

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL

CLAYTON HILGEMBERG DA COSTA

CLASSIFICAÇÃO DE FALHAS EM PLANTAS FOTOVOLTAICAS
USANDO APRENDIZADO DE MÁQUINA

DISSERTAÇÃO

CURITIBA

2020

CLAYTON HILGEMBERG DA COSTA

**CLASSIFICAÇÃO DE FALHAS EM PLANTAS FOTOVOLTAICAS
USANDO APRENDIZADO DE MÁQUINA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial (CPGEI) da Universidade Tecnológica Federal do Paraná (UTFPR) como requisito à obtenção do título de “Mestre em Ciências”. Área de Concentração: Engenharia De Automação E Sistemas.

Orientador: Prof. Dr. André Eugenio Lazzaretti
Coorientador: Prof. Dr. Guilherme Luiz Moritz

CURITIBA

2020

Dados Internacionais de Catalogação na Publicação

Costa, Clayton Hilgemberg da

Classificação de falhas em plantas fotovoltaicas usando aprendizado de máquina [recurso eletrônico] / Clayton Hilgemberg da Costa. -- 2020.

1 arquivo eletrônico (77 f.): PDF; 12,3 MB.

Modo de acesso: World Wide Web.

Texto em português com resumo em inglês.

Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. Área de Concentração: Engenharia de Automação e Sistemas. Linha de Pesquisa: Sistemas de Energia, Curitiba, 2020.

Bibliografia: f. 73-77.

1. Engenharia elétrica - Dissertações. 2. Células fotovoltaicas. 3. Geração de energia fotovoltaica - Aspectos ambientais. 4. Sistemas de energia fotovoltaica - Monitorização. 5. Localização de falhas (Engenharia). 6. Aprendizado do computador. 7. Redes neurais (Computação). 8. Modelos matemáticos. 9. Simulação (Computadores). 10. Validação. I. Lazzaretti, André Eugênio, orient. II. Moritz, Guilherme Luiz, coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. IV. Título.

CDD: Ed. 23 -- 621.3

Biblioteca Central do Câmpus Curitiba - UTFPR
Bibliotecária: Luiza Aquemi Matsumoto CRB-9/794



Ministério da Educação
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DIRETORIA DE PESQUISA E POS-GRADUACAO-CT
SECRETARIA GERAL DO STRICTO SENSU SEDE CENTRO - CT



TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 871

A Dissertação de Mestrado intitulada "Classificação de Falhas em Plantas Fotovoltaicas Usando Aprendizado de Máquina", defendida em sessão pública pelo(a) candidato(a) **Clayton Hilgemberg da Costa**, no dia 13 de julho, foi julgada para a obtenção do título de Mestre em Ciências, área de concentração Engenharia de Computação, linha de pesquisa Sistemas de Energia, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial.

Este Termo de Aprovação encontra-se na pasta do aluno na Secretaria do Programa, contendo a assinatura da Coordenação realizada após a entrega da versão final, incluindo correções necessárias, permitindo o encaminhamento para análise e publicação no Repositório Institucional.

Prof(a). Dr(a). André Eugenio Lazzaretti - UTFPR

Prof(a). Dr(a). Myriam Regattieri De Biase da Silva Delgado - UTFPR

Prof(a). Dr(a). Robson Ribeiro Linhares - UTFPR

(Participação por videoconferência) Prof(a). Dr(a). Alexandre Rasi Aoki - UFPR

Assinado eletronicamente por

Coordenador do Programa
André Schneider de Oliveira

AGRADECIMENTOS

Ao meu orientador André Eugênio Lazzaretti e coorientador Guilherme Luiz Moritz pela paciência, apoio, motivação e ensinamentos.

Aos meus amigos e colegas de mestrado Marcelo Paludetto Rodrigues, Bruna Machado Mulinari, Guilherme Dan Yamada e Hellen Cristina Ancelmo, e ao Gilberto Lexinoski e demais amigos e colegas de laboratório, por toda ajuda e necessários momentos de descontração.

À minha mãe Silvia Hilgemberg da Costa, à minha irmã Aline Hilgemberg da Costa e à minha companheira, melhor amiga e amada Nailyn Andrade Moço, por toda a paciência, carinho e compreensão.

À Agência Nacional de Energia Elétrica (ANEEL) e à Companhia Paranaense de Energia Elétrica (COPEL) pelo suporte. Este trabalho teve financiamento via projeto de Pesquisa & Desenvolvimento dentro do programa da ANEEL, código PD2866-0464/2017.

RESUMO

COSTA, Clayton Hilgemberg da. **CLASSIFICAÇÃO DE FALHAS EM PLANTAS FOTOVOLTAICAS USANDO APRENDIZADO DE MÁQUINA**. 2020. 77 f. Dissertação (Mestrado em Engenharia Elétrica e Informática Industrial) – Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Com o aumento progressivo no interesse pela energia solar fotovoltaica, cresce também o interesse no monitoramento das plantas, bem como no desenvolvimento de métodos práticos para se detectar e até mesmo classificar falhas que possam ocorrer em qualquer parte dos sistemas fotovoltaicos. Dentre alguns dos motivadores para este crescimento, pode ser citado o fato de que grande parte destes sistemas fica em ambiente aberto e desprotegido, ficando suscetível ao efeito de eventos meteorológicos, tais quais ventos e chuvas fortes, o que acarreta numa maior propensão à falhas. Outro fator é que a falha em um elemento do sistema, como um módulo fotovoltaico por exemplo, pode prejudicar o funcionamento dos demais, além do fato de que um sistema operando com falhas por muito tempo pode reduzir a vida útil dos equipamentos e até gerar riscos maiores como a ocorrência de incêndios. Desta forma, neste trabalho é apresentado um sistema de classificação de falhas de curto-circuito, degradação, circuito aberto e sombreamento em plantas fotovoltaicas ligadas à rede elétrica. Este sistema foi construído utilizando modelos de aprendizado de máquina, a partir das medições de irradiância local e temperatura, bem como das tensões e correntes geradas pelos módulos fotovoltaicos. Ainda, o modelo de classificação foi treinado com dados provenientes de uma simulação construída e, foi ainda, implementado em um sistema de monitoramento desenvolvido, o qual foi utilizado para coletar as variáveis elétricas e ambientais em torno da planta fotovoltaica. O trabalho foi validado em uma planta real de 5kWp e levou em conta duas abordagens, sendo uma com e outra sem detecção das falhas. O sistema final apresentou um resultado competitivo com os encontrados na literatura, alcançando acurácias superiores a 90%, apresentando ainda vantagens relevantes como: a classificação não intrusiva, classificador embarcado no sistema de monitoramento, o treinamento com dados exclusivos de simulação e validação com dados reais, incluindo o sombreamento natural da planta.

Palavras-chave: Detecção de Falhas. Classificação de Falhas. Energia Solar Fotovoltaica. Aprendizado de Máquina.

ABSTRACT

COSTA, Clayton Hilgemberg da. **FAULT CLASSIFICATION FOR PHOTOVOLTAIC PLANTS USING MACHINE LEARNING**. 2020. 77 p. Dissertation (Master's Degree in Graduate Program in Electrical and Computer Engineering) – Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

With the progressive increase in interest in solar photovoltaic energy, there is also an growing interest in monitoring plants, as well as the development of practical methods to detect and even classify faults that may occur in any part of photovoltaic systems. Among some of the motivators for this growth, it can be mentioned the fact that most of these systems are in an open and unprotected environment, being susceptible to the effect of meteorological events, such as strong winds and storms, which leads to a greater tendency to failures. Another factor is that the failure of a single system element, such as a photovoltaic module for instance, can impair the functioning of the others, in addition to the fact that a system operating with failures for a long time can reduce the lifespan of the equipment and even generate risks as the occurrence of fires. Thus, this work presents a classification system for short-circuit, degradation, open circuit and shading failures in on grid photovoltaic plants. This system was built using machine learning models, using measurements of local irradiance and temperature, as well as the voltages and currents generated by the photovoltaic modules. Also, the classification model was trained with data gathered from simulation and was also implemented in a monitoring system, which was used to collect the electrical and weather variables around the photovoltaic plant. The work was validated in a real 5kWp plant and took into account two approaches, one with and other without fault detection. The final system presented a competitive result with those found in the literature, reaching over 90% accuracy, while still presenting relevant advantages such as: non-intrusive classification, classifier embedded in the monitoring system, training with simulation data only and validation with data including the natural shading on the photovoltaic plant.

Keywords: Fault Detection. Fault Classification. Photovoltaic Solar Energy. Machine Learning.

LISTA DE ILUSTRAÇÕES

Figura 1 – Curva IV e MPP	18
Figura 2 – Modelo de um diodo	19
Figura 3 – Representação Gráfica de um exemplo de uma Árvore de Decisão de cinco classes	25
Figura 4 – Exemplo de aplicação do SVM em um problema de classificação	26
Figura 5 – Modelo de um neurônio artificial	30
Figura 6 – Autoria Própria	32
Figura 7 – Resumo das Falhas no lado de corrente contínua	36
Figura 8 – Visão Geral do Sistema	41
Figura 9 – Diagrama Elétrico do Sistema	43
Figura 10 – Quadro com dispositivos de manobra	43
Figura 11 – Quadro com dispositivos de proteção	44
Figura 12 – Planta Fotovoltaica Utilizada	44
Figura 13 – Sistema no laboratório que inclui: (a) Inversor; (b) Quadro elétrico com equipamentos de proteção; (c) Monitor utilizado para apresentar a interface homem-máquina do sistema de monitoramento; (d) CompactRIO; (e) Sensores de corrente CC e CA; (f) Banco de resistores utilizado para gerar a falha de degradação; e (g) Analisador de qualidade de energia	47
Figura 14 – Estação Meteorológica	47
Figura 15 – Medição de Temperatura	48
Figura 16 – Medição de Irradiância	48
Figura 17 – Diagrama da instrumentação para coleta das variáveis elétricas	49
Figura 18 – Fluxo lógico do sistema de monitoramento	50
Figura 19 – Visão geral do circuito simulado	51
Figura 20 – Circuito Boost implementado	52
Figura 21 – Inversor implementado	52
Figura 22 – Diagrama de blocos do sistema de controle utilizado	53
Figura 23 – Dispositivos para introdução de falhas na simulação.	54
Figura 24 – Exemplo de dia do banco de dados	56
Figura 25 – Esquema de Treinamento e Teste dos Classificadores	57
Figura 26 – Esquema de Classificação	57
Figura 27 – Processo de Validação Cruzada	58
Figura 28 – Interface do sistema de monitoramento.	60
Figura 29 – Comparação entre potência real e simulada.	62
Figura 30 – Corrente Alternada - Osciloscópio e Sistema de Monitoramento (Calibrado)	63
Figura 31 – Matriz de confusão da RNA para a abordagem de classificação.	65
Figura 32 – Matriz de confusão da RNA para a abordagem de classificação com detecção.	66

LISTA DE TABELAS

Tabela 1 – Dados Elétricos do Módulo CS6U-330P	42
Tabela 2 – Sensores da estação meteorológica	47
Tabela 3 – Parâmetros do modelo utilizados	53
Tabela 4 – Dados para teste do classificador	56
Tabela 5 – Resultado das validações	61
Tabela 6 – Resultado dos classificadores para abordagem com apenas classificação . .	64
Tabela 7 – Resultado dos classificadores para abordagem com detecção e classificação .	66

LISTA DE SIGLAS

CA	Corrente Alternada
CC	Corrente Contínua
DT	<i>Decision Tree</i>
FPGA	<i>Field Programmable Gate Array</i>
kNN	<i>k-Nearest-Neighbors</i>
MLP	<i>Multilayer Perceptron</i>
MPP	<i>Maximum Power Point</i>
MPPT	<i>Maximum Power Point Tracker</i>
P&O	Perturba e Observa
PI	Proporcional-Integral
PWM	<i>Pulse Width Modulation</i>
RBF	<i>Radial Basis Function</i>
RNA	Rede Neural Artificial
STC	<i>Standard Test Conditions</i>
SVM	<i>Support Vector Machine</i>

SUMÁRIO

1	INTRODUÇÃO	12
1.1	OBJETIVOS	13
1.1.1	Objetivo Geral	13
1.1.2	Objetivos Específicos	13
1.2	ESTRUTURA DO TRABALHO	14
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	ENERGIA SOLAR FOTOVOLTAICA	15
2.1.1	Células Solares e Módulos Fotovoltaicos	15
2.1.2	Impacto de variáveis ambientais	16
2.1.3	Sistemas conectados à rede e sistemas isolados	17
2.1.4	Rastreadores de Máxima Potência	17
2.2	MODELO DE UM DIODO	19
2.3	CLASSIFICADORES	22
2.3.1	K-vizinhos-mais-próximos	23
2.3.2	Árvores de Decisão	24
2.3.3	Máquina de Vetores Suporte	26
2.3.3.1	Máquina de Vetores Suporte Não-Linear	28
2.3.3.2	Casos de classificação multiclasse	29
2.3.4	Redes Neurais Artificiais	30
2.3.4.1	Redes Perceptron Multicamadas	31
2.3.4.2	Algoritmo da Retropropagação do Erro	31
3	TRABALHOS CORRELATOS	34
3.1	MONITORAMENTO DE PLANTAS FOTOVOLTAICAS	34
3.2	FALHAS EM SISTEMAS FOTOVOLTAICOS	34
3.3	MÉTODOS VISUAIS	35
3.4	MÉTODOS POR TERMOGRAFIA	36
3.5	MÉTODOS MATEMÁTICOS	37
3.6	INTELIGÊNCIA ARTIFICIAL	38
3.7	CONTRIBUIÇÕES	39
4	METODOLOGIA	41
4.1	SISTEMA FOTOVOLTAICO	42
4.1.1	Falhas no Sistema Fotovoltaico	44
4.1.1.1	Módulos em Curto-Circuito	45
4.1.1.2	Degradação	45
4.1.1.3	Circuito Aberto	46
4.1.1.4	Sombreamento	46
4.2	SISTEMA DE MONITORAMENTO	46
4.2.1	Variáveis Ambientais	46
4.2.2	Variáveis Elétricas	48
4.2.3	Lógica do sistema de monitoramento	49
4.3	SISTEMA FOTOVOLTAICO SIMULADO	50
4.3.1	Geração de falhas Simuladas	53

4.4	DETECÇÃO E CLASSIFICAÇÃO DE FALHAS	54
4.4.1	Determinação dos hiperparâmetros	57
4.4.2	Implementação dos classificadores no sistema de monitoramento	59
5	RESULTADOS E ANÁLISES	61
5.1	VALIDAÇÃO DA SIMULAÇÃO	61
5.2	VALIDAÇÃO DO SISTEMA DE MONITORAMENTO	62
5.3	CLASSIFICAÇÃO DE FALHAS	63
5.3.1	Abordagem 1 - Somente classificação	64
5.3.2	Abordagem 2 - Inclusão da detecção	65
5.3.3	Comparação com resultados da literatura	67
6	CONCLUSÃO E TRABALHOS FUTUROS	69
6.1	TRABALHOS FUTUROS	70
6.2	TRABALHOS PUBLICADOS	70
	REFERÊNCIAS	73

1 INTRODUÇÃO

Visando reduzir drasticamente o uso de combustíveis fósseis, especialmente com o objetivo de se atenuar o impacto das mudanças climáticas, nas últimas décadas observou-se um crescente interesse pelas chamadas energias renováveis. As energias renováveis, tal qual sua denominação sugere, são derivadas de fontes renováveis de energia, tais quais a energia solar, hídrica, eólica, maremotriz e de biomassa, por exemplo. Dentre estas, a energia solar e, mais especificamente, a energia solar fotovoltaica, vem apresentando uma evolução (e conseqüente uso) cada vez maior na matriz energética mundial, sendo que o ano de 2019 foi o nono ano consecutivo em que este foi o tipo de energia renovável que mais recebeu investimentos, alcançado 42,5% de todo investimento destinado à categoria (JÄGER-WALDAU, 2019). Isto se deve principalmente a algumas vantagens relevantes em relação as demais energias renováveis. Uma destas vantagens é que sua implementação causa um impacto ambiental significativamente menor do que a implementação das usinas hidrelétricas, por exemplo, as quais tendem a mudar drasticamente o ecossistema em seus arredores. Outro ponto é que durante a geração, a planta fotovoltaica gera uma quantidade quase nula de ruídos e vibração, diferente da energia eólica.

Ainda assim, a energia solar fotovoltaica apresenta problemas importantes. Um destes problemas é que boa parte de uma planta fotovoltaica fica diretamente exposta ao ambiente externo. Este fato acarreta numa probabilidade considerável da ocorrência de falhas que podem reduzir significativamente (ou até mesmo anular) a geração de energia e, ainda, aumentar o risco de problemas sérios como incêndios, por exemplo.

Desta maneira, o monitoramento de plantas fotovoltaicas, aliado à utilização de técnicas para se identificar e classificar estas falhas, pode permitir um acompanhamento mais apropriado de uma planta deste tipo, bem como a aceleração de um processo de manutenção.

Neste contexto, várias propostas de sistemas de detecção e classificação de falhas podem ser encontradas na literatura. Algumas destas propostas possuem um caráter mais visual, seja por inspeção realizadas por humanos, como o método apresentado em (PACKARD *et al.*, 2012) ou por meio de termografia, como os métodos apresentados em (TSANAKAS; BOTSARIS, 2012), (HU *et al.*, 2014), (TSANAKAS *et al.*, 2015), (SINHA *et al.*, 2016) e (TSANAKAS *et al.*, 2017). Ambos os métodos, porém, são muito limitados em relação às falhas que podem ser detectadas e, principalmente, em relação a escalabilidade para plantas fotovoltaicas de porte médio ou grande.

Ainda, existe uma série de propostas com métodos matemáticos, como os apresentados em (CHINE *et al.*, 2014), (MADETI; SINGH, 2017b) e (MANSOURI *et al.*, 2018). Diferente dos métodos puramente visuais e dos métodos termográficos, os matemáticos podem ser aplicados, sem maiores problemas, em plantas de maior porte, porém ainda são limitados em relação à gama de falhas que podem ser abordadas, além de serem mais voltadas para detecção de falhas apenas.

Também existem as propostas baseadas em técnicas de inteligência artificial, como as apresentadas em (LI *et al.*, 2017), (CHINE *et al.*, 2016), (CHOUAY; OUASSAID, 2018) e (GAROUJJA *et al.*, 2017). Em sua maioria, estes métodos também podem ser empregados em plantas fotovoltaicas maiores e, ainda, demonstram capacidade para abordar uma quantidade maior de falhas e, em muitos casos, classificá-las além de detectar sua ocorrência. Nestes trabalhos, porém, as soluções não foram embarcadas junto à sistemas de monitoramento, sendo então necessária a utilização de um computador sempre ligado e próximo ao sistema para realizar a classificação. Ainda, de maneira geral os métodos podem carecer de outras características relevantes. Muitos destes métodos, por exemplo, utilizam a curva $V \times I$ da planta fotovoltaica em operação, o que inviabiliza a realização da detecção e classificação das falhas de maneira não-intrusiva, visto ser desnecessária a desconexão da planta fotovoltaica do resto do sistema para geração da curva $V \times I$.

1.1 OBJETIVOS

Dadas as limitações apresentadas na literatura e a necessidade técnica observada, os objetivos deste trabalho são detalhados a seguir.

1.1.1 Objetivo Geral

Desenvolver um sistema de monitoramento inteligente da operação de sistemas fotovoltaicos.

1.1.2 Objetivos Específicos

- Construir e validar uma simulação que represente uma planta fotovoltaica real instalada.
- Desenvolver um sistema de monitoramento de sistemas de energia solar fotovoltaica.

- Construir uma base de dados simulados e reais de falhas em uma planta fotovoltaica.
- Desenvolver modelos de aprendizado de máquina capazes de detectar e classificar falhas em uma planta fotovoltaica real utilizando, em seus treinamentos, dados exclusivamente simulados.
- Comparar o desempenho de diferentes modelos de aprendizado de máquina na classificação de falhas em um sistema fotovoltaico.
- Embarcar um método de classificação de falhas junto à um sistema de monitoramento de sistemas fotovoltaicos.

1.2 ESTRUTURA DO TRABALHO

Seguinte a esta introdução, o trabalho é estruturado da seguinte forma: No Capítulo 2, são explicados alguns dos principais fundamentos teóricos utilizado na elaboração do presente trabalho. Em seguida, o Capítulo 3 apresenta alguns trabalhos relacionados diretamente ao tema de detecção e classificação de falhas em sistemas fotovoltaicos. O Capítulo 4 segue com a apresentação da metodologia elaborada e seguida para realização do trabalho. Na sequência, Capítulo 5, são apresentados os principais resultados obtidos em sua realização, enquanto no Capítulo 6 são discutidas as conclusões, bem como trabalhos futuros, além ainda, da citação de trabalhos correlatos publicados durante esta pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

Este trabalho é baseado em princípios de duas áreas do conhecimento: energia solar fotovoltaica e aprendizado de máquina. Tendo isto em vista, o presente Capítulo apresentará os principais aspectos teóricos utilizados. Inicialmente, serão discutidos os sistemas fotovoltaicos, incluindo seus principais componentes e tipos de instalação, a influência de variáveis externas em seu funcionamento e um de seus modelos representativos. Na sequência, serão apresentados os quatro classificadores utilizados neste trabalho para detecção e classificação das falhas.

2.1 ENERGIA SOLAR FOTOVOLTAICA

Com a crescente busca por formas de se reduzir a emissão de carbono e poluição do ar, há um aumento na busca por energias renováveis eficientes e com custo acessível. Entre os mais diversos tipos de energia renovável, tal qual a energia eólica, maremotriz e de biomassa, por exemplo, a energia solar fotovoltaica tem demonstrado o maior potencial de crescimento atualmente. Um fator importante neste significativo crescimento da energia solar fotovoltaica é a redução de preços pela qual a tecnologia passou nos últimos anos, com uma diminuição em 77% no seu custo global médio ponderado (IRENA, 2019).

2.1.1 Células Solares e Módulos Fotovoltaicos

A energia solar fotovoltaica, tal como sugere sua denominação, é a energia elétrica obtida pela conversão da luz solar por meio do efeito fotovoltaico. O efeito fotovoltaico foi descoberto em 1839 por Alexandre-Edmond Becquerel e resume-se no aparecimento de uma corrente elétrica contínua em materiais semicondutores quando expostos à luz (ZAIDI, 2018).

Um dos principais componentes na geração prática de energia elétrica por meio do efeito fotovoltaico é a célula solar, a qual consiste na junção de duas regiões semicondutoras com concentrações distintas de elétrons, podendo ainda ser fabricada com diversos tipos de materiais, entre os quais destacam-se o silício monocristalino, silício policristalino e o telureto de cádmio (SÁNCHEZ *et al.*, 2019).

As células solares, porém, não são utilizadas isoladamente, visto que geram uma tensão geralmente menor que 1V, além de serem muito frágeis. Devido a estes fatos, são construídos

os módulos fotovoltaicos, estruturas compostas por um arranjo de diversas células conectadas em série, para que haja um aumento na tensão elétrica gerada (PINHO; GALDINO, 2014). Em alguns casos estes arranjos de células conectadas em série são também conectados em paralelo com outros arranjos, promovendo também um aumento da corrente gerada. Ainda, além de permitir a elevação da tensão e corrente geradas, estes módulos contém diversas estruturas, tais como vidro temperado e moldura de alumínio, que provém diversos benefícios como proteção e facilidade de instalação, por exemplo. Além disto, é comum que os módulos fotovoltaicos sejam fabricados com uma caixa de conexão para instalação de diodos de desvio (do inglês, *bypass*), os quais são responsáveis por impedir que um módulo com problemas influencie negativamente outros módulos conectados em série no sistema (PINHO; GALDINO, 2014). Assim como várias células são conectadas para a composição dos módulos, em sistemas fotovoltaicos também é comum que se realize a conexão entre diversos módulos, formando-se então um painel ou *string*.

2.1.2 Impacto de variáveis ambientais

Muito embora possa parecer comum se pensar apenas na luz do sol ao se considerar o impacto de variáveis ambientais na geração fotovoltaica, os módulos fotovoltaicos (e consequentemente as células solares) estão constantemente sujeitos à influência de diversas outras, muitas das quais podem ter um impacto direto na geração ou mesmo integridade do módulo afetado. Entre estas variáveis, a que possui maior impacto na geração, além da irradiância solar, é a temperatura, tendo um impacto já bem estabelecido na literatura (SKOPLAKI; PALYVOS, 2009), ao ponto que a informação da quantidade de perda de eficiência por acréscimo de temperatura é disponibilizada pelos fabricantes, nas folhas de dados da maioria dos módulos fotovoltaicos atualmente. O aumento da temperatura da célula solar causa vibrações moleculares que dificultam a passagem dos portadores de carga no efeito fotovoltaico (HERSCH; ZWEIBEL, 1982), reduzindo assim sua eficiência.

Entre as demais variáveis que podem impactar na geração fotovoltaica estão incluídos o vento, a umidade e a deposição de poeira (MEKHILEF *et al.*, 2012). Porém, ao contrário dos casos da temperatura e da irradiância, estas variáveis estão mais ligadas a uma correlação que possuem entre si e com a própria irradiância e temperatura, do que com um efeito físico. A poeira, por exemplo, tende a reduzir a área do módulo que é atingida pela luz solar, bem como a intensidade da luz que atinge sua superfície, enquanto a velocidade do vento impacta na temperatura do módulo, podendo ainda carregar poeira até o mesmo ou ainda limpá-lo, dependendo de condições

como sua velocidade e direção. A umidade pode impactar na temperatura do módulo e, caso consiga alcançar as células solares, pode causar degradação nas mesmas (MEKHILEF *et al.*, 2012).

2.1.3 Sistemas conectados à rede e sistemas isolados

Os sistemas exclusivamente fotovoltaicos podem ser divididos em duas categorias básicas: sistemas conectados à rede elétrica e os sistemas isolados.

Mais comuns em regiões rurais e regiões com pouco ou nenhum acesso à rede elétrica, os sistemas isolados são aqueles nos quais a energia excedente gerada se dissipa ou, mais comumente, é armazenada localmente, geralmente com o auxílio de baterias próprias para tal função.

Nos sistemas conectados à rede elétrica, a energia excedente gerada é injetada na rede elétrica. Como estas redes são de corrente alternada e a energia fotovoltaica é em corrente contínua, é necessário o uso de um inversor para que esta injeção seja possível. Além disso, devido às características próprias das redes, as instalações fotovoltaicas conectadas devem atender alguns parâmetros como fator de potência, frequência da corrente injetada e nível de tensão, por exemplo. Ainda, torna-se necessária a implementação de um sistema anti-ilhamento, de forma que nos casos em que ocorram a interrupção do fornecimento de energia, o sistema fotovoltaico não injete corrente na rede. Isto ocorre para evitar acidentes na execução de manutenções. Os sistemas conectados à rede elétrica são, atualmente, os mais comuns no mundo (IRENA, 2019), sendo então o tipo estudado neste trabalho.

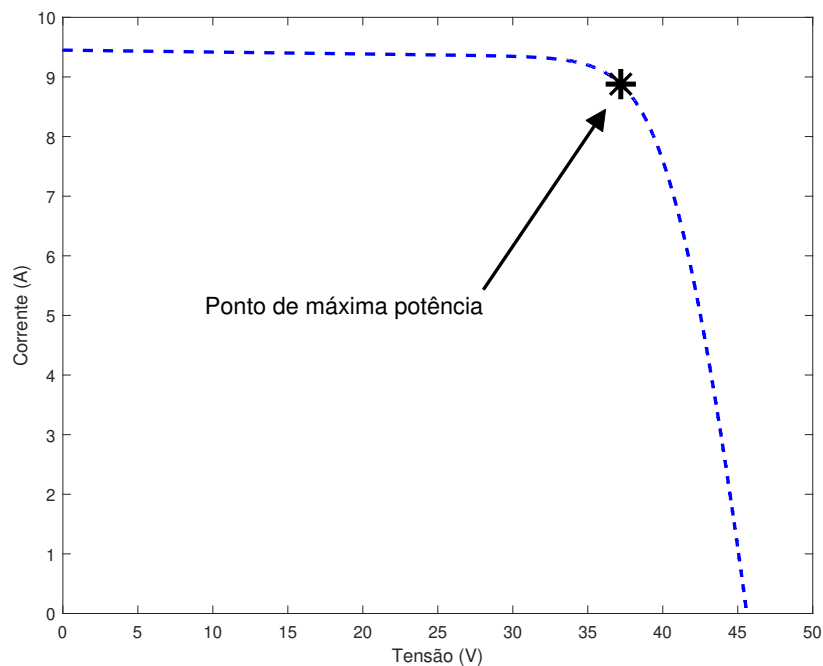
2.1.4 Rastreadores de Máxima Potência

Um ponto de significativa importância em relação aos módulos fotovoltaicos é o fato da corrente gerada por estes variar de acordo com a tensão elétrica entre seus terminais. O fator que acarreta nesta importância é o chamado ponto de máxima potência ou MPP (do inglês, *Maximum Power Point*, sendo este o ponto onde o par tensão e corrente acarretam na máxima geração de energia. É importante notar que esta variação é dependente da irradiância e temperatura sob as quais o módulo está operando, isto é, uma variação na irradiância e/ou temperatura, faz com que o ponto de máxima potência se desloque.

A Figura 1 apresenta a curva da variação de tensão e da corrente de um módulo

fotovoltaico, bem como, em destaque, o ponto de máxima potência para uma irradiância e temperatura arbitrária. Desta forma, dada a característica variável das saídas dos sistemas fotovoltaicos, sobretudo pela dependência da irradiância solar incidente, temperatura e nível de carga, torna-se interessante o uso de algum método para se extrair a máxima potência dos módulos fotovoltaicos. Isto é obtido pelo uso dos MPPTs (do inglês, *Maximum Power Point Trackers*), sistemas de controle que podem oferecer um ganho relevante de energia gerada (PINHO; GALDINO, 2014), sendo o método de controle mais usual nestes sistemas. Diversos algoritmos de MPPT tem sido propostos na literatura, sendo um dos mais populares o do tipo P&O (Perturba e Observa), principalmente pela sua simplicidade de entendimento e implementação, além de seu baixo custo (EL-KHOZONDAR *et al.*, 2016).

Figura 1 – Curva IV e MPP



Fonte: Autoria Própria

O método P&O pode ser resumido em quatro passos básicos:

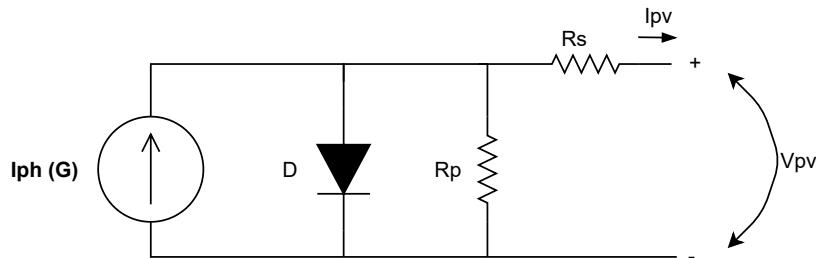
- Mede-se a potência de saída do sistema fotovoltaico.
- "Perturba-se" o sistema, geralmente com uma elevação ou redução da corrente do mesmo.
- Mede-se novamente a potência e verifica se houve aumento ou diminuição na potência de saída.

- Caso tenha havido uma elevação na potência no passo anterior, continua-se a perturbar o sistema na mesma direção, caso contrário, passa-se a perturbá-lo no sentido contrário, reduzindo a tensão caso estava sendo elevada e vice-versa.

2.2 MODELO DE UM DIODO

Por se tratar de uma tecnologia de comportamento extremamente variável, até por sua dependência das características ambientais, é comum que se trabalhe com modelos para representatividade de células solares e, por consequência, módulos fotovoltaicos. Um dos modelos mais comuns na literatura é o modelo de um diodo (IZADIAN *et al.*, 2012), representado na Figura 2, na qual a I_{pv} é a corrente gerada fornecida pela célula, V_{pv} é a tensão da célula, a fonte de corrente $I_{ph}(G)$ representa a célula solar em si (a corrente gerada, dada uma certa irradiância incidente G , desconsiderando as perdas), o diodo D é utilizado para caracterizar a curva V-I da célula, o resistor R_s caracteriza as perdas geradas pelas camadas de materiais construtivos e conexões e o resistor R_p caracteriza perdas por correntes de fuga e pequenos curtos que possam ocorrer internamente à célula (IZADIAN *et al.*, 2012).

Figura 2 – Modelo de um diodo



Fonte: Autoria Própria

No modelo de um diodo, a corrente gerada pela célula é dada por (1), na qual I_s é a corrente de saturação do diodo, n é o seu fator de idealidade e V_{th} é a tensão térmica da junção.

$$I_{pv} = I_{ph} - I_s \left(e^{\frac{V_{pv} + R_s I_{pv}}{n V_{th}}} - 1 \right) - \frac{V_{pv} + R_s I_{pv}}{R_p}. \quad (1)$$

A tensão térmica da junção é dada por (2), sendo k a constante de Boltzmann ($k = 1,381 \times 10^{-23} \frac{J}{K}$), T a temperatura, em graus celsius, da junção PN e q a carga do elétron ($q = 1,60218 \times 10^{-19} C$).

$$V_{th} = \frac{k (T + 273,15)}{q}. \quad (2)$$

O modelo de um diodo também pode ser utilizado para representar módulos fotovoltaicos ou mesmo *strings* completas. Para isto, (1) é adaptada, de forma que inclua as quantidades de células que estão em paralelo e série. Esta adaptação pode ser vista em (3), na qual N_s é a quantidade de células em série e N_p a quantidade de células em paralelo.

$$I_{pv} = N_p \left(I_{ph} - I_s (e^\theta - 1) - \frac{\frac{N_s}{N_p} V_{pv} + R_s I_{pv}}{R_p} \right), \quad (3)$$

$$\theta = \frac{\frac{V_{pv}}{N_s} + \frac{R_s I_{pv}}{N_p}}{n V_{th}}.$$

Como pode ser visto, (3) possui cinco variáveis que devem ser encontradas para o uso do modelo: (I) I_{ph} ; (II) I_s ; (III) R_s ; (IV) R_p ; e (V) n . Para encontrar os valores destas variáveis para um determinado módulo, uma das possibilidades é o uso de informações fornecidas nas folhas de dados dos módulos, sobretudo o comportamento do módulo operando sob as chamadas 'Condições-padrão de teste' ou STC (do inglês, *Standard Test Conditions*), as quais estabelecem uma irradiância de $1000 \frac{W}{m^2}$ e temperatura de $25^\circ C$. A partir destes fatores, pode-se então desenvolver novas equações para elaboração de um sistema para que se encontrem as cinco variáveis necessárias para o modelo. Uma destas, apresentada em (4), é obtida com a substituição, em (3), dos valores de geração em STC com o módulo estando em curto-circuito. Nestas condições, $V_{pv} = 0$ e $I_{pv} = I_c$, sendo I_c a corrente de curto-circuito em STC do módulo em questão.

$$I_c = N_p \left(I_{ph} - I_s (e^{\theta'} - 1) \right) - \frac{I_c R_s}{R_p}, \quad (4)$$

$$\theta' = \frac{R_s I_c}{N_p n V_{th}}.$$

Da mesma forma, substituindo em (3) os valores de geração em STC com o módulo em circuito aberto, isto é, $V_{pv} = V_a$ e $I_{pv} = 0$, sendo V_a a tensão de circuito aberto em STC, obtêm-se (5).

$$N_p \left(I_{ph} - I_s (e^{\theta''} - 1) - \frac{V_a}{N_s R_p} \right) = 0, \quad (5)$$

$$\theta'' = \frac{V_a}{N_s n V_{th}}.$$

Uma terceira equação, apresentada em (6), pode ser obtida a partir da substituição, em (3), da tensão (V_m) e corrente (I_m) geradas pelo módulo operando no ponto de máxima potência

sob STC.

$$I_m = N_p \left(I_{ph} - I_s (e^{\theta'''} - 1) - \frac{\frac{N_s}{N_p} V_m + R_s I_m}{R_p} \right), \quad (6)$$

$$\theta''' = \frac{V_m + R_s I_m}{n V_{th}}.$$

Outra equação é obtida a partir do fato que a variação da potência, em relação à tensão, é nula no ponto de máxima potência, conforme (7).

$$\left. \frac{d(V \cdot I)}{dV} \right|_{V=V_m, I=I_m} = 0. \quad (7)$$

Derivando a multiplicação e substituindo a corrente e tensão no ponto de máxima potência em (7), obtém-se (8):

$$\left. \frac{dI}{dV} \right|_{V=V_m, I=I_m} = -\frac{I_m}{V_m}. \quad (8)$$

Visto a quantidade de variáveis que precisam ser determinadas, uma quinta equação é necessária. Diversas equações têm sido propostas na literatura para este fim. Tian *et al.* (2012), por exemplo, apresenta uma solução baseada na operação do módulo em circuito aberto sob condições distintas das STC, conforme (9), na qual β é um coeficiente de temperatura para a tensão de circuito aberto, a qual é geralmente fornecida na folha de dados dos módulos fotovoltaicos, T_b é uma temperatura arbitrária e V_b é a tensão de circuito aberto para a temperatura T_b .

$$V_a = V_b (1 + \beta(T_b - 25)). \quad (9)$$

Substituindo (9) em (5), obtêm-se (10).

$$N_p \left(I_{ph} - I_s (e^{\theta''} - 1) - \frac{V_b (1 + \beta(T_b - 25))}{N_s R_p} \right) = 0, \quad (10)$$

$$\theta'' = \frac{V_b (1 + \beta(T_b - 25))}{N_s n V_{th}}.$$

Assim, pode-se construir um sistema com cinco equações ((4), (5), (6), (8) e (10)) e cinco variáveis (I_{ph} , I_s , R_s , R_p e n). Convém apontar porém, que a solução deste sistema não é trivial, visto tratar de um sistema de equações transcendentais e não-lineares, sendo então utilizado um método numérico para sua solução.

Uma vez solucionado o sistema de equações, obtém-se então os parâmetros sob STC do módulo desejado. O próximo passo é generalizar o modelo para quaisquer valores de irradiância (G) e temperatura (T). A Equação (11), na qual α é uma constante de temperatura da corrente do módulo, apresenta a corrente I_{ph} em função da irradiância e temperatura.

$$I_{PH}(G,T) = \frac{G}{1000} I_{ph}(\alpha(T - 25) + 1) . \quad (11)$$

Da mesma forma, a corrente de saturação I_s pode ser escrita em função da irradiância e temperatura conforme (12), na qual E_g é a energia de banda proibida do silício sob uma temperatura arbitrária (T), o qual é dado por (13) e E_{ref} é a energia de banda proibida do silício em STC.

$$I_S(G,T) = I_s \left(\frac{T + 273.15}{298.15} \right)^3 e^{\frac{E_{ref}}{298.15 \cdot k} - \frac{E_g}{T + 273.15}} , \quad (12)$$

$$E_g = 1.16 - 7.02 \times 10^{-4} \frac{(T + 273.15)^2}{T - 834.85} . \quad (13)$$

Vale ressaltar que os valores utilizados nas equações apresentadas para construção do modelo, isto é: α , β , V_m , I_m , N_s , N_p , V_a e I_c , podem ser encontrados nas folhas de dados das maiorias dos módulos disponíveis comercialmente. Neste trabalho, os valores utilizados para o modelo (I_{ph} , I_s , R_s , R_p e n) serão detalhados no Capítulo 4.

2.3 CLASSIFICADORES

A teoria de aprendizado de máquinas contempla diversos tipos de classificadores, modelos que buscam imitar a capacidade humana para reconhecer padrões (CHERKASSY; MULIER, 2007). Estes modelos variam em abordagem, custo computacional e eficácia para cada caso em que são aplicados. Seu funcionamento consiste em aprender, por meio da observação de determinadas características de diversos exemplos, a classificar amostras em uma ou mais classes. Ainda, estes classificadores podem ser distribuídos em três subcategorias do aprendizado de máquinas: aprendizado supervisionado, não supervisionado e aprendizado por reforço. Os classificadores utilizados no presente trabalho estão contidos na subcategoria de aprendizado supervisionado, o mais estudado atualmente (SUTTON *et al.*, 2015), e que consiste em aprender de exemplos cuja a classe pertencente é previamente conhecida.

Dois conceitos importantes, quando do treinamento dos classificadores, são os parâmetros e hiperparâmetros. Os parâmetros são valores que são determinados pelo próprio modelo durante sua fase de treinamento, em outras palavras, são os valores que definem o aprendizado em si. Os hiperparâmetros, por outro lado, são valores que devem ser determinados previamente ao treinamento do classificador, não podendo ser aprendidos pelo mesmo durante seu treinamento e, portanto, sua escolha pode se tornar difícil e demorada, sobretudo para classificadores com múltiplos hiperparâmetros dependentes.

Convém apontar ainda que a escolha de valores adequados para os hiperparâmetros é um passo essencial para o desenvolvimento, com boa taxa de sucesso, de um classificador. Isso se dá pois, além de influenciar na taxa acerto no processo de classificação durante o treinamento, eles podem impactar diretamente na capacidade de generalização de um modelo. Esta capacidade de generalização está relacionada com a habilidade de um modelo classificar corretamente amostras que não foram utilizadas durante a etapa de treinamento. Existem dois conceitos ligados diretamente a esta capacidade de generalização: o sobreajuste (do inglês, *overfitting*) e subajuste (do inglês, *underfitting*). O primeiro ocorre quando um modelo se ajustou excessivamente aos exemplos de treinamento, o que acarreta em um acerto grande na classificação destas amostras, mas com alta taxa de erro para novas amostras (DUDA *et al.*, 2000). Já o subajuste ocorre quando um modelo não possui uma taxa aceitável de acerto nem para os exemplos de treinamento, tão pouco para novos exemplos. Diferentes estratégias, além da seleção de hiperparâmetros, podem ser utilizadas para diferentes tipos de classificadores, entre as quais destaca-se: a interrupção prematura de treinamento (limite de ciclos de treinamento), aumento do número de exemplos utilizados para treinamento e o uso de técnicas de regularização.

A seguir, serão detalhados aspectos teóricos mais relevantes dos classificadores usados neste trabalho.

2.3.1 K-vizinhos-mais-próximos

Um dos classificadores mais simples, o K-vizinhos-mais-próximos ou kNN (do inglês, *k-nearest-neighbors*) é um algoritmo que consiste em classificar exemplos com base na classe dos exemplos mais próximos de si, na maioria das vezes, utilizando a distância euclidiana como método de determinação de proximidade (CHERKASSY; MULIER, 2007). Em outros termos, o kNN avalia a quais classes pertencem os exemplos mais próximos da amostra a ser classificada e determina que esta deve pertencer a mesma classe da maioria dos seus vizinhos. O

hiperparâmetro utilizado pelo kNN é o valor 'k', o qual refere-se a quantidade, a ser avaliada, de vizinhos mais próximos das amostras a serem classificadas.

Por utilizar apenas dados de distância entre a amostra a ser classificada e exemplos próximos, esse método não requer um treinamento em si, o que torna rápida sua implementação. Apesar disto, a utilização deste algoritmo requer o armazenamento de todos os seus exemplos, o que acarreta em uso potencialmente elevado de memória de armazenamento.

2.3.2 Árvores de Decisão

As Árvores de Decisão ou DT (do inglês, *Decision Trees*), são classificadores construídos pela sucessiva divisão das entradas, ou características, em determinados nós com base em condições dicotômicas (se e senão) (BREIMAN *et al.*, 1984)f. O principal mote da árvore de decisão é a definição de suas divisões. De modo simplista, o objetivo é que se façam divisões de forma que haja uma redução de impureza de um nó para seus nós subsequentes. Para tal, define-se uma função custo, a qual representa este grau de impureza a qual é então minimizada, de forma que se obtenha uma redução ótima. Ainda, esta função custo deve atender três critérios fundamentais (BREIMAN *et al.*, 1984):

- A função custo deve ser um máximo apenas no ponto $(\frac{1}{j}, \frac{1}{j}, \dots, \frac{1}{j})$;
- A função custo deve ser um mínimo apenas nos pontos $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, \dots, 0, 1)$;
- A função custo deve ser simétrica de $p(1), p(2), \dots, p(j)$.

Nestes critérios, j refere-se a uma classe $j \in (1, 2, 3, \dots, J)$ em um problema com J classes e $(p(1), p(2), \dots, p(J))$ representam as probabilidades de ocorrência de cada uma das J classes na árvore de decisão em questão.

Uma das funções que atendem a estes critérios e que é utilizada em implementações práticas de árvores de decisão é a Função de Entropia (CHERKASSY; MULIER, 2007) que pode ser vista em (14), na qual $E(t)$ é a função de entropia e $p(j|t)$ é a probabilidade de ocorrência de uma das J classes no nó t .

$$E(t) = - \sum_j p(j|t) \ln(p(j|t)). \quad (14)$$

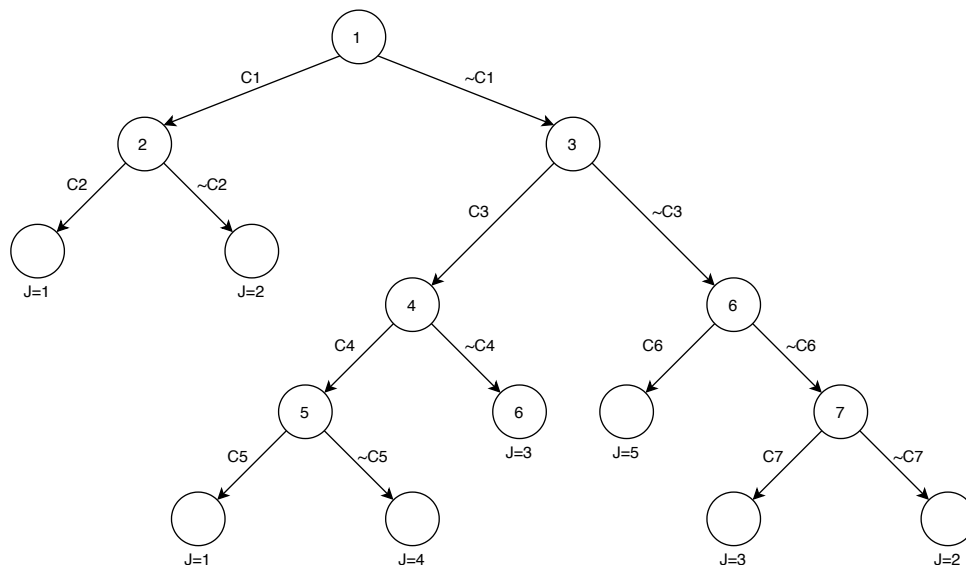
O processo de construção, ou crescimento, de uma Árvore de Decisão ocorre até que um critério de parada, previamente estabelecido, seja atingido. Entre alguns critérios de parada, cita-se um tamanho máximo para árvore (quantidade máxima de nós) e a diminuição do erro de classificação para um valor abaixo de um certo limiar.

Uma possibilidade após o treinamento de uma árvore de decisão é a ocorrência do sobreajuste. Uma das estratégias que podem ser usadas para contrapor esta ocorrência é a poda da árvore, isto é, utilizar alguma técnica de se reduzir o tamanho da árvore de decisão, mantendo uma taxa alta de acerto de classificação, porém aumentando sua capacidade de generalização.

Entre os hiperparâmetros que podem ser utilizados pelas árvores de decisão, encontram-se o, já citado, tamanho (ou profundidade) máxima da árvore, a quantidade mínima de exemplos *contidos* em um nó e o limiar de impureza para realização de uma nova divisão.

A Figura 3 apresenta a representação gráfica de uma árvore de decisão de cinco classes, na qual C_n e $\sim C_n$, com $n \in 1,2,\dots,5$ representam, respectivamente, a condição e sua negação para cada nó n .

Figura 3 – Representação Gráfica de um exemplo de uma Árvore de Decisão de cinco classes



Fonte: Autoria Própria

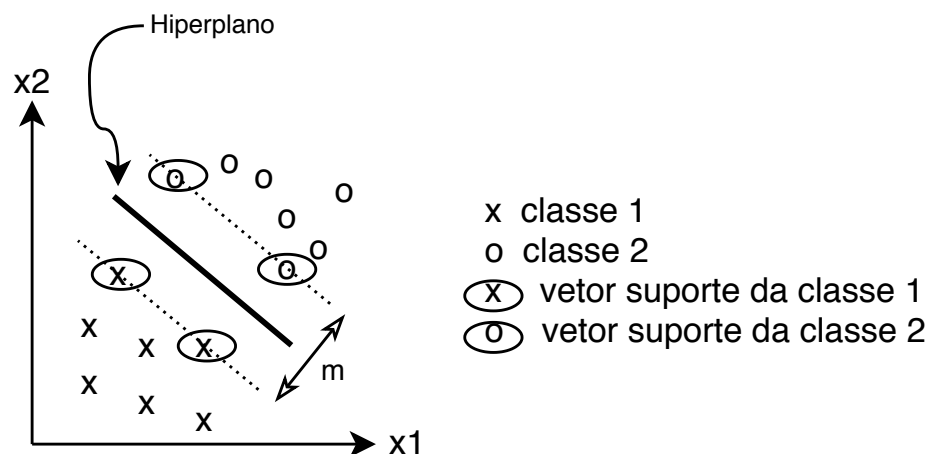
Como exemplo, na árvore apresentada na Figura 3, se uma amostra percorre, o caminho dado pelos nós 1, 3, 4 e 5 e a condição C_5 é atendida, considera-se que a amostra em questão pertence a classe $J = 1$. Por outro lado, se o mesmo caminho for percorrido e a condição C_5 não for atendida, a amostra é classificada como pertencente a classe $J = 4$.

2.3.3 Máquina de Vetores Suporte

A Máquina de Vetores Suporte, ou SVM (do inglês, *Support Vector Machine*), é um classificador que consiste em encontrar um hiperplano ótimo, que gere a máxima margem possível entre os pontos, estes chamados vetores suporte, mais próximos da superfície que separa duas classes linearmente separáveis (VAPNIK, 1998). A Figura 4 apresenta um exemplo de hiperplano que maximiza a margem m em um problema de classificação linear com características x_1 e x_2 .

O hiperplano que maximiza a margem entre os vetores suporte pode ser encontrado através da solução do problema de otimização apresentado em (15), na qual \mathbf{w} e w_0 são os parâmetros do hiperplano, \mathbf{x}_i é o vetor de entradas (características) dos N exemplos utilizados para treinamento e y_i é o indicador da classe do exemplo i , sendo $y_i \in \{-1,1\}$.

Figura 4 – Exemplo de aplicação do SVM em um problema de classificação



Fonte: Autoria Própria

$$\min J(\mathbf{w}, w_0) = \frac{1}{2} \|\mathbf{w}\|^2, \quad (15)$$

$$\text{Sujeito a: } y_i(\mathbf{w}^T \mathbf{x}_i + w_0) \geq 1, \quad i = 1, 2, \dots, N.$$

O problema apresentado em (15) é um problema de otimização convexa, podendo ser utilizado para sua solução, o Método dos Multiplicadores de Lagrange, conforme apresentado em (BOYD; VANDENBERGHE, 2001), resultando no problema de otimização equivalente

apresentado em (16), na qual λ é o vetor com os multiplicadores de Lagrange.

$$\max\left(\sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j\right),$$

$$\text{Sujeito a: } \begin{cases} \sum_{i=1}^N \lambda_i y_i = 0, \\ \lambda \geq 0. \end{cases} \quad (16)$$

A saída do classificador construído com base no hiperplano encontrado pela solução da (16) é dada pela (17), na qual N_s indica o número de vetores suporte e o sinal da saída $f(x)$ determina a qual classe pertence a amostra cujas características sejam x .

$$f(x) = \sum_{i=1}^{N_s} \lambda_i y_i \mathbf{x}_i^T x + w_0. \quad (17)$$

Vale ressaltar que o problema apresentado em (15) é aplicável aos casos em que existem apenas duas classes linearmente separáveis, visto que a solução não admite erros de classificação. Porém, é possível estabelecer uma certa tolerância a erros de classificação, o que permite a utilização desse classificador linear mesmo em casos em que as classes não atendam ao critério de separação linear (THEODORIDIS; KOUTROUMBAS, 2003).

Esta tolerância aos erros de classificação é estabelecida pela introdução de novas variáveis no problema, estas chamadas de variáveis de folga (do inglês, *slack variables*) e representadas por ξ_i , sendo $i = 1, 2, \dots, N$ e N o número de exemplos de treinamento. Estas variáveis recebem este nome pois tem por função, o relaxamento das restrições na determinação do hiperplano na forma de admitir alguns erros de classificação. Ainda, estas variáveis podem assumir valores em três regiões básicas: (i) $\xi_i = 0$, que ocorre quando o exemplo i situa-se no lado correto do hiperplano e fora da margem; (ii) $0 < \xi_i \leq 1$, que ocorre quando o exemplo i encontra-se do lado correto ao hiperplano, porém está dentro da margem; e (iii) $\xi_i > 1$, o qual ocorre quando o exemplo é classificado incorretamente.

A maximização da margem, com a adição do conceito supracitado, pode ser obtida com a solução do problema de otimização expresso em (18), que inclui, além das variáveis de folga, um parâmetro de regularização (HAYKIN, 2001) identificado por C . Este termo determina o balanço entre a capacidade de generalização do classificador, que aumenta conforme C é reduzido, e sua taxa de erro no treinamento, e deve ser determinado pelo usuário previamente ao

treinamento do classificador, sendo portanto um de seus hiperparâmetros.

$$\min J(\mathbf{w}, w_0, \boldsymbol{\xi}) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i,$$

$$\text{Sujeito a } \begin{cases} y_i[\mathbf{w}^T \mathbf{x}_i + w_0] \geq 1 - \xi_i, & i = 1, 2, \dots, N, \\ \xi_i \geq 0, & i = 1, 2, \dots, N. \end{cases}$$

(18)

Utilizando o Método dos Multiplicadores de Lagrange em (18), obtêm-se o problema de otimização equivalente expresso em (19).

$$\max \left(\sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \right),$$

$$\text{Sujeito a: } \begin{cases} \sum_{i=1}^N \lambda_i y_i = 0, \\ 0 \leq \lambda_i \leq C, & i = 1, 2, \dots, N. \end{cases}$$

(19)

2.3.3.1 Máquina de Vetores Suporte Não-Linear

Outra possibilidade para abordar problemas de classificação nos quais as classes não podem ser separadas linearmente é a construção de uma máquina de vetores suporte não-linear. Isto pode ser alcançado com o uso de *kernels*, funções relacionadas com o mapeamento das entradas do classificador para um espaço n -dimensional de alta dimensionalidade, chamado espaço de características. O uso de um espaço de maior dimensionalidade se justifica pelo fato que acarreta também no aumento da probabilidade de que o problema torne-se linear (VAPNIK, 1998). É importante que se note que um hiperplano que separa linearmente os exemplos em duas classes num espaço de características, pode tornar-se uma superfície de decisão não-linear no espaço de entradas. O hiperplano ótimo neste caso, para o espaço de características, pode ser obtido com a solução do problema de otimização apresentado em (20), na qual $K(\mathbf{x}_i, \mathbf{x}_j)$ é a

função *kernel*.

$$\max \left(\sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \right),$$

$$\text{Sujeito a: } \begin{cases} \sum_{i=1}^N \lambda_i y_i = 0, \\ 0 \leq \lambda_i \leq C, \quad i = 1, 2, \dots, N. \end{cases} \quad (20)$$

Em (20) é possível observar que a função kernel substituiu o termo $\mathbf{x}_i^T \mathbf{x}_j$, o que ocorre também na função que rege a classificação em si, (17), que para o caso não linear torna-se a (21), cujo sinal do resultado determina a qual classe pertence a amostra \mathbf{x} :

$$f(x) = \sum_{i=1}^{N_s} \lambda_i y_i K(\mathbf{x}_i^T \mathbf{x}) + w_0. \quad (21)$$

Um *kernel* muito comum na literatura é o Kernel RBF (do inglês, *Radial Basis Function*) (HAYKIN, 2001), o qual é expresso em (22). Em termos práticos o *kernel* RBF gera fronteiras de decisão em formato circular, cujas áreas estão diretamente relacionadas com o valor de γ , que deve ser escolhido pelo usuário antes que se incie o treinamento do classificador, tornando-o outro hiperparâmetro das Máquinas de Vetores Suporte.

$$K(\mathbf{x}, \mathbf{z}) = e^{-\gamma \|\mathbf{x} - \mathbf{z}\|^2}. \quad (22)$$

2.3.3.2 Casos de classificação multiclasse

As Máquinas de Vetores Suporte apresentadas até o momento neste trabalho têm seu uso limitado aos casos em que existam apenas duas classes. Porém, existem estratégias que permitem o uso deste, e demais classificadores binários, para casos multiclasse (que possuem três ou mais classes), sobretudo as estratégias *um-contra-um* e *um-contra-todos* (BISHOP, 2006).

Em um problema com J classes, a estratégia *um-contra-todos* consiste em treinar J classificadores binários, onde uma de suas classes é composta por todos os exemplos de uma das classes de J e outra por todas as demais. A classificação de novas amostras se dá com base na classe que tiver resultado positivo entre todos os classificadores binários.

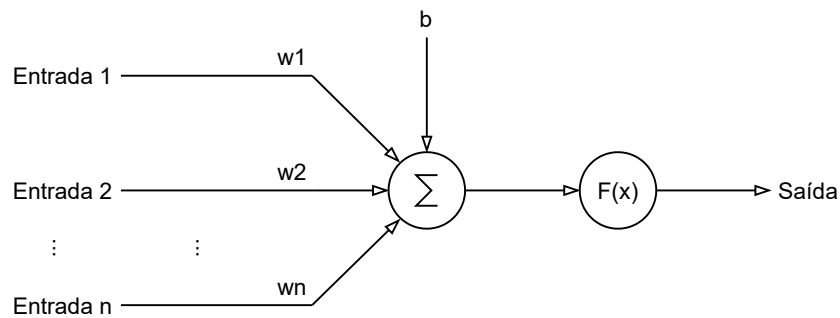
Já na estratégia *um-contra-um*, também em um problema com J classes, treinam-se $\frac{J-1}{2}$ classificadores, cujas classes são formadas com os exemplos de duas das J , de forma que sejam formados todos os pares possíveis das J classes. Neste caso, a determinação da classe

pertencente a uma nova amostra se da por votação, isto é, a classe escolhida será a que mais aparece como resultado entre todos os classificadores binários.

2.3.4 Redes Neurais Artificiais

As Redes Neurais Artificiais, ou RNAs, são modelos computacionais que visam imitar a capacidade de aprendizado humana, utilizando para isto a interligação de unidades de processamento denominadas neurônios (HAYKIN, 2001), sendo este nome derivado no neurônio biológico, o qual serviu de leve inspiração para estas unidades de processamento (DUDA *et al.*, 2000). A Figura 5 apresenta um modelo não-linear desses neurônios artificiais.

Figura 5 – Modelo de um neurônio artificial



Fonte: Autoria Própria

Na Figura 5, (w_1, w_2, \dots, w_n) são conhecidos como pesos, sendo estes valores pelos quais suas respectivas entradas são multiplicadas; $F(\mathbf{X})$ é uma função conhecida como Função de Ativação, a qual tem por objetivo restringir a saída de seu neurônio em um determinado limite; e b é um valor conhecido como viés (do inglês, *bias*), o qual amplia o grau de liberdade com o qual os pesos do neurônio podem se ajustar. A saída do modelo da Figura 5 é apresentado em (23), na qual y é a saída do neurônio e x_k com $k \in 1, 2, \dots, n$ são suas N entradas.

$$y = F(x_1w_1 + x_2w_2 + \dots + x_nw_n + b). \quad (23)$$

Diversas funções de ativações têm sido propostas na literatura, dentre as quais citam-se: A função Logística (LECUN *et al.*, 1998):

$$F(x) = \frac{1}{1 + e^{-x}}, \quad (24)$$

a função tangente hiperbólica (LECUN *et al.*, 1998):

$$F(x) = \tanh(x), \quad (25)$$

a função de ativação linear retificada (GLOROT *et al.*, 2011):

$$F(x) = \max(0, x). \quad (26)$$

2.3.4.1 Redes Perceptron Multicamadas

Do ponto de vista de classificação, um dos primeiros modelos práticos de RNAs foi o Perceptron, uma rede formada por um único neurônio como o apresentado na Figura 5. Este tipo de classificador é, porém, restrito aos casos com apenas duas classes, as quais devem ser linearmente separáveis (HAYKIN, 2001). Baseado neste conceito, surgem então as Redes Perceptron Multicamadas ou MLP (do inglês, *Multilayer Perceptron*), redes neurais artificiais compostas por diversos neurônios dispostos em camadas sequenciais, de forma que todos os neurônios de uma camada tornam-se as entradas da sua camada subsequente.

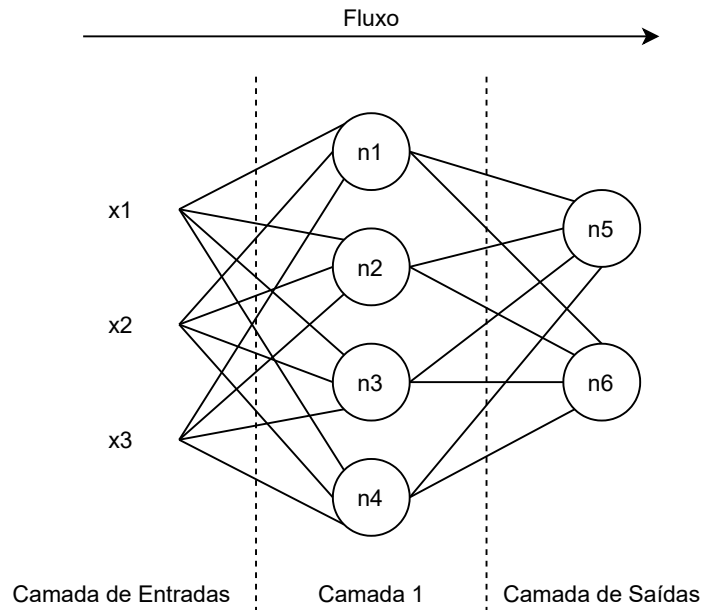
Nesta arquitetura, as camadas existentes entre as entradas e a última camada da rede, são chamadas de camadas ocultas e os neurônios destas camadas são denominados unidades ocultas. A Figura 6 apresenta um exemplo de uma arquitetura MLP com três entradas (x_1, x_2, x_3) , quatro neurônios (n_1, n_2, n_3, n_4) na sua camada oculta e 2 neurônios (n_5, n_6) na camada de saída. Ainda, o fluxo de operações nesta arquitetura é sequencial, partindo das entradas, passando por cada camada oculta e terminando na camada de saídas e é, portanto, chamado de alimentação adiante (do inglês, *feedforward*).

Devido à esta topologia mais robusta, as MLPs são capazes de reconhecer padrões muito mais complexos do que as de uma rede perceptron comum. Segundo (CHERKASSY; MULIER, 2007) as MLPs são capazes de aproximar qualquer função contínua, com uma acurácia pré-especificada.

2.3.4.2 Algoritmo da Retropropagação do Erro

Em essência o aprendizado de uma rede neural está ligado com o ajuste de seus parâmetros (pesos e *bias*). Este aprendizado ocorre com o auxílio dos chamados algoritmos de aprendizado, os quais definem como irão ocorrer os ajustes e reajustes dos parâmetros de

Figura 6 – Autoria Própria



Fonte: Exemplo de uma MLP

uma rede neural artificial durante seu treinamento (HAYKIN, 2001). No caso das MLPs, um algoritmo muito comum para aprendizado é chamado de retropropagação do erro (CHERKASSY; MULIER, 2007) (do inglês, *error backpropagation*). Em suma, o algoritmo da retropropagação do erro é um método iterativo que consiste em:

- 1 - Inicializar os pesos da rede neural (geralmente inicialização aleatória).
- 2 - Apresentar exemplos para a rede, os quais à percorreram até a camada de saídas.
- 3 - Aplicar uma função custo e calcular seu gradiente em função dos parâmetros da rede.
- 4 - Recursivamente, utilizando a regra da cadeia, atualizar os parâmetros da rede neural.
- 5 - Caso um critério de parada pré-estabelecido não tenha sido atingido, repetir a partir do Passo 2.

A atualização de cada um dos pesos, neste caso, se dá pela aplicação da chamada Regra Delta, apresentada em (27), na qual em uma iteração k , $w(k + 1)$ é o peso atualizado, $w(k)$ é o peso atual, δ é o erro e β é um valor chamado de taxa de aprendizado que determina o passo de atualização dos pesos da rede.

$$w(k + 1) = w(k) - \beta\delta(k) \quad (27)$$

Como possíveis hiperparâmetros para as MLPs, tem-se a quantidade de camadas ocultas,

a quantidade de neurônios em cada uma destas camadas, a taxa de aprendizado β , a quantidade máxima de iterações do algoritmo de aprendizado (esta geralmente chamada de épocas), a função de ativação utilizada nas camadas ocultas e, quando desejado, um parâmetro de regularização que pode ser incluso na função custo utilizada.

3 TRABALHOS CORRELATOS

O presente Capítulo apresentará alguns dos principais trabalhos da literatura a respeito do monitoramento de sistemas fotovoltaicos e de métodos de detecção e classificação de falhas, em especial os que utilizam técnicas de inteligência artificial, apresentando ainda uma comparação com o presente trabalho, de forma a apontar e destacar quais as principais contribuições oferecidas.

3.1 MONITORAMENTO DE PLANTAS FOTOVOLTAICAS

Na área de monitoramento de plantas fotovoltaicas, diversos sistemas têm sido propostos na literatura. JUCÁ *et al.* (2011), por exemplo, apresentam um sistema de monitoramento de baixo custo, de variáveis elétricas e ambientais, visando a aplicação na geração descentralizada de energia por fontes renováveis. No trabalho em questão, o sistema foi ainda implementado, como estudo de caso, em um sistema de bombeamento de água alimentado por uma pequena planta fotovoltaica. Em outra proposta de sistema de baixo custo, Eke *et al.* (2012) apresentam um sistema de monitoramento específico para plantas fotovoltaicas, com a aplicação em um único módulo fotovoltaico. Ainda nesta linha, em (SANCHEZ-PACHECO *et al.*, 2014) é proposta uma técnica de transmissão de dados obtidos em uma planta fotovoltaica por meio de um sistema de monitoramento de baixo custo, enquanto Han *et al.* (2015) utilizam técnica semelhante para monitorar plantas fotovoltaicas residenciais. Já no contexto de plantas de maiores potências, Moreno-Garcia *et al.* (2016) propõe um sistema de supervisão em tempo real de usinas fotovoltaicas de larga escala por meio de redes de sensores sem fio, que inclui, além da medição de variáveis elétricas e ambientais, a análise de qualidade da energia gerada. Este sistema foi demonstrado com a aplicação em um subsistema de geração fotovoltaica de 400kW.

3.2 FALHAS EM SISTEMAS FOTOVOLTAICOS

Um ponto importante para o estudo e avaliação da ocorrência de falhas, bem como de seus impactos, é o levantamento e detalhamento de, ao menos, as falhas mais comuns em sistemas fotovoltaicos como um todo. Neste contexto, Madeti e Singh (2017a) apresentam um extensivo estudo sobre estas falhas, dividindo-as em duas categorias: falhas do lado de corrente

contínua e falhas do lado de corrente alternada. As falhas do lado de corrente alternada são as decorrentes de problemas no(s) inversor(es) do sistema ou na rede elétrica em si. Já as falhas do lado de corrente contínua são mais numerosas e incluem: problemas com o algoritmo buscador de máxima potência, ou MPPT, falhas no diodo de *bypass*, falha de aterramento, falhas de arco, descasamento de células ou módulos (podendo este ser temporário ou permanente), circuito aberto e curto-circuito, dentre os quais, as três últimas fazem parte do escopo do presente trabalho.

A falha de circuito aberto ocorre quando, em algum ponto do sistema, há uma desconexão, fazendo com que a circulação da corrente elétrica seja interrompida. Do ponto de vista da geração de energia, esta falha é uma das que tem maior impacto, visto que pode afetar desde uma única *string*, ou conjunto, de módulos, até o sistema todo, dependendo da localização da desconexão e da topologia do sistema fotovoltaico.

O descasamento de células ou módulos ocorre quando existem, no sistema fotovoltaico, células ou módulos com propriedades elétricas muito distintas da maioria, atrapalhando seu funcionamento. A falha de descasamento pode ser dividida em duas subcategorias: temporário e permanente. O descasamento temporário é aquele causado por eventos como a deposição de poeira ou neve e por sombreamento causado por prédios ou linhas de transmissão. Já o descasamento permanente ocorre devido à degradação e danificação das células e módulos afetados. No presente trabalho, ambos os casos são levados em consideração, utilizando o sombreamento e a degradação como casos de descasamento temporário e permanente respectivamente, ambos ocorrendo em nível de módulo.

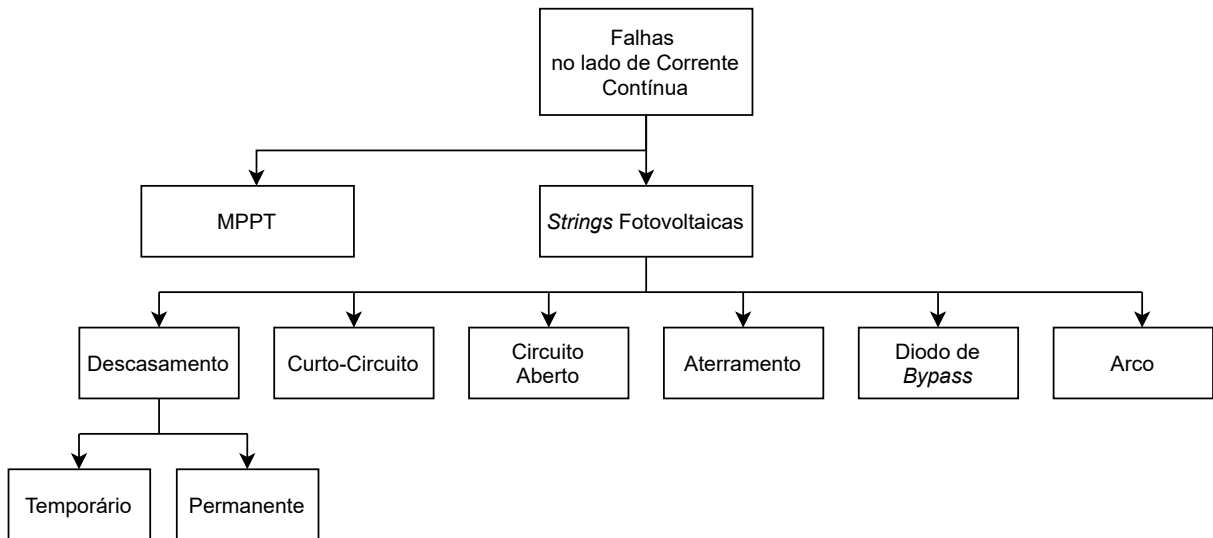
A falha de curto-circuito ocorre quando surge um caminho de baixa impedância no meio do sistema. No caso dos sistemas fotovoltaicos, isto pode ocorrer em diversos pontos como entre dois terminais de um mesmo módulo, entre dois pontos de uma mesma *string*, de duas *strings* distintas e entre uma *string* e o terra, por exemplo. Particularmente neste trabalho, será levada em consideração a ocorrência entre dois pontos de uma mesma *string*, mais especificamente entre o terminal negativo de um módulo e o terminal positivo de seu módulo adjacente.

A Figura 7 apresenta um resumo das falhas do lado de corrente contínua.

3.3 MÉTODOS VISUAIS

Métodos visuais para detecção e classificação de falhas em sistemas fotovoltaicos têm sido brevemente discutidos na literatura. Em (PACKARD *et al.*, 2012), por exemplo, foi desen-

Figura 7 – Resumo das Falhas no lado de corrente contínua



Fonte: Adaptado de (MADETI; SINGH, 2017b)

volvido um roteiro para inspeção visual de cerca de 60 módulos fotovoltaicos, visando encontrar pontos de degradação nos mesmos. Estes métodos tendem a ser baratos, visto dispensarem a compra de equipamentos. Por outro lado, estes não são facilmente escaláveis para plantas de grande porte, uma vez que a inspeção de muitos módulos pode levar um tempo considerável e necessitar de aumento de mão de obra especializada. Além disto, este método não é capaz de detectar muitos tipos de falhas, estando mais restrito a casos de degradação e sombreamento, dado que outras falhas podem não apresentar características facilmente identificáveis a olho nu.

3.4 MÉTODOS POR TERMOGRAFIA

Métodos por análise de imagens termográficas, em contraponto aos métodos visuais, vêm sendo abordados de maneira um pouco mais extensiva na literatura. Em (TSANAKAS; BOTSARIS, 2012), é apresentada uma metodologia baseada em termografia para detecção de algumas falhas em módulos fotovoltaicos. A metodologia consiste em, após calibrar a câmera termográfica com dados ambientais, tais quais a umidade e velocidade do vento no momento da coleta, verificar pontos quentes, células que apresentam temperatura consideravelmente superiores as demais. Estes pontos indicam que a célula em questão está, potencialmente, com algum tipo de falha. Seguindo a mesma linha, Hu *et al.* (2014) avaliam o comportamento e descasamento de células sob influência de falhas de circuito aberto e curto-circuito. Em (TSANAKAS *et al.*, 2015), a termografia é utilizada em conjunto com técnicas de processamento de imagens, sobretudo detecção de bordas, na detecção e localização de pontos quentes. Já em (SINHA *et al.*,

2016), a técnica de termografia é utilizada para detectar, especificamente, casos de delaminação em módulos fotovoltaicos por meio de comparação com simulações.

A utilização de termografia neste âmbito torna-se interessante por se tratar de uma técnica não intrusiva e não destrutiva, que pode detectar a presença e localização de algumas falhas, mesmo que não sejam detectáveis a olho nu. Porém, esta técnica possui algumas desvantagens relevantes. Apesar de detectar e localizar algumas falhas, utilizando apenas a termografia, não é possível classificá-las. Outro ponto é que câmeras termográficas tendem a possuir um preço elevado, além de serem sensíveis à variáveis como umidade e vento, o que pode acarretar na necessidade de reconfigurações constantes durante uma inspeção prolongada. Ainda, é preciso utilizar a câmera apontada para cada módulo da planta o que, para plantas maiores, pode gerar a necessidade de pessoas se locomoverem por toda a planta ou, ainda, de estratégias como o uso de drones, como apresentado em (TSANAKAS *et al.*, 2017).

3.5 MÉTODOS MATEMÁTICOS

Outra abordagem comum no contexto de detecção e classificação de falhas em sistemas fotovoltaicos, é a utilização de métodos matemáticos. Em (CHINE *et al.*, 2014), por exemplo, é apresentado um método de detecção de falhas em sistemas fotovoltaicos conectados à rede elétrica por meio da análise dos resultados de cálculos de índice de desempenho comuns em sistemas fotovoltaicos, tais quais a taxa de desempenho (do inglês, *performance ratio*) e a produtividade.

O algoritmo proposto consiste na comparação entre os índices de desempenho da planta estudada com de uma simulação, podendo alarmar que a planta está com: falha no inversor, falha em uma *string*, falha em módulos ou erro no MPPT. Seguindo uma linha similar, Madeti e Singh (2017b) também propõem a comparação de valores medidos com simulados, porém utilizando diretamente os valores de tensão, corrente e potência em vez de índices de desempenho, além de expandir a quantidade de falhas detectáveis. As falhas visadas são: curto-circuito, diodo de *bypass* invertido, circuito aberto (com e sem diodos de *bypass*), sombreamentos parciais (com e sem diodos de *bypass* e com diodo invertido), bem como casos de circuito aberto e curto-circuito no inversor. Já em (MANSOURI *et al.*, 2018), é proposta a utilização de uma versão otimizada da média móvel exponencial combinada com a representação por *wavelet*. O método foi testado em um sistema de módulo único, verificando o seu comportamento sobre efeito de sombreamento, descasamento de células e problemas no diodo de *bypass*.

Os métodos matemáticos apresentados para detecção de falhas em sistemas fotovoltaicos, em contrapartida aos métodos por termografia, não necessitam de equipamentos caros e complexos, podendo ser implementados, geralmente, utilizando apenas medições de tensão e de corrente. Estes métodos podem ser aplicados em plantas maiores sem a necessidade de pessoal adicional ou deslocamento, além de demonstrarem bons resultados para detecção de falhas. Uma das maiores desvantagens é que os métodos apresentados carecem de boa capacidade de classificação destas falhas. Mesmo casos em que a classificação de falhas foi levada em consideração, os métodos tendem a conseguir reduzir a situação da planta a um grupo de possíveis falhas em vez de determinar qual falha está de fato ocorrendo.

3.6 INTELIGÊNCIA ARTIFICIAL

Outra forma atingir classificação e detecção de falhas que vem recebendo crescente atenção e popularidade, é a utilização de modelos de inteligência artificial, sobretudo de aprendizado de máquinas, sendo esta também a forma proposta por este trabalho.

Li *et al.* (2017), por exemplo, apresentam a utilização de redes neurais artificiais para classificação da operação de um sistema fotovoltaico em quatro estados possíveis: Normal (operação sem falha), degradação, curto-circuito e sombreamento. Este método foi treinado e testado em ambiente de simulação e obteve uma acurácia de, aproximadamente, 88,89% das novas amostras de teste. Em um outro trabalho, Chine *et al.* (2016) apresentam um sistema em dois estágios, sendo o primeiro para detecção de falhas e o segundo para classificação dos seguintes casos: circuito aberto, degradação, curto-circuito e sombreamento, considerando ainda estes casos com e sem diodo de *bypass*. Este sistema se baseia na comparação da potência da planta fotovoltaica com sua correspondente simulada para detecção de falhas e uma rede neural artificial do tipo Perceptron multicamadas para classificação, alcançando uma acurácia de 90,3%. Ainda, este sistema utiliza apenas dados simulados para o treinamento da rede e é testado com uma planta real e baseia-se na curva $V \times I$ do sistema, o que inviabiliza a classificação imediata de falhas, visto que para gerar a curva $V \times I$ é preciso desconectar o sistema e conectá-lo a um equipamento próprio. Seguindo a mesma linha, Chouay e Ouassaid (2018) apresentam um sistema similar, obtendo uma acurácia de 94,1%, porém com testes em ambiente de simulação apenas. Seguindo ainda a ideia de métodos inteligentes em dois estágios, em (GAROUDDJA *et al.*, 2017) é apresentado um sistema para detecção de três casos: Normal, circuito aberto e curto-circuito (este com dois subcasos, um com três módulos em curto-circuito e outro com

10). A abordagem de dois estágios neste trabalho se dá com a utilização de duas redes neurais probabilísticas, uma para detecção e outra para classificação das falhas citadas. Foram realizados dois testes para validação do sistema proposto, um com e outro sem ruído. No caso sem ruídos e em ambiente simulado, as acurácias dos dois estágios foram de 100%. Já para o caso com ruído artificial, as acurácias de detecção e classificação foram, respectivamente, 82,34% e 98,19%.

Em uma abordagem mais específica, em (YI; ETEMADI, 2017) é apresentado um método especializado na detecção de falhas de curto-circuito em diversas situações distintas, incluindo diferentes quantidades de módulos em situação de curto-circuito. Este sistema utiliza uma Máquina de vetores suporte, a qual é treinada com dados simulados e testada com uma planta fotovoltaica real. A validação deste método se deu com testes com 25% e 50% dos módulos da planta em curto-circuito, obtendo acurácias de 91,4% e 94,74% respectivamente. Chen *et al.* (2017), por outro lado, propõem um método utilizando uma variação de redes neurais artificiais chamada *Kernel Extreme Learning Machine* para, com base em dados da curva VxI de um sistema fotovoltaico, classificar o sistema em um dos seguintes casos: Circuito Aberto, sombreamento, curto-circuito e degradação. Para avaliação do desempenho do sistema, foram realizados três estudos de caso, um que utiliza apenas dados de simulação para treino e teste, outro com apenas dados reais e um último com dados simulados para treinamento e dados reais para teste. No caso que utiliza apenas dados simulados, o sistema proposto alcançou acurácia de 100%. No caso com dados reais, as acurácias encontradas variaram entre 97,9% e 99% e, no caso com dados misturados, a acurácia final foi de 98,9%. Em um outro trabalho, (KURUKURU *et al.*, 2019), uma rede neural artificial é combinada com a técnica de termografia para tentar identificar módulos fotovoltaicos com falhas. O trabalho leva em conta sete falhas, sem citar especificamente quais são, e alcançou 92,8% de acurácia para as falhas propostas.

3.7 CONTRIBUIÇÕES

Os trabalhos apresentados na Seção 3.6 apresentam acurácias superiores a 90%, porém possuem certas limitações. A comum utilização de dados da curva VxI do sistema estudado como características para os classificadores, auxilia no alcance de altas taxas de acurácia, porém acarretam na necessidade de que a planta seja desconectada e ligada em um equipamento próprio para análise de curva VxI para então realizar a classificação. Outra limitação presente em alguns trabalhos é a utilização de dados simulados para teste final, os quais podem ser distintos o suficiente de dados reais para impactar no desempenho do classificador numa aplicação prática.

Ainda, estes métodos não são embarcados, sendo necessário o uso constante de um computador para utilizá-los.

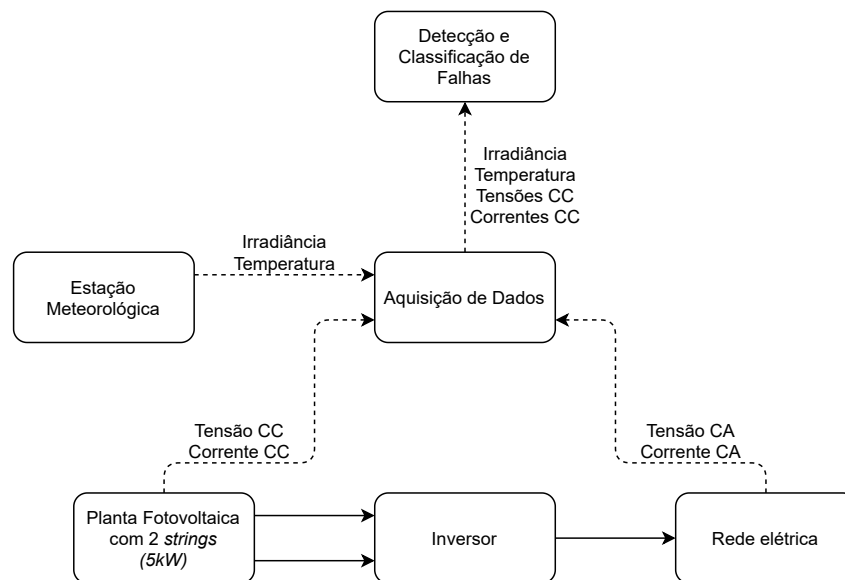
Assim, o presente trabalho propõe um método de detecção e classificação de falhas em sistemas fotovoltaicos, tendo ainda como suas principais contribuições:

- Realização da detecção e classificação não intrusiva das falhas de curto-circuito, degradação, circuito aberto e sombreamento, em sistemas fotovoltaicos.
- Uso exclusivo de dados simulados para treinamento e validação com dados reais da planta.
- Classificador embarcado em um sistema de monitoramento desenvolvido.

4 METODOLOGIA

O presente Capítulo abordará a metodologia utilizada na realização deste trabalho. Cinco classes foram escolhidas para serem abordadas: caso normal (sistema sem influência de qualquer falha), caso com módulos em curto-circuito, módulos sombreados, circuito aberto (*string* desconectada do sistema) e degradação. O caso normal foi escolhido para que houvesse também a detecção de falhas, além da classificação. Já os demais casos foram escolhidos devido ao fato de serem algumas das falhas mais abordadas na literatura, conforme pode ser visto em (CHEN *et al.*, 2017), (CHINE *et al.*, 2016), (CHOUAY; OUASSAID, 2018) e demais trabalhos apresentados no Capítulo 3. A Figura 8 apresenta uma visão geral do sistema utilizado.

Figura 8 – Visão Geral do Sistema



Fonte: Autoria Própria

Além disto, o uso dos métodos de aprendizado de máquina, apresentados no Capítulo 2, requer o uso de características que, de alguma forma, representem o caso no qual a planta fotovoltaica se encontra. Desta forma foram escolhidas, como possíveis características, dados elétricos e ambientais, de forma similar a alguns dos trabalhos correlatos apresentados no Capítulo 2. Ainda, optou-se por utilizar exclusivamente dados simulados para treinamento dos modelos, utilizando dados reais apenas para avaliação do desempenho dos classificadores. Esta decisão se dá, primariamente, por dois fatores: (i) O uso de dados de simulação acarreta na redução do tempo que a planta operará sob influência de falhas, sendo este tempo reservado apenas para coleta de dados para validação; e (ii) Simulações podem ser parametrizadas para

gerar dados com quaisquer combinações de irradiância e temperatura, muitas das quais poderiam não ser coletadas durante a elaboração deste trabalho. Além disto, o uso de dados provenientes exclusivamente de simulação, pode facilitar a implementação destes classificadores para outras plantas fotovoltaicas de forma que em vez de colocar a nova planta sob influência de cada uma das falhas e coletar dados para treinamento, seja necessária apenas uma simples reparametrização da simulação.

Dadas estas características metodológicas, a estrutura deste Capítulo se dá com a apresentação, inicialmente, do sistema fotovoltaico utilizado como objeto de estudo, bem como a forma utilizada para colocá-lo sob a influência de cada uma das falhas selecionadas. Na sequência, será descrito um sistema de monitoramento desenvolvido para aquisição de dados da planta e para implementação dos classificadores. Também será exibida a simulação criada para geração dos dados de treinamento. Finalmente, é apresentada a forma como os dados de treinamento e teste foram divididos e como foram determinados os hiperparâmetros utilizados por cada classificador.

4.1 SISTEMA FOTOVOLTAICO

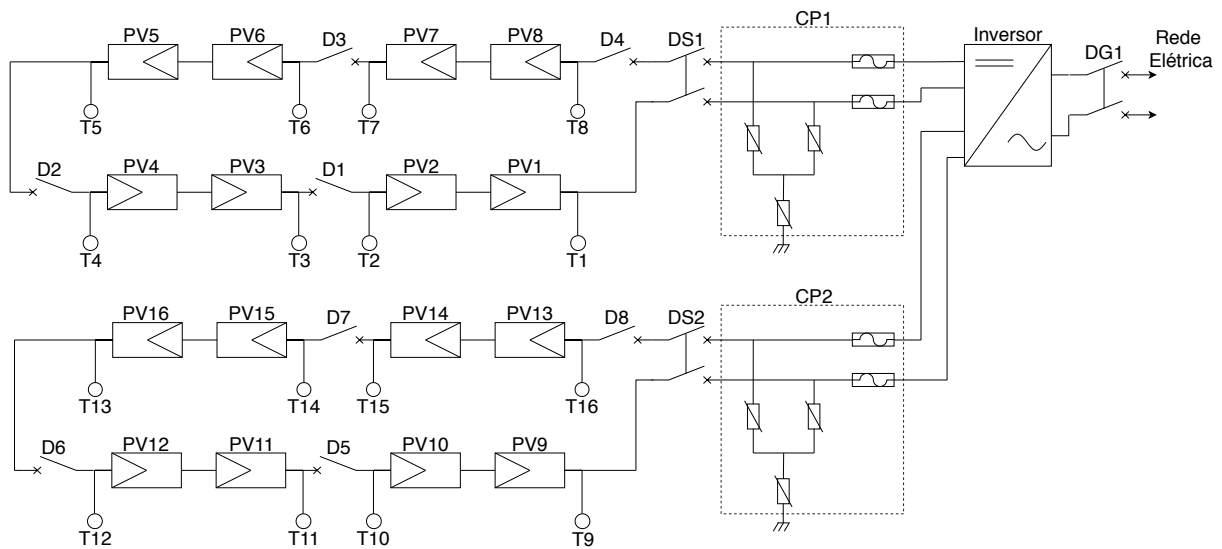
Como objeto de estudo para o este trabalho, foi utilizada uma planta fotovoltaica instalada no telhado do bloco E da sede Centro do Campus Curitiba da Universidade Tecnológica Federal do Paraná. A Figura 9 apresenta um diagrama desse sistema, o qual é composto por duas *strings* de oito módulos fotovoltaicos cada. Estes módulos (representados no diagrama por PV1, PV2, ..., PV8) são de Silício Policristalino, da fabricante Canadian Solar e modelo CS6U-330P (CANADIAN SOLAR INC, 2018). A Tabela 1 apresenta os principais dados elétricos deste módulo.

Tabela 1 – Dados Elétricos do Módulo CS6U-330P

Parâmetro	Valor
Potência nominal	330W
Tensão operacional ideal (V_m)	37.2V
Tensão de circuito aberto (V_a)	45.6V
Corrente operacional ideal (I_m)	8.88A
Corrente de curto-circuito (I_c)	9.45A
Coefficiente de temperatura p/ tensão (β)	-0.31%/°C
Coefficiente de temperatura p/ corrente (α)	0.053%/°C

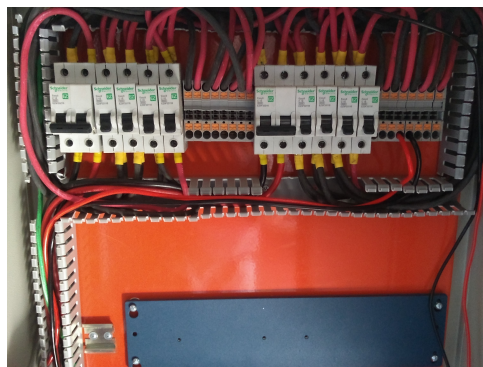
Fonte: (CANADIAN SOLAR INC, 2018)

As *strings* estão conectadas em duas entradas de um inversor 5K-GDM1 (NHS SISTEMAS ELETRÔNICOS, 2018) da fabricante NHS, sendo este próprio para instalações foto-

Figura 9 – Diagrama Elétrico do Sistema

Fonte: Autoria Própria

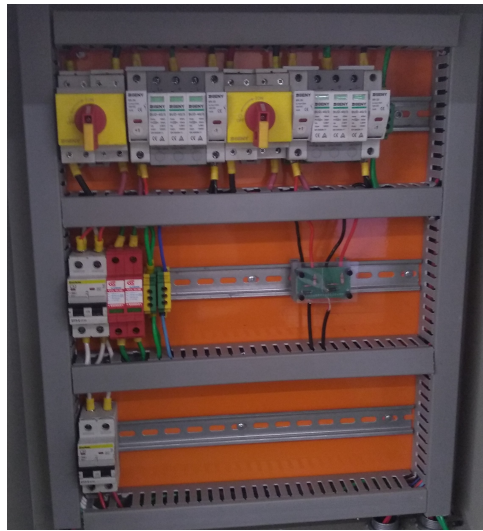
voltaicas. O inversor em questão possui 5kW de potência nominal, é do tipo conectado à rede elétrica e possui duas entradas com rastreadores individuais de máxima potência (um para cada *string*). A instalação dispõe ainda de dois quadros elétricos. O primeiro quadro, que pode ser visto na Figura 10, foi instalado próximo à planta fotovoltaica e possui disjuntores (no diagrama representados por D1, D2, ..., D8) e terminais de conexão (T1, T2, ..., T16 no diagrama) extras entre cada par de módulos, sendo estes equipamentos utilizados para isolamento dos módulos e introdução de falhas no sistema, o que será detalhado na Seção 4.1.1. Um segundo quadro, apresentado na Figura 11, foi instalado junto ao inversor e possui circuitos de proteção do sistema (CP1 e CP2 no diagrama), além de dos disjuntores gerais de cada *string* (representados por DS1 e DS2 no diagrama).

Figura 10 – Quadro com dispositivos de manobra

Fonte: Autoria Própria

A Figura 12 apresenta a planta fotovoltaica utilizada.

Figura 11 – Quadro com dispositivos de proteção



Fonte: Autoria Própria

Figura 12 – Planta Fotovoltaica Utilizada



Fonte: Autoria Própria

4.1.1 Falhas no Sistema Fotovoltaico

Dado o fato que neste trabalho a utilização dos métodos de aprendizado de máquina e a validação de seus resultados requer dados coletados com a planta fotovoltaica operando sob influência de cada uma das falhas estudadas, torna-se interessante o desenvolvimento de uma metodologia para implementação das falhas no sistema. Como já apresentado, junto à

planta foi instalado um quadro elétrico com disjuntores e terminais de conexão entre pares de módulos, de forma a facilitar a introdução das falhas de maneira segura, fácil e rápida. Em seguida será descrita a metodologia utilizada para introdução de cada falha estudada: (1) Módulos em Curto-Circuito; (2) Degradação; (3) Sombreamento; (4) Circuito Aberto.

4.1.1.1 Módulos em Curto-Circuito

A introdução da falha de curto-circuito é feita diretamente, colocando um condutor conectado entre os terminais dos módulos. Para se garantir a segurança, os módulos, nos quais o curto-circuito será introduzido, são isolados do resto do circuito com a abertura dos disjuntores adjacentes. Em seguida, os módulos em questão são cobertos com tecidos escuros, para que sua capacidade de fornecimento de corrente elétrica seja reduzida. Por fim, um condutor é conectado entre os terminais de manobra adjacentes aos módulos, os disjuntores são rearmados e a cobertura dos módulos é removida. Exemplificando, para colocar em curto o par de módulos PV5-PV6 no diagrama da Figura 9, primeiro abre-se os disjuntores D2 e D3, em seguida cobre-se os módulos PV5 e PV6 e conecta-se os terminais T5 e T6. Finalmente, os disjuntores D2 e D3 são rearmados e os tecidos removidos.

4.1.1.2 Degradação

A falha de degradação foi representada pelo aparecimento de uma resistência no meio do sistema, de maneira similar ao apresentado em (CHEN *et al.*, 2017). Para aplicação na planta foi utilizado um banco de resistores em série com as *strings*. O método para inserção é similar ao do curto-circuito. O ponto de falha é isolado utilizando os disjuntores de manobra, os módulos adjacentes ao ponto de degradação são cobertos e a associação de resistores é conectada entre dois terminais de manobra. Em seguida os disjuntores laterais são rearmados, deixando aberto apenas o disjuntor entre os terminais aos quais os resistores foram conectados, de forma que a corrente circule pelo banco de resistores e, por último, são removidas as coberturas dos módulos. Novamente utilizando o diagrama da Figura 9 para exemplificação, caso deseje-se um ponto de degradação entre os pares de módulos PV5-PV6 e PV7-PV8, abre-se os disjuntores D2, D3 e D4, cobre-se os módulos PV5, PV6, PV7 e PV8 e conecta-se a associação de resistores entre os terminais T6 e T7. Por último, os disjuntores D2 e D4 são rearmados e os tecidos removidos. A Figura 13, em (f), apresenta o banco de resistores utilizado.

4.1.1.3 Circuito Aberto

A introdução da falha de circuito aberto pode ser feita com o simples desarme manual do disjuntor (DS1 ou DS2 no diagrama) correspondente à *string* que se deseja desconectar do sistema.

4.1.1.4 Sombreamento

Diferentemente das demais falhas, o sombreamento não foi gerado artificialmente para validar os classificadores desenvolvidos para este trabalho. Isto se dá pois foram identificados, em alguns períodos do dia, sombreamentos naturais ocasionados pelos prédios que se encontram ao redor da planta fotovoltaica instalada. Com isto foi possível utilizar dados reais de sombreamento, o que permitiu, além de possuir um dado mais realista do que das falhas anteriores, geradas artificialmente, a coleta de uma quantidade muito maior de amostras deste caso.

4.2 SISTEMA DE MONITORAMENTO

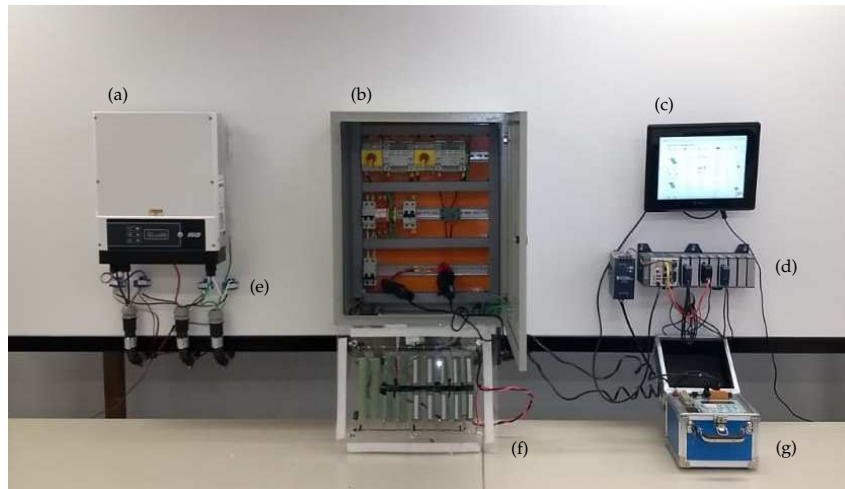
Junto à instalação da planta fotovoltaica descrita, foi desenvolvido um sistema de monitoramento das variáveis elétricas e ambientais que poderiam vir a servir na classificação das falhas selecionadas. Esse sistema foi construído utilizando um controlador CompactRIO, da National Instruments, o qual é programado utilizando o software LabVIEW. Também foram instaladas duas ferramentas adicionais na CompactRIO, sendo a primeira um banco de dados para posterior salvamento dos dados coletados e calculados e a segunda um interpretador Python para auxiliar na implementação dos classificadores no sistema embarcado. A Figura 13 apresenta a parcela do sistema de monitoramento que fica dentro do laboratório em que foi desenvolvido e implantado.

Esta seção abordará alguns dos principais aspectos, físicos e lógicos, acerca do sistema supracitado.

4.2.1 Variáveis Ambientais

Além da planta fotovoltaica em si, dado o já citado impacto de variáveis ambientais na geração fotovoltaica, também foi instalada uma estação meteorológica, a qual pode ser vista

Figura 13 – Sistema no laboratório que inclui: (a) Inversor; (b) Quadro elétrico com equipamentos de proteção; (c) Monitor utilizado para apresentar a interface homem-máquina do sistema de monitoramento; (d) CompactRIO; (e) Sensores de corrente CC e CA; (f) Banco de resistores utilizado para gerar a falha de degradação; e (g) Analisador de qualidade de energia



Fonte: Autoria Própria

na Figura 14, que mede a temperatura ambiente, umidade relativa do ar, além de velocidade e direção do vento. Além destas variáveis, também são medidas as temperaturas de dois módulos fotovoltaicos de cada *string*, utilizando sensores acoplados à suas faces inferiores, conforme apresentado na Figura 15, bem como a irradiância solar, medida com o sensor apresentado na Figura 16. A Tabela 2 apresenta os sensores utilizados na estação meteorológica.

Figura 14 – Estação Meteorológica



Fonte: Autoria Própria

Tabela 2 – Sensores da estação meteorológica

Variável	Sensor
Irradiância Solar	SP110 (APOGEE INSTRUMENTS, 2019)
Temperatura Ambiente	MeteoTemp RH+T (BARANI DESIGN, 2017)
Umidade Relativa do Ar	MeteoTemp RH+T (BARANI DESIGN, 2017)
Velocidade do Vento	FA01 (NANHUA ELECTRONICS, 2018)
Direção do Vento	FA02 (NANHUA ELECTRONICS, 2019)
Temperatura dos Módulos	SFCSD 50 (KIMO INSTRUMENTS, 2012)

Figura 15 – Medição de Temperatura



Fonte: Autora Própria

Figura 16 – Medição de Irradiância



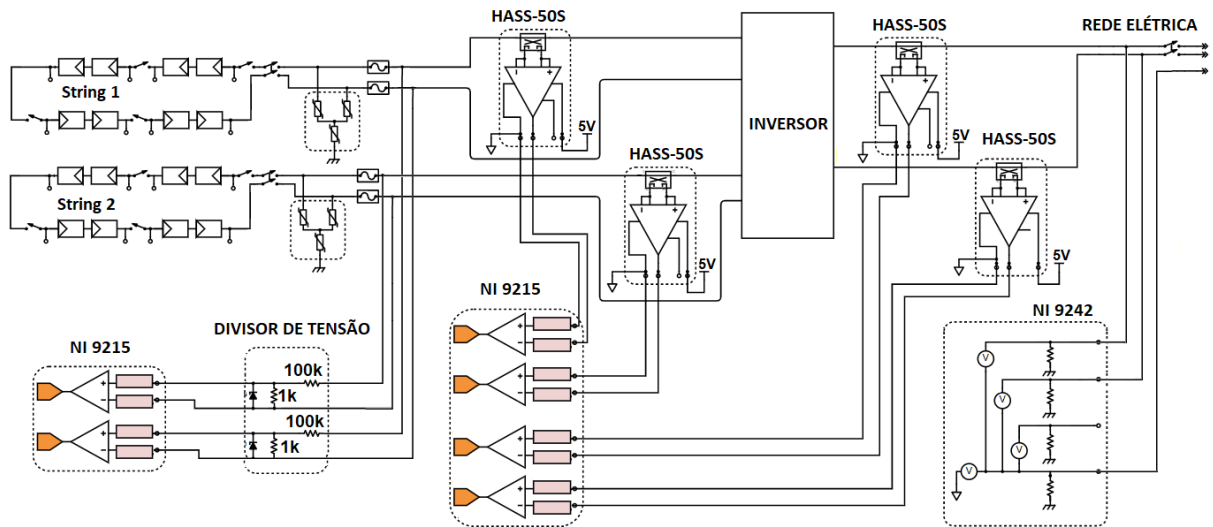
Fonte: Autoria Própria

4.2.2 Variáveis Elétricas

As variáveis elétricas coletadas pelo sistema de monitoramento desenvolvido se resumem em tensões e correntes da geração (tensões e correntes de cada *string*) e de saída do inversor (tensões e correntes de cada fase conectada à rede elétrica local).

Para as medições de tensão das *strings* foi utilizado um cartão de entradas de tensão NI9215 (NATIONAL INSTRUMENTS, 2018a) em conjunto com um circuito de condicionamento e proteção. Já para a coleta dos valores de corrente, tanto na entrada quanto saída do inversor, foram utilizados quatro transdutores de corrente HASS-50S (LEM, 2019), sendo estes também conectados a um cartão de entrada de tensão NI9215. Por fim, para a tensão de saída do inversor, esta uma tensão alternada que varia entre -180V e +180V, foi utilizado um cartão NI9242 (NATIONAL INSTRUMENTS, 2018b), o qual opera na faixa de -250V a +250V. A Figura 17 apresenta uma vista geral desta instrumentação para as variáveis elétricas.

Figura 17 – Diagrama da instrumentação para coleta das variáveis elétricas



Fonte: Adaptado de (COSTA *et al.*, 2019)

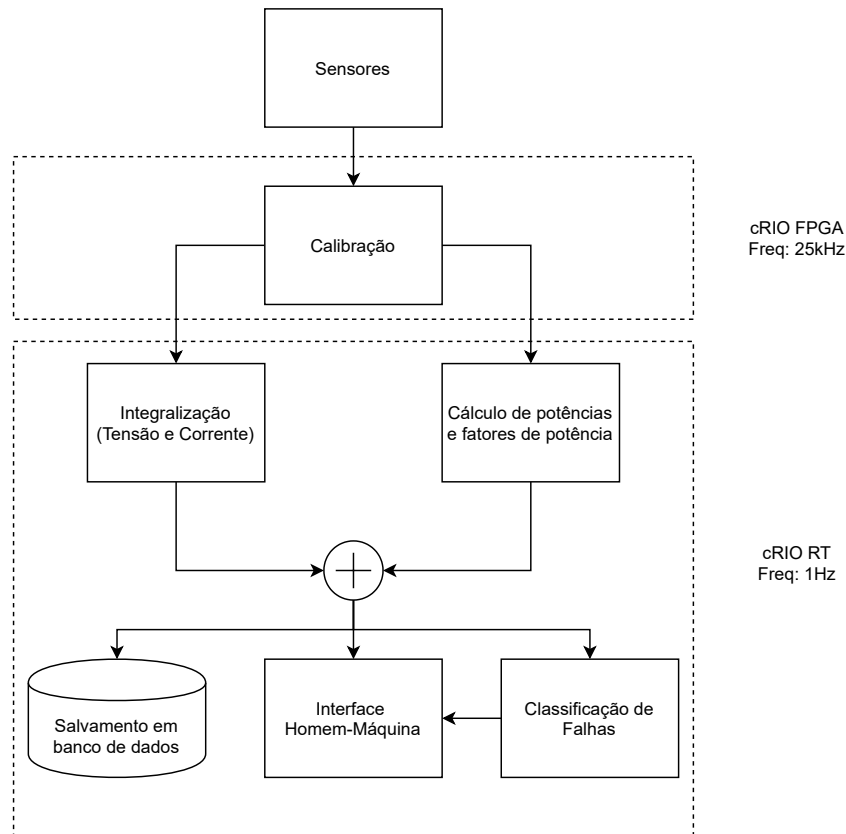
4.2.3 Lógica do sistema de monitoramento

O sistema de monitoramento possui duas partes principais, do ponto de vista lógico, sendo a primeira a parte que funciona em *FPGA* (do inglês *Field Programmable Gate Array*), e a segunda no sistema em tempo real (RT) da CompactRIO. A parte que funciona em *FPGA* é executada em uma frequência de 25kHz e consiste em, essencialmente, coletar oito sinais (tensões e correntes de cada *string* e tensões e correntes em cada fase nas saídas do inversor para a rede), executar interpolação (com base em valores determinados na etapa de calibração do sistema) e dispô-los em uma fila para posterior manipulação. Vale o apontamento de que a escolha dos 25kHz na coleta se deve a uma limitação de hardware do controlador utilizado.

Já a segunda parte é executada com uma frequência de 1Hz e é responsável por quatro macroatividades: (i) Integralização; (ii) Cálculos de potência; (iii) Salvamento; e (iv) Interface. A primeira macroatividade consiste em, tal como sua denominação sugere, integralizar as vinte e cinco mil amostras de cada um dos oito sinais coletados. Esta integralização se dá por meio de cálculos de média, para os valores em corrente contínua, e de valor eficaz, no caso dos valores em corrente alternada. Na segunda macroatividade, são realizados os cálculos de potência para cada par de tensão-corrente coletados. Para cada um dos pares de tensões da rede e correntes injetadas, por se tratarem de valores em corrente alternada, são calculados quatro valores: potência ativa, potência reativa, potência aparente e fator de potência. A terceira macroatividade é composta pelo salvamento das variáveis calculadas nas macroatividades anteriores e das variáveis ambientais no banco de dados instalado previamente na CompactRIO. A última macroatividade resume-se

em mostrar em uma interface homem-máquina todos os valores medidos, além do resultado da classificação de falhas, o que será discutido mais a frente neste trabalho. A Figura 18 apresenta o fluxo descrito.

Figura 18 – Fluxo lógico do sistema de monitoramento



Fonte: Autoria Própria

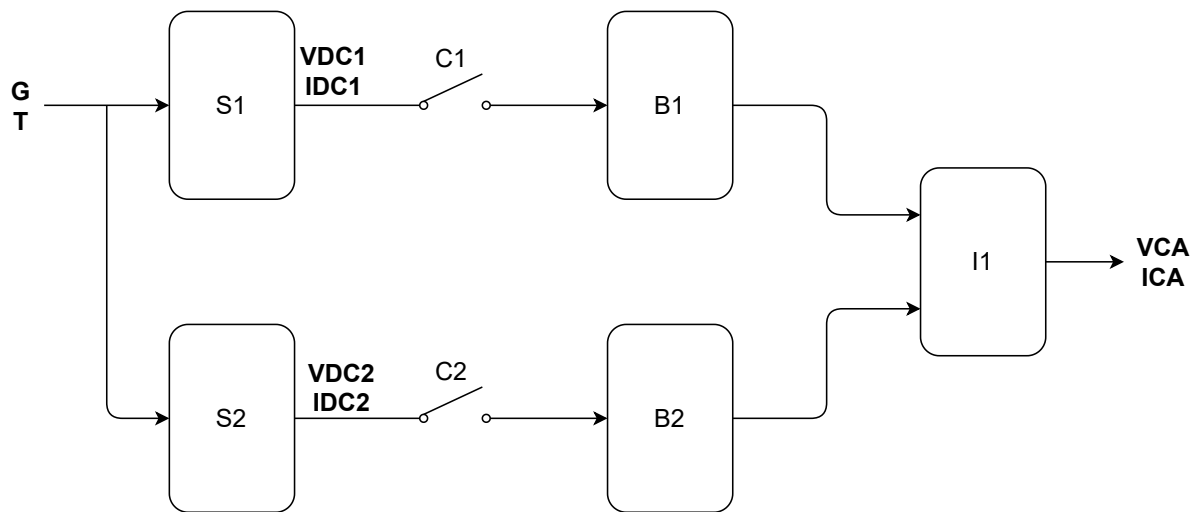
4.3 SISTEMA FOTOVOLTAICO SIMULADO

Conforme citado previamente, optou-se por desenvolver uma simulação com um circuito equivalente ao instalado fisicamente. Com isto, torna-se possível simular a planta sob diversas situações de irradiância e temperatura para cada um dos casos de falha estudados.

A citada simulação foi implementada utilizando o software PSIM em co-simulação com MATLAB - Simulink®. Em ambiente PSIM, foi construído o circuito elétrico que simula a planta descrita na Seção 4.1, enquanto o MATLAB foi utilizado para controle das variáveis e arquivos envolvidos nas simulações. A Figura 19 apresenta uma vista geral do circuito implementado na simulação.

Na Figura 19, os blocos S1 e S2, correspondem às *strings*, cada qual composta de oito módulos fotovoltaicos de 330W. Estes blocos possuem como entradas a irradiância solar (G) e

Figura 19 – Visão geral do circuito simulado



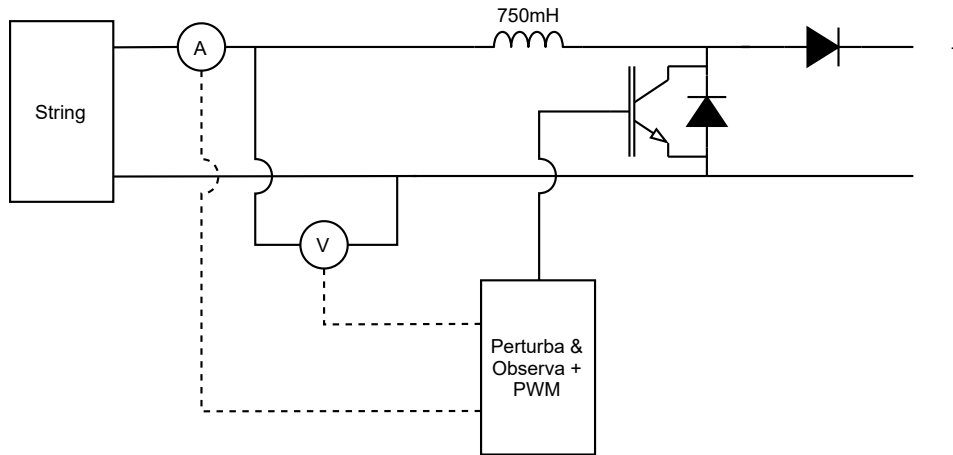
Fonte: Autoria Própria

temperatura dos módulos (T). As saídas do bloco são os valores de tensão (V) e corrente (I) da *string*. Os dois blocos posteriores, B1 e B2, são compostos por conversores do tipo elevador de tensão (*boost*) controlados por algoritmos rastreadores de máxima potência do tipo *Perturba e Observa*, conforme apresentado no Capítulo 2. A Figura 20 apresenta o circuito do conversor implementado e os valores de seus componentes. É importante notar que são utilizados os valores da tensão e da corrente da *string* pelo algoritmo do rastreador de máxima potência, o qual controla o chaveamento do conversor através da técnica de Modulação de Largura de Pulso, ou PWM (do inglês, *Pulse Width Modulation*).

O último bloco, I1, trata-se de um inversor do tipo ponte-H, conectado à rede elétrica. No presente trabalho, a rede elétrica é representada por uma fonte de tensão em corrente alternada, com tensão eficaz de 127V e frequência de 60Hz e o inversor é controlado por duas malhas de controle Proporcional-Integral, sendo uma para manter uma tensão estável entre a saída dos conversores *boost* e a entrada do inversor e outra para garantir a sincronia da corrente gerada com a tensão da rede elétrica em si. Na Figura 21 é apresentado o inversor em questão, enquanto a Figura 22 apresenta o diagrama de blocos da malha de controle, na qual *Vlink* é a tensão atual presente no capacitor existente entre os conversores *boost* e a ponte-H, o valor de 7.5, presente no bloco *Vref*, representa a tensão que se deseja neste capacitor (dividida por 100), os blocos C1 e C2 representam os controladores Proporcional-Integral (PI), os blocos M1 e M2 são multiplicadores, *Ica* e *Vca* são, respectivamente, a corrente injetada e tensão da rede elétrica. Ainda, a partir destas malhas, um PWM (bloco S) controla a abertura e fechamento das chaves da ponte-H, de forma que as chaves superior esquerda e inferior direita estejam fechadas quando

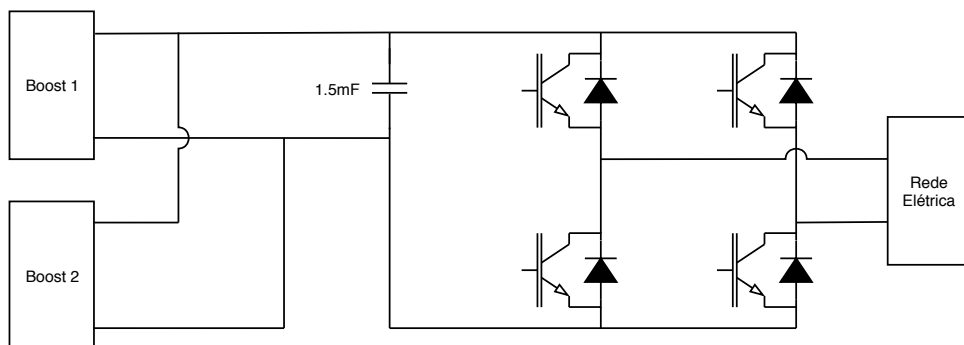
as demais estão abertas e vice-versa. A Equação 28 apresenta a função de transferência dos controladores PI utilizados, na qual k é o ganho e T é a constante de tempo do controlador.

Figura 20 – Circuito Boost implementado



Fonte: Autoria Própria

Figura 21 – Inversor implementado



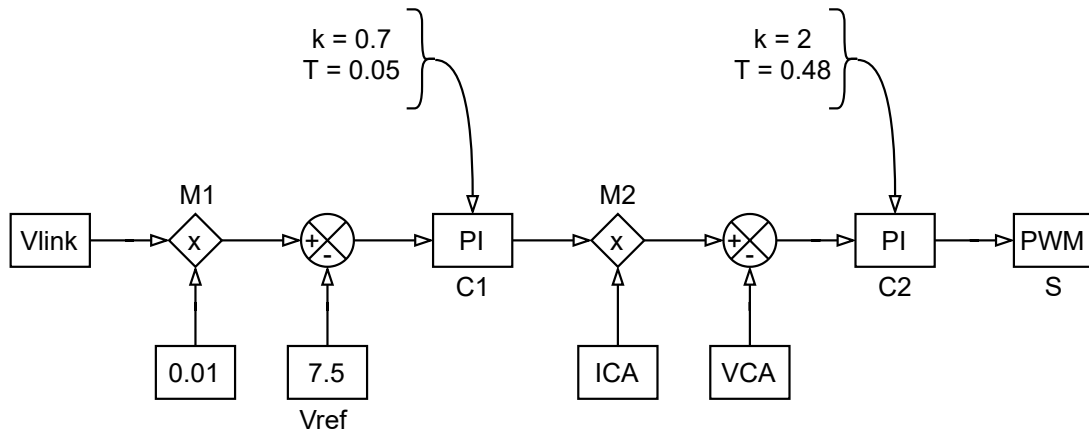
Fonte: Autoria Própria

$$G(s) = k \frac{1 + sT}{sT}. \quad (28)$$

O desenvolvimento deste controlador, bem como do inversor em si, foi baseado em alguns dos trabalhos correlatos apresentados no Capítulo 3, no contato realizado com a fabricante do inversor utilizado na planta real de 5kW e na validação que será apresentada no Capítulo 5

Para simular os módulos fotovoltaicos, foi utilizado um modelo disponibilizado pelo próprio PSIM, o qual é baseado no modelo de um diodo, apresentado na Seção 2.2. Junto ao modelo o PSIM disponibiliza uma ferramenta que pode ser utilizada, com base nas especificações técnicas do módulo fotovoltaico desejado, para calcular os parâmetros necessários para o modelo. Assim, utilizando as informações da Tabela 1 e seguindo as instruções de (POIWERSIM INC.,

Figura 22 – Diagrama de blocos do sistema de controle utilizado



Fonte: Autoria Própria

2016), os cinco parâmetros do modelo de um diodo da célula fotovoltaica foram definidos. A Tabela 3 apresenta estes parâmetros.

Tabela 3 – Parâmetros do modelo utilizados

Variável	Valor
N_s	72
n	0.95
R_s^1	$4.8265\text{m}\Omega$
R_p^1	4.3308Ω
I_s	0.04878nA
I_{ph}	9.46A

Fonte: Autoria Própria

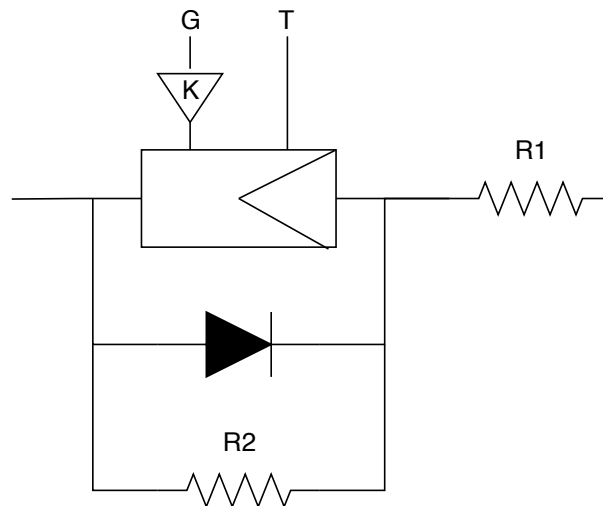
4.3.1 Geração de falhas Simuladas

Tal como na planta fotovoltaica utilizada, foram desenvolvidos procedimentos para a introdução dos casos de falha na planta simulada. Para auxiliar nesta tarefa, alguns dispositivos foram incluídos na simulação desenvolvida em PSIM, os quais foram controlados via MATLAB. Para a falha de circuito aberto, foram adicionadas chaves nas saídas de cada *string*. Estas chaves estão representadas por C1 e C2 na Figura 19. Durante a simulação, se alguma das chaves estiver aberta, a *string* respectiva estará desconectada do sistema, da mesma forma que foi realizado com o disjuntor na planta real. Para os demais casos, foram adicionados dispositivos de falha em cada um dos módulos. Para efeito de visualização, a Figura 23 apresenta os dispositivos em um módulo, na qual G e T representam as entradas de irradiância e temperatura, respectivamente. Os elementos adicionados foram: (i) o ganho K na entrada de irradiância, sendo este um valor entre

¹ Os valores da Tabela 3 referem-se ao modelo da célula. Para utilizar como modelo do módulo, é necessário que se multiplique as resistências R_s e R_p por $N_s = 72$

0 e 1 e inversamente proporcional à parcela do módulo que esta sombreada; (ii) R1, representando o banco de resistores no caso da planta real, para inserção da falha de degradação; e (iii) R2 para representação de curto-circuito. Ademais, para este trabalho foram considerados e simulados quatro casos de sombreamento, variando entre 5 e 15% de sombreamento em uma *string*.

Figura 23 – Dispositivos para introdução de falhas na simulação.



Fonte: Autoria Própria

4.4 DETECÇÃO E CLASSIFICAÇÃO DE FALHAS

Neste trabalho foram realizadas duas abordagens quanto à classificação de falhas. A primeira abordagem assume a existência de um detector ideal de falhas, realizando então somente a classificação entre as quatro falhas possíveis. Para este caso, foram excluídos os dados gerados enquanto a planta operava sem falhas em nenhuma das strings, tanto em treinamento quanto em teste. Já a segunda consiste em utilizar os classificadores para, além de classificar, detectar a ocorrência de uma falha, classificando o sistema entre os estados: Normal, Curto-Circuito, Degradação, Sombreamento e Circuito Aberto. Para ambas as abordagens foram avaliados diferentes algoritmos de aprendizado de máquina, sendo eles: kNN, Árvore de Decisão, Máquina de Vetores Suporte e Rede Neural Artificial, conforme descritos no Capítulo 2. O kNN foi escolhido por ser um dos classificadores fundamentais de aprendizado de máquina. Por se tratar de um algoritmo tão simples, permite que se avalie se é ou não necessário o investimento em classificadores mais complexos para a atividade de classificação de falhas neste contexto. Os demais classificadores foram escolhidos por serem alguns dos mais comuns para este tipo de tarefa, tendo então, em certo nível, sua aplicabilidade já comprovada.

Ainda, como características para serem utilizadas pelos classificadores, foram selecionados os dados elétricos do lado de corrente contínua, isto é, tensão da *string* 1, tensão da *string* 2, corrente da *string* 1 e corrente da *string* 2, a temperatura em nível de módulo e a irradiância incidente na planta. A escolha de se utilizar a temperatura e irradiância se deu pelo, já citado, alto impacto destas duas variáveis ambientais em sistemas fotovoltaicos. Já a escolha das tensões e correntes do lado de corrente contínua se deram por serem as que caracterizam o comportamento do sistema, tanto em operação normal quanto sob as quatro falhas selecionadas.

O processo de treinamento dos classificadores foi realizado utilizando apenas dados gerados por simulação e avaliados utilizando dados coletados na planta fotovoltaica real. Para a geração dos dados de simulação, tornou-se necessário estabelecer valores de irradiância e temperatura para a planta, de forma a abranger a maior quantidade de situações que poderiam ocorrer na planta real. Tendo isso em vista, foram utilizados valores de irradiância e temperatura compreendidos em intervalos compatíveis com a cidade de Curitiba-PR e com os módulos utilizados na planta. O extremo inferior do intervalo de irradiância foi determinado por, aproximadamente, o menor valor de irradiância em que o inversor operou desde o início da coleta de dados pelo sistema de monitoramento ($100\text{W}/\text{m}^2$), enquanto o extremo superior foi de $1000\text{W}/\text{m}^2$, que equivale à irradiância utilizada como condição padrão de ensaio para determinação da potência nominal de módulos fotovoltaicos (PINHO; GALDINO, 2014).

O intervalo de temperatura utilizado ficou entre -5°C e $+85^\circ\text{C}$ que equivalem a, respectivamente, aproximadamente a menor temperatura absoluta registrada em Curitiba entre 1981 e 2010 (INMET, 2018) e ao maior valor de temperatura de operação do módulo fotovoltaico utilizado (CANADIAN SOLAR INC, 2018). Como resultado, foram obtidos 19 valores possíveis de temperaturas (intervalo de -5°C a $+85^\circ\text{C}$ com passo 5°C) e 19 valores possíveis de irradiância (intervalo de 100 à 1000 com passo 50), culminando em um banco de dados com 361 instâncias (valores de irradiância, temperatura, correntes e tensões) para cada caso e *string*.

Já para o teste dos classificadores, foram coletados os dados da planta operando em diversas épocas do ano e em dias ensolarados, visto que em dias chuvosos e nublados não há formas de garantir que sombreamentos naturais estejam influenciando nos dados coletados. Ao todo foram coletados dezesseis dias, dentre os quais, cinco dias tiveram as falhas de degradação, curto-circuito e circuito aberto geradas, cada qual ocorrendo duas vezes no dia. Ainda, em dois outros dias apenas a falha de degradação foi gerada. A situação normal e o sombreamento natural ocorrem em todos os dezesseis dias. A Tabela 4 apresenta a quantidade de amostras obtidas

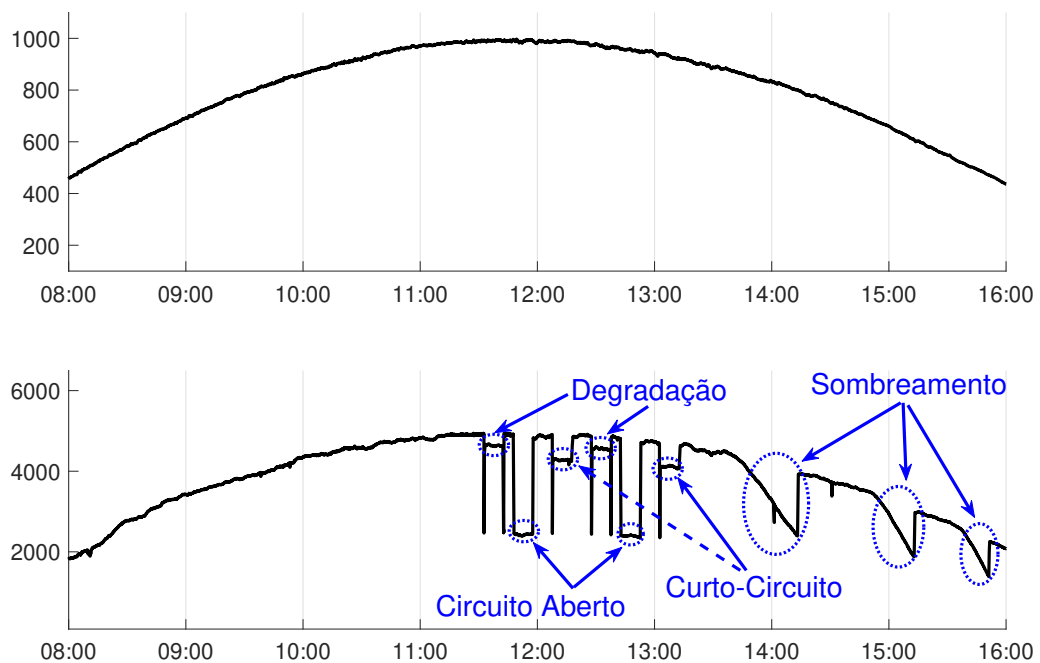
com este banco de dados e a Figura 24 apresenta um exemplo de um dos dias coletados, com a irradiância e a soma das potências das *strings*.

Caso	Quantidade de amostras [s]
Normal	309253
Curto-Circuito	5999
Degradação	10371
Circuito Aberto	6024
Sombreamento	184311

Tabela 4 – Dados para teste do classificador

Fonte: Fonte: Autoria Propria

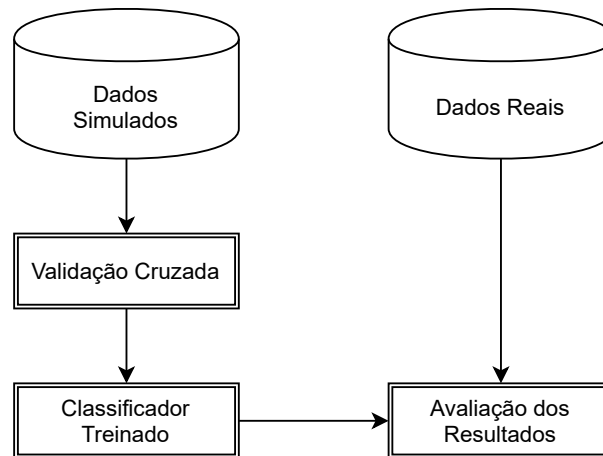
Figura 24 – Exemplo de dia do banco de dados



Fonte: Autoria Própria

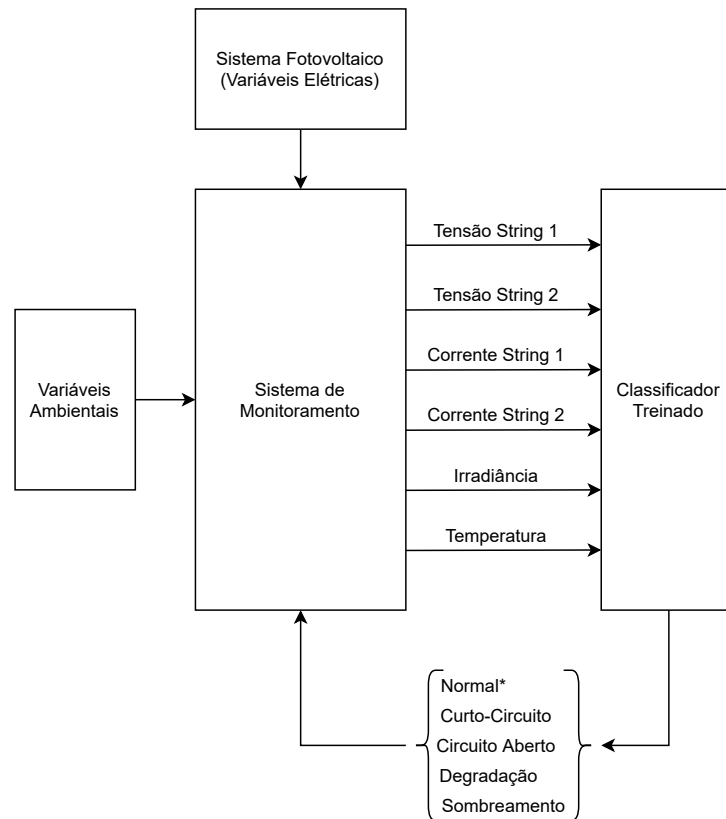
A Figura 25 apresenta o esquema de treinamento e teste dos classificadores, com o treinamento dos modelos utilizando dados simulados e o método de validação cruzada e avaliação com dados reais, enquanto a Figura 26 apresenta o esquema de classificação implementado. É possível observar que o sistema de monitoramento lê as variáveis elétricas e ambientais, transfere as tensões e correntes das strings, além da irradiância e temperatura para o classificador treinado que retorna para o sistema de monitoramento qual o estado atual da planta.

Figura 25 – Esquema de Treinamento e Teste dos Classificadores



Fonte: Autoria Própria

Figura 26 – Esquema de Classificação



*Apenas na abordagem com detecção

Fonte: Autoria Própria

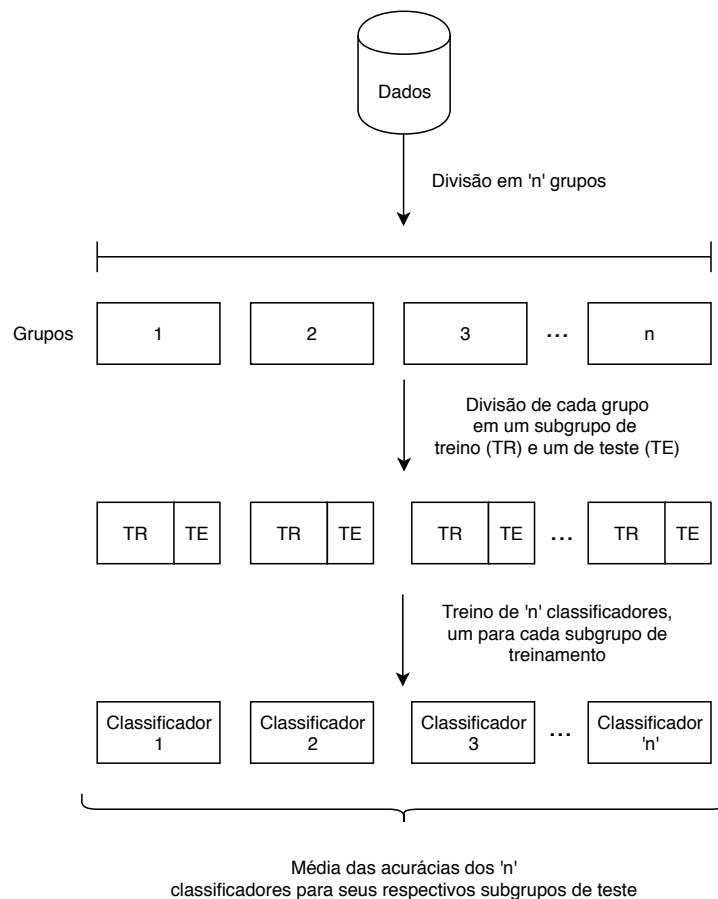
4.4.1 Determinação dos hiperparâmetros

Como já apontado, um ponto de suma importância no treinamento dos classificadores é a determinação dos hiperparâmetros a serem utilizados. Neste trabalho os hiperparâmetros

foram determinados utilizando um método de busca exaustiva conhecido por *Grid Search* (HSU *et al.*, 2010), que consiste em treinar um tipo de classificador com todos os valores, dentro de um determinado intervalo, no seu hiperparâmetro. Em casos com dois ou mais hiperparâmetros, o treinamento é realizado com todas as combinações possíveis dos valores dentro dos seus respectivos intervalos.

Para avaliar o desempenho dos classificadores durante o processo de busca citado, foi utilizado um método de Validação Cruzada (CHERKASSY; MULIER, 2007), ilustrado na Figura 27. A Validação Cruzada consiste em dividir os dados em diversos grupos, os quais são novamente divididos em dois subgrupos, um para treino dos classificadores e outro para avaliação de seus desempenhos. Por último, utiliza-se uma métrica, geralmente a média das acurácias dos classificadores treinados com os dados de cada grupo, para avaliar o resultado final para os hiperparâmetros escolhidos.

Figura 27 – Processo de Validação Cruzada



Fonte: Autoria Própria

Neste trabalho foi realizada a divisão dos dados simulados em 10 grupos, cada qual com um subgrupo com 80% dos dados para treino e outro complementar para teste. Valem os

apontamentos de que foram mantidas as quantidades de dados de cada classe tanto na divisão nos 10 grupos quanto da divisão em subgrupos de treino e teste. Como métrica de desempenho foi utilizada a média das acurácias por classe de cada grupo.

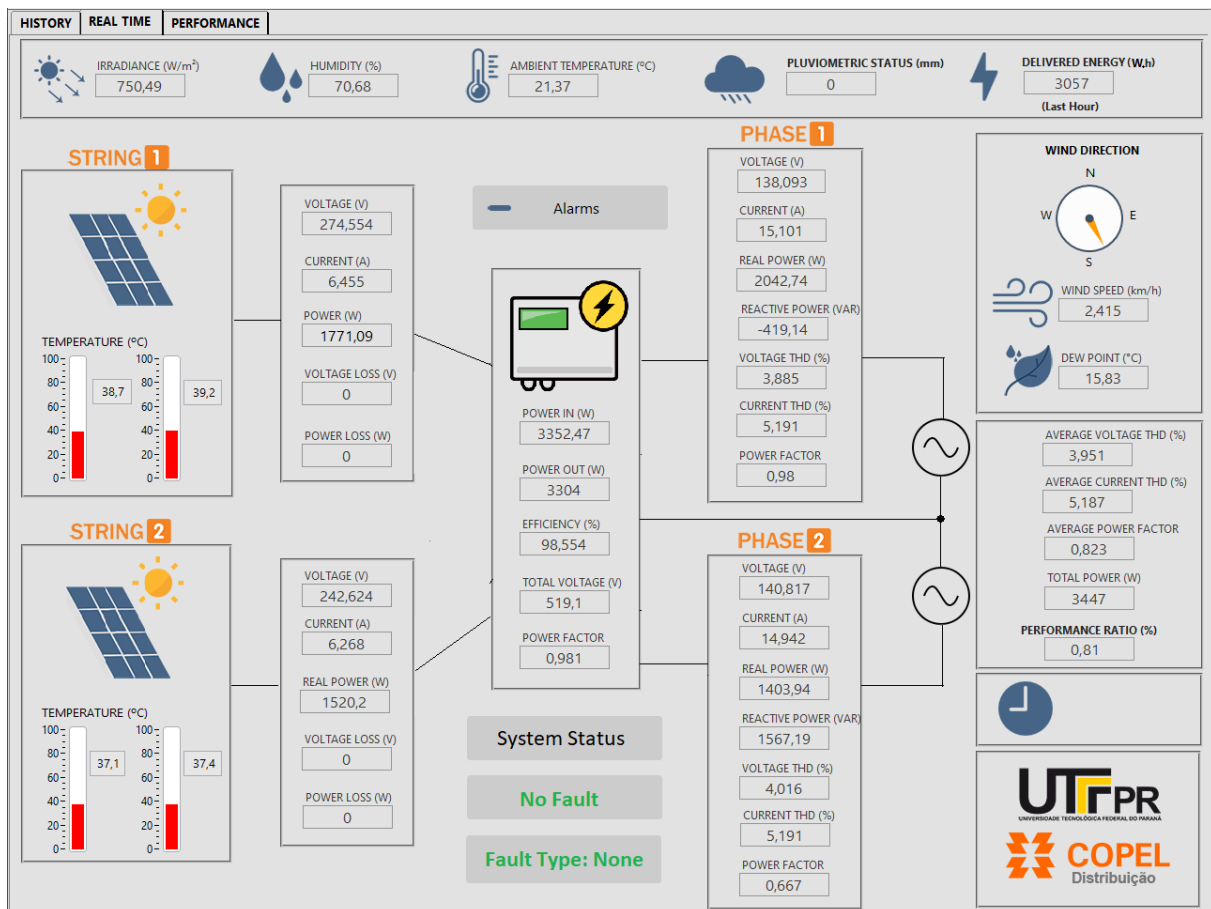
4.4.2 Implementação dos classificadores no sistema de monitoramento

A implementação destes classificadores foi realizada em ambiente 'scikit-learn', em sua versão 0.18.3, uma biblioteca Python para auxílio no desenvolvimento de modelos de aprendizagem de máquina. Primeiramente foram escritos os códigos para divisão dos dados de treino (simulados) em seus grupos/subgrupos para validação cruzada. Na sequência, os dados de treino foram normalizados e os de teste foram normalizados utilizando os mesmos parâmetros de normalização do grupo de treino. Posteriormente foram estabelecidos os intervalos dos hiperparâmetros de cada classificador para aplicação da técnica de *Gridsearch* que, foi então utilizada em conjunto com a validação cruzada, treinando cada classificador com cada combinação de seus hiperparâmetros.

Finalmente, o modelo com maior acurácia por grupo foi selecionado para ser implementado. Com isto, foi salvo um arquivo contendo modelo e os parâmetros de normalização utilizados, os quais foram transferidos para a CompactRIO. Ainda, a cada segundo, após integralização dos dados coletados, é criado um vetor de características a partir dos dados integralizados naquele segundo, o qual é normalizado e apresentado para o modelo do classificador salvo. O resultado do classificador é então apresentado na interface homem-máquina do sistema de monitoramento, indicado o caso atual. Ainda, como na prática não há um sistema ideal de identificação de falhas, apenas o resultado da abordagem que realiza tanto a detecção quanto a classificação foi implementada no sistema de monitoramento. A Figura 28¹ apresenta a interface homem-máquina do sistema de monitoramento proposto com o classificador embarcado.

¹ Esta tela é parte do desenvolvimento do sistema de monitoramento do projeto de PD 2866-0464/2017

Figura 28 – Interface do sistema de monitoramento.



Fonte: Autoria Própria

5 RESULTADOS E ANÁLISES

O presente Capítulo apresentará os principais resultados obtidos no desenvolvimento deste trabalho. Inicialmente será apresentada a validação da simulação desenvolvida, em relação à planta real instalada. Na sequência, será brevemente explicado como foi realizada a calibração do sistema de monitoramento e então demonstrado seu resultado. Por fim, serão apresentados os resultados da classificação de falhas, incluindo uma avaliação comparativa dos desempenhos dos classificadores e a escolha do classificador final para ser embarcado.

5.1 VALIDAÇÃO DA SIMULAÇÃO

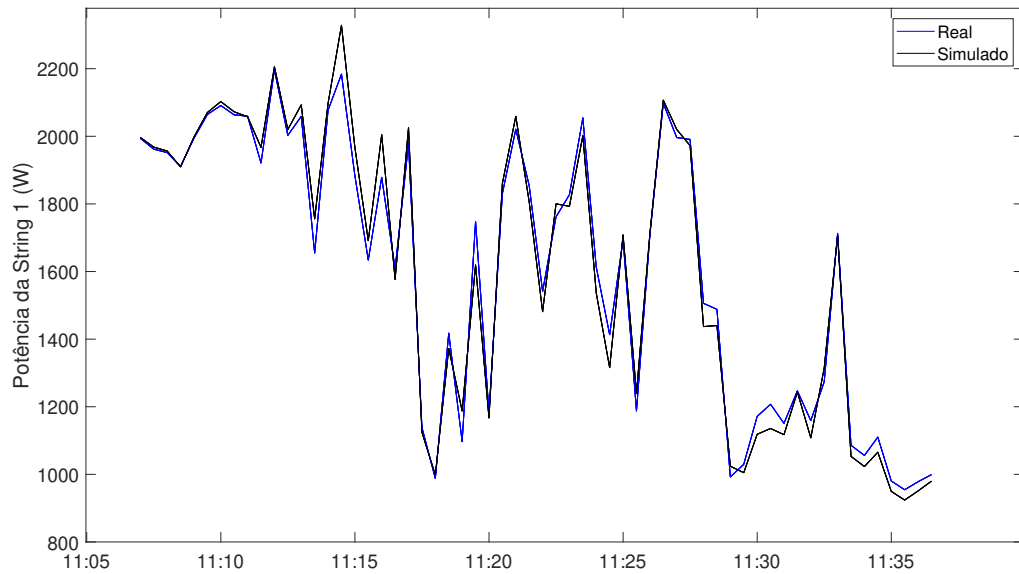
Uma vez construídas as simulações, torna-se necessário verificar se seu comportamento condiz com a planta fotovoltaica real. Este processo de validação foi realizado da seguinte maneira: Foram coletados os dados de irradiância, temperatura, tensões e correntes das *strings* com a planta real operando em cada um dos casos estudados. Com os dados de irradiância e temperatura, foram simulados os mesmos casos e os dados elétricos simulados foram comparados com os coletados na planta real. Com isto, o erro percentual médio absoluto da potência de todos os casos ficou inferior à 3%. A Figura 29 apresenta a comparação entre a potência real da *string* 1 e sua equivalente simulada para o pior caso (erro percentual médio absoluto = 2,588%), enquanto a Tabela 5 apresentam os erros para cada caso, com exceção do circuito aberto, visto sua potência ser essencialmente zero.

Tabela 5 – Resultado das validações

Caso	Erro - Potência CC
Normal	2.588%
Curto-Circuito	0.5746%
Degradação	0.3501%
Sombreamento	2.3354%

É possível observar na Figura 29 que em alguns pontos há um erro um pouco mais relevante. Diversos fatores podem ocasionar estes acontecimentos. Um destes fatores, sobretudo em pontos de maior variação de irradiância, pode ser a diferença entre os algoritmos de MPPT implementados no inversor real e simulado. Outros fatores que podem contribuir são: distribuição desigual da irradiância na planta (alguns módulos com menos ou mais irradiância dos que a lida pelo sensor de irradiância), distribuição desigual de temperatura nos módulos da planta e

Figura 29 – Comparação entre potência real e simulada.



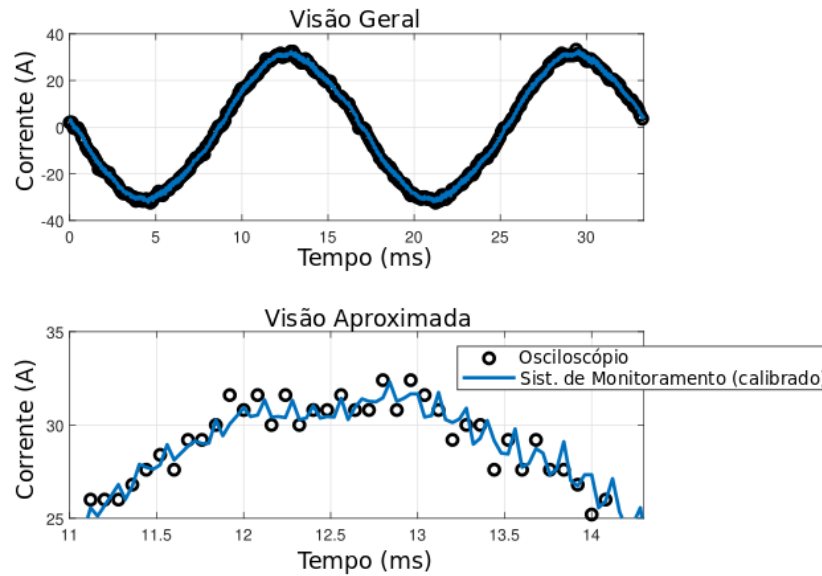
Fonte: Autoria Própria

sensibilidade dos sensores

5.2 VALIDAÇÃO DO SISTEMA DE MONITORAMENTO

Ao termino da instrumentação do sistema, o mesmo foi calibrado utilizando um osciloscópio Tektronix MDO3014 com certificado de calibração rastreável válido. Tal procedimento seguiu os seguintes passos (para cada variável medida): Utilizando um cartão de saída de tensão da CompactRIO foram iniciadas as medições sincronizadas da variável a ser calibrada, tanto no osciloscópio quanto no sistema de monitoramento. Estas medições foram então importadas para o MATLAB onde foi encontrada, via pesquisa binária, a constante que ao ser multiplicada pela medição do sistema de monitoramento, minimizava o erro quadrático médio entre a leitura dos equipamentos. Vale ressaltar que ambos os sinais foram filtrados com um filtro passa-baixa com frequência de corte de 1kHz. O procedimento foi repetido cinco vezes para cada variável, e a média destas repetições foram tomadas como fator de calibração de cada grandeza. A Figura 30 apresenta 40ms de dados da corrente alternada, tanto leitura do osciloscópio quanto do sistema de monitoramento calibrado.

Figura 30 – Corrente Alternada - Osciloscópio e Sistema de Monitoramento (Calibrado)



Fonte: Autoria Própria

5.3 CLASSIFICAÇÃO DE FALHAS

Conforme citado previamente, foram realizadas duas abordagens neste trabalho. A primeira abordagem leva em conta a detecção de falhas na planta, além da classificação, enquanto na segunda foi considerado um caso em que a detecção teria sido previamente realizada, restando ao classificador apenas o trabalho de classificar qual a falha estaria ocorrendo. Para ambos os casos, os hiperparâmetros e os seus intervalos de busca, via a técnica de *Gridsearch* foram os mesmos. Para o k-Vizinhos mais próximos foi utilizada a distância euclidiana e o hiperparâmetro k , que determina quantos vizinhos mais próximos devem ser considerados durante a classificação foi variada entre 2 e 150, com passo 1. Já a árvore de decisão foi treinada com a combinação de dois hiperparâmetros: a profundidade máxima da árvore e a quantidade máxima de 'nós folha', isto é, nós que não possuem novos caminhos a seguir, determinando então uma classe. A variação de ambos os parâmetros foi a mesma do hiperparâmetro do k-Vizinhos mais próximos (2, 3, ..., 150).

Na máquina de vetores suporte, foi utilizado um *kernel* RBF, conforme explicado na Seção 2.3.3. O hiperparâmetro C foi variado em: (0.5, 1, 2, 5, 10, 15, 20, 25), enquanto γ em: (0.1, 0.5, 1, 2, 3, 5, 12, 15). Finalmente, a rede neural artificial foi construída com uma única camada oculta, com a quantidade de neurônios variando entre 1 e 30 e com a função de ativação linear retificada. A escolha desta função de ativação se deu por se tratar de uma função que faz com que a rede neural tenda a convergir mais rapidamente.

Após os processos conjuntos de validação cruzada e *gridsearch* a melhor instância de cada tipo de classificador, isto é, a que obteve a maior acurácia por classe com os dados de treino (dados simulados), foi testada com os dados reais da planta fotovoltaica.

5.3.1 Abordagem 1 - Somente classificação

Para a abordagem que assume a detecção ideal, realizada previamente à classificação, os dados gerados com a planta operando com ausência de qualquer falha foram desconsiderados, tanto em treino quanto em teste. Sendo assim o banco de dados utilizado possui 206705 amostras, dentre as quais 5999 pertencem à classe de curto-circuito, 10371 à de degradação, 6024 à de circuito-aberto e as demais à classe de sombreamento. Os valores dos hiperparâmetros para os melhores classificadores foram: (i) $k = 13$, no k-Vizinhos mais próximos; (ii) Profundidade máxima = 17 e Quantidade máxima de nós folhas = 144, na árvore de decisão; (iii) $C = 25$ e $\gamma = 5$ na Máquina de vetores suporte; e (iv) 21 Neurônios na Rede neural artificial. A Tabela 6 apresenta, o melhor classificador de cada tipo, incluindo suas acurácias médias por classe de treino e teste.

Classificador	Acurácia de Treino	Acurácia de Teste
k-Vizinhos mais próximos	86.16%	89.96%
Árvore de Decisão	85.59%	69.70%
Máquina de Vetores Suporte	93.55%	94.59%
Rede Neural Artificial	97.19%	95.44%

Tabela 6 – Resultado dos classificadores para abordagem com apenas classificação

Fonte: Fonte: Autoria Propria

É possível observar alguns pontos relevantes na Tabela 6. Um destes pontos é o fato de que, para o k-Vizinhos mais próximos e a Máquina de vetores suporte obteve um desempenho superior com os dados de teste do que com os dados de treino. Uma hipótese que justificaria este acontecimento é que há, nas simulações, dados em condições mais extremas de temperatura, como por exemplo a planta operando com temperaturas muito baixas (até -5°C) e muito altas (até $+85^{\circ}\text{C}$), o que não ocorre nos dados coletados na planta real.

Outro ponto relevante é que a Árvore de Decisão teve um desempenho muito inferior ao segundo pior classificador (k-Vizinhos mais próximos) quando aplicado ao grupo de teste, ainda que ambos tenham tido um desempenho similar no treinamento. Ainda, a Máquina de vetores suporte e a Rede neural artificial demonstram desempenhos similares, sendo o segundo o superior, com acerto de 97% no grupo de treino e um acerto superior a 95% no de testes. A

matriz de confusão desta rede neural, para os dados de teste, pode ser vista na Figura 31.

Figura 31 – Matriz de confusão da RNA para a abordagem de classificação.

		Valor Previsto				Acurácia da Classe
		Curto-Circuito	Degradação	Circuito Aberto	Sombreamento	
Valor Verdadeiro	Curto-Circuito	5960	1	4	34	99.34%
	Degradação	600	9664	24	83	93.18%
	Circuito Aberto	0	0	6024	0	100%
	Sombreamento	10892	8864	65	164490	89.24%
Média						95.44%

Fonte: Autoria Própria

Pode ser visto na matriz de confusão que a maior parte de erros está relacionada à classe de sombreamento, sobretudo com o classificador indicando amostras com sombreamento como pertencente às classes de curto-circuito e degradação. Isto pode ser explicado pela imprevisibilidade e alto grau de variabilidade no sombreamento, visto que existem diversas formas que pode ocorrer, afetando diferentes quantidade de células e em diferentes módulos. Com isto, é possível que o aumento de casos de sombreamento sendo simulados acarrete também no aumento da acurácia para este classificador.

5.3.2 Abordagem 2 - Inclusão da detecção

Visto não haver um detector ideal de falhas, foi considerado também a utilização dos mesmos classificadores para a detecção das falhas, considerando para isso a existência da classe normal, ou sem falhas. Nesta abordagem os hiperparâmetros para os melhores classificadores foram: (i) $k = 17$, no k -Vizinhos mais próximos; (ii) Profundidade máxima = 25 e Quantidade máxima de nós folhas = 149, na árvore de decisão; (iii) $C = 25$ e $\gamma = 5$ na Máquina de vetores suporte; e (iv) 24 Neurônios na Rede neural artificial. A Tabela 7 apresenta os resultados de cada classificador para esta abordagem.

Pode se observar que o melhor caso continua sendo a rede neural, com um decréscimo

Classificador	Acurácia de Treino	Acurácia de Teste
k-Vizinhos mais próximos	74.60%	76.48%
Árvore de Decisão	84.78%	68.09%
Máquina de Vetores Suporte	93.16%	80.89%
Rede Neural Artificial	97.02%	92.08%

Tabela 7 – Resultado dos classificadores para abordagem com detecção e classificação

Fonte: Fonte: Autorial Propria

de 0.17% no grupo de treinamento e de aproximadamente 3% para o grupo de testes. O pior classificador do caso anterior, a árvore de decisão, teve um decréscimo de 0.81% para o grupo de treino e de 1.61% para o grupo de testes, mantendo-se o pior caso. O k-Vizinhos mais próximos teve decréscimos superiores à 10% para ambos os grupos e a máquina de vetores suporte manteve uma acurácia similar de treino, porém com uma queda de pouco menos de 15% para teste. Por se manter o melhor classificador para esta abordagem, a rede neural artificial com 24 neurônios em sua camada oculta, foi escolhida e implementada no sistema de monitoramento. A Figura 32 apresenta a matriz de confusão para este classificador quanto aplicado aos dados de teste.

Figura 32 – Matriz de confusão da RNA para a abordagem de classificação com detecção.

		Valor Previsto					Acurácia da Classe
		Normal	Curto-Circuito	Degradação	Circuito Aberto	Sombreamento	
Valor Verdadeiro	Normal	239522	2887	11048	2	55794	77.45%
	Curto-Circuito	2	5952	16	0	29	99.21%
	Degradação	15	33	10221	24	78	98.55%
	Circuito Aberto	0	0	0	6024	0	100%
	Sombreamento	4821	4317	18074	101	156998	85.18%
Média						92.08%	

Fonte: Autorial Própria

Percebe-se na matriz de confusão, que neste caso, além dos erros por se predizer amostras de sombreamento como pertencentes às classes de degradação e curto-circuito, observados na primeira abordagem, há também 2887 e 11048 amostras normais que foram preditas como de curto-circuito e degradação respectivamente e 55794 amostras normais preditas como sombrea-

mento. Além da já citada variabilidade do sombreamento, outro fator que pode explicar este fato é que, quando a planta está em baixa geração, o impacto de uma falha pode ser muito pequeno o que dificulta a diferenciação da operação normal. Isto pode valer especialmente para os casos de um sombreamento pequeno e degradação, que por se tratar de uma resistência, tem seu impacto (queda de potência) dependente da corrente sendo gerada. Outro aspecto a ser observado é que esta proporção de falsos negativos, sobretudo no caso de sombreamento, pode problematizar a aplicação prática do classificador, que pode acabar disparando o alarme de falhas muitas vezes. Uma das formas que pode ser utilizada para atenuar esta ocorrência é considerar um tempo maior de dados para fazer a detecção e classificação, visto que pode reduzir a ocorrência de, por exemplo, alertas de sombreamentos quando há apenas a ocorrência de uma nuvem passageira.

5.3.3 Comparação com resultados da literatura

Um ponto que deve ser considerado para a comparação dos resultados obtidos neste trabalho com os disponíveis na literatura é que não há uma padronização de elementos e configurações essenciais nos sistemas. As diferenças em plantas, seja pelo uso de diferentes inversores, tipos e quantidades de módulos e/ou topologia, bem como as condições ambientais do local de instalação e a não padronização da forma como são geradas as falhas, acarretam no uso de diferentes bancos de dados por cada autor.

Desta forma, para realizar uma comparação mais adequada dos métodos, foram selecionados trabalhos que realizam a detecção e classificação na mesma gama de falhas do que o presente trabalho. O primeiro destes, apresentado em (CHINE *et al.*, 2016), utilizou um sistema com dois estágios, sendo um para detecção e outro para classificação de falhas. Utilizando exclusivamente dados simulados para treino e dados reais para teste, a rede neural utilizada atingiu 90.3% de acurácia de classificação, desconsiderando a de detecção. Este sistema, porém, requer a desconexão da planta para realizar a classificação, visto que utiliza dados da curva V_{xI} da planta.

Em um segundo trabalho, (CHEN *et al.*, 2017), alcançou acurácias que variaram entre 97.9 e 99%. Porém, apesar do excelente desempenho, este trabalho também utiliza a curva V_{xI} da planta fotovoltaica, inviabilizando a detecção e classificação não intrusiva das falhas. Outro ponto relevante é que são utilizados casos controlados de sombreamento.

No caso do presente trabalho, a acurácia alcançada foi de 92.08%, considerando tanto detecção quanto classificação das falhas. Além disto, estas podem ser executadas em conjunto

com a operação normal da planta, sem que haja necessidade de desconexões e simulações adicionais. Ainda, diferentemente dos casos apresentados em (CHEN *et al.*, 2017) e (CHINE *et al.*, 2016), que utilizam um sombreamento controlado, o presente trabalho teve seu classificador testado com sombreamentos naturais que ocorreram na planta, o que dificulta consideravelmente sua caracterização, visto haver centenas de possibilidades de sombreamento, variando em intensidade e em quantidade de células, módulos e *strings* afetadas.

6 CONCLUSÃO E TRABALHOS FUTUROS

A maior parte dos sistemas de energia solar fotovoltaica ficam instalados em ambiente desprotegido, estando então suscetíveis às influências de diversos fatores que aumentam a chance de ocorrência de danos e falhas. Estas falhas, além de prejudicar a geração da energia elétrica, podem acarretar em riscos de redução da vida útil dos equipamentos e até mesmo incêndios. Com isto, torna-se cada vez mais importante que se encontrem e desenvolvam formas efetivas de se localizar e identificar estas falhas. Isto pode ser evidenciado pelo significativo avanço verificado na literatura acerca do assunto, com diversas metodologias propostas e aprimoradas para este fim. Neste contexto, o presente trabalho visou a união de algumas das principais vantagens apresentadas na literatura, além de contribuição própria, com um método embarcado de classificação não intrusivo de falhas em sistemas fotovoltaicos conectados à rede elétrica.

Inicialmente foi desenvolvida e validada uma simulação capaz de representar o funcionamento da planta fotovoltaica utilizada como objeto de estudo, tanto em operação normal quanto sob influência de quatro falhas comuns na literatura. Na sequência, modelos de k-Vizinhos mais próximos, árvore de decisão, máquinas de vetores suporte e redes neurais artificiais (perceptron multicamadas), foram treinados utilizando dados gerados pela simulação supracitada e validados a partir de dados reais coletados na planta instalada, havendo ainda a comparação de seus desempenhos. Também foi desenvolvido um sistema de monitoramento para o sistema fotovoltaico, no qual foi embarcada a rede neural artificial, classificador que obteve 97.02% de acurácia para os dados de treino e 92.08% para teste.

Em relação a outros métodos propostos na literatura, o sistema do presente trabalho apresentou um desempenho competitivo, com acurácia de 92.08% para dados reais, englobando ainda vantagens e características importantes, dentre as quais destacam-se: classificação não intrusiva - não há necessidade de desconexão da planta para realizar a classificação; método embarcado; detecção e classificação no mesmo modelo; utilização de dados exclusivamente simulados para o treinamento - não é preciso coletar dados com a planta real operando sob influência de falhas, fazendo com que o desenvolvimento seja mais seguro, menos trabalhoso e escalável para outras plantas, sendo necessário para isto apenas uma reparametrização da simulação.

Por fim, os dados utilizados para teste também foram disponibilizados em:

https://github.com/clayton-h-costa/pv_fault_dataset ; incluindo os dados de irradiância,

temperatura, correntes e tensões de ambas *strings* e as anotações de instantes e classificação real de falhas. Esta base de dados contém um total de 1373798 amostras, sendo estas 1162931 do caso Normal, 5999 de Curto-Circuito, 10371 de Degradação, 6024 de Circuito Aberto e 188473 do caso de sombreamento, sendo o período de coleta: um segundo.

6.1 TRABALHOS FUTUROS

Como continuação para o presente trabalho, propõe-se:

- Localizar em qual(is) módulo(s) as falhas de curto-circuito e sombreamento ocorreram, utilizando sensores de temperatura acoplados a cada um dos módulos e/ou o cálculo de impedância até o ponto da falha;
- Acrescentar novas falhas a serem detectadas e classificadas, tais quais as apresentadas na Figura 7;
- Expandir o sistema para casos com múltiplas falhas ocorrendo simultaneamente via classificadores do tipo *multi-label*.;
- Expandir o sistema para uma planta de maior porte (aproximadamente 100kW), instalada na UTFPR.

6.2 TRABALHOS PUBLICADOS

Durante a realização deste trabalho foram publicados dois artigos em congresso internacional, sendo ambos diretamente relacionados ao presente trabalho:

COSTA, C. H.; MORITZ, G.; LAZZARETTI, A.; MULINARI B. M.; ANCELMO, H. C.; RODRIGUES, M. P.; OROSKI, E.; GOES, R., *A Comparison of Machine Learning-Based Methods for Fault Classification in Photovoltaic Systems*. In: **IEEE PES Innovative Smart Grid Technologies Conference - Latin America, 2019**, 2019. p. 1–6.

COSTA, C. H.; RODRIGUES, M. P.; YAMADA, G. D.; RODRIGUES, G.; JIAYU, X.; MORITZ, G. L.; GOES, R.; LAZZARETTI, A.; OROSKI, E.; STADZISZ, P. C., *Monitoring and Management System for Medium-Scale Photovoltaic Plants*. In: **IEEE PES Innovative Smart Grid Technologies Conference - Latin America, 2019**, 2019. p. 1–6.

Também foi feita a submissão de um artigo para a revista Sensors MDPI (Fator de Impacto 3,031), o qual foi aceito em agosto de 2020:

André E. Lazzaretti, **Clayton H. da Costa**, Marcelo P. Rodrigues, Guilherme D. Yamada, Gilberto Lexinoski, Guilherme L. Moritz, Elder Oroski, Rafael E. de Góes, Robson R. Linhares, Paulo C. Stadzisz, Júlio S. Omori, and Rodrigo B. dos Santos, *A Monitoring System for Online Fault Detection and Classification in Photovoltaic Plants*.

Ademais, também foram publicados quatro artigos, em congressos internacionais, que não incluem resultados do presente trabalho, mas que permeiam as mesmas áreas de conhecimento, sobretudo sistemas fotovoltaicos e aprendizado de máquina aplicado à sistemas elétricos. Os citados artigos são os seguintes:

ANCELMO, H. C.; GRANDO, F. L.; **COSTA, C. H.**; MULINARI, B. M.; OROSKI, E.; LAZZARETTI, A. E.; POTTKER, F.; RENAUX, D. P. B., *Automatic Power Signature Analysis using Prony's Method and Machine Learning-Based Classifiers*. In: **2018 2nd European Conference on Electrical Engineering and Computer Science (EECS)**, IEEE, 2018. p. 65–70.

MULINARI, B. M.; CAMPOS, D. P.; **COSTA, C. H.**; ANCELMO, H. C.; LAZZARETTI, A. E.; OROSKI, E.; LIMA, C. R. E.; RENAUX, D. P. B.; POTTKER, F.; LINHARES, R. R., *A New Set of Steady-State and Transient Features for Power Signature Analysis Based on V-I Trajectory*. In: **2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)**, 2019. p. 1-6.

ANCELMO, H. C.; GRANDO, F. L.; MULINARI, B. M.; **COSTA, C. H.**; LAZZARETTI, A. E.; OROSKI, E.; RENAUX, D. P. B.; POTTKER, F.; LIMA, C. R. E.; LINHARES, R. R., *A Transient and Steady-State Power Signature Feature Extraction Using Different Prony's Methods*. In: **2019 20th International Conference on Intelligent System Application to Power Systems (ISAP)**, 2019. p. 1-6.

RODRIGUES, M. P.; OROSKI, E.; STADZISZ, P. C.; **COSTA, C. H.**; LAZZARETTI, A.

E.; Linhares, R. R.; GOES, R. *A MISO Nonlinear Model of Photovoltaic Panel Based on System Identification*. In: **IEEE PES Innovative Smart Grid Technologies Conference - Latin America, 2019**, 2019. p. 1–6.

REFERÊNCIAS

APOGEE INSTRUMENTS. **Apogee Pyranometers - Silicon-Cell Pyranometers | SP-100 & SP200 Series**. [S.l.], 2019.

BARANI DESIGN. **MeteoTemp TH+T & Pressure**. [S.l.], 2017.

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

BOYD, Stephen; VANDENBERGHE, Lieven. **Convex Optimization**. Cambridge: Cambridge University Press, 2001.

BREIMAN, Leo; FRIEDMAN, Jerome H.; OLSHEN, Richard A.; STONE, Charles J. **Classification and Regression Trees**. Boca Raton: Taylor & Francis, 1984.

CANADIAN SOLAR INC. **Canadian Solar - Max Power CS6U-315|320|325|330P**. [S.l.], 2018.

CHEN, Zhicong; WU, Lijun; CHENG, Shuying; LIN, Peijie; WU, Yue; LIN, Wencheng. Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics. **Applied Energy**, v. 204, p. 912–931, 2017.

CHERKASSY, Vladimir; MULIER, Filip. **Learning from Data - Concepts, Theory and Methods**. 2. ed. New Jersey: WILEY-INTERSCIENCE, 2007.

CHINE, W.; MELLIT, A.; LUGHI, V.; MALEK, A.; SULLIGOI, G.; Massi Pavan, A. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. **Renewable Energy**, Elsevier Ltd, v. 90, p. 501–512, 2016.

CHINE, W.; MELLIT, A.; PAVAN, A. Massi; KALOGIROU, S. A. Fault detection method for grid-connected photovoltaic plants. **Renewable Energy**, Elsevier Ltd, v. 66, p. 99–110, 2014.

CHOUAY, Yassine; OUASSAID, Mohammed. An intelligent method for fault diagnosis in photovoltaic systems. **Proceedings of 2017 International Conference on Electrical and Information Technologies, ICEIT 2017**, v. 2018-Janua, p. 1–5, 2018.

COSTA, C. H.; Rodrigues, M. P.; Yamada, G. D.; Rodrigues, G. V.; Jiayu, X.; Moritz, G. L.; Oroski, E.; de Góes, R. E.; Lazzaretti, A. E.; Stadzisz., P. C. A monitoring and management system for medium-scale photovoltaic plants. *In*: **2019 IEEE PES Innovative Smart Grid**

Technologies Conference - Latin America (ISGT Latin America). Gramado, RS, BR: IEEE, 2019. p. 1–6.

DUDA, Richard O.; HART, Peter E.; STORK, David G. **Pattern Classification**. 2. ed. New York: Wiley-Interscience, 2000.

EKE, Rustu; KAVASOGLU, A. Sertap; KAVASOGLU, Nese. Design and implementation of a low-cost multi-channel temperature measurement system for photovoltaic modules. **Measurement**, v. 45, n. 6, p. 1499 – 1509, 2012. ISSN 0263-2241.

EL-KHOZONDAR, Hala J.; EL-KHOZONDAR, Rifa J.; MATTER, Khaled; SUNTIO, Teuvo. A review study of photovoltaic array maximum power tracking algorithms. **Renewables: Wind, Water, and Solar**, Springer Singapore, v. 3, n. 1, 2016.

GAROUDJA, Elyes; CHOUDER, Aissa; KARA, Kamel; SILVESTRE, Santiago. An enhanced machine learning based approach for failures detection and diagnosis of PV systems. **Energy Conversion and Management**, Elsevier, v. 151, n. September, p. 496–513, 2017.

GLOROT, Xavier; BORDES, Antoine; BENGIO, Yoshua. Deep sparse rectifier neural networks. *In*: GORDON, Geoffrey; DUNSON, David; DUDÍK, Miroslav (Ed.). **Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics**. Fort Lauderdale, FL, USA: PMLR, 2011. (Proceedings of Machine Learning Research, v. 15), p. 315–323.

HAN, Jinsoo; LEE, Ilwoo; KIM, Sang-Ha. User-friendly monitoring system for residential PV system based on low-cost power line communication. **IEEE Transactions on Consumer Electronics**, v. 61, p. 175–180, 05 2015.

HAYKIN, Simon. **Redes Neurais - Princípios e prática**. 2. ed. [S.l.]: Bookman, 2001.

HERSCH, Paul; ZWEIBEL, Kenneth. **Basic Photovoltaic Principles and Methods**. Golden: Technical Information Office, 1982.

HSU, Chih Wei; CHANG, Chih Chung; LIN, Chih Jen. **A Practical Guide to Support Vector Classification**. Taipei, 2010.

HU, Yihua; CAO, Wenping; MA, Jien; FINNEY, Stephen J.; LI, David. Identifying PV module mismatch faults by a thermography-based temperature distribution analysis. **IEEE Transactions on Device and Materials Reliability**, IEEE, v. 14, n. 4, p. 951–960, 2014.

INMET. **Instituto Nacional de Meteorologia - Normais Climatológicas - Temp. Min. Absoluta**. 2018. Disponível em: <http://www.inmet.gov.br/portal/index.php?r=clima/normaisclimatologicas>. Acesso em: 14 maio 2019.

IRENA. **Future of Solar Photovoltaic - Deployment, investment, technology, grid integration and socio-economic aspects (A Global Energy Transformation: paper)**. Abu Dhabi, 2019.

IZADIAN, Afshin; POURTAHERIAN, Arash; MOTAHARI, Sarasadat. Basic model and governing equation of solar cells used in power and control applications. *In: 2012 IEEE Energy Conversion Congress and Exposition, ECCE 2012*. Raleigh: IEEE, 2012. p. 1483–1488.

JÄGER-WALDAU, Arnulf. **PV Status Report 2019**. Luxembourg, 2019.

JUCÁ, Sandro C. S.; CARVALHO, Paulo C. M.; BRITO, Fábio T. A low cost concept for data acquisition systems applied to decentralized renewable energy plants. *Sensors*, v. 11, n. 1, p. 743–756, 2011.

KIMO INSTRUMENTS. **Surface contact wire temperature probe - SFCS50 / SFCS50**. [S.l.], 2012.

KURUKURU, V. S.Bharath; HAQUE, Ahteshamul; KHAN, Mohammed Ali; TRIPATHY, Arun Kumar. Fault classification for Photovoltaic Modules Using Thermography and Machine Learning Techniques. **2019 International Conference on Computer and Information Sciences, ICCIS 2019**, IEEE, p. 1–6, 2019.

LECUN, Yann; BOTTOU, Leon; ORR, Genevieve B.; MULLER, Klaus-Robbert. Efficient BackProp. *In: Neural Networks: Tricks of the Trade*. [S.l.]: Springer, 1998.

LEM. **Current Transducer HASS 50 ... 600-S**. [S.l.], 2019.

LI, Zhihua; WANG, Yuanzhang; ZHOU, Diqing; WU, Chunhua. An Intelligent Method for Fault Diagnosis in Photovoltaic Array Configuration of the Proposed System. **2017 International Conference on Electrical and Information Technologies**, p. 10–16, 2017.

MADETI, Siva Ramakrishna; SINGH, S. N. A comprehensive study on different types of faults and detection techniques for solar photovoltaic system. *Solar Energy*, Elsevier, v. 158, p. 161–185, 2017.

MADETI, Siva Ramakrishna; SINGH, S. N. Online modular level fault detection algorithm for grid-tied and off-grid PV systems. *Solar Energy*, Elsevier, v. 157, n. May, p. 349–364, 2017.

MANSOURI, Majdi; AL-KHAZRAJI, Ayman; HAJJI, Mansour; HARKAT, Mohamed Faouzi; NOUNOU, Hazem; NOUNOU, Mohamed. Wavelet optimized EWMA for fault detection and application to photovoltaic systems. *Solar Energy*, Elsevier, v. 167, n. November 2017, p. 125–136, 2018.

MEKHILEF, S.; SAIDUR, R.; KAMALISARVESTANI, M. Effect of dust, humidity and air velocity on efficiency of photovoltaic cells. **Renewable and Sustainable Energy Reviews**, Elsevier Ltd, v. 16, n. 5, p. 2920–2925, 2012.

MORENO-GARCIA, Isabel M.; PALACIOS-GARCIA, Emilio J.; PALLARES-LOPEZ, Victor; SANTIAGO, Isabel; GONZALEZ-REDONDO, Miguel J.; VARO-MARTINEZ, Marta; REAL-CALVO, Rafael J. Real-time monitoring system for a utility-scale photovoltaic power plant. **Sensors**, v. 16, n. 6, 2016. ISSN 1424-8220.

NANHUA ELECTRONICS. **Wind Speed Sensor FA01**. [S.l.], 2018.

NANHUA ELECTRONICS. **FA02 Wind Direction Sensor**. [S.l.], 2019.

NATIONAL INSTRUMENTS. **NI9215 Datasheet**. [S.l.], 2018.

NATIONAL INSTRUMENTS. **NI9242 Datasheet**. [S.l.], 2018.

NHS SISTEMAS ELETRÔNICOS. **Inversores On Grid NHS SOLAR | Inversor On Grid GSM / GDM**. [S.l.], 2018.

PACKARD, Corinne E; WOHLGEMUTH, John H; KURTZ, Sarah R. Development of a Visual Inspection Data Collection Tool for Evaluation of Fielded PV Module Condition. **NREL Technical Report**, n. August, p. 10, 2012.

PINHO, João Tavares; GALDINO, Marco Antonio. **Manual de Engenharia para Sistemas Fotovoltaicos**. Rio de Janeiro: CEPEL - CRESESB, 2014.

POIWERSIM INC. **PSIM Tutorial - How to Use Solar Module Physical Model**. [S.l.], 2016. Disponível em: <https://powersimtech.com/drive/uploads/2016/04/Tutorial-Solar-Module-physical-model.pdf>. Acesso em: 19 abr. 2019.

SÁNCHEZ, José Antonio Luceño; DÍEZ-PASCUAL, Ana María; CAPILLA, Rafael Peña. Materials for photovoltaics: State of art and recent developments. **International Journal of Molecular Sciences**, v. 20, n. 4, 2019.

SANCHEZ-PACHECO, Francisco J.; SOTORRIO-RUIZ, Pedro Juan; HEREDIA-LARRUBIA, Juan Ramón; PÉREZ-HIDALGO, Francisco; CARDONA, Mariano Sidrach de. PLC-based PV plants smart monitoring system: Field measurements and uncertainty estimation. **IEEE Transactions on Instrumentation and Measurement**, v. 63, p. 2215–2222, 2014.

SINHA, A.; SASTRY, O. S.; GUPTA, R. Detection and characterisation of delamination in PV modules by active infrared thermography. **Nondestructive Testing and Evaluation**, Taylor & Francis, v. 31, n. 1, p. 1–16, 2016.

SKOPLAKI, E.; PALYVOS, J. A. Operating temperature of photovoltaic modules: A survey of pertinent correlations. **Renewable Energy**, v. 34, n. 1, p. 23–29, 2009.

SUTTON, Richard S.; BARTO, Andrew G.; BACH, Francis. **Reinforcement Learning: An Introduction**. Cambridge: MIT Press, 2015.

THEODORIDIS, Sergios; KOUTROUMBAS, Konstantinos. **Pattern Recognition**. 2. ed. Londres: Academic Press, 2003.

TIAN, Hongmei; MANCILLA-DAVID, Fernando; ELLIS, Kevin; MULJADI, Eduard; JENKINS, Peter. A cell-to-module-to-array detailed model for photovoltaic panels. **Solar Energy**, v. 86, n. 9, p. 2695 – 2706, 2012.

TSANAKAS, John A.; BOTSARIS, P N. An infrared thermographic approach as a hot-spot detection tool for photovoltaic modules using image histogram and line profile analysis. **International Journal of Condition Monitoring**, v. 2, n. 1, p. 22–30, 2012.

TSANAKAS, J. A.; CHRYSOSTOMOU, D.; BOTSARIS, P. N.; GASTERATOS, A. Fault diagnosis of photovoltaic modules through image processing and Canny edge detection on field thermographic measurements. **International Journal of Sustainable Energy**, v. 34, p. 351–372, 2015.

TSANAKAS, John A.; HA, Long D.; Al Shakarchi, F. Advanced inspection of photovoltaic installations by aerial triangulation and terrestrial georeferencing of thermal/visual imagery. **Renewable Energy**, Elsevier Ltd, v. 102, p. 224–233, 2017.

VAPNIK, Vladimir N. **Statistical Learning Theory**. New York: Wiley-Interscience, 1998.

YI, Zhehan; ETEMADI, Amir H. Line-to-line fault detection for photovoltaic arrays based on multi-resolution signal decomposition and two-stage support vector machine. **IEEE Transactions on Industrial Electronics**, v. 64, n. 11, p. 8546–8556, 2017.

ZAIDI, Beddiaf. Solar Panels and Photovoltaic Materials. *In: Solar Panels and Photovoltaic Materials*. : Intech, 2018. cap. 1.