

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



Aplicação de técnicas de aprendizagem profunda estruturada para diagnóstico de funcionamento de centrais fotovoltaicas.

David da Silva Moreira Freire

INICIAÇÃO À INVESTIGAÇÃO

MESTRADO EM ENGENHARIA ELETROTÉCNICA E COMPUTADORES

Orientador: Cláudio Domingos Martins Monteiro

December 31, 2022

**Aplicação de técnicas de aprendizagem profunda
estruturada para diagnóstico de funcionamento de
centrais fotovoltaicas.**

David da Silva Moreira Freire

MESTRADO EM ENGENHARIA ELETROTÉCNICA E COMPUTADORES

December 31, 2022

Resumo

Este documento ilustra o formato a usar em dissertações na Faculdade de Engenharia da Universidade do Porto. São dados exemplos de margens, cabeçalhos, títulos, paginação, estilos de índices, etc. São ainda dados exemplos de formatação de citações, figuras e tabelas, equações, referências cruzadas, lista de referências e índices. Este documento não pretende exemplificar conteúdos a usar. É usado o *Loren Ipsum* para preencher a dissertação.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam vitae quam sed mauris auctor porttitor. Mauris porta sem vitae arcu sagittis facilisis. Proin sodales risus sit amet arcu. Quisque eu pede eu elit pulvinar porttitor. Maecenas dignissim tincidunt dui. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec non augue sit amet nulla gravida rutrum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nunc at nunc. Etiam egestas.

Donec malesuada pede eget nunc. Fusce porttitor felis eget mi mattis vestibulum. Pellentesque faucibus. Cras adipiscing dolor quis mi. Quisque sagittis, justo sed dapibus pharetra, lectus velit tincidunt eros, ac fermentum nulla velit vel sapien. Vestibulum sem mauris, hendrerit non, feugiat ac, varius ornare, lectus. Praesent urna tellus, euismod in, hendrerit sit amet, pretium vitae, nisi. Proin nisl sem, ultrices eget, faucibus a, feugiat non, purus. Etiam mi tortor, convallis quis, pharetra ut, consectetur eu, orci. Vivamus aliquet. Aenean mollis fringilla erat. Vivamus mollis, purus at pellentesque faucibus, sapien lorem eleifend quam, mollis luctus mi purus in dui. Maecenas volutpat mauris eu lectus. Morbi vel risus et dolor bibendum malesuada. Donec feugiat tristique erat. Nam porta auctor mi. Nulla purus. Nam aliquam.

Abstract

The increase in photovoltaic power plants has led to the need for effective methods for detecting and addressing component faults, which can have significant economic impacts. In this work, we will explore the current state of fault detection and state estimation tools in the field of PV systems, with a focus on understanding how these tools work and identifying their strengths and limitations. We will also propose improvements to existing approaches or develop a novel approach to address this issue. By examining the most successful tools to date and offering new solutions, we aim to help PV plant operators improve the reliability and efficiency of their systems.

Agradecimentos

Aliquam id dui. Nulla facilisi. Nullam ligula nunc, viverra a, iaculis at, faucibus quis, sapien. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Curabitur magna ligula, ornare luctus, aliquam non, aliquet at, tortor. Donec iaculis nulla sed eros. Sed felis. Nam lobortis libero. Pellentesque odio. Suspendisse potenti. Morbi imperdiet rhoncus magna. Morbi vestibulum interdum turpis. Pellentesque varius. Morbi nulla urna, euismod in, molestie ac, placerat in, orci.

Ut convallis. Suspendisse luctus pharetra sem. Sed sit amet mi in diam luctus suscipit. Nulla facilisi. Integer commodo, turpis et semper auctor, nisl ligula vestibulum erat, sed tempor lacus nibh at turpis. Quisque vestibulum pulvinar justo. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nam sed tellus vel tortor hendrerit pulvinar. Phasellus eleifend, augue at mattis tincidunt, lorem lorem sodales arcu, id volutpat risus est id neque. Phasellus egestas ante. Nam porttitor justo sit amet urna. Suspendisse ligula nunc, mollis ac, elementum non, venenatis ut, mauris. Mauris augue risus, tempus scelerisque, rutrum quis, hendrerit at, nunc. Nulla posuere porta orci. Nulla dui.

Fusce gravida placerat sem. Aenean ipsum diam, pharetra vitae, ornare et, semper sit amet, nibh. Nam id tellus. Etiam ultrices. Praesent gravida. Aliquam nec sapien. Morbi sagittis vulputate dolor. Donec sapien lorem, laoreet egestas, pellentesque euismod, porta at, sapien. Integer vitae lacus id dui convallis blandit. Mauris non sem. Integer in velit eget lorem scelerisque vehicula. Etiam tincidunt turpis ac nunc. Pellentesque a justo. Mauris faucibus quam id eros. Cras pharetra. Fusce rutrum vulputate lorem. Cras pretium magna in nisl. Integer ornare dui non pede.

O Nome do Autor

*“You should be glad that bridge fell down.
I was planning to build thirteen more to that same design”*

Isambard Kingdom Brunel

Contents

1	Introduction	1
2	Fault detection in Utility Scale PV Plants	5
2.1	Photovoltaic Plant's architecture and components	5
2.2	Types of faults	6
2.3	Modeling photovoltaic's physical/electrical behavior	7
2.3.1	The five-parameter model	8
2.4	Fault detection in Photovoltaic Systems	9
2.4.1	Statistical and Signal Processing Algorithms	9
2.4.2	Machine Learning Algorithms	9
2.4.3	Deep Learning Algorithms	10

List of Figures

2.1	PV inverter topologies: (a) Central, (b) String, (c) Multi-string, (d) Module integrated ?	6
2.2	Classification of faults in PV Systems, from ??	7

List of Tables

Abreviaturas e Símbolos

PV	Photovoltaic
DC	Direct current
STC	Standard Test Conditions
MCD	Minimum Covariance Determinant
WWW	<i>World Wide Web</i>

Chapter 1

Introduction

The XIX century marked a significant shift in the world's perception of energy resources as the desire to invest in renewable energy sources to power modern societies grew. This transition was driven by the need to reduce dependency on fossil fuels, mitigate the effects of global warming, and slow climate change. Solar photovoltaic energy is a desirable renewable energy source due to its abundance, accessibility, and environmental benefits. Renewable energy sources offer a range of benefits, including reduced greenhouse gas emissions, improved air quality, and increased energy security. While solar photovoltaic energy has proven to be both cost-efficient and environmentally friendly, it also comes with unprecedented challenges, such as its intermittent nature, low electrical inertia, complex forecasting, and geographic-dependent operating conditions. Despite these challenges, recent reports [?] show that the economic benefits of investing in renewable energy outweigh the complications, as there is an increasing global investment trend in these sources.

The general construction of PV farms, particularly on the utility-scale, has led to a need for effective maintenance and monitoring to ensure maximum efficiency and operational reliability. To achieve this, various algorithms and routines are used to monitor the state of PV farms and identify any potential issues that may arise. Fault detection is a crucial aspect of this process, as it allows PV farm operators to identify and address any problems that may occur quickly. Given the importance of maintaining high levels of operation, knowing if action is needed to restore or fix components from an anomalous scenario is desirable. By detecting faults and identifying the necessary steps, PV farm operators can prevent or minimize downtime and ensure optimal performance.

Integrating intermittent energy resources into modern electric grids has led to stricter requirements for connecting such power systems to ensure safe grid operating conditions. As a result, companies that own or plan to build photovoltaic farms must comply with these requirements and have adequate power electronics and monitoring/control capabilities. Failure to meet these requirements can result in sanctions or fines for the responsible party, as well as potential impacts on system availability, asset value, and disturbance propagation to the grid. To minimize these risks and maximize the value of their assets, companies may opt to implement fault detection and state estimation tools. These tools allow for the early detection and resolution of potential issues

and can prevent or minimize downtime. The need to create or improve existing fault detection and state estimation tools, and the search for the most effective methodologies for addressing these issues, drive research in this field.

Having laid the basis for why there must be system behavior assessment in utility-scale PV plants, it is necessary to understand what business concepts are crucial to this field. In the course of this work, the presented topics will go over the following questions:

- What components mostly fail in photovoltaic power systems?
- What is the average frequency of faults?
- What fault detection/state estimation tools exist for photovoltaic power systems?
- What are the most successful ones?
- What's their structure? Are they mostly centralized or decentralized?
- What are their computational costs/efficiency?
- What is the expected magnitude of precision and confidence?
- Which key performance indicators can evaluate the success of these tools?
- What are their implementation difficulties?

With these questions uncovered, the main objective is to adapt or design a novel algorithm/approach to fault detection based on modern artificial intelligence solutions. However, this can be split up into finer goals:

- Identify and study existing fault detection tools for photovoltaic power systems.
- Adapt or develop a new tool.
- Apply and test the new tool in real case study PV assets.
- Validate the developed methodologies by comparison to reference tools.

Before reviewing state-of-the-art fault detection tools, types of failures in photovoltaic systems need to be understood: find which components usually fail, which ones fail more often, and how often. For this, it is necessary to understand such components' physical and electrical properties and the modeling techniques used to characterize them. There will be an assessment of utility-scale power plants architecture through literature, alongside the detection objective of state-of-the-art fault detection tools applied in this field. Then, there shall be an extensive analysis and review of what tools have been designed and used in this field. In this step, critical evaluation of the literature is a must for understanding the tool's scope, ease of implementation, and understanding that the data sets available for this work are compatible. Having selected the most prominent ones, they're

to be qualitatively and quantitatively compared to each other in their application context so that the results allow objective evaluations. This process requires implementing these tools, following the guidelines in the respective article/book/report, verifying their metrics, and checking if the achieved results resemble the same as the literature suggests. It will require gathering data sets, which can either be artificially generated through simulation or provided by an enterprise that services photovoltaic plant owners.

There's a desire that, in the end, the developed work helps achieve an improved method for fault detection and state estimation in photovoltaic power systems, resulting in a production-ready software application agile enough to deploy for any PV assets. Depending on the new algorithm's characteristics, it could result in an approach capable of generalization and application to other engineering systems, benefiting more than just PV systems. No matter the chosen methodology, fault detection will, in most cases, result in an economic benefit, catastrophe prevention, and safety increase. Those are a few positive contributions that are possible for this work's outcome.

Chapter 2

Fault detection in Utility Scale Photovoltaic Plants

2.1 Utility-Scale Photovoltaic System's Architecture

Utility-scale photovoltaic (PV) power plants are large-scale PV systems that are connected to the electrical grid and have installed capacities ranging from kilowatts peak (kWp) to megawatts peak (MWp). These systems typically consist of many PV panels interconnected through power electronics to aggregate and inject active power into the grid. The number and type of components in a PV power plant depend on the plant's scale and topology, with different configurations possible for large-scale applications, including central inverters, string inverters, and multi-string inverters. Understanding the architecture and components of PV power plants is vital for designing, operating, and maintaining these systems, as it helps optimize their performance and reliability.

In figure ??, the fourth configuration presented (d) will not be considered for utility-scale plants due to its expensive nature. After DC/AC conversion, another voltage conversion step usually establishes the grid connection: an AC/AC transformer.

For completeness, the physical installation of PV modules can include solar tracking apparatuses, such as single and dual-axis trackers, which add to system complexity and change production behavior. Nonetheless, and turning the focus back toward the electrical components, the main ones are the following:

- Solar photovoltaic panels.
- Electric cabling.
- Inverter(s) (mostly with Max Power Point Trackers).
- AC Transformer(s).
- Protection components (circuit breakers, fuses, surge protectors, etc.)

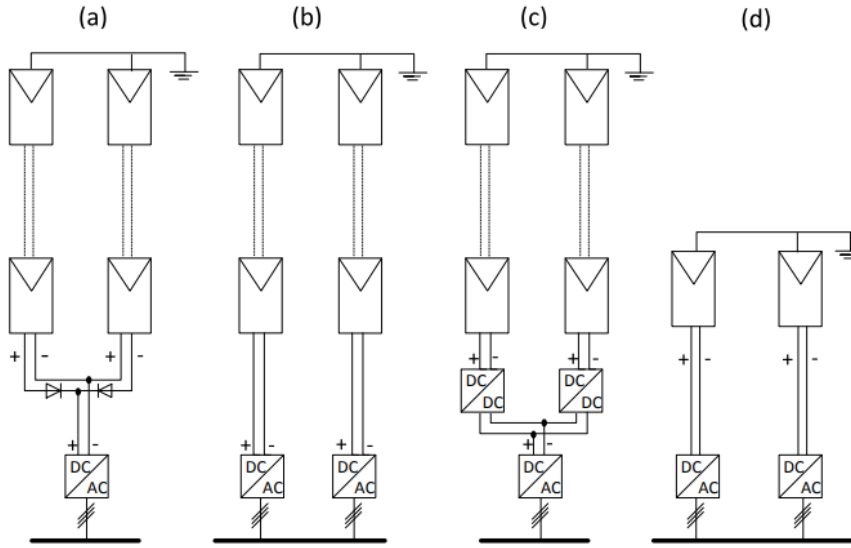


Figure 2.1: PV inverter topologies: (a) Central, (b) String, (c) Multi-string, (d) Module integrated ?.

Most of these components have intrinsic variables, such as voltage and current values, that can help determine their operation states. Given that the utility grids (and the associated electricity market) integrate large-scale PV assets, some of the before-mentioned components require constant monitoring and control, achieved with adequate embedded systems and sensor infrastructure ?. Thanks to the continuous advancements in communication technologies, namely in IoT (Internet Of Things), data acquisition is becoming faster, more reliable, and more precise. Not only is this fundamental for real-time asset assessment, but it also allows better training of prediction algorithms. However, on the industrial scale (in the order of MWp production), having sensors embedded in every PV module comes with a high economical cost, and inverters are the components that usually possess monitoring capabilities. Since monitoring utility-scale PV assets rely on the investment and technologies employed for this purpose, engineers must consider data availability when developing data-driven algorithms, such as forecasting and fault detection.

2.2 Types of faults

Several types of faults can occur in utility-scale photovoltaic (PV) power plants, which can negatively impact the performance and reliability of the system, possibly causing safety hazards ?. Some are challenging to detect and protect the electrical installation against, thus requiring sophisticated detection algorithms ?. According to ?, these faults can fit into three categories: electrical, mechanical, and environmental. Electrical faults include short circuits, open circuits, and inverter failure, affecting the PV panels' power output and the system's overall efficiency. Mechanical faults include broken panels, damaged cables, and defective inverters, which can lead to system downtime and reduced performance. Environmental faults include extreme weather events, such

as hail or strong winds, which can damage the PV panels and other components ?. Figure ?? illustrates a more extensive summary of fault types.

figures/chapter2/types_of_faults.pdf

Figure 2.2: Classification of faults in PV Systems, from ??.

Throughout the literature ?, some of the most studied faults in the context of fault detection are:

- Shading: partial coverage, temporary or not, of a PV array or module. Might result in Hot Spot fault.
- Soiling: dirt accumulation, blocking sunlight from reaching PV Cells. Might also result in Hot Spot fault.
- Short circuit: either line-line or line-ground.
- Open circuit: connection breakage between modules.
- DC arc fault: eletricity plasma arc formed on broken connections.

Due to the difficulty of classifying some of these faults, given their similarity on the consequent effect in the system, it will be seen in further sections that most fault detection algorithms only endeavor to classify between two to five types of faults.

2.3 Modeling photovoltaic's physical/electrical behavior

Photovoltaic cells are the fundamental components of photovoltaic panels. They are made from semiconductor materials, such as silicon, and absorb photons that generate an electric current. Their electrical behavior is characterizable using the current-voltage (I-V) equation. This equation, which represents a fundamental relationship governing the operation of PV cells, can be used to predict their performance under various operating conditions such as solar irradiance and temperature.

For state estimation, it is crucial to accurately model PV modules' performance from the DC side of power converters. This information is vital for designing and optimizing PV power systems, as it enables predicting PV module performance under different conditions, as mentioned before. Accurate PV module models are also essential for state estimation and fault detection, as they provide critical information about the health and performance of PV modules, allowing for early identification of potential issues. In addition, they can be used to optimize the control and operation of PV power systems, which can improve the overall efficiency and reliability of the system ?.

Physical and empirical models broadly categorize the several state-of-the-art methods for modeling photovoltaic modules. Physical models lie on the fundamental physical principles governing PV modules' operation. They typically require detailed knowledge of the PV module's electrical and optical properties, such as its current-voltage (I-V) characteristics, spectral response, and temperature dependence. These models can accurately predict the PV module's performance under a wide range of operating conditions, but they may be complex and computationally intensive to implement. On the other hand, empirical models are based on experimental data and are typically more straightforward to implement. However, they may not be as accurate as physical models, especially under conditions that differ significantly from those used to generate the experimental data (usually STC). Some examples of state-of-the-art physical models for PV modules include the single-diode model (also known as the five-parameter model), and the two-diode model. In contrast, one of the most used state-of-the-art empirical models is the Sandia model. The choice of modeling method will depend on the specific application and the required level of accuracy and complexity; in some cases, there can be a combination of physical and empirical models.

Suppose the need arises to model PV modules in this work. In that case, it is critical to select a simple methodology so that the module's datasheet characteristics are sufficient to model the PV arrays accurately. In the case of utility-scale PV systems, detailed knowledge of the module's electrical and optical properties of empirical data may be limited, and building a model is only possible by using datasheet data. A complex model that requires more detailed information may not be feasible in such cases, and a simpler model that relies on fewer input parameters is more appropriate. The single-diode model seems appropriate for this use case, given the excellent trade-off between complexity and accuracy.

2.3.1 The five-parameter model

The single-diode model representation of the photovoltaic module is presented in figure ???. The I-V curve can be obtained by numerically solving the following equation:

$$I = I_L - I_0 \times \left(e^{\frac{V + I \times R_s}{n \times V_T}} - 1 \right) \quad (2.1)$$

where I_L is the light-generated current, I_0 is the reverse saturation current, R_s is the series resistance, n is the ideality factor, and V_T is the thermal voltage. These parameters are determined by fitting the model to experimental data or using data from the PV module's datasheet. The single-diode model can predict the PV module's performance under a wide range of operating conditions while maintaining reasonable accuracy. However, remembering that the single-diode model is a simplified representation of the PV module, it will have poor accuracy under certain conditions compared to the more complete two-diode model.

2.4 Literature on Fault detection for Photovoltaic Systems

The parent field of fault detection is anomaly detection (also known as outlier detection), a highly studied subject in the scope of statistics [1], applied in many scientific areas. The tools dedicated to fault detection and state estimation mostly come from mathematical/statistical methodologies, machine learning, and deep learning applications [2]. Parting from the three general problem-solving principles mentioned before, machine learning and deep learning are the most popular and successful ones for recent applications that ought to solve a plethora of complex problems. Such potential has led to an interest in their implementation for fault detection algorithms.

2.4.1 Statistical and Signal Processing Algorithms

Statistical methodologies look into historical data to find the characteristics of how samples relate to the population (interpolation). These methodologies yield good results in case studies of PV farms that have been logging data for a considerable time, suffering in the cases that do not. Therefore, they are limited in that it is required to have curated data sets of historical significance for relevant features of the studied systems.

For the field of fault detection in PV systems, the literature on statistical and signal processing algorithms is mostly quite dated [3, 4, 5], given that more recent machine learning methods became increasingly interesting for this matter. Nonetheless, anomaly (or outlier) detection statistical algorithms can be used for fault detection in PV systems by identifying unusual patterns or deviations from normal behavior in the data collected from the PV system. For this, we might employ distance-based methods, such as Euclidean, Mahalanobis, and MCD-based distances [6]. Although simple, distance-based methods only might work for detecting outliers in the context of PV systems if they are scale-invariant (due to the different magnitude in the system's state variables) and resilient to outlier contamination (only with MCD-based distance). In [7], the authors applied Analysis of Variance (ANOVA) and Kruskal-Wallis test for inverter failure detection, with the downside of only being able to identify outliers in a sub-array resolution, i.e., not for specific string or module failures.

2.4.2 Machine Learning Algorithms

Machine learning came to solve some of the complications referred to in the two past subjects, as neural networks (or other learning structures) are easily capable of modeling complex, non-trivial, and nonlinear relations between data. Still, they are as good as the training data, with many structures requiring many representative learning examples to achieve good results. Their output can also be very obfuscated, meaning that many methods do not allow a direct interpretation of the relationship between inputs and outputs. This "black-box" characteristic, specifically of neural networks, is considered a disadvantage. Besides, extrapolating data remains a challenge when classically using these structures. Still, they have immense applications for PV systems, from MPP (Max Power Point) estimation to power forecasting, soiling, and fault prediction.

2.4.3 Deep Learning Algorithms

The field of deep learning branches off from machine learning, with the term "deep" referring to amplified machine learning structures that ought to understand data patterns. A simple example would be the design of an artificial neural network with multiple hidden layers, with the intuition that each of these "extra" layers achieves feature/pattern recognition in a cascade. They have been explored alongside machine learning techniques for PV systems, although the known disadvantage is a usually high computational cost and relatively tricky implementation.

While classical fault detection lies in the synchronous and direct evaluation of state estimation variables, fault prediction requires the input of time-series features. Although relatively simple, some classical time-series forecasting and analysis tools can be of great support to help design a fault prediction algorithm, such as Box-Jenkins methods and the Partial Auto-correlation function. Still, the majority of modern prediction tools comprise neural networks and variations. With this in mind, recent developments in the intelligent composition of learner structures spark some interest in the application to this field, such as the new deep learning technique named Cell Complex Neural Networks ?. Further investigation of such modern practices will unroll throughout the development of this work.