximo
    erro_maximo = c
    print("[LOG] Erro Máximo. ", erro_maximo)
    print("")
def plotDados(d):
    if(len(d) > 0):
        df_dados = d.reset_index()
        plt.figure(figsize=(15, 5))
        plt.plot(df_dados["RADIACAO GLOBAL (KJ/m²)"], label='Irradiância')
        plt.show(block=False)
        plt.legend(loc='best')
        plt.xlabel('Dias', fontweight='bold', fontsize=15)
        plt.ylabel('IRRADIAÇÃO GLOBAL (kWh/m²)', fontweight='bold',
fontsize=15)

    else:
        print("[LOG] Dados vazios.")
def plotLoss(a, b):
    if(len(a) > 0):
        plt.figure(figsize=(15, 5))
        plt.plot(a, label='Treinamento')
        plt.plot(b, label='Validação')
        plt.show(block=False)
        plt.legend(loc='best')
        plt.xlabel('Épocas', fontweight='bold', fontsize=15)

```

```

plt.ylabel('RMSE (Escala 0 a 1)', fontweight='bold', fontsize=15)
else:
    print("[LOG] Dados vazios.")
def plotResultadoTeste(X_test, y_test):
    if(len(X_test) > 0):
        global melhor_rede_modelo
        Xt = melhor_rede_modelo.predict(X_test)
        plt.figure(figsize=(15, 5))
        for i in range(0, len(Xt)):
            plt.plot([x + i*forward_days for x in range(len(Xt[i]))],
                    scl.inverse_transform(Xt[i].reshape(-1, 1)), color='r')
        plt.plot(0, scl.inverse_transform(
            Xt[i].reshape(-1, 1))[0], color='r', label='Previsto') # only to place the
label

plt.plot(scl.inverse_transform(y_test.reshape(-1, 1)), label='Real')
plt.show(block=False)
plt.legend(loc='best')
plt.xlabel('Dias', fontweight='bold', fontsize=15)
plt.ylabel('IRRADIAÇÃO GLOBAL (kWh/m²)', fontweight='bold',
fontsize=15)
else:
    print("[LOG] Dados vazios.")
def plotResultado(resultado):
    if(len(resultado["teste_x"]) > 0):
        plt.figure(figsize=(15, 5))
        plt.plot(resultado["treino_x"],
                resultado["treino_y"], color='r', label='Treino')
        plt.plot(resultado["teste_x"], resultado["teste_y"],
                color='g', label='Teste')
        plt.plot(resultado["dados_x"], resultado["dados_y"],
                color='b', label='Real')
        plt.show(block=False)
        plt.legend(loc='best')
        plt.xlabel('Dias', fontweight='bold', fontsize=15)

```

```

plt.ylabel('IRRADIAÇÃO GLOBAL (kWh/m²)', fontweight='bold',
fontsize=15)
else:
    print("[LOG] Dados vazios.")
def salvarModelo(modelo):
    global cidade, mes_atual_nome, look_back, forward_days, EPOCHS,
NUM_N_P, NUM_N_S
    now = datetime.now()
    dt_string = now.strftime("%d-%m-%Y %H-%M-%S")
    file_name =
'./RedesSolar/{}/LSTM_CIDADE_{}_MES_{}_A{}_P{}_E{}_NP{}_NS{}_{}.h5'.format(
    cidade, cidade, mes_atual_nome, look_back, forward_days, EPOCHS,
NUM_N_P, NUM_N_S, dt_string)
    modelo.save(file_name)
    print("Modelo salvo no disco: {}".format(file_name))
# Função que faz a separação dos dados para o treino, sendo os X dias no
passado para a entrada e Y dias no futuro para a previsão.
def processData(data, look_back, forward_days, jump=1):
    X, Y = [], []
    for i in range(0, len(data) - look_back - forward_days + 1, jump):
        X.append(data[i:(i+look_back)])
        Y.append(data[(i+look_back):(i+look_back+forward_days)])
    return np.array(X), np.array(Y)
def treinamento():
    global df, scl, array_test, array_train, melhor_rede, melhor_rede_modelo,
lista_redes, medido_real_calculado, previsto_pvsyst
    previsto_pvsyst = pvsyst[mes_atual-1]
    medido_real_calculado = real_calculado[mes_atual-1]
    # Preparando dados para a normalização
    array = df.values.reshape(df.shape[0], 1)
    # print(array[:5])
    # Normalizando os dados para o intervalo de 0 a 1
    array = scl.fit_transform(array)
    # print(array[:5])

```

# Dividir os dados em Treino/Validação para o modelo LSTM e os dados de Teste # O ultimo mes será utilizados para o teste. O resto será utilizado para o treinamento (Treino e Validação).

```

division = len(array) - forward_days
array_test = array[division-look_back:]
array_train = array[:division]
# print(array_test.shape)
# print(array_train.shape)
## Separa os dados de teste
X_test, y_test = processData(
    array_test, look_back, forward_days, forward_days)
y_test = np.array([list(a.ravel()) for a in y_test])
# Separa os dados de treino
X, y = processData(array_train, look_back, forward_days)
y = np.array([list(a.ravel()) for a in y])

# Separa os dados de treino em Treino e Validação (20%)
X_train, X_validate, y_train, y_validate = train_test_split(
    X, y, test_size=0.20, random_state=42)
# Tamanho de cada vetor Treino/Teste/Validação
# print(X_train.shape)
# print(X_validate.shape)
# print(X_test.shape)
# print(y_train.shape)
# print(y_validate.shape)
# print(y_test.shape)
limpaMelhorRede()
limpaListaRede()
limpaListaRedeMelhor()
window['-PROGRESS BAR-'].UpdateBar(0, repeticoes)
# LSTM (Treinando a LSTM)
for x in range(repeticoes):
    sg.Print('Treinando Modelo %d \n' %
        (x+1), background_color='blue', text_color='white', end=")

```

```

# Construindo o modelo
model = Sequential()
model.add(LSTM(NUM_N_P, input_shape=(
    look_back, 1), return_sequences=True))
model.add(LSTM(NUM_N_S,
    input_shape=(NUM_N_P, 1)))
model.add(Dense(forward_days))
model.compile(loss='mean_squared_error', optimizer='adam')
history = model.fit(X_train, y_train, epochs=EPOCHS, validation_data=(
    X_validate, y_validate), shuffle=False, batch_size=1, verbose=2)
loss_treinamento = history.history['loss']
loss_validacao = history.history['val_loss']

# Teste da previsão realizada pelo modelo
Xt = model.predict(X_test)
for i in range(0, len(Xt)):
    y_teste_previsto = scl.inverse_transform(Xt[i].reshape(-1, 1))
y_treino_real = scl.inverse_transform(y_train.reshape(-1, 1))
y_teste_real = scl.inverse_transform(y_test.reshape(-1, 1))
testScore = math.sqrt(mean_squared_error(
    y_teste_real[:, 0], y_teste_previsto[:, 0]))
sg.Print('Modelo %d Treinado \n' %
    (x+1), background_color='yellow', text_color='black', end=")
print("#####")
print(hora(), " RESULTADO")
print("#####")
print('\n Modelo %d - Teste: %.2f RMSE' % (x+1, testScore))
# Usa o modelo utilizando os dados de treino e teste
division = len(array) - forward_days
leftover = division % forward_days+1
array_test = array[division-look_back:]
array_train = array[leftover:division]
Xtrain, ytrain = processData(
    array_train, look_back, forward_days, forward_days)

```

```

Xtest, ytest = processData(
    array_test, look_back, forward_days, forward_days)
Xtrain = model.predict(Xtrain)
Xtrain = Xtrain.ravel()
Xtest = model.predict(Xtest)
Xtest = Xtest.ravel()
y = np.concatenate((ytrain, ytest), axis=0)

# Treino
treino_x = [x for x in range(
    look_back+leftover, len(Xtrain)+look_back+leftover)]
treino_y = scl.inverse_transform(Xtrain.reshape(-1, 1))
plt.plot(treino_x, treino_y, color='r', label='Treino')

# Teste
teste_x = [x for x in range(
    look_back + leftover + len(Xtrain),
len(Xtrain)+len(Xtest)+look_back+leftover)]
teste_y = scl.inverse_transform(Xtest.reshape(-1, 1))
dados_x = [x for x in range(
    look_back+leftover, look_back+leftover+len(Xtrain)+len(Xtest))]
dados_y = scl.inverse_transform(y.reshape(-1, 1))

# Realiza a previsão com o dado de teste
teste_y = scl.inverse_transform(Xtest.reshape(-1, 1))

# Calcula o valor mensal da previsão
previso_dp = np.sum(teste_y)/1000

# Calcula os erros em relação ao PVSyst e Medido
erro_pvsyst_medido = abs(
    medido_real_calculado - previsto_pvsyst) / previsto_pvsyst * 100
erro_pvsyst_previsto = abs(
    previso_dp - previsto_pvsyst) / previsto_pvsyst * 100

```

```

erro_medido_previsto = abs(
    previso_dp - medido_real_calculado) / medido_real_calculado * 100

lista_redes[x] = [x+1, round(mes_atual, 2), round(testScore, 2),
round(previso_dp, 2),
    round(erro_pvsyst_medido, 2), round(erro_pvsyst_previsto,
2), round(erro_medido_previsto, 2)]
window.FindElement("-LISTA-").Update(lista_redes)
# Grava os dados da melhor rede
if(melhor_rede['erro_medido'] > erro_medido_previsto or
melhor_rede['valor_previsto'] == 0):
    melhor_rede_modelo = model
    lista_redes_melhor[0] = [x+1, round(mes_atual, 2), round(testScore,
2), round(previso_dp, 2),
        round(erro_pvsyst_medido, 2),
round(erro_pvsyst_previsto, 2), round(erro_medido_previsto, 2)]
window.FindElement("-LISTA2-").Update(lista_redes_melhor)
melhor_rede = {
    "id": x+1,
    "mes_nome": mes_atual_nome,
    "mes_atual": mes_atual,
    "rsme": testScore,
    "valor_previsto": previso_dp,
    "valor_pvsyst": previso_pvsyst,
    "valor_medido": medido_real_calculado,
    "erro_pvsyst": erro_pvsyst_previsto,
    "erro_medido": erro_medido_previsto,
    "loss_treinamento": loss_treinamento,
    "loss_validacao": loss_validacao,
    "X_test": X_test,
    "y_test": y_test,
    "teste_x": teste_x,
    "teste_y": teste_y,
    "treino_x": treino_x,

```

```

        "treino_y": treino_y,
        "dados_x": dados_x,
        "dados_y": dados_y
    }
    resultados = [
        ["Variavel", "Valor", "Previsto", "Erro (%)"],
        #          ["Medido          x          Calculado", round(medido_real,
2), round(medido_real_calculado, 2), round(erro_real_calculado, 2)],
        ["Pvsyst x Medido", round(previsto_pvsyst, 2), round(
            medido_real_calculado, 2), round(erro_pvsyst_medido, 2)],
        ["Pvsyst x Previsto", round(previsto_pvsyst, 2), round(
            previso_dp, 2), round(erro_pvsyst_previsto, 2)],
        ["Medido x Previsto", round(medido_real_calculado, 2), round(
            previso_dp, 2), round(erro_medido_previsto, 2)],
    ]
    print(tabulate(resultados,
                    headers="firstrow"))
    print("")
    print("")
    window['-PROGRESS BAR-'].UpdateBar((x+1), repeticoes)
    # Sai do loop de treinamento caso o erro atinja o desejado
    if(erro_maximo > 0):
        if(erro_medido_previsto < erro_maximo):
            break
    sg.Print('Treinamento Concluído \n', background_color='green',
            text_color='white', end=")

    sg.Print("Mês", melhor_rede['mes_nome'], end='\n')
    sg.Print("RSME", round(
        melhor_rede['rsme'], 2), end='\n')
    sg.Print("Valor previsto", round(
        melhor_rede['valor_previsto'], 2), end='\n')
    sg.Print("Erro: PVsyst x Previsto", round(
        melhor_rede['erro_pvsyst'], 2), end='\n')

```

```

sg.Print("Erro: Medido x Previsto", round(
    melhor_rede['erro_medido'], 2), end='\n')
print("")
print("#####")
print(hora(), " MELHOR RESULTADO")
print("#####")
resultados = [
    ["Variavel", "Valor", "Previsto", "Erro (%)",
    #      ["Medido      x      Calculado",round(medido_real,
2),round(medido_real_calculado, 2),round(erro_real_calculado, 2)],
    ["Pvsyst x Medido", round(melhor_rede['valor_pvsyst'], 2), round(
        melhor_rede['valor_medido'], 2), round(erro_pvsyst_medido, 2)],
    ["Pvsyst x Previsto", round(melhor_rede['valor_pvsyst'], 2), round(
        melhor_rede['valor_previsto'], 2), round(melhor_rede['erro_pvsyst'], 2)],
    ["Medido x Previsto", round(melhor_rede['valor_medido'], 2), round(
        melhor_rede['valor_previsto'], 2), round(melhor_rede['erro_medido'],
2)]]

print(tabulate(resultados, headers="firstrow"))

salvarModelo(melhor_rede_modelo)
print("#####")

```