



UNIVERSIDAD MICHOACANA DE SAN NICOLÁS DE
HIDALGO

DIVISIÓN DE ESTUDIOS DE POSGRADO DE LA FACULTAD DE
INGENIERÍA ELÉCTRICA

**El Problema del Diagnóstico de Fallas en
Sistemas Fotovoltaicos basado en Análisis
Topológico de Datos.**

T E S I S

QUE PARA OBTENER EL GRADO DE:

Maestro en Ciencias en Ingeniería Eléctrica

PRESENTA:

Brayan Hernández Calvillo

DIRECTOR:

Dr. Juan Anzures Marín

CO-DIRECTOR

Dr. Juan Ahtziri González Lemus



Morelia, Michoacán, México.

septiembre de 2025

Al Brayan del 2023, quien estuvo a punto de rendirse.

Agradecimientos

En primer lugar, quiero expresar mi más profundo agradecimiento a mi familia, que han sido un pilar fundamental a lo largo de toda mi formación. A pesar de los tropiezos y dificultades que he enfrentado, siempre han estado a mi lado brindándome su apoyo incondicional. A mis padres, gracias por sus consejos, su paciencia y por acompañarme en cada etapa de mi vida, transmitiéndome valores que han guiado mi camino. A mis hermanas, gracias por su comprensión y por soportar mis malos momentos, pero sobre todo por estar presentes cuando más lo necesitaba, ofreciéndome su apoyo y cariño. Sin ellos, este logro no habría sido posible.

A mis asesores, el Dr. Juan y el Dr. Ahtziri, les agradezco no solo por su orientación académica, sino también por la confianza y la paciencia para el desarrollo de esta tesis. Sus conocimientos y experiencia aportados fueron de gran importancia para lograr este trabajo y para fortalecer mi formación académica.

Quiero agradecer al Dr. Eduardo, pues durante el tiempo que hemos colaborado me ha enseñado la importancia de realizar cada tarea con el mayor compromiso posible. También agradezco mucho su apoyo económico brindado. Estas cosas me han ayudado a seguir creciendo como estudiante.

Agradezco profundamente a mis amigas Lizeth, Vinely y Yesenia, por estar a mi lado, escuchándome y brindándome su apoyo en cada momento en que lo necesité, incluso en aquellos problemas que parecían pequeños pero que para mí fueron significativos. Su compañía y palabras de aliento fueron una parte importante para mantenerme motivado y con la fuerza necesaria para continuar este camino.

A mis compañeros de maestría, Sebastián y Gochi, por la paciencia al explicarme temas con los que no estaba familiarizado durante mi maestría. A mi compadre y a Reynaldo, por tener siempre la confianza de acercarse y por las anécdotas que hemos generado a lo largo de nuestra amistad. A todos los compañeros que he conocido en el posgrado, gracias por las vivencias compartidas durante estos dos últimos años.

Por último, agradezco a SECIHTI por haberme proporcionado el apoyo económico para poder realizar mis estudios de maestría y a la División de Estudios de Posgrado de la Facultad de Ingeniería Eléctrica por haberme permitido usar sus instalaciones durante mi estadía y para la realización de este trabajo.

Índice general

Resumen	xv
Abstract	xvii
1. Introducción	1
1.1. Descripción del problema	1
1.2. Objetivos	2
1.2.1. Objetivo general	2
1.2.2. Objetivos particulares	2
1.3. Hipótesis	2
1.4. Preguntas de investigación	2
1.5. Antecedentes	3
1.6. Metodología	5
1.7. Justificación	6
1.8. Descripción por capítulos	6
2. Fundamentos de sistemas fotovoltaicos	9
2.1. Principios básicos de la energía solar	9
2.1.1. Efecto fotovoltaico	9
2.2. Componentes de un Sistema Fotovoltaico	11
2.2.1. Paneles Solares Fotovoltaicos	11
2.2.2. Inversores	12
2.2.3. Baterías	12
2.2.4. Estructura y conexión	13
2.3. Energía fotovoltaica y otras formas de energía renovable	15
2.4. Tendencias y Futuro de la Tecnología Fotovoltaica	16
3. Métodos de diagnóstico	19
3.1. Métodos basados en el modelo	19
3.1.1. Análisis de Curva I-V	19
3.1.2. Modelos Eléctricos Equivalentes	21
3.1.3. Modelos de Rendimiento	22
3.1.4. Modelos Térmicos	23
3.2. Métodos Basados en Datos	24
3.2.1. Monitoreo en Tiempo Real	24
3.2.2. Aprendizaje Automático	25
3.2.3. Análisis de Series Temporales	26

3.2.4. Visión Artificial y Drones	26
3.3. Análisis topológico de datos	27
3.4. Comparación de Métodos	28
4. Fundamentos del Análisis Topológico de Datos	29
4.1. Complejos simpliciales	29
4.2. Alfa complejos	30
4.3. Hoyos en complejos simpliciales	32
4.4. Homología y Homología persistente	33
4.5. Distancia Cuello de Botella	35
5. Aplicación del TDA	43
5.1. Sistema fotovoltaico analizado	43
5.2. Extracción de la topología de la potencia	44
5.3. Comparación entre módulos	47
5.4. Modelo de detección	49
5.4.1. Definiciones de fallas	51
6. Resultados	53
6.1. Datos reales	54
6.2. Simulación de Fallas	57
6.2.1. Simulación considerando diodos <i>bypass</i>	62
6.3. Discusión	65
7. Conclusiones y trabajo futuro	69
A. Códigos implementados, Python	73
B. Meses analizados	77

Lista de Figuras

2.1. Tipos de incidencia solar sobre la superficie.	10
2.2. Efecto fotovoltaico	11
2.3. Tipos de paneles	12
2.4. Estructuras de paneles	14
2.5. Conexión completa de un sistema	15
3.1. Ejemplos de curvas I-V	20
4.1. Simplejos	30
4.2. Complejo simplicial	30
4.3. Diagrama de Voronoi	31
4.4. Filtración Alfa	32
4.5. $\mathbb{S}^j$ en complejos simpliciales	33
4.6. Diagrama de persistencia	35
4.7. Emparejamiento de diagramas	36
4.8. Búsqueda del mejor emparejamiento	36
4.9. Distancia entre círculos	37
4.10. Distancia entre un ocho y un círculo	38
4.11. Plantas fotovoltaicas en México	39
4.12. Plantas hidroeléctricas en el país	40
4.13. Distancias entre plantas solares e hidroeléctricas	40
5.1. Potencia en tres días	44
5.2. Adecuación de nube de datos	45
5.3. Topología nube ideal	45
5.4. Topología nube 2	46
5.5. Topología nube 3	46
5.6. Topología nube 4	47
5.7. Distancia entre diagramas ideales	48
5.8. Distancia entre diagrama ideal y diagrama con ruido	49
5.9. Distancia entre Potencia en continua y potencia en alterna	49
5.10. Diagrama de flujo del modelo	50
6.1. Porcentaje de la topología generada en Febrero	54
6.2. Distancia entre módulos, Febrero	55
6.3. Distancias entre PCC y PCA	56
6.5. Distancias entre módulos, falla por sombreado	58

6.4. Porcentaje de la topología simulando falla por sombreado	58
6.6. Distancias entre PCC y PCA, falla por sombreado	59
6.7. Porcentajes de topología simulando pérdida de un módulo	60
6.8. Distancias entre módulos del sistema, pérdida de módulo	61
6.9. Distancias entre PCC y PCA al simular pérdida de un módulo	61

Lista de Tablas

2.1. Costos de generación	16
5.1. Valores almacenados	44
5.2. Ejemplo falla por clima, distancias ideales	51
5.3. Ejemplo falla por clima, distancias módulos	51
5.4. Ejemplo falla módulo, distancias ideales	52
5.5. Ejemplo falla módulo, distancias módulos	52
6.1. Umbrales entre módulos	54
6.2. Umbrales entre potencias	54
6.3. Fallas del sistema en módulos, Febrero de 2025	56
6.4. Fallas del sistema en inversores, Febrero de 2025	56
6.5. Falla en módulos por rendimiento	59
6.6. Falla en inversores por bajo rendimiento en módulo	59
6.7. Fallas por pérdida de módulo	62
6.8. Fallas en inversores por pérdida de módulo	62
6.9. Fallas en módulos por pérdida de un panel	63
6.10. Fallas en inversores por pérdida de un panel.	63
6.11. Fallas en módulos por pérdida de dos paneles	64
6.12. Fallas en inversores por pérdida de dos paneles	64
6.13. Fallas en módulos por pérdida de un inversor.	65
6.14. Fallas en inversores por pérdida de un inversor.	65
6.15. Tabla de fallas	66
B.1. Fallas Potencia continua, Agosto 2021	77
B.2. Fallas Potencia alterna, Agosto 2021	77
B.3. Fallas Potencia continua, Septiembre 2021	78
B.4. Fallas Potencia alterna, Septiembre 2021	78
B.5. Fallas Potencia alterna, Septiembre 2024	79
B.6. Fallas Potencia alterna, Septiembre 2024	79
B.7. Fallas Potencia continua, Octubre 2024	80
B.8. Fallas Potencia alterna, Octubre 2024	80
B.9. Fallas Potencia continua, Abril 2025	81
B.10. Fallas Potencia alterna, Abril 2025	81
B.11. Fallas Potencia continua, Mayo 2025	82
B.12. Fallas Potencia alterna, Mayo 2025	82

Lista de Acrónimos

AI	Artificial Intelligence (Inteligencia Artificial)
A	Ampere
ARIMA	Media Móvil Integrada Autorregresiva
BD	Distancia tipo Cuello de Botella (Bottleneck Distance)
BIVP	Sistemas FV integrados en la arquitectura
CA	Corriente Alterna
CC	Corriente Continua
CPE	Costo Promedio de la Electricidad
CNN	Convolutional Neural Network (Red Neuronal Convolutacional)
FV	Fotovoltaico
GIS	Sistemas de información geográfica
Hz	Hertz
IoT	Internet of Things (Internet de las cosas)
IRT	Item Response Theory (Teoría de Respuesta al Ítem)
kWh	Kilowatt-hora
LSTM	Long Short-Term Memory
ML	Machine Learning (Aprendizaje Automático)
NOCT	Nominal Operating Cell Temperature
PCA	Principal Component Analysis (Análisis de Componentes Principales)
PR	<i>Performance Ratio</i>
RGB	Red, Green, Blue
RFE	Recursive Feature Elimination (Eliminación Recursiva de Características)

SARIMA Media Móvil Integrada Autorregresiva Estacional

SCADA Supervisory Control and Data Acquisition

STC Condiciones de Prueba Estándar (*Standard Test Conditions*)

SVM Support Vector Machine (Máquina de Vectores de Soporte)

TDA Topological Data Analysis (Análisis Topológico de Datos)

V Volt

W Watt

Tabla de Símbolos

Símbolo	Descripción
A	Ampere
$BD(Pac_{(i,0),i})$	Distancia cuello de botella entre el diagrama ideal y el módulo i (CA)
$BD(Pac_{(i,j)})$	Distancia cuello de botella entre diagramas de potencia CA de los módulos i y j
$BD(Pfv_{(i,0),i})$	Distancia cuello de botella entre el diagrama ideal y el módulo i (FV)
$BD(Pfv_{(i,j)})$	Distancia cuello de botella entre diagramas de potencia FV de los módulos i y j
$B_r(x)$	Bola abierta de radio r y centrada en x
$\mathcal{D}$	Diagrama de persistencia
d_B	Distancia cuello de botella
$Del(N)$	Triangulación de Delaunay sobre el conjunto de datos N
ΔT	Incremento de temperatura bajo irradiancia estándar
E_{AC}	Energía en corriente alterna entregada
f	Frecuencia eléctrica (Hz)
G	Irradiancia solar incidente
G_{ref}	Irradiancia de referencia
H_{POA}	Irradiancia global sobre el plano del arreglo
I	Corriente eléctrica
I_{CA}	Corriente en corriente alterna
I_{CC}	Corriente en corriente continua
I_0	Corriente de saturación del diodo
I_{mp}	Corriente en el punto de máxima potencia
I_{ph}	Corriente fotogenerada
I_{sc}	Corriente de cortocircuito
K	Complejo simplicial

Continúa en la siguiente página

Tabla 1 continua de la página anterior

Símbolo	Descripción
$m(h)$	Radio de muerte del hoyo h
m^2	Metro cuadrado
N	Conjunto de datos
$n(h)$	Radio de nacimiento del hoyo h
$\ \cdot\ $	Distancia euclidiana
$\ \cdot\ _\infty$	Distancia infinito
P	Potencia eléctrica
P_{CA}	Potencia en corriente alterna
P_{CC}	Potencia en corriente continua
P_{max}	Potencia máxima
P_{STC}	Potencia nominal bajo condiciones estándar
$P(t)$	Potencia generada por el sistema FV en el instante t
q	Carga del electrón
R_s	Resistencia serie
R_{sh}	Resistencia en paralelo (shunt)
σ	Simplejo
$\mathbb{S}^j$	Esfera de dimensión j
T	Temperatura absoluta / en el inversor
T_{amb}	Temperatura ambiente
T_{mod}	Temperatura del módulo
V	Voltaje eléctrico
V_{CA}	Voltaje en corriente alterna
V_{CC}	Voltaje en corriente continua
V_{mp}	Voltaje en el punto de máxima potencia
V_{oc}	Voltaje de circuito abierto
$V(x)$	Región de Voronoi de x
W	Watt
$\mathcal{A}(N, \varepsilon)$	Alfa complejo de radio ε sobre N
ε	Número real positivo
$^\circ C$	Grados Celsius (temperatura)
$\mathbb{R}^n$	Espacio euclidiano n -dimensional

Resumen

Este trabajo aborda el problema del diagnóstico de fallas en sistemas fotovoltaicos, basándose en la importancia de garantizar un correcto funcionamiento de los mismos. Se parte de la premisa de que fallas como sombreado, suciedad, desconexiones o degradación de componentes impactan negativamente en el rendimiento, por lo cual resulta fundamental detectarlas oportunamente. Para ello, se propone un enfoque innovador basado en el Análisis Topológico de Datos, una herramienta que permite extraer información topológica de las señales eléctricas generadas por el sistema. A diferencia de los métodos tradicionales estadísticos o de aprendizaje automático, esta técnica analiza la forma global de los datos, lo que la hace robusta frente al ruido y a pequeñas variaciones operativas.

La idea general se basa en la comparación de las topologías de los conjuntos de datos obtenidos día con día contra un ideal. Los resultados obtenidos al analizar los datos reales y comparar las fallas provocadas en los mismos demuestran que la herramienta propuesta es capaz de detectar anomalías incluso sutiles, alcanzando una efectividad del 95.5 %, en las pruebas realizadas, en la discriminación entre modos de operación normales y operación con anomalías. Además, las propiedades extraídas proporcionan una manera de diferenciar entre pérdidas energéticas por condiciones ambientales y fallas internas en los módulos. El estudio concluye que el enfoque topológico ofrece un marco prometedor, escalable y automatizable para el diagnóstico de fallas en sistemas fotovoltaicos, contribuyendo al mantenimiento predictivo, ayudando en la transición hacia sistemas energéticos inteligentes y sostenibles.

Palabras clave: Diagnóstico de fallas; Sistemas fotovoltaicos; Complejos simpliciales; distancia cuello de botella; Diagramas de persistencia.

Abstract

This work addresses the problem of fault diagnosis in photovoltaic systems, based on the importance of ensuring their proper operation. It is based on the premise that faults such as shading, dirt, disconnections, or component degradation negatively impact performance, making their timely detection essential. To this end, we propose an innovative approach based on Topological Data Analysis, a tool that extracts topological information from the electrical signals generated by the system. Unlike traditional statistical or machine learning methods, this technique analyzes the overall shape of the data, making it robust to noise and small operational variations.

The general idea is based on comparing the topologies of data sets obtained daily against an ideal. The results obtained with real data and simulations demonstrate that the proposed tool is capable of detecting even subtle anomalies, achieving an effectiveness of 95.5% in discriminating between normal operating modes and anomalies. Furthermore, the extracted properties provide a way to differentiate between energy losses due to environmental conditions and internal module failures. The study concludes that the topological approach offers a promising, scalable, and automatable framework for fault diagnosis in photovoltaic systems, contributing to predictive maintenance and aiding in the transition to smart and sustainable energy systems.

Capítulo 1

Introducción

1.1. Descripción del problema

La necesidad de reducir las emisiones de gases de efecto invernadero y minimizar la dependencia de los combustibles fósiles se ha convertido en una prioridad. La energía solar es una fuente limpia que no produce contaminación ni emisiones, lo que la convierte en una alternativa para el cambio hacia un modelo energético más verde, por lo que los sistemas fotovoltaicos (FV) se presentan como una de las soluciones más efectivas para satisfacer las necesidades energéticas sin comprometer el medio ambiente. Sin embargo, el rendimiento de estos sistemas puede verse afectado por diversas fallas, tales como sombreado parcial, suciedad, desconexiones, defectos en módulos o degradación de componentes electrónicos. Estas anomalías afectan de manera directa a la conversión energética del sistema, así como a su vida útil y costos de mantenimiento. La detección temprana y precisa de fallas es fundamental para asegurar una operación continua y eficiente de los sistemas FV.

El diagnóstico de fallas es un campo de gran importancia que se ha encargado de crear herramientas que detecten, de manera temprana, las posibles fallas que se generan en los sistemas. La idea general de este proceso es la identificación y el diagnóstico de problemas con el objetivo de corregirlos antes de que se conviertan en problemas de gran magnitud o irreparables. Sin embargo, detectar las fallas no siempre es sencillo debido a la diversidad de estas y la gran cantidad de componentes en los sistemas modernos. Durante sus inicios, la detección de fallas comenzó con métodos manuales en donde los operadores revisaban y evaluaban los sistemas haciendo uso de su experiencia en los procesos para identificar algunas irregularidades. Con la llegada de la informática y la electrónica, se comenzaron a introducir métodos basados en modelos matemáticos que representan el comportamiento ideal del sistema, logrando identificar fallas mediante las diferencias entre el modelo y el sistema real, [1]. Algunos otros modelos que se han desarrollado a lo largo de la historia son basados en modos deslizantes y análisis de señales físicas tales como vibraciones, ruido y temperatura [2], [3], [4].

Actualmente, los métodos basados en datos se han presentado como una alternativa prometedora. Estos métodos, que incluyen técnicas de aprendizaje automático, minería de datos y análisis estadístico, buscan identificar patrones anómalos directamente a partir de datos históricos o en tiempo real de operación. Gracias al creciente acceso a sensores, sistemas de adquisición de datos y almacenamiento digital, es posible recolectar grandes volúmenes de información eléctrica que pueden ser utilizados para diagnóstico y pronóstico de fallas. Sin embargo, muchos de estos enfoques requieren una gran cantidad de datos etiquetados y pueden ser sensibles al ruido o a variaciones en las condiciones de operación. Además, se enfocan en la búsqueda de relaciones entre los datos,

dejando de lado su estructura.

El **Análisis Topológico de Datos** (TDA, por sus siglas en inglés) surge con un enfoque matemático robusto para extraer información estructural a partir de datos, capturando características como la forma y su conectividad, más allá de lo que permiten algunos métodos estadísticos tradicionales. El uso del TDA representa una oportunidad para identificar y caracterizar fallas de manera no supervisada o semi-supervisada, incluso en presencia de ruido o pequeños cambios operativos. Por ello, en este trabajo se propone desarrollar un enfoque basado en TDA para la detección de fallas en sistemas FV, con el objetivo de contribuir a una supervisión robusta, eficiente y automatizada de estos sistemas.

1.2. Objetivos

1.2.1. Objetivo general

Desarrollar un enfoque de detección de fallas en sistemas FV utilizando técnicas de TDA, con el fin de identificar patrones anómalos en señales eléctricas que indiquen condiciones de falla o funcionamiento deficiente.

1.2.2. Objetivos particulares

- Analizar el sistema FV que se trabajará e identificar la manera en la que está estructurado, así como el conjunto de datos que se obtiene del mismo.
- Implementar técnicas de TDA como diagramas de persistencia y métricas entre ellos para realizar un análisis cuantitativo entre sus diferencias.
- Diseñar y evaluar un método de detección de fallas basado en características extraídas mediante TDA.
- Identificar las ventajas y desventajas del método frente a otras técnicas con el objetivo de promover el TDA como enfoque innovador en la detección de fallas de sistemas FV.

1.3. Hipótesis

El análisis topológico de datos permite detectar de manera robusta y eficiente las fallas en sistemas FV; la presencia de alteraciones en el sistema o pérdida de componentes se verá reflejada en los diagramas extraídos, proporcionando características topológicas anómalas.

1.4. Preguntas de investigación

1. ¿Qué mediciones se pueden utilizar para la identificación de una falla?
2. ¿Cuál es la topología de una nube de datos sin fallas?
3. ¿El comportamiento es similar en todo el sistema FV?
4. ¿Qué diferencias entre las topologías extraídas se corresponden con una falla?

1.5. Antecedentes

La teoría de diagnóstico de fallas es el área enfocada en el desarrollo de métodos capaces de detectar y aislar anomalías en sistemas dinámicos. Esta disciplina se ha consolidado como una de las piezas principales en la operación de sistemas complejos, permitiendo mejorar la seguridad, reducir tiempos de inactividad y optimizar el mantenimiento. Su aplicación en sistemas FV ha tomado relevancia en los últimos años, dado el crecimiento acelerado de esta tecnología como fuente de energía limpia y renovable. La detección de fallas en sistemas FV representa un campo de investigación en constante expansión, impulsado por la necesidad de garantizar un funcionamiento eficiente, seguro y sostenible. A lo largo del tiempo, la teoría del diagnóstico de fallas ha experimentado avances significativos, impulsados en gran parte por el desarrollo de nuevas tecnologías. En sus primeras etapas, las técnicas empleadas se basaban fundamentalmente en la experiencia del personal técnico; el proceso de diagnóstico era manual y consistía en inspecciones visuales y auditivas de los componentes accesibles, lo que limitaba la efectividad y la confiabilidad del proceso.

Con la necesidad de mejorar la capacidad operativa y reducir el tiempo de inactividad surgen los métodos basados en modelos matemáticos y físicos, que permiten predecir el comportamiento esperado de un sistema bajo condiciones normales y detectar desviaciones que pudieran indicar la presencia de fallas, [5]. Estos enfoques aportan mayor precisión al diagnóstico, siendo técnicas más sistemáticas y automatizadas, pero también requieren un conocimiento detallado del sistema y un modelo preciso, lo que no siempre es viable en entornos complejos o dinámicos. Existen distintos tipos de métodos basados en el modelo que se abordan en capítulos posteriores, el modelo clásico más utilizado está basado en el uso de observadores, [6], [7], [8]. Estos son capaces de simular algunos de los estados que no son medibles en el sistema físico. Existen análisis basados en el modelo que explican la manera en que se comportan los sistemas FV en una misma región, pero con diferentes características. En Reino Unido, por ejemplo, se analizaron dos sistemas con distinta orientación a lo largo de dos años, [9], en el sistema *A* se observaron pérdidas anuales promedio de 3.3 % y 6.6 % respectivamente, mientras que en el sistema *B* se observaron pérdidas anuales promedio de 18.9 %. El análisis se complementó con la clasificación de fallas usando técnicas basadas en datos, logrando caracterizar cuatro tipos distintos: fallas sostenidas de eficiencia cero; fallas breves de eficiencia cero; sombreado; y fallas sin sombreado de eficiencia distinta de cero; sin embargo, esta etapa no será abordada en esta tesis.

El avance en tecnologías trajo consigo la implementación de herramientas como sensores, para lograr una adquisición de datos y procesamiento computacional lo cual dio paso a los métodos basados en datos. Estos enfoques, apoyados en técnicas estadísticas, algoritmos de aprendizaje automático y minería de datos, permiten detectar patrones anómalos a partir del análisis de grandes volúmenes de información histórica o en tiempo real, [10]. Su principal ventaja radica en su capacidad de adaptación y su menor dependencia de modelos físicos explícitos, lo que los hace especialmente útiles en sistemas de alta variabilidad como los FV.

Los sistemas de potencia son una base fundamental, por lo que estudios como [11] presentan un enfoque para el diagnóstico y la clasificación de fallas en líneas de transmisión trifásicas utilizando técnicas de aprendizaje automático, específicamente mediante máquinas de soporte vectorial (SVM) y selección de características basada en eliminación recursiva de características (RFE). Este enfoque demuestra una alta eficiencia y robustez, destacándose como una solución prometedora para sistemas inteligentes de diagnóstico en redes eléctricas modernas. En la actualidad, los enfoques híbridos, que combinan el conocimiento físico del sistema con herramientas de inteligencia

artificial, representan una línea de investigación destacada. Estos métodos buscan aprovechar las fortalezas de ambos paradigmas para lograr una detección de fallas robusta, precisa y eficiente. En este contexto, la aplicación de estas técnicas a sistemas FV ha cobrado especial relevancia, dada la creciente adopción de esta tecnología a nivel global y la necesidad de mantener su operación continua y confiable, [12], [13].

La inspección visual conjunta con herramientas modernas como drones equipados con cámaras de alta resolución es uno de los enfoques más tradicionales y efectivos. Estos dispositivos permiten examinar grandes instalaciones de manera rápida. El proceso se reduce al sobrevuelo de los drones sobre el sistema y este envía las imágenes que captura a un servidor para su posterior proceso. El análisis térmico ha sido uno de los más estudiados; ayuda en la detección de grietas, decoloraciones o acumulación de suciedad que pueda reducir la producción energética de los paneles. Todas estas anomalías térmicas pueden indicar fallas en celdas individuales, en el sistema completo o incluso indicar un bajo rendimiento sin catalogarse como falla. El análisis realizado mediante cámaras fijas a un sistema fotovoltaico instalado en el Cairo, Egipto, [14], analiza la eficiencia del sistema al estar cubierto por suciedad debido a los factores climáticos, se analizaron casos de entre 6 y 12, 18 y 24g/m² de polvo sobre los paneles, y se concluyó que éstos tuvieron pérdidas en la generación con valores de un 13 %, un 19 %, un 25 % y un 38 %, respectivamente en comparación con un sistema limpio. Una mejora en el análisis al usar cámaras termográficas es presentada en [15], el estudio analiza imágenes proporcionadas por drones y después son procesadas mediante aprendizaje automático usando métodos de teoría de respuesta al ítem, IRT, por sus siglas en inglés. Un análisis más completo se presenta en [16], el sistema utilizado cuenta con un dron que sobrevuela el sistema, envía imágenes térmicas que son preprocesadas en escala de grises, para después analizarlas mediante algoritmos de detección de fallas y enviar una alerta junto con la latitud y longitud del módulo fotovoltaico defectuoso.

El análisis de curvas corriente-voltaje (I-V) permite evaluar el rendimiento de los módulos y detectar distintos tipos de problemas como conexión o desgaste en estos. El proceso consiste en analizar el conjunto de datos generado por el sistema; en el caso de [17] se utilizan técnicas para generar sombreados y posteriormente, se realiza un análisis de componentes principales (PCA). Al completar el estudio se observó que el método logró una tasa de clasificación del 97 % en pruebas.

Los modelos de Machine Learning como el usado en [18], son capaces de mejorar la precisión en la detección de fallas en sistemas. El uso de árboles de decisión mejorados con modelos como el Light Gradient Boosting logran una mejora en el aislamiento de fallas. La propuesta fue evaluada frente a cinco modelos existentes. En comparación con el mejor de ellos, el Gradient Boosting Regressor, se observó una mejora del 12.8 % en el error cuadrático medio, del 6.8 % en la raíz del error cuadrático medio y del 11.08 % en el coeficiente de determinación. La metodología propuesta en [19] emplea imágenes aéreas y de electroluminiscencia para la aplicación de Machine Learning, logrando efectividades de detección del 91 % al 93 %. Mientras que, en [20] se introduce un método de diagnóstico de fallas basado en aprendizaje automático no invasivo (NML-FD) para módulos FV. Sin embargo, el Machine Learning no es la única tecnología emergente. El uso de redes neuronales también ha demostrado ser de gran utilidad. Por ejemplo, la red convolucional utilizada en [21] muestra una fiabilidad del 95.78 % al detectar 20 tipos distintos de fallas. Los autores en [22] sugirieron un método llamado Convolution Ghost con BottleneckCSP, cuya meta principal es mejorar la precisión en objetivos de múltiples estados, utilizando la Convolution Ghost para aumentar la velocidad de inferencia del modelo y disminuir la cantidad de parámetros.

Otro método creado para la detección de fallas se basa en la teoría de lógica difusa. Una aplicación es mostrada en [23], al realizar el procesamiento de imágenes térmicas, la detección y

clasificación de fallas se realizó mediante la unión del método k-means y lógica difusa. Otro análisis es mostrado en [24], éste realiza simulaciones de una serie de atributos, como la relación de voltaje y la relación de potencia con límites de detección altos y bajos. Una vez simuladas, las muestras que se encuentran fuera de los límites de detección se procesan mediante un sistema de clasificación de lógica difusa, obteniendo una efectividad del 98.8 %. En simulaciones, este enfoque permite una detección precisa aunque nos muestra que los métodos de detección más actuales no siempre han sido probados a nivel práctico.

En la literatura no se encuentran demasiadas aplicaciones directas del TDA en el área de diagnóstico de fallas, sin embargo existen aplicaciones en otras áreas que presentan la identificación de patrones anómalos mediante esta técnica. En el ámbito médico, el procesamiento de imágenes es una de las técnicas más utilizadas para analizar enfermedades internas. Todos los aportes en favor de la identificación temprana del cáncer son bien recibidos, área donde se ha logrado la identificación de microcalcificaciones, [25], y la identificación de carcinomas, [26], mediante anomalías topológicas en las imágenes procesadas. Algoritmos de TDA con enfoques de clustering tales como el mapper, ha sido de ayuda en identificación de características en sistemas complejos sobre un conjunto de datos del proceso de fabricación de productos químicos, [27], la idea principal es lograr una reducción de dimensión mientras se mantienen las estructuras principales, reduciendo así el gasto computacional. Más específicamente en el diagnóstico de fallas, es posible unir métodos tradicionales de análisis como lo es la transformada de Wavelet con el TDA, [28], se encarga de realizar el diagnóstico en rodamientos mecánicos, al mismo tiempo que compara el modelo propuesto con algunos otros, observando una mejoría significativa. La industria privada también ha comenzado a incorporar enfoques topológicos en sus investigaciones. Por ejemplo, en [29] se analiza la excentricidad en motores eléctricos a partir de las corrientes registradas, representadas en un espacio euclidiano. Este enfoque permite extraer patrones que caracterizan el comportamiento ideal del motor y diferenciarlos de aquellos que presentan fallas. Las herramientas que ofrece el TDA facilitan significativamente estos estudios, al detectar patrones que comienzan a desviarse del comportamiento esperado.

A pesar de estos avances, aún existen desafíos importantes en la detección de fallas en sistemas FV, particularmente en la fiabilidad de los datos, la identificación de anomalías sutiles y la integración de los distintos enfoques de diagnóstico. No obstante, el comportamiento dinámico y complejo de los sistemas FV exige nuevas metodologías capaces de ofrecer una visión robusta y menos dependiente de supuestos. En este contexto, la principal aportación de la presente tesis consiste en la aplicación del TDA como un enfoque innovador basado en datos para la detección de fallas en sistemas fotovoltaicos. Esta metodología permite identificar propiedades de los conjuntos de datos operativos y revelar patrones asociados al comportamiento normal y anómalo del sistema, ofreciendo así una alternativa nueva y complementaria a los métodos tradicionales revisados en este capítulo.

1.6. Metodología

La aplicación del TDA sobre los conjuntos de datos obtenidos de la generación eléctrica está basada en herramientas topológico-geométricas que serán abordadas con profundidad en el Capítulo 4, debido a su contenido; en cambio, esta sección proporciona una idea intuitiva del proceso dividido en cuatro pasos principales.

Paso 1: recolección y procesamiento de información. El conjunto de datos almacenado se

encuentra almacenado en la nube y este se va actualizando día con día, generando un archivo por día de información recolectada. Estos se descargan de manera manual y contienen mediciones como voltaje, corriente, temperatura, irradiancia solar, entre otras variables. Es fundamental garantizar que los datos sean recolectados en diferentes condiciones operativas y ambientales para obtener un conjunto de datos representativo; en este caso particular se analiza un mes completo donde se tuvieron distintos comportamientos. Una vez recolectados, los datos deben ser procesados para eliminar archivos vacíos o datos con inconsistencias debido a la comunicación. Además, se realiza la eliminación de variables irrelevantes. De esta manera, el conjunto de datos analizado contiene los valores de interés que representan el comportamiento general del sistema.

Paso 2: aproximación del conjunto de datos. Una vez que se ha realizado el correcto procesamiento del conjunto de datos, se procede a la aplicación de las herramientas topológico-geométricas. Al representar los datos en un espacio n -dimensional, es posible generar una aproximación de la forma (estructura geométrica) que estos tienen. Para ello, se realizan construcciones llamadas complejos simpliciales que se basan en la distancia euclidiana de los puntos.

Paso 3: representación de propiedades topológicas. La aproximación de la forma del conjunto de datos permite extraer las propiedades topológicas que caracterizan a la misma y que serán las que diferencien el comportamiento del sistema. Al representarlas mediante diagramas dedicados, es posible identificar diferencias entre conjuntos de datos que representan un comportamiento sano y conjuntos de datos que representan comportamientos anómalos.

Paso 4: generación del modelo. Una vez identificadas las características principales mediante el TDA, se utiliza una métrica entre los diagramas utilizados, de manera que se puedan cuantificar las fluctuaciones presentadas día con día. Posterior a la cuantificación, se propone un modelo de detección de fallas basado en las distancias calculadas, el cual se valida utilizando un conjunto de datos con alteraciones mayores.

1.7. Justificación

El uso de energías limpias para la generación de energía eléctrica ha sido indispensable durante los últimos años, y es de gran importancia que los centros de generación se mantengan en funcionamiento óptimo. Por ello, es importante la detección temprana de fallas que puedan ser catastróficas. Los sistemas actuales están siendo monitoreados de distintas maneras y diariamente generan una gran cantidad de datos que se almacenan para su posterior estudio; estos grandes conjuntos de datos generan la necesidad de contar con centros de cómputo avanzados que procesen la información. La técnica propuesta promete filtrar y eliminar algunos de los datos que no proporcionen información relevante para la detección de fallas, optimizando así el proceso de análisis y reduciendo la carga de procesamiento.

1.8. Descripción por capítulos

Este capítulo desarrolla una introducción al problema de la detección de fallas en sistemas fotovoltaicos; se expone la relevancia de esta energía limpia en la actualidad, así como algunas de las técnicas y resultados que han sido expuestos en algunos artículos a lo largo de los años. Se muestran los propósitos de este trabajo al proponer un nuevo enfoque en este ámbito y se expone

la manera en que se abordará el problema. El contenido de los siguientes capítulos en esta tesis se expone brevemente a continuación:

El Capítulo 2 describe el funcionamiento de los sistemas FV y sus principales componentes.

En el Capítulo 3 se presentan de forma breve las metodologías existentes para la detección de fallas en sistemas FV.

El Capítulo 4 introduce el TDA desde un enfoque intuitivo y práctico, concluyendo con un ejemplo de su aplicación.

El Capítulo 5 detalla la composición del sistema analizado, muestra la aplicación del TDA al conjunto de datos obtenido y define el modelo de detección propuesto.

Finalmente, los Capítulos 6 y 7 exponen los resultados alcanzados mediante la técnica desarrollada, las conclusiones derivadas del estudio y las posibles líneas de trabajo futuro.

Capítulo 2

Fundamentos de sistemas fotovoltaicos

Las energías limpias han tomado una gran relevancia durante la actualidad, cumpliendo con la necesidad de combatir el cambio climático; en el capítulo anterior se ha mencionado brevemente la manera en que la energía solar puede ser aprovechada para reducir las emisiones de carbono y disminuir la dependencia de los combustibles fósiles. Por ello, es importante conocer la manera en que los sistemas FV actúan, abordando su arquitectura y los principios físicos en los que se basan. Además, es de suma importancia detectar las principales diferencias entre esta energía y algunas de las más importantes de la actualidad. El aprovechamiento de la energía solar se puede realizar de tres maneras posibles: la conversión térmica, en donde la energía solar es transformada al calentar un fluido hasta altas temperaturas; la conversión foto-química, donde se genera una reacción de CO_2 y H_2O para generar carbohidratos; y la conversión fotovoltaica, la cual busca transformar la radiación en energía eléctrica mediante placas de células fotovoltaicas [30]. Centraremos este estudio en el caso de placas, mejor conocidas como paneles solares.

2.1. Principios básicos de la energía solar

Los sistemas FV se encargan de recibir la radiación electromagnética proveniente del sol que está compuesta por tres principales regiones de la luz: visible, infrarroja y ultravioleta. La unión de estos tres tipos de luz es conocida como radiación solar y la manera en que incide sobre un sistema FV se puede dar de tres maneras: la radiación directa incide en el sistema sin tocar algún otro medio en el entorno; radiación difusa, generada por los efectos de reflexión y de refracción de la atmósfera; y, la radiación reflejada, se genera cuando la radiación incide en el sistema después de haber incidido sobre algún objeto en el entorno, ver Figura 2.1. La recepción de estas regiones dependerá, principalmente, de factores como el horario, el clima o la ubicación del sistema.

La cantidad de radiación solar que se mide en un metro cuadrado de superficie es llamada irradiación solar y es medida sobre la distancia media entre la Tierra y el Sol, en la parte superior de la atmósfera, se conoce como la constante solar y su valor es aproximado a $1367 \frac{\text{W}}{\text{m}^2}$. En un día despejado, la radiación directa en la superficie terrestre puede alcanzar aproximadamente los $1000 \frac{\text{W}}{\text{m}^2}$ [31].

2.1.1. Efecto fotovoltaico

A mediados del siglo XIX, el físico francés Edmond Becquerel descubrió que la luz podía inducir un flujo de corriente en un material conductor. Mientras realizaba experimentos con electrodos

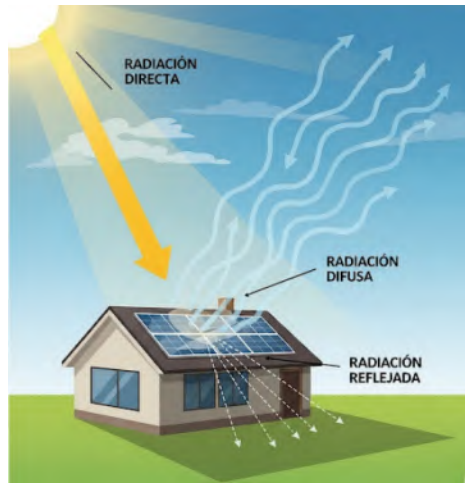


Figura 2.1: Tipos de incidencia solar sobre la superficie.

sumergidos en soluciones, observó que la luz aumentaba la corriente eléctrica generada. Debido a su época, este descubrimiento no fue desarrollado de manera práctica, pero se siguieron haciendo investigaciones sobre la teoría. Fue en 1905 cuando Albert Einstein describió cómo la luz podría ser vista como partículas, las cuales, al interactuar con los electrones de un material, generan una corriente eléctrica. Fue hasta la década de 1950 cuando el efecto fotovoltaico comenzó una nueva etapa al lograr la invención de lo que hoy se conoce como la primera célula solar. En 1954, los científicos Russell Ohl, Gerald Pearson y Daryl Chapin lograron desarrollar una célula solar de silicio que mostraba eficacia en la conversión de luz solar en electricidad. Este logro científico dio inicio al desarrollo de la teoría y la tecnología para la creación de sistemas FV, [32]. Las células solares son el principal dispositivo electrónico en un sistema FV. La composición de una célula se basa en el aprovechamiento de dos semiconductores, comúnmente una capa positiva de silicio y una capa negativa, los cuales están separados por una junta neutra y son colocados entre dos electrodos y una capa antirreflejante, Figura 2.2(a). El silicio es el semiconductor más utilizado debido a su dispersión de electrones en cada átomo y a que es abundante en nuestro entorno.

Cuando la luz solar incide sobre la célula solar, los fotones de la luz interactúan con los electrones en el material semiconductor, lo que “libera” algunos electrones de sus átomos. Al dejar el semiconductor, los electrones libres se encargan de generar la corriente eléctrica, Figura 2.2(b). Sin embargo, para que esta sea útil, es necesario que los electrones se muevan en una dirección específica. Esto se logra al generar un campo eléctrico mediante las dos capas de silicio dopado, una capa de silicio tipo N, con electrones libres adicionales y otra capa de silicio tipo P, con espacios que sirven como portadores de carga positiva. Cuando estas dos capas se combinan, se forma una unión P-N, que crea un campo eléctrico en la interfaz entre las dos capas y funciona como una barrera. Este campo eléctrico orienta a los electrones liberados por los fotones hacia un lado de la célula y los huecos hacia el otro, lo que crea un flujo de corriente eléctrica en una dirección determinada. El circuito eléctrico se completa cuando los electrones fluyen a través de un conductor externo, como un cable, hacia un dispositivo que necesita energía eléctrica, como una lámpara o un aparato electrónico. La electricidad generada por las células solares es corriente continua (CC), lo que significa que los electrones fluyen en una sola dirección. En la mayoría de las ocasiones, es necesario transformar esta electricidad en corriente alterna (CA) mediante un inversor, para que sea compatible con el sistema eléctrico.

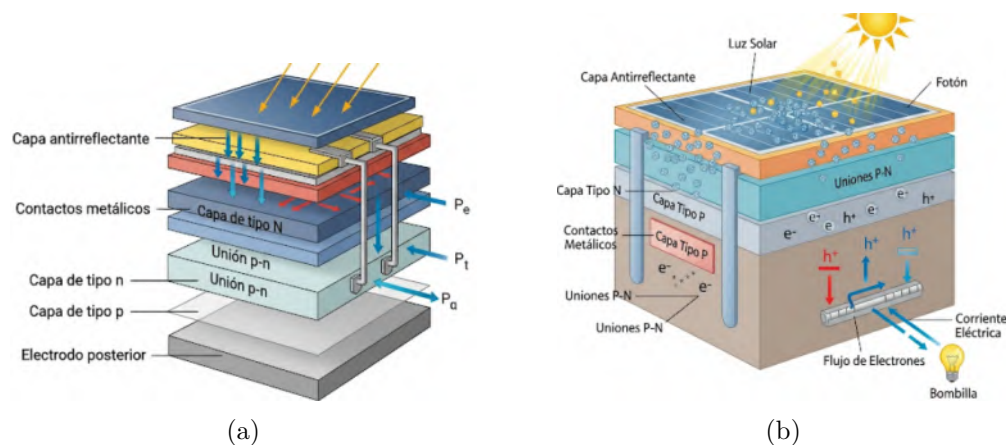


Figura 2.2: Célula solar (a) y representación del efecto fotovoltaico (b).

2.2. Componentes de un Sistema Fotovoltaico

Un sistema FV se compone de varios elementos que ayudan en la conversión de la energía solar en electricidad. A continuación se describen de manera breve los elementos principales en los sistemas, así como su función dentro de los mismos.

2.2.1. Paneles Solares Fotovoltaicos

Cada célula fotovoltaica produce una pequeña cantidad de electricidad, por lo que se agrupan en paneles solares para generar una cantidad significativa de energía. Existen distintos tipos de materiales para su construcción, [33], sin embargo son tres los principales:

- Los paneles fabricados con silicio monocristalino son fácilmente reconocibles por su color negro y su forma de células redondeadas, ver Figura 2.3(a). El silicio se obtiene a partir de un solo cristal, lo que permite una mayor eficiencia en la conversión, reportando entre el 18 y el 22 %, lo que los hace ideales para instalaciones con espacio limitado.
- Comúnmente se observa en instalaciones comerciales y residenciales paneles azules de silicio policristalino, ver Figura 2.3(b), estos se fabrican a partir de varios cristales de silicio fundidos, lo que genera una menor pureza y su costo. La eficiencia con la que cuentan ronda entre el 15 % y el 17 %.
- Por último, los paneles de silicio amorfo son los más económicos; sin embargo, este material es un tipo de silicio no cristalino, lo que significa que no tiene una estructura ordenada a nivel atómico. Su eficiencia es considerablemente menor, esta ronda el 6 % y el 9 %. Su principal ventaja es la flexibilidad, lo que los hace útiles en estructuras curvadas, ver Figura 2.3(c).

Normalmente, los paneles se agrupan en módulos para aumentar la capacidad de producción según las necesidades del sistema. Así mismo, éstos ayudan a reducir el gasto de instalación, el cual suele ser relativamente alto. Existen algunos otros tipos de paneles solares que se diferencian en su rendimiento y costo. Los paneles convencionales son los que se observan con mayor frecuencia, pero también existen paneles solares de película delgada, que son más ligeros y flexibles, y paneles solares bifaciales, que capturan luz en ambos lados, aumentando el aprovechamiento de la energía solar.

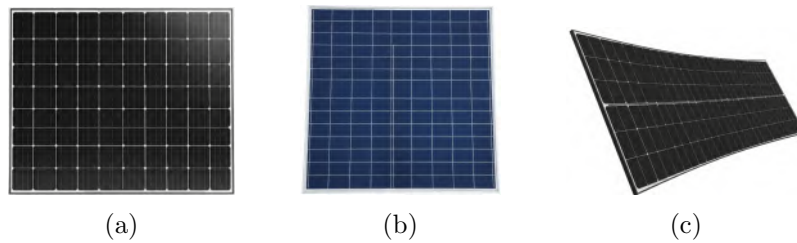


Figura 2.3: Tipos de paneles solares más comunes: monocristalino (a), policristalino (b) y amorfo (c).

2.2.2. Inversores

La energía eléctrica generada por los paneles solares es corriente continua. Los inversores son dispositivos clave en un sistema FV, ya que convierten ésta en corriente alterna (CA), el tipo de electricidad utilizada en la mayoría de los hogares y sistemas eléctricos. Existen diferentes tipos de inversores, cada uno con características y aplicaciones específicas, entre los principales se encuentran, [34]:

Los inversores centrales son los más grandes y potentes. Son utilizados en grandes instalaciones fotovoltaicas, como plantas solares. Los paneles solares son conectados en módulos, y todo el sistema de paneles está vinculado a un único inversor central. Su principal ventaja es la eficiencia en instalaciones a gran escala, y entre sus desventajas principales, requieren un espacio considerable y reducen en mayor proporción la producción de energía en condiciones de sombreado parcial o cambios en la orientación de los paneles.

Inversores más pequeños son los inversores de cadena, son adecuados para instalaciones residenciales o comerciales medianas. En estos sistemas, varios paneles solares se conectan en serie en cadenas y cada una está conectada a un inversor. Sus principales ventajas son la flexibilidad y el control, ya que es posible analizar cada cadena por separado.

No siempre se cuenta con una gran cantidad de paneles solares, los Microinversores son una opción dedicada a este tipo de sistemas, son pequeños y su instalación es directamente en cada panel. Esto ayuda a optimizar el rendimiento de cada panel individualmente, evitando las pérdidas asociadas a sombreados o fallas en un solo panel.

Los inversores no solo están dedicados a la conversión de corriente continua en alterna, existen inversores que hacen el proceso inverso, sin embargo, no son considerados en este trabajo. También, éstos realizan funciones adicionales, como la optimización del rendimiento de los paneles y el monitoreo del sistema.

2.2.3. Baterías

En el rubro de los sistemas FV existen dos tipos: los sistemas interconectados a la red eléctrica, que generan energía a lo largo del día, la cual puede ser compartida con la red eléctrica pública, en caso de no consumirla por completo. Por la noche o en días nublados, los edificios o estructuras con estos sistemas requieren de un suministro externo de energía, debido a que no se genera energía o se hace de manera intermitente. Por su parte, los sistemas autónomos están dedicados a la generación así como al almacenamiento de energía sobrante, de manera que esté disponible durante la noche o en días nublados. La manera en que éstos almacenan la energía es mediante baterías y su principal ventaja es la capacidad de ofrecer energía estable, independiente y de manera continua, esto genera una seguridad en situaciones de emergencia o cortes de electricidad.

En los sistemas de almacenamiento de energía se consideran tres tipos principales de baterías: las baterías monoblock son utilizadas para instalaciones de autoconsumo de baja y media potencia, esto se refiere a casas donde solamente se utilizan para iluminación y electrodomésticos. Las baterías monoblock son las baterías más económicas, pero también son las que cuentan con la menor duración en términos de ciclos de carga y descarga. Con el correcto mantenimiento, una batería de este tipo podría durar entre los 4 y los 10 años de vida. Para instalaciones donde el consumo es medio o alto, se sugiere el uso de baterías estacionarias, normalmente se requieren en talleres con maquinaria o casas con electrodomésticos de mayor potencia. Su duración es mayor a las baterías anteriores, prometen un aumento el cual puede ser de hasta 20 años, aunque esto implica un aumento en el coste. Las baterías más utilizadas en la actualidad son las baterías de litio, su baja necesidad de mantenimiento y gran durabilidad las han colocado como una de las mejores opciones del mercado en relación coste-capacidad, [35].

Al elegir una batería para paneles solares, es importante considerar distintos factores más allá del costo.

- a) **Capacidad:** Es importante realizar un estudio de consumo para determinar la capacidad necesaria según la demanda. La cantidad de energía que puede almacenar la batería es determinada por el amperaje, el voltaje y la velocidad de descarga.
- b) **Potencia nominal:** Indica cuánta electricidad puede suministrar la batería en un momento dado, medida en kilowatts.
- c) **Nivel de autodescarga:** Una autodescarga lenta en una batería es esencial para aprovechar mejor la energía almacenada, ya que una rápida autodescarga reduce la eficiencia de la batería.
- d) **Mantenimiento requerido:** El mantenimiento adecuado es importante para prolongar la vida útil de la batería. Esto incluye limpiar conexiones y terminales, entre otros cuidados específicos según el tipo de batería y panel solar.
- e) **Vida útil y ciclos:** La vida útil depende de su rendimiento y ciclos de carga y descarga. Las baterías de litio, por ejemplo, tienen alrededor de 6000 ciclos, lo que equivale a una duración de entre 10 a 15 años.
- f) **Certificación y seguridad:** Es importante que las baterías cumplan con las normas de seguridad y certificaciones para asegurar su correcta instalación y uso, evitando riesgos debido a las sustancias químicas que contienen.

Estos son los aspectos primordiales a considerar para asegurar una elección informada para un sistema de energía solar eficiente y duradero.

2.2.4. Estructura y conexión

La estructura de montaje de un sistema FV es fundamental para su estabilidad y rendimiento. Existen diferentes tipos de estructuras dependiendo de la ubicación de los paneles y el tipo de instalación.

Las estructuras fijas son las más comunes y están diseñadas para mantener los paneles solares en un ángulo fijo, el cual dependerá de la ubicación del sistema. Las bases de estos sistemas pueden ser de aluminio, acero galvanizado u otros materiales resistentes a los problemas climáticos. Gracias

al uso de materiales sencillos, son las estructuras con menor costo en el mercado y pueden ser instaladas casi en cualquier lugar. Los modelos más actuales en el mercado están preparados para aprovechar de mejor manera la energía solar; las estructuras móviles, también conocidas como seguidores solares, ajustan la orientación de los paneles solares según sea el movimiento del sol durante el día. Contrario a las estructuras fijas, éstas generan una inversión mayor para poder aprovechar más la energía y son utilizadas en lugares donde la radiación solar es muy variable. Al igual que la elección de baterías, la elección de estructuras de montaje debe considerar diversos factores, como la orientación, la inclinación y la ubicación geográfica de los paneles; una correcta elección maximiza su rendimiento a lo largo del año.



Figura 2.4: Estructuras de paneles solares: fija (a) y móvil (b).

El sistema de cableado y conexión es crucial para garantizar la seguridad y la eficiencia de un sistema FV, ya que los cables permiten el flujo de electricidad entre los paneles solares, los inversores y las baterías. Es fundamental que los cables sean seleccionados y dimensionados correctamente para soportar las cargas eléctricas y las condiciones ambientales a las que estarán expuestos. Los paneles solares se conectan entre sí en serie y en paralelo para formar un sistema de cadena, cuyas conexiones van hacia un inversor. En los sistemas con almacenamiento de energía, los cables también conectan los paneles al banco de baterías.

En sistemas FV, los paneles suelen incorporar diodos *bypass* para evitar que la falla de una celda o sección afecte de manera significativa la operación de toda la cadena. Estos dispositivos permiten que la corriente se desvíe alrededor de las partes defectuosas, mitigando las pérdidas de potencia y manteniendo la continuidad del suministro energético. Además, para proteger el sistema de posibles sobrecargas o cortocircuitos, se emplean fusibles, interruptores y reguladores de carga. Los fusibles protegen los cables contra sobrecargas, los interruptores permiten desconectar el sistema durante su mantenimiento, y los reguladores de carga aseguran que las baterías no sufran sobrecargas ni descargas profundas, lo que podría reducir su vida útil. La correcta instalación y el mantenimiento adecuado de todos los componentes, incluido el cableado, son esenciales para asegurar tanto el rendimiento óptimo del sistema como su seguridad.

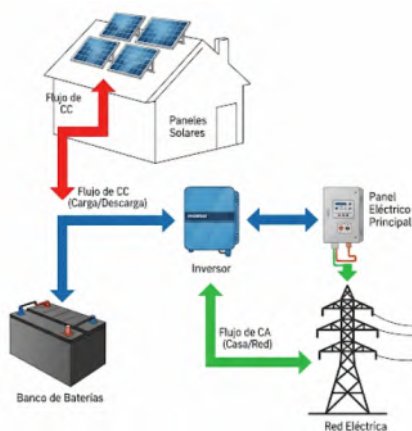


Figura 2.5: Conexión completa de un sistema FV

2.3. Energía fotovoltaica y otras formas de energía renovable

La energía fotovoltaica es una de las tecnologías más utilizadas dentro del grupo de energías renovables, pero existen otras formas de energía limpia, como la energía eólica y la hidroeléctrica. A continuación, se presenta una breve comparación entre la energía fotovoltaica y algunas de estas fuentes en términos de eficiencia, costos y accesibilidad.

La eficiencia de un sistema es un parámetro clave que mide, en porcentaje, la cantidad de energía eléctrica que genera con base en la energía que recibe; por ejemplo, un panel con una eficiencia del 18 % convierte el 18 % de la energía solar en electricidad, mientras que el resto se pierde en forma de calor o se refleja. La producción de los sistemas FV es típicamente baja en comparación con otras fuentes de energía renovable. Los paneles solares comerciales tienen un factor de conversión promedio de entre 15 % y 22 %, mientras que en las turbinas eólicas, la eficiencia puede alcanzar hasta un 40 % en condiciones óptimas. En cuanto a la energía hidroeléctrica, logran una capacidad de conversión de energía potencial en energía eléctrica de alrededor del 36 %.

Para evaluar la efectividad de los paneles solares, se utilizan diversos métodos. Uno de los más comunes es las condiciones de prueba estándar (STC), bajo las cuales se evalúa el rendimiento de los paneles en un ambiente controlado con una irradiancia de 1000 W/m^2 , una temperatura de 25°C y un ángulo de incidencia óptimo. Además de las STC, otra forma de medir la capacidad es a través de la eficiencia de conversión global, que considera las condiciones reales de funcionamiento del panel, teniendo en cuenta factores como la irradiancia variable, la temperatura y otros elementos ambientales que pueden afectar el rendimiento [36]. El *performance ratio* (PR) es una técnica que mide el rendimiento anual de un sistema FV; es un factor que tiene en cuenta las pérdidas que ocurren en el sistema, tales como las debidas a sombras, suciedad en los paneles, pérdidas en el inversor, entre otras. El valor de este ratio suele variar entre 0.75 y 0.85, lo que significa que, en promedio, un sistema FV produce entre el 75 % y el 85 % de la energía que sería posible obtener en condiciones ideales [37]. En principio, la eficiencia de la energía fotovoltaica es mínima, pero ésta ha venido mejorando gracias a los avances tecnológicos, y la ventaja de los sistemas FV es su capacidad para generar electricidad sin necesidad de equipos grandes o instalaciones costosas, lo que la hace más accesible en muchos lugares.

En términos de costos, se puede comparar entre dos tipos: costos de instalación, en donde la

energía fotovoltaica ha logrado reducirlos en más de un 70 % en los últimos 10 años, lo que ha facilitado su adopción a gran escala. A pesar de que la inversión inicial en sistemas FV sigue siendo alta, ésta es la más barata entre las energías renovables. La energía eólica también ha disminuido sus costos, existen turbinas de generación que se adaptan a necesidades menores, sin embargo sigue siendo bastante alto su valor. La energía hidroeléctrica es la más cara, al requerir una inversión significativa en infraestructura, especialmente para grandes proyectos. En términos de costos operativos a largo plazo, la energía fotovoltaica es más barata debido a la falta de combustibles y las bajas necesidades de mantenimiento. El principal proceso de mantenimiento es la limpieza de éstos y no necesita de químicos especiales, basta con el uso de agua tibia [38]. La Tabla 2.1 es extraída de [39], se presenta la comparación en algunos de los costos de las principales fuentes de energía renovable, así como el cambio desde el año 2010.

Tabla 2.1: Costo total instalado, factor de conversión y costo promedio de la electricidad (CPE) por tecnología, 2010 y 2022

Tecnología	Costo de instalación			Factor de conversión			CPE	
	2010	2022	% cambio	2010	2022	% cambio	2010	2022
Bioenergía	2 904	2 162	-26 %	72	72	0 %	0.082	0.061
Geotermia	2 904	3 478	20 %	87	85	-2 %	0.053	0.056
Hidroeléctrica	1 407	2 881	105 %	44	46	5 %	0.042	0.061
Solar FV	5 124	876	-83 %	14	17	23 %	0.445	0.049
CSP	10 082	4 274	-58 %	30	36	19 %	0.380	0.118
Eólica terrestre	2 179	1 274	-42 %	27	37	35 %	0.107	0.033
Eólica marina	5 217	3 461	-34 %	38	42	10 %	0.197	0.081

La accesibilidad de la energía fotovoltaica es una de sus principales ventajas. Los sistemas FV pueden instalarse en una variedad de ubicaciones, desde techos residenciales hasta grandes instalaciones en terrenos deshabitados. Esta flexibilidad permite que la energía fotovoltaica sea una opción viable en áreas rurales y remotas. Por su parte, la energía eólica requiere grandes áreas de terreno y condiciones de viento favorables, lo que limita su accesibilidad en ciertas regiones. La energía hidroeléctrica, aunque muy productiva, requiere la construcción de presas y una ubicación cercana a fuentes de agua, lo que puede ser costoso y ambientalmente impactante.

En resumen, la energía fotovoltaica se destaca por su alta accesibilidad y bajos costos operativos, pero su eficiencia y rendimiento pueden ser inferiores a los de otras fuentes de energía renovable, especialmente en grandes instalaciones. Sin embargo, su capacidad para ser instalada en una variedad de lugares la hace especialmente atractiva para aplicaciones residenciales y pequeñas comunidades.

2.4. Tendencias y Futuro de la Tecnología Fotovoltaica

La tecnología fotovoltaica ha experimentado uno de los crecimientos más rápidos en las últimas décadas, y se espera que continúe evolucionando de esta manera. Las tendencias actuales en la industria están centradas en la necesidad urgente de soluciones energéticas limpias, accesibles y

sostenibles para combatir el cambio climático. A medida que los costos de producción disminuyen y las eficiencias mejoran, los sistemas FV se están convirtiendo en una opción cada vez más viable para el suministro de energía en una amplia gama de aplicaciones.

Uno de los desarrollos más prometedores en la tecnología fotovoltaica es la mejora en la eficiencia de los paneles solares. Esto ha venido en aumento gracias a avances en materiales y diseños, como las células solares de perovskita y las células solares de doble y triple unión. Estas tecnologías permiten captar una mayor cantidad de energía solar, lo que resulta en paneles más pequeños y potentes. Los resultados de análisis en células solares de perovskita han mostrado un gran potencial para ofrecer paneles más baratos y productivos [40]. Este material es más fácil de producir en masa y tiene un mayor rendimiento con una menor incidencia de energía solar.

Otro avance importante está relacionado con los sistemas FV integrados en la arquitectura (BIPV, por sus siglas en inglés). Estos sistemas permiten la integración de paneles solares en edificios y casas de manera estética y funcional. En lugar de instalar paneles solares tradicionales sobre estructuras existentes, los sistemas BIPV convierten partes del edificio, como ventanas o tejados, en generadores de energía. Esto no solo ayuda con la disminución de huella de carbono, sino que también logra una generación de energía en lugares donde no se cuenta con el espacio necesario, como ciudades [41].

Las mejoras en el almacenamiento de energía no pueden quedarse atrás. El uso de baterías ha sido una limitación en los sistemas FV, especialmente en áreas sin conexión a la red eléctrica. Sin embargo, los avances en baterías de iones de litio, así como en tecnologías emergentes como las baterías de sodio-azufre y las baterías de flujo, están ayudando a superar este obstáculo. Estas tecnologías de almacenamiento mejoradas no solo permiten a los usuarios almacenar energía para su uso nocturno o en días nublados, sino que también contribuyen a la estabilidad de las redes eléctricas al equilibrar la oferta y la demanda de electricidad [42].

El uso de nuevas tecnologías como la inteligencia artificial y el análisis de datos ha comenzado a tomar un papel importante en la optimización de los sistemas FV. Mediante el uso de algoritmos avanzados, los sistemas FV pueden predecir la producción de energía en función de las condiciones meteorológicas, ajustar la orientación de los paneles para maximizar la captación solar y realizar mantenimientos predictivos para evitar fallas en los componentes. Esto no solo mejora la eficiencia operativa de los sistemas FV, sino que también reduce los costos asociados con el mantenimiento y la gestión de grandes instalaciones solares.

Las redes eléctricas inteligentes (*smart grids*) permiten la integración de fuentes de energía renovable, como la solar, de una manera más eficiente, gestionando de manera más efectiva la distribución de la electricidad en función de la oferta y la demanda. Las redes inteligentes también permiten a los consumidores monitorear su consumo y producción de energía en tiempo real, lo que facilita la adopción de sistemas FV domésticos y comerciales [43]. En el futuro, la combinación de paneles solares, almacenamiento y redes inteligentes podría crear una infraestructura energética más descentralizada, donde los hogares y las empresas no solo consumen energía, sino que también pueden contribuir activamente a la estabilidad de la red.

El futuro de la tecnología fotovoltaica está marcado por un panorama de constantes innovaciones tecnológicas que mejorarán la eficiencia, la accesibilidad y la integración de la energía solar. Con la creciente inversión en investigación y el apoyo de políticas públicas, se espera que la energía solar juegue un papel central en la transición hacia un sistema energético global más sostenible y resiliente.

Comentarios finales

En este capítulo se ofrece una perspectiva general de la energía solar como una alternativa estratégica frente al uso de combustibles fósiles, haciendo énfasis en la tecnología fotovoltaica. Se abordan los principios físicos que permiten transformar la radiación solar en electricidad y se describen los componentes principales que conforman un sistema FV. Tras introducir los conceptos básicos de la energía solar, se profundiza en el efecto fotovoltaico, explicando de forma clara cómo la luz produce corriente eléctrica dentro de las células solares. Asimismo, se contextualiza la energía fotovoltaica frente a otras fuentes renovables, destacando su relevancia dentro del panorama global de sostenibilidad. Finalmente, se revisan las tendencias actuales y los avances tecnológicos que apuntalan su papel en la transición hacia un modelo energético más limpio, eficiente y descentralizado. Esta base conceptual servirá, en los capítulos siguientes, para profundizar en el análisis técnico y metodológico de los sistemas FV, particularmente en lo referente a su diagnóstico y mejora operativa.

Capítulo 3

Métodos de Diagnóstico de Fallas en Sistemas Fotovoltaicos

A lo largo del capítulo anterior se ha mostrado la arquitectura de los sistemas FV y la importancia de mantener una correcta conversión energética para mantenerse al margen de otras fuentes de energía limpias. Por ello, la teoría de diagnóstico se ha dedicado a identificar y solucionar los problemas a tiempo, mejorando la eficiencia del sistema, y previniendo costosos daños a largo plazo. En este capítulo se introducen los principales métodos de diagnóstico utilizados para detectar las posibles fallas, divididas en dos categorías principales: permanentes, como delaminación, burbujas, amarilleamiento de celdas, rayones y celdas o inversores quemados; y temporales, las cuales se deben principalmente a efectos de sombreado parcial, y suciedad por acumulación de polvo u otros elementos. Diagnosticar fallas en un sistema FV implica una combinación de técnicas de inspección visual, mediciones eléctricas, herramientas avanzadas como la termografía infrarroja y tecnologías más recientes como el monitoreo remoto. Cada uno de estos métodos tiene ventajas y puede proporcionar información valiosa para identificar la causa de las fallas en los componentes del sistema. A continuación, se proporciona una breve descripción de la manera en la que actúan estas técnicas dividiéndolas en dos grupos.

3.1. Métodos basados en el modelo

Los métodos basados en modelo utilizan representaciones matemáticas del comportamiento esperado de un sistema FV. La idea principal es obtener predicciones teóricas que luego se comparan con los datos reales para identificar cambios que podrían representar fallas.

3.1.1. Análisis de Curva I-V

La prueba de Curva I-V (Voltaje sobre Corriente) es una de las técnicas más utilizadas para caracterizar el rendimiento de un panel solar; esta muestra cómo varía la corriente que entrega un módulo fotovoltaico en función del voltaje que se le aplica. Para obtenerla, se mide la salida del módulo mientras se cambia la carga desde cortocircuito ($V = 0$) hasta circuito abierto ($I = 0$). La curva resultante de las mediciones reales es comparada con la curva esperada según el modelo matemático en condiciones controladas; de esta manera, se obtiene información sobre la eficiencia y la capacidad del panel para generar electricidad. Una anomalía en la forma de la curva, como

disminuciones en la corriente, puede señalar fallas en las celdas, en alguna conexión o debido a algún sombreado.

Realizar una prueba de curva I-V implica seguir varios pasos para garantizar mediciones precisas y útiles. A continuación se describe de manera breve el procedimiento a seguir en una prueba, [44].

1. **Medición de valores.** Es necesario que el sistema cuente con un trazador de curvas o sistema de medición con una buena precisión. Además, se toman en cuenta valores como la irradiancia y la temperatura, valores que afectan la dinámica.
2. **Obtención de la curva I-V.** Se realiza un barrido de voltaje desde 0 hasta el circuito abierto mediante electrónica de potencia o un circuito de carga variable, y se registra la corriente en cada punto del voltaje.
3. **Extracción de parámetros.** Una vez obtenida la curva, es de suma importancia identificar los valores característicos de la misma, los cuales son: el punto de máxima potencia (P_{max}), así como su voltaje (V_{mp}) y corriente (I_{mp}); la corriente de cortocircuito (I_{sc}), el voltaje de circuito abierto (V_{oc}), el factor de forma y la eficiencia global del panel.
4. **Comparación.** La curva obtenida es comparada con la curva ideal del fabricante o con la curva esperada en condiciones de medición, en donde se considera una irradiancia de $1000 \frac{W}{m^2}$ y una temperatura de $25^\circ C$ en el módulo. Se observan desviaciones como disminución de V_{oc} , I_{sc} o deformaciones en la curva.

Existen diferentes comportamientos característicos asociados a fallas particulares de un sistema; la Figura 3.1 extraída de [45] los ilustra. Una caída en I_{sc} es generada la mayor parte del tiempo por suciedad, sombreado o deslaminación; la caída en V_{oc} normalmente aparece cuando existe una degradación o celdas dañadas; los codos o escalones presentes en la curva I-V son presentados al activarse los diodos bypass por fallas parciales; si existe una pérdida general de forma en la curva, la causa más común es una conexión suelta o degradación uniforme en las celdas solares, y la presencia de múltiples picos en la curva surge a partir de fallas en las cadenas combinadas.

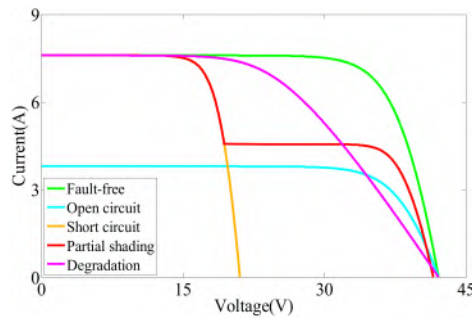


Figura 3.1: Curvas I-V representativas de distintos comportamientos en un sistema FV, incluyendo una curva estable y otras afectadas por diversas fallas.

El análisis de curvas I-V presenta varias ventajas que lo convierten en una herramienta apta para el diagnóstico del estado eléctrico de los paneles FV. Es un método no destructivo, lo que significa que no daña el módulo durante el ensayo, y permite obtener un diagnóstico preciso y cuantificable. Además, es capaz de detectar distintos tipos de fallas, como suciedad, sombreado o diodos defectuosos, y facilita la comparación directa con los valores nominales del fabricante. Otra

ventaja es su aplicabilidad en distintos niveles del sistema FV, ya sea a nivel de módulo, string o campo completo. Con el equipo adecuado, las mediciones pueden realizarse de forma rápida y relativamente sencilla. No obstante, este tipo de análisis también presenta ciertas desventajas. Su precisión depende de condiciones ambientales como la irradiancia y la temperatura, y requiere de instrumentación especializada, como trazadores de curvas y sensores. Además, no puede utilizarse como una herramienta de monitoreo continuo en tiempo real y, en algunos casos, puede ser necesario desconectar el sistema para realizar las mediciones. Aunque útil, no permite identificar todas las fallas físicas, como la delaminación o el amarillamiento, y se requiere experiencia para interpretar correctamente las curvas anómalas.

3.1.2. Modelos Eléctricos Equivalentes

Los modelos eléctricos equivalentes representan el comportamiento de los módulos fotovoltaicos mediante circuitos eléctricos conformados por componentes ideales como fuentes de corriente, diodos, resistencias en serie y paralelo. El objetivo de estos modelos es simular la respuesta eléctrica de un panel ante diferentes condiciones de operación, gracias a la precisión con la que se realiza, ésta permite predecir su desempeño y detectar desviaciones que puedan indicar fallas. El modelo más utilizado es el modelo de un diodo, aunque también existen versiones más complejas como el modelo de dos diodos, que ofrecen mayor precisión en ciertas condiciones de operación. Estos modelos se ajustan utilizando parámetros que se obtienen a partir de datos experimentales o de las hojas técnicas del fabricante. El modelo de un diodo se describe con la siguiente ecuación, [46]:

$$I = I_{ph} - I_0 \left(e^{\frac{q(V+IR_s)}{nkT}} - 1 \right) - \frac{V + IR_s}{R_{sh}} \quad (3.1)$$

donde:

- I_{ph} : corriente fotogenerada.
- I_0 : corriente de saturación del diodo.
- R_s : resistencia serie.
- R_{sh} : resistencia en paralelo (shunt).
- n : factor de idealidad del diodo.
- q : carga del electrón.
- k : constante de Boltzmann.
- T : temperatura absoluta.

La ecuación 3.1 combina el comportamiento no lineal de conducción del diodo con los efectos resistivos derivados de su estructura física y del material semiconductor. En conducción directa, la corriente del diodo depende exponencialmente del voltaje aplicado e incluye parámetros como la corriente de saturación, el factor de idealidad y el voltaje térmico, que incorporan la influencia de la temperatura. Adicionalmente, se incorpora una resistencia serie R_s , que modela las pérdidas óhmicas en los contactos metálicos y en el semiconductor; dicha resistencia reduce el voltaje efectivo aplicado a la unión PN, efecto que se vuelve especialmente relevante a altas corrientes.

El procedimiento general para el uso de modelos eléctricos equivalentes en la detección de fallas comienza con la **identificación del modelo**. Se selecciona un modelo apropiado (uno o dos diodos) en función del nivel de precisión requerido y de los datos disponibles. Posteriormente, se realiza un **ajuste de parámetros**. Se calibran los parámetros del modelo usando datos de mediciones experimentales o simulaciones en condiciones estándar. Existen algoritmos de optimización como mínimos cuadrados o métodos heurísticos que permiten este ajuste. Para obtener una base se obtiene una **simulación de comportamiento**. Se simula la curva I-V o la curva P-V bajo distintas condiciones operativas. Finalmente, se realiza una **comparación con datos reales**. Se confronta

la respuesta simulada con los datos obtenidos del sistema FV. Desviaciones significativas pueden ser indicativas de fallas como pérdidas por resistencia, defectos en los materiales, sombreado, entre otros.

Este tipo de análisis resulta bastante efectivo, entre sus principales ventajas resaltan las siguientes: permite una caracterización completa del comportamiento eléctrico del módulo, facilita la detección de fallas mediante comparación con curvas simuladas, adaptable a diferentes niveles de complejidad según el modelo utilizado, permite simulaciones bajo condiciones variadas de operación y es utilizable como base para sistemas de monitoreo automatizado. Todos estos puntos posicionan a la técnica como una de las mejores, sin embargo, una mejor eficacia en el método implica algunas desventajas en el sistema: la calibración del modelo puede ser compleja y sensible al ruido en los datos, requiere información precisa del fabricante o mediciones detalladas, por lo que se requiere de modelos más complejos y que resultan costosos computacionalmente. La calidad del diagnóstico depende de la exactitud en el ajuste de parámetros y algunos parámetros del modelo no tienen una interpretación física directa.

3.1.3. Modelos de Rendimiento

Los modelos de rendimiento buscan estimar la producción de energía de un sistema FV bajo diferentes condiciones de operación. Estos modelos se basan en expresiones matemáticas que relacionan la potencia generada con variables como la irradiancia solar, la temperatura del módulo, las pérdidas ópticas y eléctricas, y la eficiencia del inversor.

El objetivo principal es predecir la energía generada por el sistema y detectar desviaciones significativas respecto al comportamiento esperado. Estas desviaciones pueden indicar problemas como suciedad en los módulos, sombreado parcial, degradación o fallas en el sistema.

Una forma común de representar el rendimiento es mediante el *Performance Ratio* (PR), definido como, [47]:

$$PR = \frac{E_{AC}}{P_{STC} \cdot H_{POA}} \quad (3.2)$$

La ecuación 3.2 define el *performance ratio* (PR) como una métrica que relaciona la energía útil producida por el sistema FV con la energía que idealmente debería generar bajo condiciones estándar. El término E_{AC} representa la energía en corriente alterna entregada por el inversor, mientras que P_{STC} corresponde a la potencia nominal del generador medida bajo condiciones estándar de prueba (*Standard Test Conditions*). Por su parte, H_{POA} denota la irradiancia global acumulada sobre el plano del arreglo durante el periodo de evaluación. Al normalizar la producción real contra el potencial teórico, el PR permite cuantificar el impacto de pérdidas térmicas, eléctricas y operativas, proporcionando una medida objetiva del desempeño global e independiente de las variaciones climáticas. Un valor de PR consistentemente bajo puede reflejar pérdidas no esperadas en el sistema y sirve como indicador de posibles fallas.

El uso de modelos de rendimiento típicamente sigue este esquema:

1. **Definición del modelo.** Se selecciona un modelo adecuado (empírico, semiempírico o basado en simulaciones).
2. **Entrada de datos.** Se recopilan datos de entrada como irradiancia, temperatura ambiente, temperatura del módulo y potencia generada.

3. **Estimación del rendimiento esperado.** Se calcula la potencia o energía esperada con base en los datos de entrada.
4. **Comparación con datos reales.** Se contrastan los valores esperados con los medidos. Diferencias significativas pueden indicar fallas o condiciones anómalas.

Los modelos de rendimiento presentan varias ventajas significativas. En primer lugar, permiten estimar la energía generada y detectar desviaciones con relativa facilidad, lo cual los hace útiles para el monitoreo continuo del sistema. Además, su implementación requiere pocos parámetros y pueden integrarse fácilmente con sistemas SCADA (Supervisory Control and Data Acquisition) o plataformas de supervisión. También son herramientas valiosas para evaluar la eficiencia y rentabilidad del sistema FV.

Sin embargo, estos modelos también presentan algunas desventajas. Su precisión depende en gran medida de la calidad de los datos meteorológicos y operativos disponibles. No proporcionan información detallada sobre la ubicación exacta o el tipo de falla, y son sensibles a errores de calibración de sensores o entradas. Asimismo, pueden requerir ajustes periódicos para mantener su precisión, y no modelan directamente ciertos efectos como fallas eléctricas internas.

3.1.4. Modelos Térmicos

Los modelos térmicos representan la relación entre la temperatura de operación de los módulos fotovoltaicos y las condiciones ambientales, tales como la irradiancia solar y la temperatura ambiente. Esta temperatura influye directamente en el rendimiento eléctrico del módulo, ya que un aumento de temperatura generalmente reduce la eficiencia de conversión. El modelo térmico más común se expresa mediante una ecuación lineal, [48]:

$$T_{mod} = T_{amb} + \frac{G}{G_{ref}} \cdot \Delta T \quad (3.3)$$

o también:

$$T_{mod} = T_{amb} + \left(\frac{NOCT - 20}{800} \right) \cdot G \quad (3.4)$$

Las ecuaciones 3.3 y 3.4 describen el comportamiento térmico de los módulos fotovoltaicos en función de las condiciones ambientales. La primera expresión establece que la temperatura del módulo T_{mod} se obtiene al sumar a la temperatura ambiente T_{amb} un incremento proporcional a la irradiancia G , normalizada respecto a una irradiancia de referencia G_{ref} , y ponderada mediante el parámetro ΔT , que representa el aumento típico de temperatura del módulo bajo irradiancia estándar. La segunda forma utiliza el parámetro $NOCT$ (*Nominal Operating Cell Temperature*), definido como la temperatura alcanzada por el módulo bajo condiciones nominales de operación, para expresar directamente la variación térmica como una función lineal de la irradiancia G . En ambos casos, se refleja que la temperatura del módulo aumenta con la irradiancia y puede exceder significativamente la temperatura ambiente, lo cual impacta negativamente la eficiencia eléctrica del dispositivo.

Los modelos térmicos permiten estimar las pérdidas térmicas y ajustar los modelos eléctricos o de rendimiento según la temperatura real de operación.

1. **Determinación de condiciones ambientales.** Se miden o estiman la irradiancia y la temperatura ambiente.

2. **Estimación de la temperatura del módulo.** Se aplica el modelo térmico para obtener la temperatura real del módulo.
3. **Corrección de parámetros eléctricos.** Se ajustan los parámetros del modelo eléctrico para reflejar el efecto de la temperatura.
4. **Detección de anomalías.** Si la temperatura observada se desvía significativamente del modelo, puede ser indicativa de obstrucciones, suciedad, sobrecalentamiento o defectos térmicos.

Los modelos térmicos ofrecen diversas ventajas en el análisis de sistemas FV. Permiten ajustar modelos eléctricos a condiciones reales de temperatura, lo que los hace útiles para estimar pérdidas térmicas en condiciones extremas. Además, apoyan la detección de fallas relacionadas con sobrecalentamiento y son fáciles de implementar cuando se dispone de sensores ambientales básicos. Estos modelos también pueden complementar otros enfoques, como los modelos eléctricos o de rendimiento, para lograr un diagnóstico más completo del sistema.

Por otro lado, presentan algunas limitaciones. Requieren sensores térmicos precisos o supuestos que en la práctica podrían no cumplirse. En muchos casos, el modelo no logra capturar efectos dinámicos ni la acumulación térmica, y su desempeño depende en gran medida de la calidad del modelo y del montaje del sistema. La precisión también se ve reducida en instalaciones con ventilación forzada o estructuras especiales. Además, estos modelos no distinguen entre distintas causas térmicas, como la suciedad o el sombreado.

3.2. Métodos Basados en Datos

Los métodos basados en datos aprovechan grandes volúmenes de información recopilada por sensores, sistemas SCADA o tecnologías como drones e IoT. Estos métodos se apoyan en estadística avanzada e inteligencia artificial para detectar anomalías.

3.2.1. Monitoreo en Tiempo Real

El monitoreo en tiempo real es una técnica que permite observar continuamente el funcionamiento de un sistema FV mediante la recopilación y el análisis inmediato de datos operativos. Esta estrategia busca identificar de manera instantánea cualquier desviación del comportamiento normal del sistema, permitiendo detectar fallas o anomalías justo en el momento en que ocurren. Esto es crucial para maximizar la eficiencia energética y minimizar pérdidas económicas asociadas con tiempos prolongados de inactividad o mal funcionamiento, [49].

La implementación de esta técnica depende de la instalación de sensores en diversos puntos del sistema FV, tales como módulos, inversores y conexiones eléctricas. Estos sensores miden variables como voltaje, corriente, potencia, temperatura y niveles de irradiancia. Los datos generados se transmiten automáticamente a una unidad central, local o remota, que se encarga de procesarlos utilizando algoritmos de análisis, entre los que destacan la comparación con modelos físicos o con algoritmos de inteligencia artificial que son entrenados con un comportamiento ideal del sistema. Este procesamiento puede llevarse a cabo mediante plataformas SCADA, software especializado en energía solar, o sistemas basados en la nube con capacidades de visualización y alerta. La supervisión se realiza de forma continua, y los datos se actualizan con alta frecuencia, permitiendo detectar cambios sutiles en el comportamiento del sistema. Algunas plataformas permiten el almacenamiento de información, logrando una evaluación histórica para determinar si un cambio

anómalo es puntual o persistente, lo que ayuda a distinguir entre fallas temporales y problemas estructurales.

El principal beneficio del monitoreo en tiempo real radica en su capacidad de respuesta rápida. Al detectar una falla en el momento en que ocurre, se pueden tomar medidas correctivas de inmediato, evitando que la situación se agrave o que se acumulen pérdidas significativas de energía. Además, permite planificar el mantenimiento con mayor precisión, reduciendo intervenciones innecesarias. Sin embargo, también existen desafíos. La instalación inicial de los equipos de monitoreo puede representar un costo elevado, y su operación depende de tecnologías digitales que requieren mantenimiento, actualizaciones y protección frente a posibles ataques informáticos. Asimismo, el análisis de los datos puede requerir personal especializado, sobre todo si se utilizan herramientas avanzadas de diagnóstico.

3.2.2. Aprendizaje Automático

El aprendizaje automático, o machine learning, se ha convertido en una herramienta poderosa para la detección de fallas en sistemas FV, gracias a su capacidad para identificar patrones complejos en grandes volúmenes de datos. A diferencia de los métodos tradicionales basados en umbrales fijos o reglas específicas, el aprendizaje automático permite desarrollar modelos que aprenden directamente del comportamiento histórico del sistema, adaptándose a las variaciones propias de cada instalación fotovoltaica, [50].

Este enfoque se basa en el entrenamiento de algoritmos que reciben como entrada datos recopilados durante el funcionamiento normal y defectuoso del sistema. Estos datos pueden incluir variables como la irradiancia solar, la temperatura ambiente, la temperatura de los módulos, la corriente y el voltaje en distintos puntos del sistema, así como la potencia generada. A partir de estos datos, el modelo aprende a reconocer qué combinaciones de valores corresponden a un funcionamiento saludable y cuáles indican la presencia de una falla, como puede ser un módulo desconectado, suciedad excesiva, sombreado parcial, fallas en los inversores o pérdidas en los cables. Una vez entrenado, el modelo se implementa para evaluar en tiempo real los nuevos datos del sistema. Si el comportamiento observado se desvía significativamente del patrón aprendido como “norma”, el modelo emite una alerta indicando la posibilidad de una falla. La ventaja de este enfoque es su capacidad para detectar fallas incipientes o sutiles que podrían pasar desapercibidos mediante métodos convencionales. Por ejemplo, una disminución gradual en la eficiencia de un módulo puede ser detectada antes de que cause una caída significativa en la producción, lo cual permite intervenir con mayor anticipación.

Existen diferentes tipos de algoritmos de aprendizaje automático utilizados en esta tarea, como árboles de decisión, máquinas de soporte vectorial (SVM), redes neuronales artificiales, modelos de random forest o métodos de aprendizaje profundo. La elección del algoritmo depende de la complejidad del sistema, la cantidad de datos disponibles y la precisión deseada. Además, en muchos casos se utilizan técnicas de clasificación supervisada, donde el modelo se entrena con datos etiquetados previamente, aunque también se aplican enfoques no supervisados para detectar anomalías sin necesidad de datos de fallas conocidos. El principal beneficio del uso del aprendizaje automático es su capacidad de mejorar continuamente a medida que se dispone de más datos, haciendo que los modelos se vuelvan más precisos y específicos para cada instalación. También permite una detección más inteligente, capaz de adaptarse a condiciones variables como cambios estacionales o degradación natural de los componentes. No obstante, su implementación conlleva ciertos desafíos, como la necesidad de disponer de grandes volúmenes de datos de alta calidad, la

complejidad en el diseño y ajuste de los modelos, y la necesidad de expertos en ciencia de datos para interpretar los resultados y mantener actualizados los sistemas.

3.2.3. Análisis de Series Temporales

El análisis de series temporales es una metodología estadística utilizada para estudiar cómo varían las magnitudes medidas en un sistema a lo largo del tiempo. En el contexto de los sistemas FV, esta técnica permite examinar el comportamiento dinámico de variables como la potencia generada, el voltaje de salida o la corriente en distintos puntos del sistema. Estas mediciones, registradas a intervalos regulares, conforman series de datos que reflejan tanto el rendimiento del sistema como las condiciones ambientales que lo afectan. El análisis de estas series permite detectar desviaciones en el comportamiento esperado, lo que resulta útil para identificar fallas o degradaciones incipientes, [51]. La utilidad del análisis temporal radica en que muchas fallas no aparecen como eventos instantáneos claramente visibles, sino como cambios graduales, recurrentes o anómalos en los patrones de operación.

Para realizar este tipo de análisis, es necesario contar con datos históricos del sistema, preferentemente de alta resolución temporal, que permitan comparar el rendimiento actual con el comportamiento habitual en condiciones similares. Se utilizan técnicas estadísticas como la descomposición de series (tendencia, estacionalidad y ruido), la detección de outliers, el análisis de correlación cruzada y modelos autoregresivos como ARIMA, SARIMA o modelos más recientes como Prophet o LSTM (Long Short-Term Memory), estos últimos basados en redes neuronales especializadas en secuencias temporales. Uno de los enfoques más comunes es la construcción de un modelo que represente el comportamiento normal del sistema, utilizando datos históricos que no contienen fallas. Luego, el modelo se aplica a los datos en tiempo real y cualquier desviación significativa entre lo observado y lo esperado se interpreta como un posible indicio de falla. Este enfoque permite generar alertas automáticas sin necesidad de observar manualmente los datos.

Entre las ventajas del análisis de series temporales se encuentra su capacidad para detectar anomalías con base en patrones reales de comportamiento del sistema, sin requerir sensores adicionales ni modificaciones físicas. Además, permite realizar predicciones del rendimiento futuro y evaluar tendencias de degradación. Sin embargo, también presenta limitaciones, como la necesidad de contar con registros de buena calidad durante un periodo prolongado y la complejidad de ajustar modelos adecuados cuando existen múltiples factores externos que influyen en el sistema, como variaciones climáticas o estacionales.

El análisis de series temporales constituye una herramienta robusta para la detección de fallas en sistemas FV. Su capacidad para modelar el comportamiento dinámico del sistema y detectar anomalías sutiles lo hace especialmente útil en el monitoreo continuo y predictivo de instalaciones solares. A medida que se incrementa la disponibilidad de datos históricos de operación, esta técnica adquiere un papel cada vez más relevante en la gestión inteligente y automatizada de la energía solar.

3.2.4. Visión Artificial y Drones

La visión artificial asistida por drones ha emergido como una solución innovadora y eficiente para la inspección y detección de fallas en sistemas FV a gran escala. Esta técnica combina tecnologías de captura de imagen aérea con algoritmos de procesamiento digital para identificar anomalías visuales o térmicas en los módulos solares sin la necesidad de inspección física directa.

Su aplicación resulta especialmente valiosa en plantas solares de gran extensión, donde los métodos tradicionales de revisión manual son lentos, costosos y a menudo ineficaces, [52].

El proceso comienza con la programación de vuelos autónomos mediante drones equipados con cámaras de alta resolución, tanto en el espectro visible como en el infrarrojo. Estas cámaras permiten capturar imágenes detalladas de los paneles desde el aire, abarcando grandes áreas en tiempos reducidos. Las imágenes térmicas, en particular, permiten detectar puntos calientes (hotspots), que son signos típicos de celdas defectuosas, conexiones sueltas, suciedad localizada o sombreados parciales. Las cámaras RGB, por su parte, son útiles para identificar daños físicos como grietas, decoloraciones, roturas o acumulación de polvo y residuos. Una vez capturadas las imágenes, estas se procesan utilizando técnicas de visión artificial y aprendizaje profundo. Los algoritmos entrenados son capaces de segmentar las imágenes, identificar regiones de interés y clasificar visualmente los distintos tipos de fallas. Este análisis se puede realizar en plataformas locales o en la nube, y en muchos casos se integra con sistemas de información geográfica (GIS) para mapear con precisión la ubicación de cada módulo afectado. La automatización de esta etapa permite obtener reportes detallados en poco tiempo, facilitando la planificación de acciones correctivas.

La principal ventaja de esta técnica radica en su eficiencia y alcance. Un sólo dron puede inspeccionar cientos de módulos en cuestión de minutos, sin contacto físico y sin interrumpir la operación del sistema. Además, el uso de visión infrarroja permite detectar fallas invisibles al ojo humano, incluso en etapas tempranas. La repetibilidad y objetividad del proceso también representa una mejora sustancial respecto a la inspección visual tradicional. Sin embargo, su implementación requiere una inversión inicial significativa en equipamiento especializado y capacitación del personal. También pueden surgir desafíos relacionados con las condiciones climáticas, la calibración de los sensores o las restricciones regulatorias del uso de drones en ciertas regiones.

3.3. Análisis topológico de datos

El análisis topológico de datos representa un enfoque emergente en la ciencia de datos que permite estudiar la forma y la estructura de los datos. En el contexto de sistemas FV, el TDA ofrece un marco robusto para la detección de fallas al identificar patrones globales, ciclos, agrupamientos y cambios estructurales en conjuntos de datos multivariados, incluso en presencia de ruido o alta dimensionalidad. Cuando el sistema funciona correctamente, los datos recolectados suelen presentar una estructura topológica de manera repetitiva, por ejemplo, ciclos diarios o agrupamientos estables en el espacio de fases. Sin embargo, la aparición de una falla altera dicha estructura, con la aparición de deformaciones o cambios en la conectividad que pueden ser detectados con facilidad, [53].

A diferencia de los métodos clásicos que operan directamente sobre las variables medidas, el TDA transforma los datos en representaciones geométricas, sobre las cuales se aplican herramientas de topología algebraica. Esto permite capturar características invariantes de la forma de los datos, como componentes conexas, agujeros y cavidades de distintas dimensiones, las cuales son luego interpretadas en relación con el comportamiento del sistema. El resultado se representa en diagramas o códigos de barras que resumen la aparición y desaparición de características topológicas a diferentes escalas. Cambios grandes o patrones anómalos en estos diagramas pueden ser indicadores tempranos de fallas o de alteraciones en la dinámica del sistema.

Entre las principales ventajas del TDA se encuentra su capacidad para manejar datos complejos, no lineales y ruidosos, así como su independencia de modelos específicos o suposiciones estadísticas. Además, permite comparar topológicamente diferentes periodos de operación del sistema, facili-

tando la detección de eventos inusuales. No obstante, su aplicación requiere un conocimiento en herramientas específicas de topología computacional y, en muchos casos, un manejo sólido de la técnica para poder hacer la interpretación de los resultados obtenidos.

3.4. Comparación de Métodos

Los métodos basados en modelo utilizan representaciones matemáticas o físicas del comportamiento esperado de los sistemas FV, permitiendo identificar fallas mediante la comparación entre el desempeño real y el simulado. Entre estos métodos se incluyen el análisis de curvas I-V, los modelos eléctricos equivalentes y los modelos de rendimiento. Estas técnicas ofrecen diagnósticos precisos y detallados, especialmente útiles para detectar fallas eléctricas internas, degradación de componentes o pérdidas por resistencias. Sin embargo, suelen requerir conocimientos técnicos avanzados, condiciones de operación específicas y una instrumentación más compleja, lo que puede limitar su implementación en tiempo real o en sistemas de gran escala.

En contraste, los métodos basados en datos se apoyan en técnicas estadísticas o de inteligencia artificial para identificar patrones anómalos directamente a partir de grandes volúmenes de datos operacionales. Estos enfoques no dependen necesariamente de un conocimiento profundo del sistema físico y pueden adaptarse a múltiples condiciones ambientales, siempre que se disponga de datos representativos y de buena calidad. Aunque su capacidad de interpretación puede ser limitada en comparación con los modelos físicos, los métodos basados en datos destacan por su escalabilidad, flexibilidad y facilidad de integración en plataformas de monitoreo continuo, siendo ideales para aplicaciones en tiempo real.

Comentarios finales

Los métodos de diagnóstico revisados en este capítulo, tanto los basados en modelos como los basados en datos, ofrecen enfoques complementarios. Los primeros utilizan una base matemática sólida que permite comparar el desempeño real del sistema con su comportamiento esperado bajo condiciones controladas; sin embargo, aunque son útiles para realizar análisis detallados, dependen de una correcta parametrización. Por otro lado, los métodos basados en datos destacan por su capacidad de aprovechar el gran volumen de información disponible, adaptándose a la variabilidad de las condiciones operativas y facilitando la detección de fallas mediante técnicas modernas como el aprendizaje automático. Es importante mencionar que la elección de un método de diagnóstico depende tanto de los objetivos de la evaluación como de los recursos disponibles y resulta evidente que los métodos basados en datos han tomado bastante relevancia durante los últimos años. Aprovechando esta tendencia, en el presente trabajo se propone el uso del TDA, constituyéndose como una alternativa innovadora para complementar los enfoques tradicionales, proporcionando una perspectiva diferente que permite la detección de fallas con datos reales de operación.

Capítulo 4

Fundamentos del Análisis Topológico de Datos

Los grandes volúmenes de información que se generan día con día han impulsado la generación de nuevas técnicas de análisis que no dependan en gran medida del modelo y se adapten de manera frecuente al sistema. A finales del siglo pasado surge una técnica capaz de extraer distintas propiedades de los conjuntos de datos. Gracias a su enfoque, el TDA es capaz de extraer la “forma” de nubes de datos con base en la cantidad de hoyos n -dimensionales, los cuales son calculados mediante construcciones geométricas. Este capítulo está dedicado a dar una idea intuitiva de las herramientas topológico-geométricas utilizadas para analizar las nubes de datos en $\mathbb{R}^2$ que son extraídas de la generación de energía. Las definiciones formales están fuera de los objetivos de esta tesis debido a las bases matemáticas necesarias; si el lector está interesado en una introducción más completa, puede consultar las siguientes referencias [54], [55], [56].

A lo largo del capítulo se denota por $N = \{y_0, y_1, \dots, y_\ell\} \subset \mathbb{R}^2$ un conjunto de datos.

4.1. Complejos simpliciales

Debido a las propiedades topológicas de una nube de datos, no es posible realizar este análisis sin aproximarla mediante objetos topológicos que estén determinados por una cantidad finita de puntos. Los complejos simpliciales son objetos de este tipo que se introducen en esta sección.

Definición 4.1.1. *Un **d -simplejo** es la estructura geométrica generada por los puntos $x_0, x_1, \dots, x_d \in \mathbb{R}^n$ que no están contenidos en un subespacio de dimensión menor (es decir, que no son colineales, coplanares, etc.), es denotado por $\sigma = \langle x_0, x_1, \dots, x_d \rangle$.*

Por ejemplo, un punto es un 0-simplejo, Figura 4.1(a); dos puntos distintos generan un 1-simplejo, Figura 4.1(b); y, tres puntos no alineados generan un 2-simplejo, Figura 4.1(c).

Los vértices de σ son los puntos $x_0, x_1, \dots, x_d$, sus caras son simplejos formados por subconjuntos del mismo y la dimensión es $d-1$. A grandes rasgos, los simplejos son las versiones generalizadas de triángulos en diferentes dimensiones.

Los simplejos son la base para la construcción de los espacios con los que se aproximan las nubes de datos.

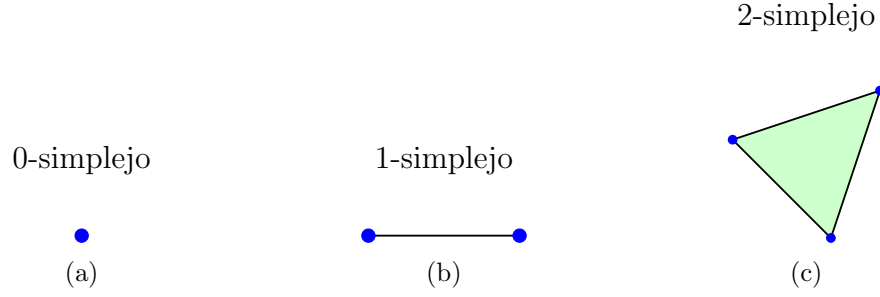


Figura 4.1: Simplejos de dimensiones 0, 1 y 2, formados por envoltorios de puntos en $\mathbb{R}^2$.

Definición 4.1.2. Un **complejo simplicial finito** K , es una colección finita de simplejos que se unen mediante sus caras.

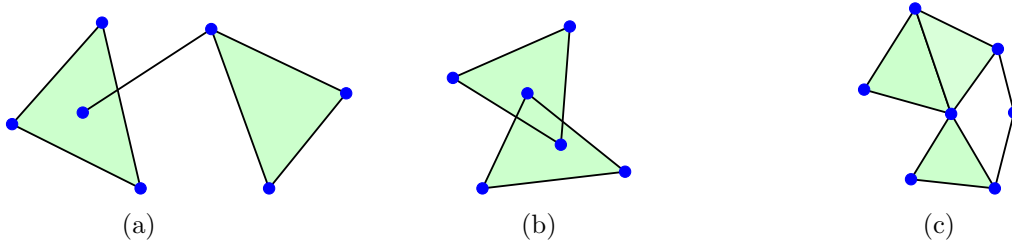


Figura 4.2: En (a) el 1-simplejo no intersecta al 2-simplejo en un vértice. De manera análoga, los 2-simplejos en (b) no se unen mediante vértices o aristas. En (c) se muestra la manera correcta de formar un complejo simplicial.

La **dimensión** de un complejo simplicial K es la mayor dimensión de sus simplejos. En la Figura 4.2(c), se muestra un complejo simplicial de dimensión 2.

4.2. Alfa complejos

Existen distintas maneras de construir complejos simpliciales que aproximan a una nube de datos [57], sin embargo, en la mayoría de los casos, el costo computacional es grande. Esta sección está dedicada a introducir la construcción de complejos simpliciales que aproximan a una nube de datos reduciendo el costo computacional al construirlos, haciendo uso de herramientas adicionales.

Definición 4.2.1. Considérese una nube de datos N , para cada $y_i \in N$, la celda o región de Voronoi de y_i es

$$V(y_i) = \{x \in \mathbb{R}^2 \mid \|x - y_i\| \leq \|x - y_j\|, \text{ para todo } j \neq i\},$$

donde $\|\cdot\|$ denota la norma euclidiana. Esto es, el conjunto de puntos en $\mathbb{R}^2$ que está más cerca de y_i que de cualquier otro punto en N .

La intersección entre las regiones de Voronoi $V(y_i)$ y $V(y_j)$ se denomina **arista de Voronoi**, y se denota por a ; esta corresponde al bisector del segmento que une los puntos y_i y y_j . Cuando tres regiones de Voronoi, $V(y_i)$, $V(y_j)$ y $V(y_k)$, se intersectan en un punto, éste se conoce como **vértice de Voronoi** y se denota por v . Este vértice es el centro del círculo circunscrito que pasa por los puntos y_i , y_j y y_k . La Figura 4.3 (a) ilustra estos elementos.

Definición 4.2.2. El **Diagrama de Voronoi** de N es la colección de celdas, aristas y vértices de Voronoi.

Definición 4.2.3. La **triangulación de Delaunay** sobre N , denotada por $\mathcal{Del}(N)$, es un complejo simplicial sobre N tal que,

- los 0-simplejos son los puntos de N .
- $\langle y_i, y_j \rangle$ es 1-simplejo si y sólo si, existe la arista de voronoi entre estos puntos.
- $\langle y_i, y_j, y_k \rangle$ es 2-simplejo si y sólo si, el vértice de Voronoi v entre estos puntos existe.

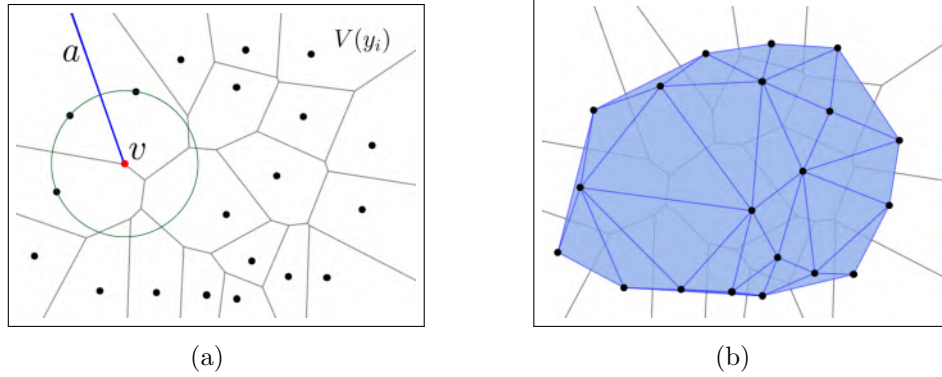


Figura 4.3: Diagrama de Voronoi (a) y triangulación de Delaunay (b) construidos sobre una nube de puntos N . En (a), las etiquetas señalan una arista, una región de Voronoi y un vértice, que corresponde al centro del círculo verde que pasa por tres puntos de N .

La triangulación de Delaunay nos proporciona una aproximación a “la forma” de la nube de datos, pero ésta no termina de ser “la más exacta”. La idea propuesta es extraer las propiedades topológicas sobre una sucesión de triangulaciones basadas en radios de distintos tamaños. Recordar que si $x_0 \in \mathbb{R}^2$ y $r > 0$, la bola abierta de radio r centrada en x_0 es $B_r(x_0) = \{x \in \mathbb{R}^2 \mid \|x_0 - x\| < r\}$.

Definición 4.2.4. Consideremos una nube de puntos N y $\varepsilon > 0$. El **Alfa complejo** de radio ε asociado a N , denotado $\mathcal{A}(N, \varepsilon)$, es el complejo simplicial tal que $\sigma \in \mathcal{A}(N, \varepsilon)$ si, y sólo si, $\sigma \in \mathcal{Del}(N)$ y $\sigma \subset \bigcup_{i=0}^{\ell} B_\varepsilon(y_i)$.

Un Alfa complejo está formado por los simplejos que pertenecen a la triangulación de Delaunay y que “aparecen” al tomar bolas de radio ε sobre N .

Definición 4.2.5. Si $\varepsilon_0 < \varepsilon_1 < \varepsilon_2 < \dots < \varepsilon_f$, entonces a la sucesión de complejos simpliciales

$$N = \mathcal{A}(N, \varepsilon_0) \subset \mathcal{A}(N, \varepsilon_1) \subset \mathcal{A}(N, \varepsilon_2) \subset \dots \subset \mathcal{A}(N, \varepsilon_f),$$

se le llama **filtración de Alfa complejos de N** .

La Figura 4.4 muestra cuatro etapas de una filtración sobre la nube de puntos N de la Figura 4.4(a). Los círculos amarillos que rodean los puntos de N se utilizan como apoyo visual para identificar cuándo se forma un simplejo entre ellos, según el valor del radio. Estos están formados por los puntos negros, los segmentos y los triángulos azules.

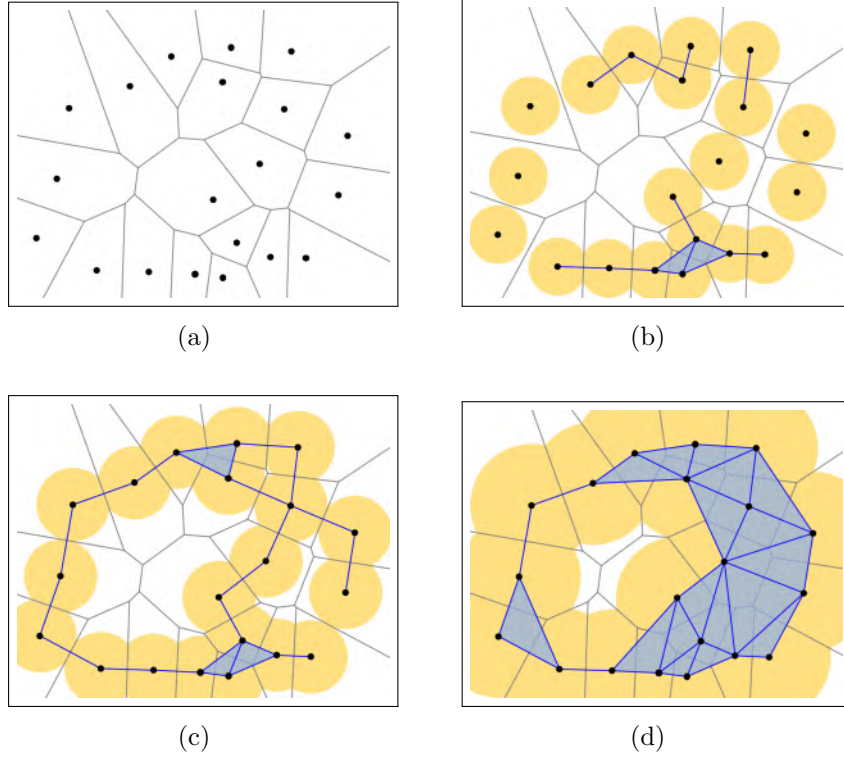


Figura 4.4: Filtración de Alfa complejos sobre la nube N mostrada en (a). Las cuatro imágenes corresponden a etapas de la filtración comenzando con $\varepsilon = 0.0001$.

4.3. Hoyos en complejos simpliciales

Los espacios topológicos se caracterizan por el número de componentes o su número de “hoyos”, éstos están definidos para cualquier dimensión mayor o igual a 0. Debido a los intereses de esta tesis, solamente se trabaja con hoyos de dimensión 0 y 1. Las ilustraciones son realizadas sobre complejos simpliciales.

Recordar que la esfera de dimensión 0 es $\mathbb{S}^0 := \{x \in \mathbb{R} : \|x\| = 1\} = \{-1, 1\}$. Análogamente, el círculo es la esfera de dimensión 1, $\mathbb{S}^1 := \{x \in \mathbb{R}^2 : \|x\| = 1\}$. Se denotará por $\mathbb{S}^j$ a estas esferas.

Definición 4.3.1. Si K es complejo simplicial, se dice que **una copia** de $\mathbb{S}^j$ en K es un subconjunto de K que “es parecido” a $\mathbb{S}^j$.

Una copia de $\mathbb{S}^0$ es una pareja de puntos $\{p, q\} \subset K$. Una copia de $\mathbb{S}^1$ es una curva cerrada o ciclo en K . Por ejemplo, el Alfa complejo de la Figura 4.4(c) contiene dos ciclos, uno de mayor tamaño. Por otro lado, el complejo de la Figura 4.4(d) cuenta con un ciclo. La curva formada por las aristas de cualquier triángulo también es una copia de $\mathbb{S}^1$.

Definición 4.3.2. Supongamos que $A \subset K$ es una copia de $\mathbb{S}^j$. Se dice que A es **contráctil** si es posible deformar continuamente a A dentro de K hasta llevarla a un punto. En caso contrario, se dice que A es no contráctil.

Ejemplo 4.3.3. Si K es el complejo simplicial de la Figura 4.5(a).

- Cualquier copia de $\mathbb{S}^0$ en K es contráctil, ya que es posible llevar un punto a otro mediante una curva o 1-simplejo.

- *Considérese la curva cerrada formada por las caras de cualquier 2-simplejo. Ésta es una copia contráctil de $\mathbb{S}^1$, basta con “aplastar” el triángulo hacia el centro o algún 0-simplejo.*
- *Los dos triángulos vacíos formados por 1-simplejos, son ejemplos de copias no contráctiles de $\mathbb{S}^1$. No existe lugar en K para poder “aplastar” a éstos.*

Ejemplo 4.3.4. *El complejo simplicial de la Figura 4.5(b) cuenta con una copia de $\mathbb{S}^0$ que no es contráctil. Para entenderlo, basta tomar los puntos rojos mostrados, uno en cada componente, éstos forman una copia de $\mathbb{S}^0$ y no es posible llevar un punto a otro dentro de K . Por otro lado, el triángulo y el cuadrilátero vacíos formados por 1-simplejos son copias de $\mathbb{S}^1$ que no son contráctiles.*

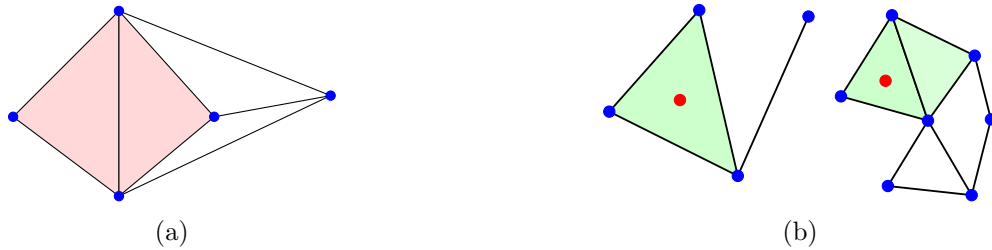


Figura 4.5: Complejos simpliciales que ejemplifican copias de $\mathbb{S}^j$ contráctiles y no contráctiles.

De los ejemplos anteriores, es posible deducir que un hoyo es un objeto que no está relleno. La siguiente definición formaliza esta idea.

Definición 4.3.5. *Si K es complejo simplicial, se dice que un **j -hoyo** en K es una copia de $\mathbb{S}^j$ que no es contráctil.*

El complejo simplicial de la Figura 4.5(a) no tiene 0-hoyos, pero sí tiene dos 1-hoyos. El complejo de la Figura 4.5(b) cuenta con un 0-hoyo y dos 1-hoyos.

En una filtración, la cantidad de 0- y 1-hoyos se modifica conforme el radio aumenta. Por otro lado, el número de 0-hoyos se relaciona con el número de componentes conexas en un complejo simplicial. El siguiente ejemplo ilustra esto.

Ejemplo 4.3.6. *En la filtración de la Figura 4.4(a), se comienza con tantos 0-hoyos como número de puntos menos 1 y no se tienen 1-hoyos. En la Figura 4.4(b) el número de 0-hoyos se reduce a 8 y aún no se generan 1-hoyos. Luego, la Figura 4.4(c) no muestra 0-hoyos, pues el complejo formado es una componente conexa, pero se muestran dos 1-hoyos. Finalmente, la etapa de la filtración mostrada, Figura 4.4(d), contiene un 1-hoyo que será “cubierto” en próximos radios.*

De manera que el número de 0-hoyos se puede calcular tomando el número de componentes conexas menos 1.

4.4. Homología y Homología persistente

La homología es una herramienta de la topología algebraica cuyo objetivo es identificar los hoyos o ciclos de un espacio topológico, proporcionando una manera de identificar la forma del mismo. A través de los llamados grupos de homología, se asocia a cada espacio una serie de

números que capturan información sobre su forma. Sin embargo, esta teoría supone que el espacio está completamente definido y no considera cómo varía su topología bajo perturbaciones.

La homología persistente extiende la homología para estudiar cómo las características topológicas aparecen y desaparecen a lo largo de la filtración de una nube de datos, lo que la hace especialmente útil para analizar datos ruidosos o discretos, como las nubes de puntos. La homología persistente se encarga de identificar cuándo “aparecen” y “desaparecen” los hoyos, así como su persistencia. A continuación, se presentan las definiciones para los conceptos de aparición, persistencia y desaparición de un j -hoyo. Si $\varepsilon > 0$, se denotará por K_ε al Alfa complejo $\mathcal{A}(N, \varepsilon)$.

Definición 4.4.1. *Supongase que $j \geq 0$.*

- Se dice que el j -hoyo h **nació** en K_{ε_0} si h “aparece” en K_{ε_0} y para todo $\varepsilon < \varepsilon_0$, h no “aparece” en K_ε . Se denota $n(h)$ al radio de nacimiento de h .*
- Si $\varepsilon_1 < \varepsilon_2$ y h “aparece” en K_{ε_1} , se dice que h **persiste** en K_{ε_2} , si h “aparece” en K_{ε_2} .*
- Se dice que h **murió** en K_{ε_m} si para todo $\varepsilon > \varepsilon_m$, h no “aparece” en K_ε . Se denota $m(h)$ al radio de muerte de h .*
- La **persistencia** de h en la filtración es el número $m(h) - n(h)$.*

La persistencia de los hoyos a lo largo de una filtración nos permite identificar las propiedades más representativas de una nube de datos. Una manera de presentar esta información es mediante diagramas de persistencia.

Definición 4.4.2. *El **diagrama de persistencia** asociado a la nube de datos N , se construye usando el primer cuadrante del plano, donde los ejes de éste representan los radios de nacimiento y de muerte de un hoyo, por lo que un hoyo h define un punto de la forma $(n(h), m(h))$.*

Algunas propiedades sobre los diagramas de persistencia son las siguientes:

- Todo punto de coordenadas $(n(h), m(h))$ se ubica arriba de la diagonal $x = y$, ya que el radio de muerte es mayor al de nacimiento.
- Los 0-hoyos se encuentran sobre el vértice de muerte, ya que sus coordenadas son de la forma $(0, m(h))$.
- No se indica el punto que representa la componente conexa final, ya que ésta persiste para todo radio. En su lugar, se coloca una línea punteada que representa el final de la filtración.
- Los puntos cercanos a la diagonal son considerados como “ruido”; estos representan pequeños hoyos que se generan debido a la construcción utilizada en la filtración.

Ejemplo 4.4.3. *La Figura 4.6 muestra el diagrama de persistencia de la filtración construida en la Figura 4.4. Los 0-hoyos son ilustrados en color azul, los 1-hoyos en naranja y la banda en color rosa sobre la diagonal identifica aquellos puntos considerados como ruido, aunque no es común observarla en los diagramas. A continuación se proporciona una breve explicación del diagrama:*

- *El radio de muerte de los 0-hoyos no es cubierto por la banda rosa, lo cual indica que los puntos de la nube están dispersos.*

- Existen tres 1-hoyos que no se consideran como ruido, dos alejados de la diagonal y el restante está justo sobre la banda. Estos puntos ilustran que la nube forma tres “círculos”, donde uno nace y muere en un intervalo pequeño, y existe uno con persistencia significativamente mayor que el restante. Para identificarlos, es necesario observar la Figura 4.4(c). En esta etapa, ya se ha generado el ciclo más pequeño y el más grande; el ciclo restante está a punto de formarse, sin embargo no aparece en la filtración.
- El radio de muerte del 0-hoyo más persistente coincide con el radio de nacimiento del 1-hoyo más persistente, esto quiere decir que, el 1-hoyo nace en cuanto todos los puntos se unen. Caso contrario, el 1-hoyo intermedio tarda un poco más en formarse, después de que el complejo de la filtración es conexo.

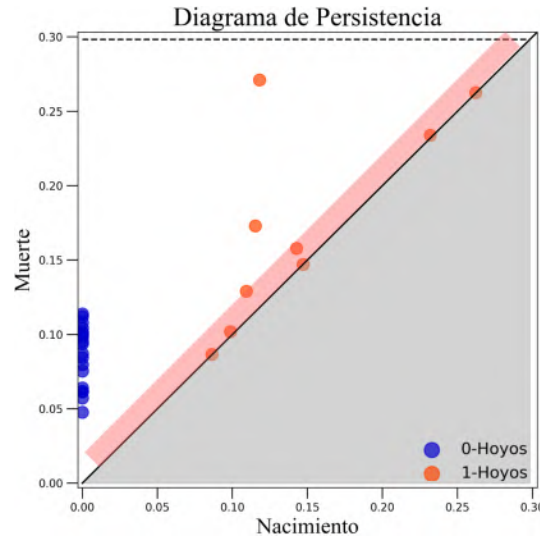


Figura 4.6: Diagrama de persistencia asociado a la filtración de la Figura 4.4. Los 0-hoyos se representan en color azul, mientras que los 1-hoyos en color naranja. La banda color rosa identifica puntos que no son considerados como características de la nube.

Los diagramas de persistencia nos proporcionan una manera eficaz de detectar las propiedades más importantes de la nube de datos, las cuales se ven representadas como puntos “alejados” de la diagonal.

4.5. Distancia Cuello de Botella

Intuitivamente, una manera “natural” de comparar la topología entre dos nubes de datos es mediante sus diagramas de persistencia. Una herramienta que ayuda en este proceso es la distancia cuello de botella, ésta proporciona una medida cuantitativa de la diferencia entre dos diagramas de persistencia. Al medir las diferencias entre diagramas de persistencia, se obtiene una idea de qué tan diferentes son dos nubes de datos. La definición formal de la distancia utilizada es compleja, por lo que primero se proporciona una idea general del procedimiento realizado para cuantificarla.

Si $\mathcal{D}_1$ y $\mathcal{D}_2$ son diagramas de persistencia de dimensión j asociados a las nubes N_1 y N_2 , respectivamente, se procede de la siguiente manera:

- I. **Emparejamiento de diagramas:** Sobreponer $\mathcal{D}_1$ y $\mathcal{D}_2$ de manera que se comparen la mayor cantidad de puntos posible entre éstos, es decir, puntos de $\mathcal{D}_1$ se medirán con puntos “cercanos” en $\mathcal{D}_2$. Los puntos que no cuenten con alguno cercano, se comparan con la diagonal, Figura 4.7.

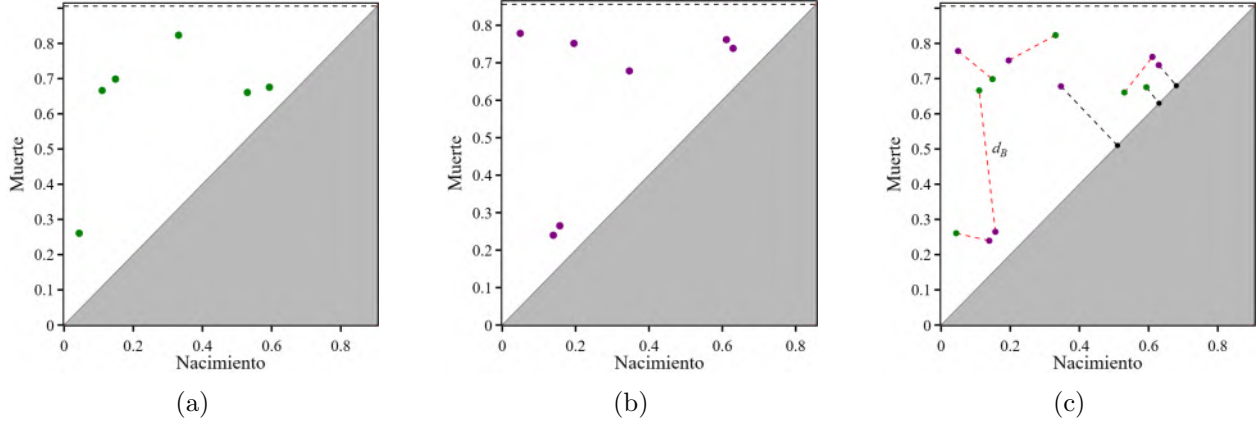


Figura 4.7: Dos diagramas de persistencia asociados a nubes distintas se muestran en (a) y (b), un emparejamiento entre ellos se da en (c).

- II. **Cálculo de distancias:** Para cada par de puntos emparejados, calcular la diferencia en sus valores de nacimiento y muerte, como el máximo de $|n_1 - n_2|$ o $|m_1 - m_2|$, la mínima distancia de todos será almacenada durante el proceso y será denotada por d_B . En el ejemplo de la Figura 4.7 la distancia entre los diagramas es $d_B = 0.4$, la cual es la primera distancia encontrada, pero no es la distancia final.

- III. **Búsqueda del “mejor” emparejamiento:** Se prueban todos los posibles emparejamientos y se elige el que tenga la menor diferencia entre las parejas de puntos. Para encontrar la mejor distancia, hay que comparar todos los emparejamientos posibles entre los dos diagramas.

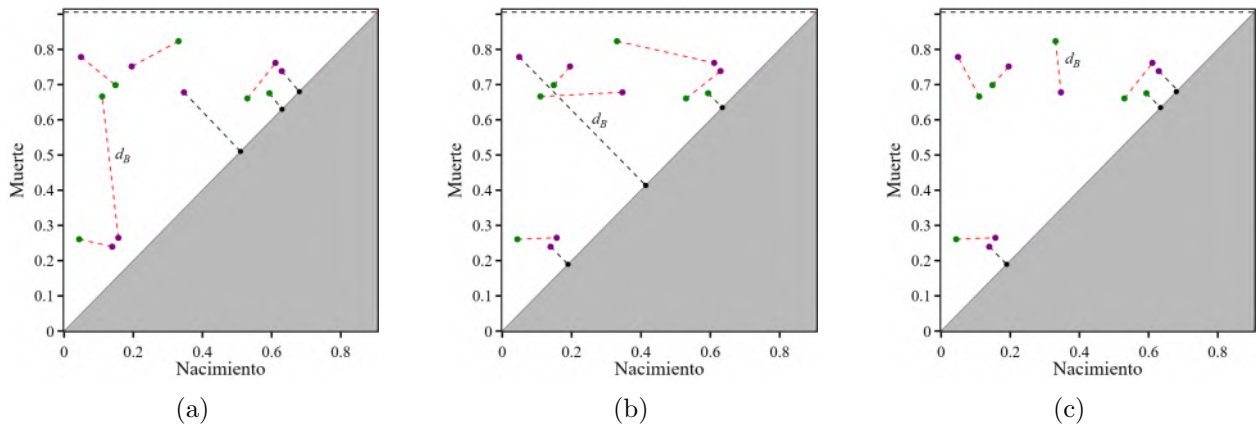


Figura 4.8: Algunos posibles emparejamientos entre los dos diagramas de persistencia.

- IV. **Cálculo de la distancia cuello de botella:** La distancia cuello de botella será la máxima distancia del emparejamiento que minimiza, en nuestro ejemplo, dicho emparejamiento se

presenta en la Figura 4.10(c). Al realizar todas las mediciones, se obtiene que la distancia cuello de botella entre los diagramas es $d_B = 0.1623$.

La distancia cuello de botella entre dos diagramas de persistencia $\mathcal{D}_1$ y $\mathcal{D}_2$ es definida considerando todas las posibles biyecciones entre los dos conjuntos de puntos, tomando puntos de la diagonal en caso de ser necesario. Las distancias entre las biyecciones son calculadas usando la norma infinito, definida como $\|p_1 - p_2\|_\infty = \max\{|n_1 - n_2|, |m_1 - m_2|\}$. Formalmente, para una biyección $\phi : D_1 \rightarrow D_2$, la distancia cuello de botella está dada por, [58],

$$d_B(D_1, D_2) = \inf_{\phi} \sup_{x \in D_1} \|x - \phi(x)\|_\infty. \quad (4.1)$$

La distancia entre dos diagramas de persistencia nos indica si estos son parecidos; distancias grandes implican nubes de datos diferentes, y distancias pequeñas se corresponden con nubes de datos similares. En comparación con los valores en los diagramas, la distancia encontrada en el ejemplo es grande, lo cual indica que las nubes de datos correspondientes son distintas.

A continuación, se ejemplifican algunos cálculos de la distancia cuello de botella. Los cálculos necesarios, como los alfa complejos y los diagramas de persistencia, se realizaron mediante software de código abierto para topología computacional; en particular, se calcularon con la biblioteca *GUDHI* ejecutada en Python [59]; el tutorial de esta librería se puede consultar en [60]. En general, el proceso se realiza para los diagramas de persistencia de todas las dimensiones que se tengan de las nubes de datos, debido a los intereses de este estudio, solamente se calcula la métrica sobre los 1-diagramas de persistencia.

Ejemplo 4.5.1. *Existen comportamientos ideales que suponen nubes de datos sin ruido; sin embargo, esto rara vez ocurre en conjuntos de datos reales. La Figura 4.10 muestra una nube de datos ideal (a) y una nube con ruido (b), ambas formando un círculo. Al calcular la distancia, (c), se observa que, a pesar de las deformaciones introducidas por el ruido, la forma general se conserva. El valor obtenido por la métrica es $d_B = 0.045$, lo que indica que ambas nubes capturan círculos similares. En este caso, la distancia se alcanza al comparar la diferencia entre sus respectivos radios de muerte.*

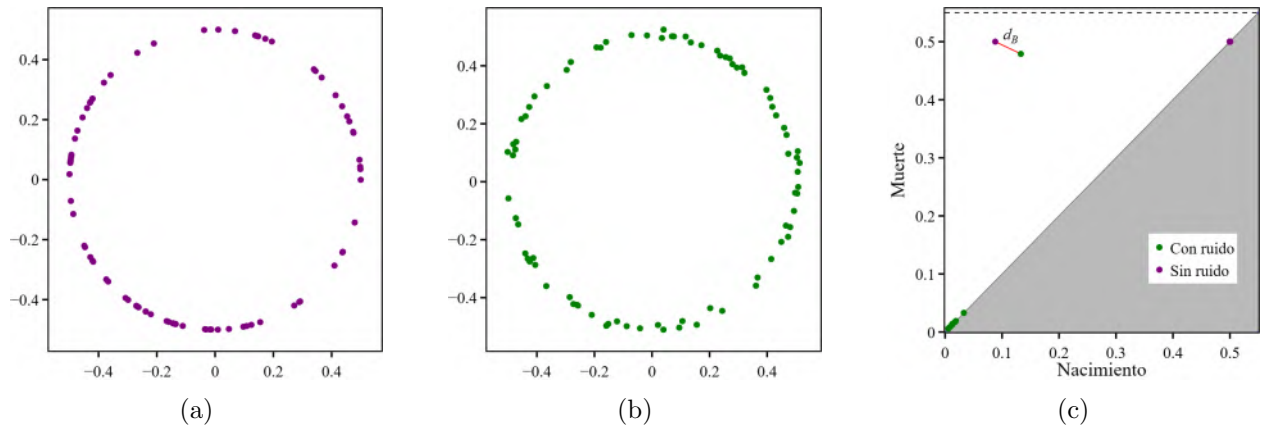


Figura 4.9: Dos nubes de datos alrededor de un círculo, una nube de datos ideal, (a), y una con ruido, (b). El emparejamiento que devuelve la distancia cuello de botella se muestra en (c).

Ejemplo 4.5.2. *Si se presentan dos nubes de puntos con topologías distintas, la métrica reflejará estas diferencias al comparar sus diagramas de persistencia. Al comparar el diagrama de una nube*

de datos alrededor de un círculo, Figura 4.10(a), con el de una nube alrededor de dos círculos, Figura 4.10(b), el emparejamiento resultante mostrado en la Figura 4.10(c), evidencia sus diferencias al emparejar un punto que representa un ciclo de la segunda nube con el punto que representa el ciclo de la primera. La distancia obtenida se corresponde a la diferencia entre los radios de nacimiento de un punto que identifica uno de los círculos en la primera nube de datos y la diagonal, el valor calculado es $d_B = 0.16$

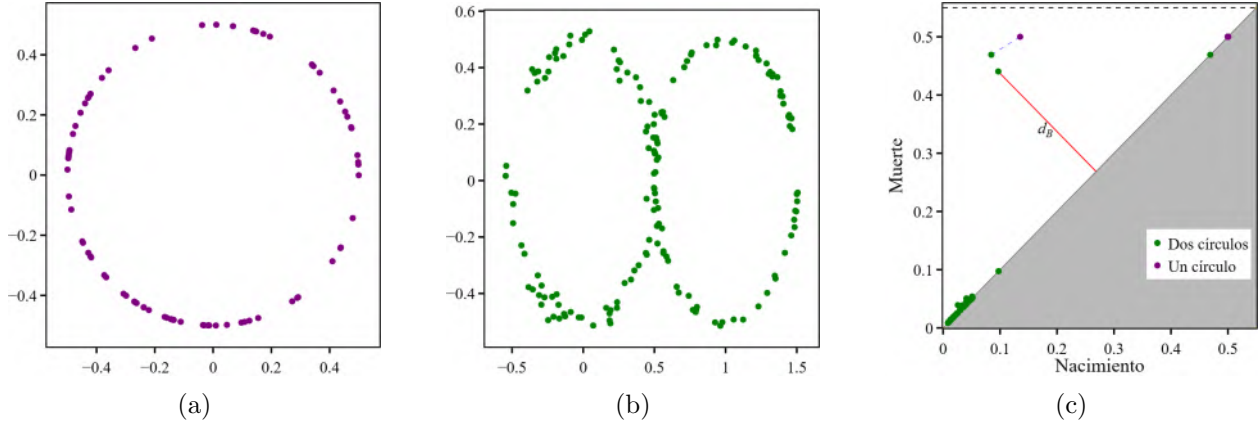


Figura 4.10: Nubes de datos con distintas propiedades topológicas. Una nube de datos alrededor de un círculo, (a), y una que forma dos círculos, (b). El emparejamiento que devuelve la distancia cuello de botella se muestra en (c).

De esta manera, la métrica cuello de botella es capaz no solamente de identificar nubes con ruido que son similares a nubes ideales, también identifica nubes con distintas propiedades. Además, la manera en que realiza el emparejamiento asegura que se tomen las características principales de cada nube de datos.

Ejemplo y conclusiones

Para finalizar el capítulo, se muestra un ejemplo de la aplicación del TDA sobre un conjunto de datos; sin embargo, este solo se muestra con fines ilustrativos y no es utilizado en capítulos posteriores. En este caso, se analizan las nubes de datos asociadas a la ubicación de las plantas de energía fotovoltaica en el país, así como la ubicación de las plantas de energía hidroeléctrica. La información es extraída de [61] y al identificar sus coordenadas de latitud y longitud, es posible representar este conjunto de datos como puntos en $\mathbb{R}^2$; los mapas funcionan de manera ilustrativa.

La Figura 4.11(a) muestra la ubicación de las plantas de energía fotovoltaica instaladas en el país hasta el año 2022. Es posible identificar una mayor concentración en el centro y norte del país con tan solo 4 plantas en el sur. Una vez identificados, se procede a la aplicación del TDA, obteniendo el diagrama de persistencia mostrado en la Figura 4.11(b). Es posible identificar algunas propiedades importantes de éste:

- i. Al analizar los radios de muerte de los 0-hoyos, es posible identificar tres cúmulos de plantas: el primero se corresponde con la zona de Sonora, así como Baja California Norte y Sur, cuyos puntos presentan radios de muerte entre 0.5 y 1; el segundo abarca la región de Chihuahua, Durango, Zacatecas y el centro del país, con radios entre 0 y 0.5; y, por último, la zona de

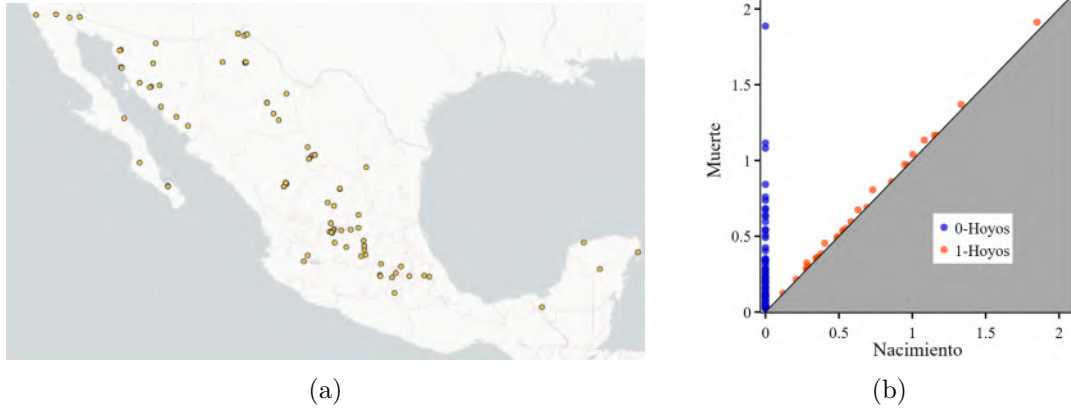


Figura 4.11: Ubicación de plantas fotovoltaicas en México, (a), y el diagrama de persistencia, (b), asociado su representación en $\mathbb{R}^2$.

Yucatán y Quintana Roo, cuyos puntos presentan radios entre 1 y 1.5. El punto cercano a 2 indica la separación entre la zona centro y la zona sur. Esta configuración permite identificar una división territorial en tres zonas: Oeste, Centro y Este.

- ii. En cuanto a los 1-hoyos, estos revelan que las plantas se encuentran distribuidas de manera relativamente “uniforme” dentro de sus agrupaciones y no forman estructuras cíclicas destacadas, ya que no se observan 1-hoyos con persistencias significativas.

La energía hidroeléctrica es una de las energías limpias que ha tenido más peso en el país a lo largo de los años. La Figura 4.12(a) muestra la distribución de las plantas en el territorio mexicano, a diferencia de los sistemas FV, la energía hidroeléctrica no puede ser utilizada en todo el país, por lo que el norte es el lugar donde se tiene menos presencia de ésta, siendo el centro donde se distribuye mayormente. Al igual que en el caso anterior, al usar las latitudes y longitudes es posible extraer el diagrama de persistencia mostrado en la Figura 4.12(b), el cual se explica a continuación:

- i. La distribución de los 0-hoyos en el diagrama refleja la organización de las plantas hidroeléctricas en el territorio, permitiendo identificar una división aproximada en tres zonas. El punto con radio de muerte cercano a 0.6 representa la conexión entre la zona sur y la zona centro. Aquellos con radios cercanos a 0.9 indican la unión entre la zona centro y el cúmulo de puntos ubicado en la región norte. Finalmente, los puntos con radios entre 1 y 1.5 se corresponden con plantas instaladas cerca de la frontera, siendo estas las más alejadas del resto.
- ii. Por otro lado, los 1-hoyos evidencian la formación de una estructura circular generada por ciertas plantas, la cual se localiza en la zona norte. La presencia de este 1-hoyo, con una persistencia relevante, respalda la hipótesis de que en estos estados no es viable la instalación de plantas hidroeléctricas, debido a condiciones climáticas poco favorables.

Al realizar la comparación de las nubes de datos mediante sus diagramas de persistencia, es posible observar la diferencia entre estas. A pesar de que los dos diagramas identifican los tres grupos de plantas, éstos se distribuyen de manera distinta, por lo que el resultado de la distancia en el diagrama de persistencia en dimensión 0 es $d_B = 0.602$, Figura 4.13(a). Por otro lado,

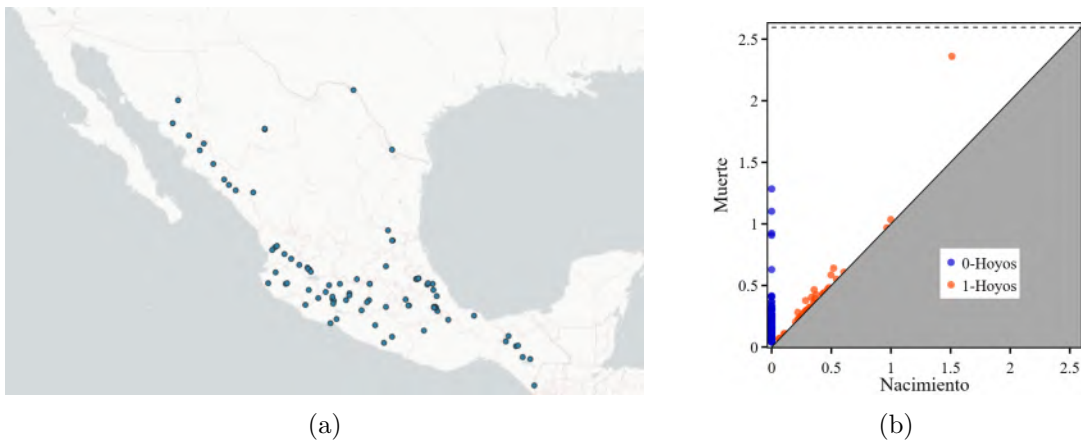


Figura 4.12: Distribución de las plantas hidroeléctricas en México, (a). Diagrama de persistencia asociado a la nube de datos extraída al usar la latitud y la longitud, (b).

la presencia del 1-hoyo en la nube de datos de las plantas hidroeléctricas genera una distancia relativamente grande entre los diagramas de los 1-hoyos, retornando un valor de $d_B = 0.424$, Figura 4.13(b).

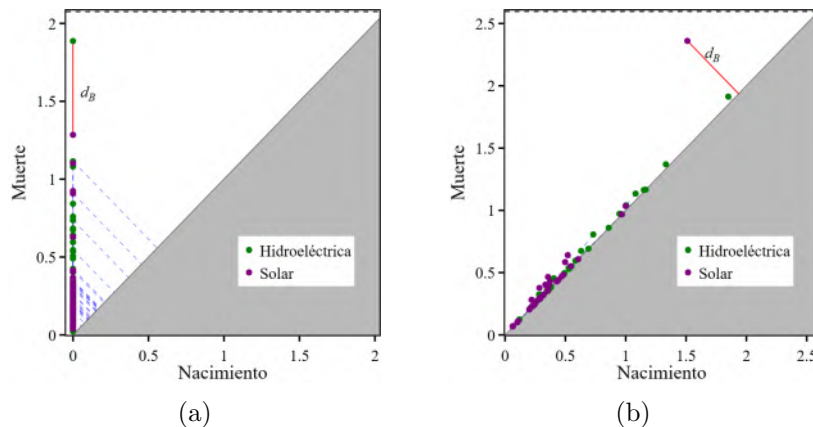


Figura 4.13: Emparejamiento entre los diagramas de persistencia de dimensión 0, (a) y 1, (b), de los diagramas asociados a las nubes de datos de las plantas solares e hidroeléctricas del país.

Comentarios finales

En esta tesis se desarrolla una introducción intuitiva al TDA, mostrando herramientas como los Alfa complejos, que capturan la forma geométrica de los datos. A través del estudio de los hoyos en estos complejos, se pueden identificar características significativas que describen la estructura global de la nube. Para distinguir estas características relevantes del ruido, se emplea la homología persistente, una herramienta que registra la aparición y desaparición de ciclos, reflejada en los diagramas de persistencia. Estos se muestran como una representación compacta y visual de la información topológica. Para comparar y cuantificar la similitud entre diferentes nubes de datos, se utiliza la distancia cuello de botella, una métrica que mide la diferencia entre sus diagramas de persistencia. Esta métrica permite reconocer cuándo dos conjuntos de datos presentan estructuras topológicas similares o distintas, incluso en presencia de ruido o deformaciones. En conjunto, estas

herramientas proporcionan un enfoque sólido para caracterizar y comparar nubes de datos desde una perspectiva topológica, lo que resulta valioso para aplicaciones en análisis de datos, ciencias de la computación y otras áreas interdisciplinarias.

Capítulo 5

Aplicación del TDA al Diagnóstico de Fallas en Sistemas Fotovoltaicos

Los conjuntos de datos generados por los sistemas FV presentan cambios a lo largo del día, debido a diversos factores como las condiciones climáticas. Estas variaciones en la generación eléctrica pueden producir cambios significativos en la topología del conjunto de datos. Los diagramas de persistencia se constituyen como una herramienta prometedora para identificar patrones característicos de funcionamiento y detectar posibles comportamientos anómalos. En este capítulo se muestra cómo la aplicación del TDA permite extraer información relevante de los datos, aportando una perspectiva innovadora para el diagnóstico de fallas en sistemas FV.

5.1. Sistema fotovoltaico analizado

El sistema analizado está instalado en la **División de Estudios de Posgrado de la Facultad de Ingeniería Eléctrica de la Universidad Michoacana de San Nicolás de Hidalgo**, éste cuenta con **tres módulos de doce paneles**, los cuales están conectados a **tres inversores** y un **sistema de almacenamiento de datos**. Debido al gasto de energía en la institución, ésta no se almacena ni se suministra a la red eléctrica pública. La Tabla 5.1 muestra los datos almacenados de cada módulo. Además, para realizar un análisis comparativo entre la potencia generada y la potencia suministrada a la red, se realiza el cálculo de la potencia en corriente continua. Es importante mencionar que los datos son medidos en el inversor, por lo que valores como la temperatura se miden sobre éste y la medición de valores tiene una frecuencia de 5 minutos, generando un archivo nuevo por cada día.

La generación eléctrica en un sistema FV depende de la intensidad de la energía solar, que varía a lo largo del día. Se espera que la potencia siga un patrón determinado, afectado principalmente por la posición del sol; por la mañana, la generación de electricidad comienza a aumentar gradualmente, durante las primeras horas la eficiencia de los paneles solares es relativamente baja debido al ángulo de incidencia de la luz solar; al mediodía, la generación alcanza su pico máximo, ya que el sol está en su punto más alto y el sistema recibe la mayor cantidad de energía solar directa; por la tarde, a medida que el sol se oculta, la generación de electricidad disminuye gradualmente para finalizar sin generación de energía durante la noche.

La propuesta de modelo de detección se basa en datos, por lo que el comportamiento ideal del sistema es tomado del conjunto de datos almacenado. La Figura 5.1 muestra la potencia generada

Tabla 5.1: Datos y unidades almacenados sobre la generación de energía de cada módulo.

Parámetro	Unidad	Parámetro	Unidad
Hora de medición	hh:mm	Corriente en alterna	A
Backup State	-	Corriente en continua	A
CO2 ahorrado	lbs	Modo	-
Error	-	Potencia en alterna	W
Energía Total	kWh	Power On	-
Frecuencia	Hz	Número de serie	-
Tipo de malla	-	Temperatura	°C
Horas encendidas	h	Voltaje en alterna	V
Horas totales	h	Voltaje en continua	V

en tres días distintos, la curva de puntos que se observa en la Figura 5.1(a) es considerada como la ideal del sistema. Casos con pérdidas de energía son registrados en las Figuras 5.1(b) y 5.1(c), éstas pueden ser causadas por factores climáticos.

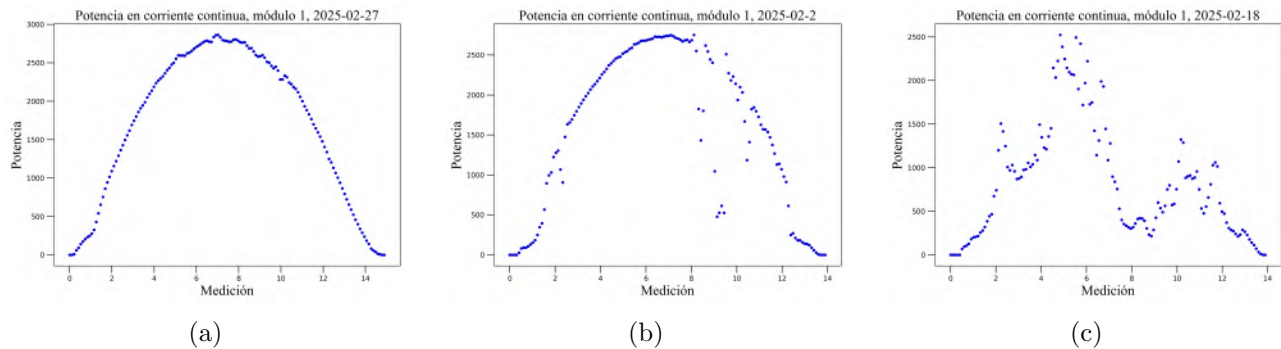


Figura 5.1: Mediciones de la potencia en corriente continua obtenidos a lo largo del día. Se muestran tres días distintos, con distinta pérdida de energía. En (a) se muestra el comportamiento ideal, mientras que en (b) y (c) valores afectados por factores climáticos.

5.2. Extracción de la topología de la potencia

En esta sección se analizan algunas nubes de datos obtenidas. Debido a las características propias de las mediciones a lo largo del día, el enfoque se basa en el análisis de la potencia generada. Se muestran ejemplos sobre la potencia en corriente continua, pero el análisis es el mismo para la potencia en corriente alterna, la cual solamente se modifica por un factor de conversión.

Las nubes de datos obtenidas tienen una frecuencia de muestreo relativamente grande, por lo que es conveniente realizar algunas adecuaciones para mejorar el análisis. Se ilustran los cambios en una nube de datos con pérdida de energía, Figura 5.2(a), para un mejor entendimiento.

El primer paso consiste en añadir puntos sobre el eje x , uno por cada medición obtenida, Figura 5.2(b). Esto genera un “ciclo”, el cual será el analizado con TDA. Para una mejor precisión

en la información, el segundo paso consiste en “rellenar” los datos faltantes debido al tiempo de medición. Para esto, se añaden puntos entre cada dos mediciones registradas, Figura 5.2(c).

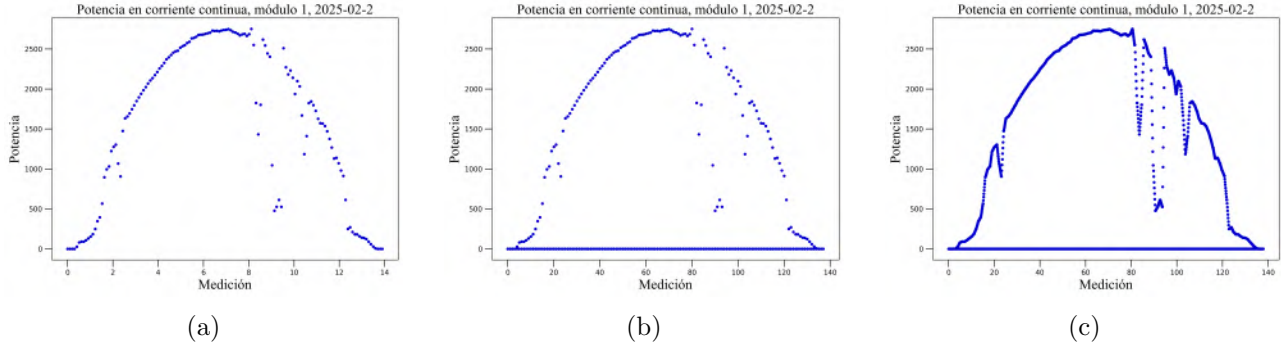


Figura 5.2: Adecuaciones a la nube de datos para la extracción de propiedades con TDA.

Una vez hechas las modificaciones convenientes, se procede a analizar las características extraídas por la técnica. A continuación se explican los diagramas de persistencia asociados a las nubes de datos de la Figura 5.1. Los ejemplos son tomados del módulo 1, las nubes de datos y diagramas correspondientes a los módulos 2 y 3 en los mismos días se comportan de manera similar.

Ejemplo 5.2.1. La nube de datos considerada como la representación de generación ideal en un día, Figura 5.3(a) fue obtenida el día 27 de Febrero del 2025. Aunque algunos puntos se desvían ligeramente de la línea de tendencia éstos no afectan el análisis. El diagrama de persistencia asociado, mostrado en la Figura 5.3(b), contiene los 0- y 1-hoyos de la nube de datos. Debido a la cercanía de los puntos en la nube, los 0-hoyos no se distinguen. Por otro lado, se muestra un 1-hoyo con una persistencia de 1163. Notar que existe ruido que se mantiene sobre la diagonal y éste viene de la construcción utilizada, por lo que se confirma que, los pequeños cambios en la nube de datos no representan alguna modificación en la topología. Se considera este diagrama como el diagrama de persistencia ideal.

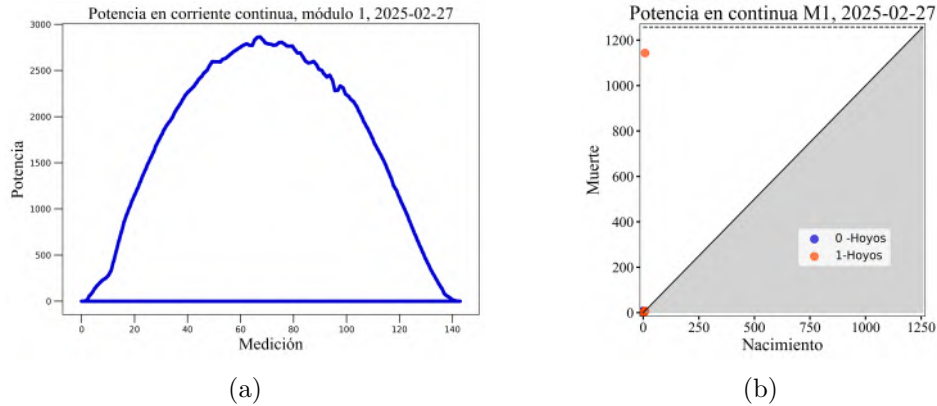


Figura 5.3: Nube de datos ideal de generación del 27/02/2025, (a) y su diagrama de persistencia asociado, (b).

Ejemplo 5.2.2. Una de las nubes de datos con pérdidas de energía se obtuvo el día 2 de Febrero, Figura 5.4(a). Es posible observar que durante la tarde de ese día existieron algunas bajas significativas de energía. El diagrama de persistencia asociado, Figura 5.4(b), ilustra estas pérdidas con dos 1-hoyos alejados de la diagonal. El 1-hoyo con persistencia de valor 783 se corresponde

al ciclo generado por la nube de datos completa. Las dos pérdidas mayores de energía generan el 1-hoyo con persistencia igual a 176. Debido al corto intervalo de tiempo durante el cual se genera la pérdida, no se generan más 1-hoyos de persistencias relevantes. Sin embargo, los 0-hoyos y el ruido generado tienen presencia en este diagrama al alejarse de la diagonal.

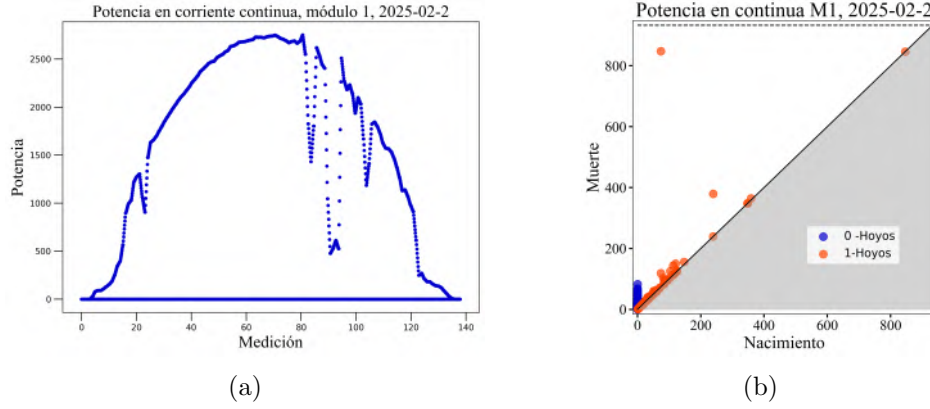


Figura 5.4: Nube de datos de generación del 02/02/2025, (a). El diagrama de persistencia asociado, (b) muestra dos 1-hoyos, de los cuales uno es generado por las pérdidas de energía.

Notar que los 1-hoyos con mayor persistencia en los Ejemplos 5.2.1 y 5.2.2 difieren por un valor cercano a 380, lo cual se debe a las pérdidas de energía. De esta manera, mientras se tengan más pérdidas significativas en la potencia generada, menos persistencia tendrán los 1-hoyos formados por la nube de datos. Por otro lado, la pérdidas “pequeñas” en la potencia, no representan grandes cambios en el diagrama de persistencia.

Ejemplo 5.2.3. El día 11 de Febrero se tuvo una gran deficiencia en la generación de energía a partir de la segunda mitad del día, Figura 5.5(a). El diagrama de persistencia asociado refleja éstas pérdidas, Figura 5.5(b). Al igual que en el ejemplo anterior se muestran dos 1-hoyos. En este caso, sus persistencias son menores al contar con una menor generación de energía. El 1-hoyo principal tiene una persistencia poco mayor a la mitad de una persistencia ideal, siendo 673 en este caso. La persistencia del segundo 1-hoyo también se reduce a 194, debido a que la pérdida de energía se presentó en un mayor intervalo de tiempo. A medida que aumentan los comportamientos anómalos, comienza a tomar aún más presencia el ruido sobre la diagonal.

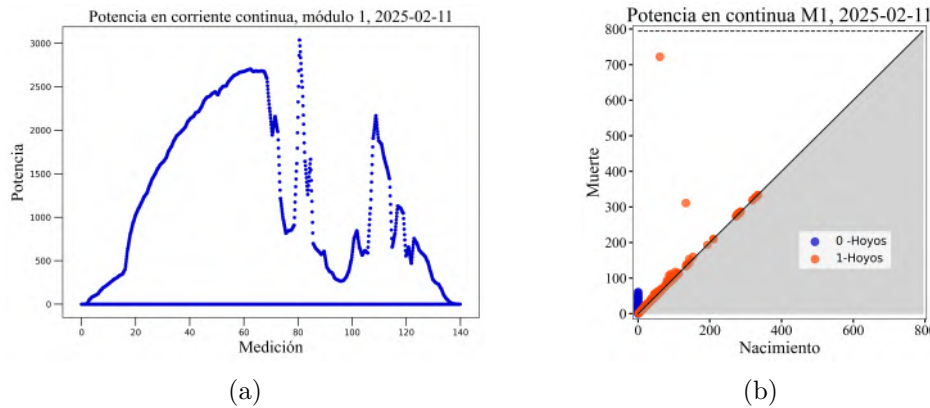


Figura 5.5: Potencia generada en corriente continua del día 11/02/2025, (a). El diagrama asociado, (b), muestra dos 1-hoyos alejados de la diagonal con un cúmulo de ruido notable.

Ejemplo 5.2.4. Existen días en los que los factores climáticos afectan de manera intermitente a lo largo del día. Un ejemplo de esto se presenta en la nube de datos obtenida el día 17 de febrero, mostrada en la Figura 5.6(a). Se observa que, a diferencia de los ejemplos anteriores, la generación de energía presenta varias interrupciones de corta duración, lo que indica la presencia de eventos breves que afectan la potencia del sistema. El diagrama de persistencia mostrado en la Figura 5.6(b), refleja estas interrupciones mediante el aumento de ruido en el diagrama. A pesar de que la curva mantiene su “forma”, las caídas en la potencia representan cambios significativos en el diagrama. Uno de los principales es el aumento en el radio de muerte de tres 0-hoyos, correspondientes a los puntos cerca de la medición 90. Además, el 1-hoyo correspondiente al ciclo principal de generación diaria, presenta una persistencia de 529. Por otro lado, se identifican dos 1-hoyos con persistencias menores generados por las caídas de energía.

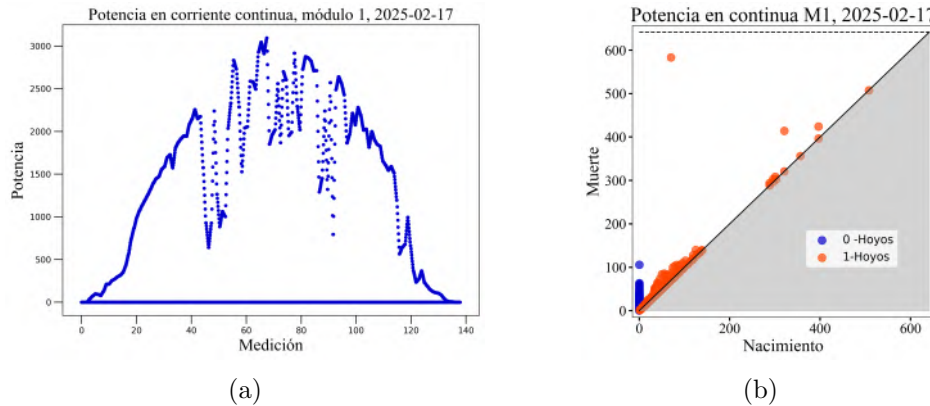


Figura 5.6: Potencia generada en corriente continua el día 17/02/2025, (a). La intermitencia en las caídas de energía genera un aumento de ruido en el diagrama, (b), así como el aumento de persistencia en los 0-hoyos.

El TDA ofrece una manera eficaz para detectar cambios en la potencia generada. A través del análisis de la persistencia de los 0-hoyos, es posible identificar las nubes de datos donde ocurrieron pérdidas significativas. Por otro lado, la cantidad y la persistencia de los 1-hoyos proporcionan información sobre la frecuencia e intensidad de las caídas en la generación eléctrica.

5.3. Comparación entre módulos

La base del método propuesto es la identificación de patrones anómalos en la topología de las nubes de datos, la métrica cuello de botella ayuda a detectarlos mediante diferencias entre los diagramas asociados a cada módulo. A continuación se presentan algunos resultados de la comparación entre diagramas de las nubes obtenidos, recordando que solamente se consideran los 1-hoyos en los diagramas y el ruido no es tomado en cuenta para el análisis.

Ejemplo 5.3.1. Al comparar los diagramas correspondientes al módulo 1, Figura 5.7(a), y al módulo 2, Figura 5.7(b), en un día de generación ideal, la métrica reporta un valor de 1.196. Éste indica una gran similitud entre las nubes de datos, las dimensiones del diagrama hacen que sea insignificante; sin embargo, refleja un cambio que no es perceptible a simple vista. El emparejamiento mostrado en la Figura 5.7(c) muestra la similitud entre éstos y la necesidad de realizar una inspección detallada mostrada en la Figura 5.7(d).

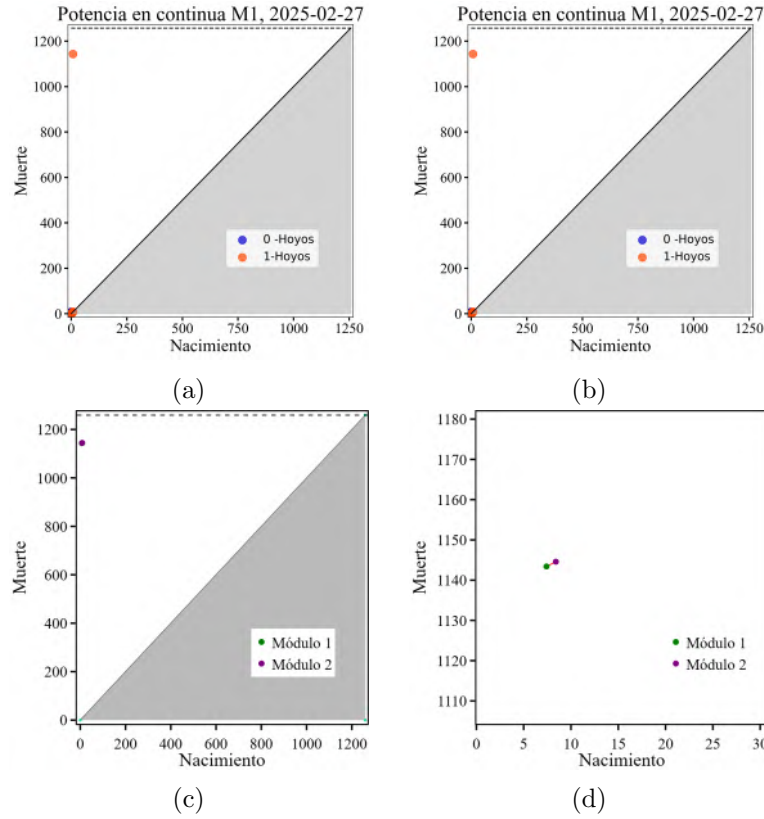


Figura 5.7: Cálculo de la distancia cuello de botella entre los diagramas asociados a la potencia de los módulos 1 y 2 el día 27 de febrero. (a) y (b) presentan los diagramas asociados, mientras que (d) muestra un acercamiento al emparejamiento presentado en (c).

Al realizar el análisis correspondiente entre los diagramas de la potencia en los módulos 1 y 3, se obtiene una distancia de 5.064. Por otro lado, la distancia entre los diagramas de los módulos 2 y 3 tiene un valor de 6.261. En un sistema ideal, se espera que el comportamiento de la generación eléctrica fuese similar en los tres módulos; la métrica cuello de botella refleja este comportamiento, indicando pequeñas variaciones en el día a día. Debido a esto, se propone realizar la comparación de manera individual por módulo. Es decir, el diagrama obtenido cada día se compara con el diagrama ideal correspondiente a ese mismo módulo para obtener un porcentaje en el cambio de la topología.

Ejemplo 5.3.2. En la sección 5.2 se analizó cómo las pérdidas de energía impactan la forma de los diagramas de persistencia. Al comparar el diagrama correspondiente al módulo 1 del día 11 de febrero, Figura 5.8(b), con el diagrama ideal, Figura 5.8(a), se evidencian modificaciones significativas que se reflejan en un valor elevado de la métrica, Figura 5.8(c), alcanzando un valor de 421.183.

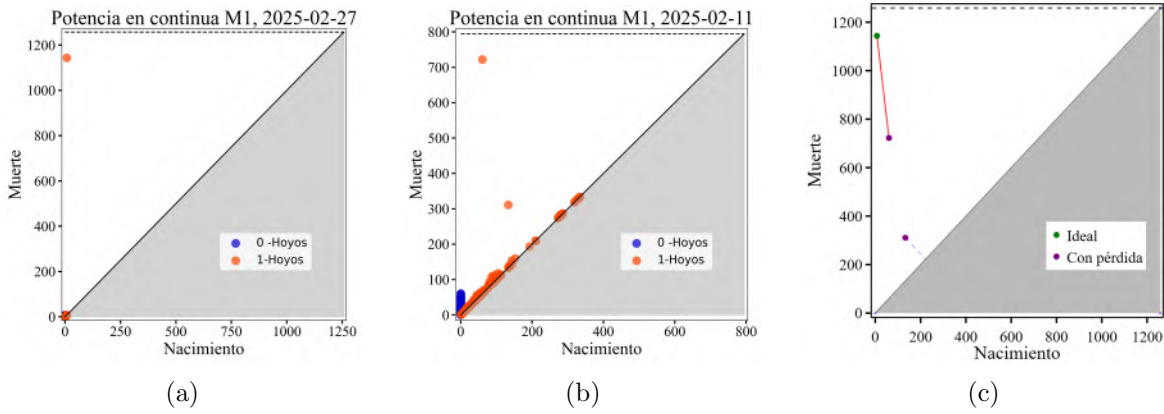


Figura 5.8: Cálculo de la distancia cuello de botella entre los diagramas asociados.

Ejemplo 5.3.3. La comparación entre los diagramas ideales asociados a la potencia en corriente continua, Figura 5.9(a), y potencia en alterna, Figura 5.9(b), cuantifica la distancia ideal entre estas potencias. La Figura 5.9(c), muestra que la diferencia es mínima, alcanzando un valor en la métrica de cuello de botella de 19.856.

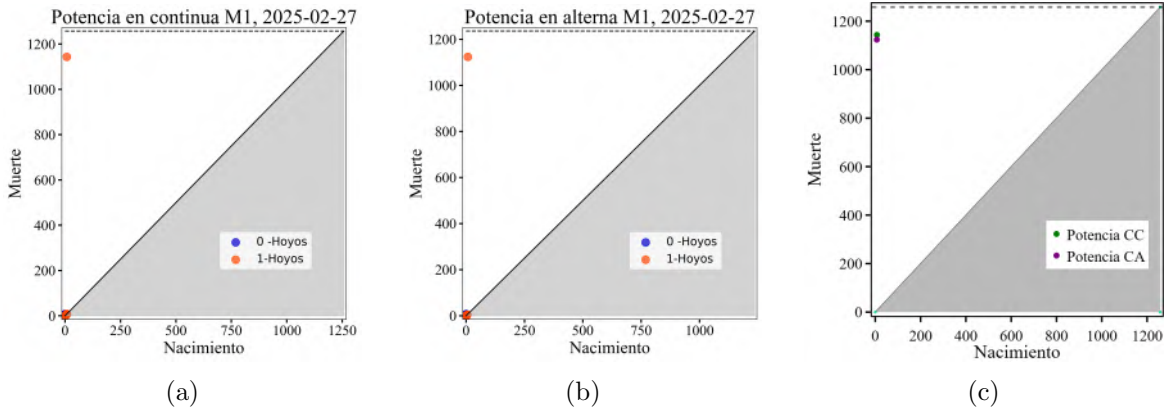


Figura 5.9: Distancia entre los diagramas de potencia en continua, (a) y potencia en alterna, (b). El emparejamiento mostrado en (c) obtiene un valor de 19.856.

La métrica de cuello de botella nos permite establecer umbrales para la detección de fallas, con base en las propiedades topológicas obtenidas.

5.4. Modelo de detección

En esta sección, se desarrolla el modelo de detección y se define la manera en que se reporta una falla de manera temporal o permanente.

El diagrama de flujo de la Figura 5.10 muestra la manera en la que actúa el modelo de detección, una vez obtenido el conjunto de datos.

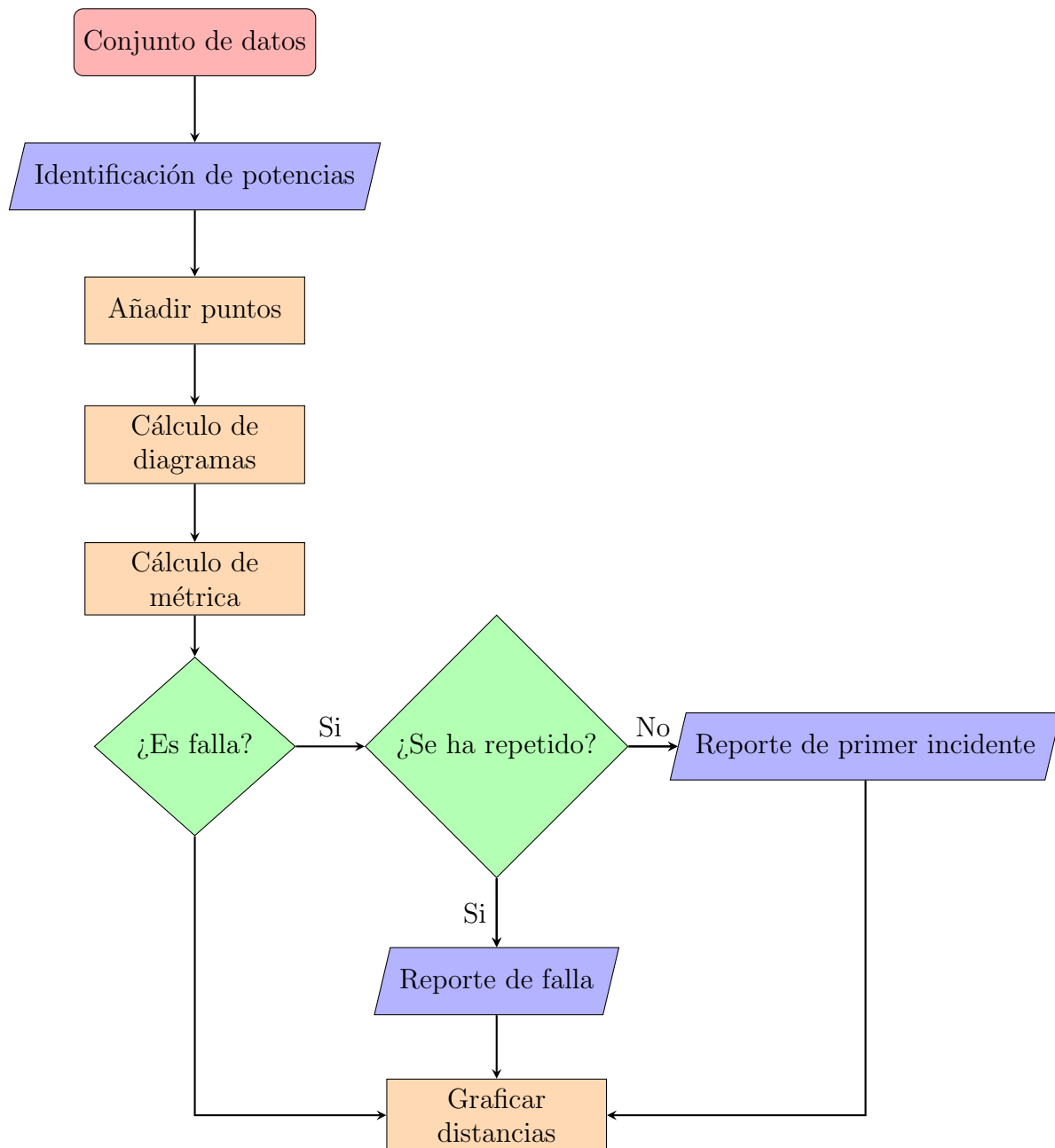


Figura 5.10: Diagrama de flujo del modelo de detección propuesto.

El procedimiento de análisis se lleva a cabo de la siguiente manera: en primer lugar, se realiza la identificación de las señales de potencia en corriente continua y en corriente alterna, lo que permite distinguir adecuadamente la potencia generada por el sistema y la suministrada por el inversor. Posteriormente, se procede al tratamiento de las nubes de datos, en donde los puntos faltantes se interpolan conforme al método ilustrado en la Figura 5.2. Esta etapa garantiza una representación “continua” de los datos.

Posteriormente, se calculan los diagramas de persistencia asociados a cada conjunto de datos. Es importante destacar que estos diagramas no se almacenan de forma permanente, sino que son desechados una vez que su proceso es completado.

El análisis se centra en el cálculo de la distancia cuello de botella. Ésta se evalúa en dos

partes: primero, comparando el diagrama de persistencia de la potencia en cada módulo obtenido diariamente, con el diagrama ideal predefinido; y segundo, comparando cada par de diagramas generados diariamente de los módulos. El propósito de esta doble comparación es cuantificar los cambios en la topología de los datos y detectar desviaciones significativas en el comportamiento de algún módulo respecto a los demás.

La detección de fallas en los inversores se basa en la comparación de potencia recibida por éstos y suministrada a la red. Finalmente, los resultados de estas comparaciones se almacenan y se representan gráficamente. En caso de que se identifiquen desviaciones que superen un umbral previamente establecido, el sistema clasifica la situación como una posible falla y se proporciona una mensaje de prevención al usuario.

5.4.1. Definiciones de fallas

Es importante definir los conceptos que identificarán una falla en el sistema de acuerdo con la técnica propuesta. A continuación, se mencionan las consideraciones para éstas, basándose en la definición de un umbral topológico.

Definición 5.4.1. *Se define una falla de detección cuando el sistema de medición no proporciona los valores de verificación requeridos.*

Para las definiciones posteriores, la distancia entre el diagrama ideal y el diagrama diario del módulo i será denotada por $BD(Pfv_{(i,0),i})$ y $BD(Pac_{(i,0),i})$, respectivamente. Por otro lado, las distancias entre los diagramas diarios generados por pares de módulos se denotarán por $BD(Pfv_{(i,j)})$ y $BD(Pac_{(i,j)})$, respectivamente. Para un mejor entendimiento, se añaden un par de tablas con ejemplos de los posibles valores que se considerarían fallas con un umbral hipotético.

La siguiente definición formaliza la idea de que los módulos pueden presentar menor potencia generada, pero tener el mismo comportamiento a lo largo de un día, lo cual es indicio de pérdidas por factores climáticos.

Definición 5.4.2. *Se dice que el conjunto de módulos presenta fallas atribuibles al clima cuando las distancias $BD(Pfv_{(i,0),i})$ exceden el umbral predefinido, y, al mismo tiempo, las distancias $BD(Pfv_{(i,j)})$ se mantienen dentro de dicho umbral.*

Tabla 5.2: Ejemplos de distancias obtenidas y umbrales ideales.

	Distancia	Umbral
$BD(Pfv_{(1,0),1})$	180	45
$BD(Pfv_{(2,0),2})$	182.5	48
$BD(Pfv_{(3,0),3})$	186	47

Tabla 5.3: Ejemplos de distancias obtenidas y umbrales entre módulos

	Distancia	Umbral
$BD(Pfv_{(1,2)})$	45	52
$BD(Pfv_{(1,3)})$	48	50
$BD(Pfv_{(2,3)})$	47	54

Definición 5.4.3. *El i -ésimo módulo presentará una falla cuando las distancias $BD(Pfv_{(i,j)})$ y $BD(Pfv_{(i,k)})$ estén fuera del umbral, así como la distancia $BD(Pfv_{(i,0),i})$.*

Tabla 5.4: Ejemplos de distancias obtenidas y umbrales ideales.

	Distancia	Umbral
$BD(Pfv_{(1,0),1})$	39	45
$BD(Pfv_{(2,0),2})$	42	48
$BD(Pfv_{(3,0),3})$	186	47

Tabla 5.5: Ejemplos de distancias obtenidas y umbrales entre módulos

	Distancia	Umbral
$BD(Pfv_{(1,2)})$	45	52
$BD(Pfv_{(1,3)})$	120	50
$BD(Pfv_{(2,3)})$	128	54

Definición 5.4.4. *Una falla en el i -ésimo inversor se presentará cuando no exista falla en el i -ésimo módulo, y la distancia $BD(Pfv_i, Pac_i)$ supere el umbral ideal.*

El comportamiento del sistema puede reflejar una falla durante un día y desaparecer al siguiente. Mientras que una pérdida de energía por clima será catalogada como una falla temporal, es necesario determinar si una falla frecuente es permanente con base en su historial.

Definición 5.4.5. *Se define una falla frecuente (permanente) en el i -ésimo módulo o inversor cuando la falla temporal se presente de manera continua durante 7 días en un mes.*

Comentarios finales

En este capítulo se ha demostrado que el TDA es una herramienta eficaz para detectar comportamientos anómalos en sistemas fotovoltaicos a partir de datos reales de generación eléctrica. Al aplicar técnicas topológico-geométricas sobre las nubes de puntos construidas a partir de la potencia generada a lo largo del día, es posible identificar diferencias significativas entre comportamientos ideales y días con pérdidas de energía. La presencia, cantidad y persistencia de los 1-hoyos en los diagramas permiten distinguir entre niveles de afectación, mostrando que una menor generación de energía se refleja directamente en una menor persistencia topológica. Así mismo, el aumento del ruido y su alejamiento de la diagonal en los diagramas proporciona una señal cuantificable del deterioro en el desempeño del sistema.

Capítulo 6

Resultados

Este capítulo presenta los resultados obtenidos tras la implementación del sistema de detección de fallas en el sistema FV. El análisis se realizó sobre conjuntos de datos correspondientes a diversas condiciones operativas del sistema, abarcando los meses de septiembre y octubre de 2024, así como febrero, abril y mayo de 2025. Adicionalmente, se consideró un conjunto de datos correspondiente a los meses de agosto y septiembre de 2021.

El capítulo se divide en dos secciones principales. En la primera, se detallan los resultados obtenidos durante el análisis del mes de febrero. En la segunda, se utilizan los mismos datos para simular fallas en el sistema, con el fin de evaluar el comportamiento del modelo y su capacidad de detección. Los resultados correspondientes a los demás meses se encuentran en el Apéndice B. Para cada conjunto de datos, se presentan tres figuras representativas y dos tablas que permiten un análisis integral. A continuación, se describe el contenido y propósito de cada elemento:

Figura 1: Porcentaje de coincidencia entre la topología generada día con día y la topología ideal. Un valor cercano al 0% indica una pérdida total de información topológica, mientras que porcentajes bajos reflejan una reducción en la potencia del sistema, posiblemente causada por condiciones climáticas adversas o la presencia de anomalías operativas.

Figura 2: Presenta las distancias calculadas entre cada par de módulos por día, junto con el umbral de referencia. Los días en los que se presenta una pérdida de conexión son codificados con un valor asignado de 100. Valores distintos corresponden a comportamientos anómalos del sistema.

Figura 3: Ilustra la distancia de la comparación de las topologías de potencia en corriente continua y en corriente alterna. En caso de existir una falla por conexión se representa mediante un valor de 100, valores diferentes que superen el umbral son asociados a problemas en el componente.

Tablas: Se incluyen dos tablas, una correspondiente a los módulos y otra a los inversores. Ambas utilizan una codificación de tres niveles para representar el estado del sistema: el valor **1**, representa el componente en funcionamiento normal; el valor **0.5**, una pérdida atribuible a factores climáticos o al proceso de conversión; y, un **0**, indica una falla detectada en el sistema. En el caso de los inversores, se utiliza el valor **-1** para representar la falla por falta de datos desde el módulo.

A continuación, se presentan dos tablas: la Tabla 6.1, muestra las distancias cuello de botella entre los diagramas ideales y los correspondientes a los módulos, y la Tabla 6.2, que presenta las distancias entre los diagramas ideales de los módulos y los diagramas de los inversores. Estos valores determinan los umbrales que utilizados en el análisis.

Tabla 6.1: Umbrales utilizados para la comparación entre los diagramas asociados a los módulos FV.

	M1 vs M2	M1 vs M3	M2 vs M3
Umbral	41.581	41.408	40.722

Tabla 6.2: Umbrales para la comparación entre los diagramas asociados a los módulos FV y los diagramas de los inversores diariamente.

	M1 vs I1	M2 vs I2	M3 vs I3
Umbral	58.994	58.939	58.214

6.1. Análisis de datos reales

El análisis de los porcentajes de potencia generada en corriente continua por los tres módulos durante el mes de Febrero revela patrones de rendimiento dinámicos, generados por las variaciones estacionales, Figura 6.1. Durante la primera mitad del mes, la generación se mantiene consistentemente por encima del 80 %, con picos cercanos al 95 %. Sin embargo, la segunda quincena de Febrero presenta un descenso notable y sincronizado en el rendimiento de todos los módulos, evidenciando una posible disminución en la irradiancia solar o condiciones climáticas adversas a mediados de mes.

Existen algunas características a notar sobre la gráfica; la primera es el inicio de mes del módulo 2, éste cuenta con la menor coincidencia obteniendo un valor aproximado del 45 %, pero corrige su comportamiento para mantenerse cercano a los otros dos durante todo el resto de mes. Por otro lado, el día 24 se observa que el porcentaje excedió el 100 % de coincidencia, lo que indica una mayor generación. Este fenómeno sucede debido a la presencia de menor temperatura en algunos puntos del mediodía. lo cual permite que el sistema se aproveche de manera más óptima.



Figura 6.1: Porcentaje de la topología extraída de la potencia generada día con día, contrastado con la topología ideal a lo largo del mes de Febrero de 2025.

Las diferencias cuantificadas entre las potencias generadas por las parejas de módulos son clave para determinar si algún módulo exhibió un comportamiento atípico durante el mes. En este sentido, la Figura 6.2 ilustra las variaciones observadas en el Módulo 2 al inicio de Febrero. Específicamente, las distancias calculadas entre el Módulo 1 y Módulo 2, y entre el Módulo 2 y Módulo 3,

exceden el umbral preestablecido de manera significativa. No obstante, lejos de atribuir esto directamente a una falla, es fundamental considerar la escala de los valores de los diagramas asociados, los cuales se aproximan a 1150, Figura 5.3, una distancia de 107 corresponde aproximadamente a un 10 % de diferencia en el porcentaje de generación, el cual se visualiza en la Figura 6.1. El día 3, se presentó una anomalía más que no es observable en la figura anterior, en este caso las tres distancias entre los módulos superan el umbral definido, indicando algún comportamiento a revisar. Para el resto del mes, las distancias entre módulos permanecen dentro del umbral, lo que indica un patrón de operación similar entre ellos.

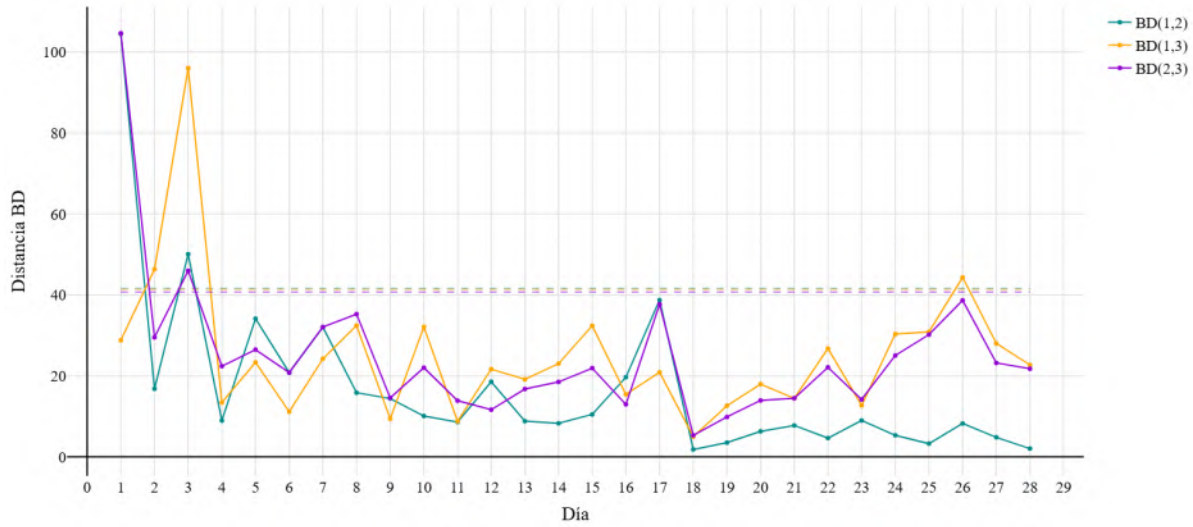


Figura 6.2: Distancias entre los diagramas asociados a los módulos, PCC.

La comparación entre los diagramas de PCC y PCA debe mantenerse dentro de un umbral preestablecido para asegurar que los inversores realicen la conversión de potencia de forma eficiente. La Figura 6.3 muestra el comportamiento de estas distancias a lo largo del mes. Se identifican dos posibles anomalías: el día 17, el inversor 3 presentó una distancia significativamente mayor en comparación con los otros dos, aunque esta discrepancia se estabilizó al día siguiente, lo cual sugiere una anomalía temporal. Por otro lado, el día 24 se observó un pico en las distancias entre los tres diagramas de PCC y PCA, superando el umbral establecido; este comportamiento está relacionado con un incremento en la potencia generada.

Las Tablas 6.3 y 6.4 contienen los valores asignados a los tipos de comportamientos a lo largo del mes. La notación de M_1 , M_2 y M_3 corresponde a los módulos, mientras que I_1 , I_2 e I_3 corresponde a los inversores. La potencia generada por los módulos mantiene un rendimiento, en general, por debajo del umbral a lo largo del mes. Esto es reflejado al solamente contar con 5 días de generación ideal. La anomalía observada en la Figura 6.2 el día 3, se representa con valores de 0 en la fila correspondiente de la Tabla 6.3. La Tabla 6.4 presenta un valor de 0 solamente para el primer módulo, confirmando que éste fue el de mayor fluctuación. A pesar de que los módulos presentan pérdidas en la potencia por factores climáticos, la conversión de potencia en los inversores se realiza de manera óptima, por lo que solamente se reporta un valor de 0 en la segunda tabla, indicando que la falla se presentó de manera temporal.

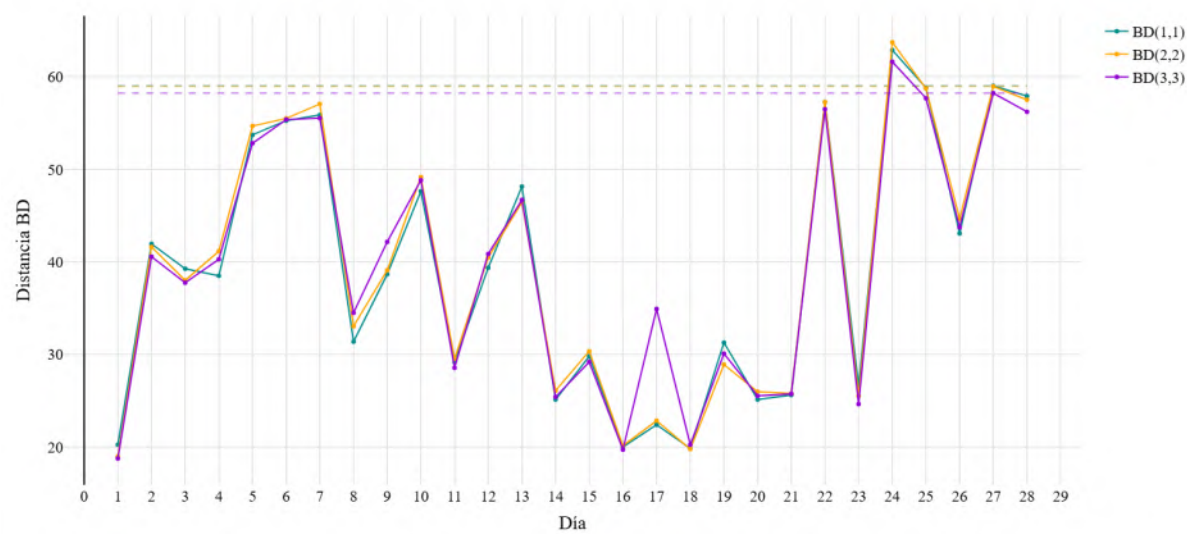


Figura 6.3: Distancias entre los diagramas asociados a la PCC y la PCA durante el mes de Febrero.

Tabla 6.3: Fallas del sistema en módulos, Febrero de 2025

D	M_1	M_2	M_3
1	0.5	0	0.5
2	0.5	0.5	0.5
3	0	0	0
4	0.5	0.5	0.5
5	0.5	0.5	0.5
6	1	1	1
7	0.5	0.5	0.5
8	0.5	0.5	0.5
9	0.5	0.5	0.5
10	0.5	0.5	0.5
11	0.5	0.5	0.5
12	0.5	0.5	0.5
13	0.5	0.5	0.5
14	0.5	0.5	0.5
15	0.5	0.5	0.5
16	0.5	0.5	0.5
17	0.5	0.5	0.5
18	0.5	0.5	0.5
19	0.5	0.5	0.5
20	0.5	0.5	0.5
21	0.5	0.5	0.5
22	0.5	0.5	0.5
23	0.5	0.5	0.5
24	1	1	1
25	1	1	1
26	0.5	0.5	0.5
27	1	1	1
28	1	1	1

Tabla 6.4: Fallas del sistema en inversores, Febrero de 2025

D	I_1	I_2	I_3
1	1	1	1
2	1	1	1
3	0	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	1	1	1
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1

Es evidente que, si bien el sistema muestra una alta eficiencia en condiciones óptimas, también es susceptible a disminuciones sincronizadas en el rendimiento. Las distancias entre módulos y la comparación PCC y PCA son métricas fundamentales para la detección de anomalías. Si bien se identificaron desviaciones, como el comportamiento inicial del Módulo 2 y los cambios en los

inversores, estas no siempre se traducen en fallas críticas. Es crucial considerar las escalas para evitar interpretaciones erróneas y entender que pequeñas variaciones son normales dentro de un sistema complejo. La conversión óptima de potencia en los inversores, a pesar de las pérdidas en los módulos, resalta la robustez de esta etapa del proceso.

6.2. Simulación de Fallas

Con el objetivo primordial de evaluar la capacidad del modelo propuesto para detectar fallas, se generaron simulaciones utilizando el conjunto de datos almacenado. Estas simulaciones se concentraron en explorar dos escenarios fundamentales que representan condiciones críticas para un sistema FV. En primer lugar, se analizó un sistema operando sin diodos de bypass. Bajo esta condición, la simulación se centró en el impacto de la pérdida total de un panel, así como la afectación de un módulo por sombreado parcial a lo largo del día o desgaste. El segundo escenario abordado fue el de un sistema que sí incorpora diodos de bypass. El análisis se expandió para incluir no solo el comportamiento ante la falla de los módulos, sino también la detección de fallas en los inversores, permitiendo así una evaluación más comprensiva del modelo bajo distintas condiciones operativas y de fallo.

Fallas en sistema sin bypass

Se simularon condiciones como sombreado parcial y desconexión de una cadena o fallos internos en los paneles. En los casos más extremos, la potencia de una cadena se redujo a cero, representando una falla total debido al daño de un panel o desconexión. Para fallas parciales, se aplicaron reducciones aleatorias entre el 30 % y el 70 % de la potencia nominal, simulando situaciones de degradación o sombreado parcial.

Falla por sombreado parcial

Se consideraron fallas por sombreado en los días 7 y 26 en el módulo 1. El valor asignado fue aleatorio para los dos casos. La pérdida de potencia es reflejada en la Figura 6.4, las diferencias entre los porcentajes de generación difieren por un 50 y 22 %, respectivamente. La presencia de la falla en el módulo 1 es reflejada en la Figura 6.5, las distancias entre el módulo 1 y 2, así como entre el módulo 1 y 3, superan significativamente el umbral, mientras que la distancia entre los módulos 2 y 3 se mantiene por debajo de éste.

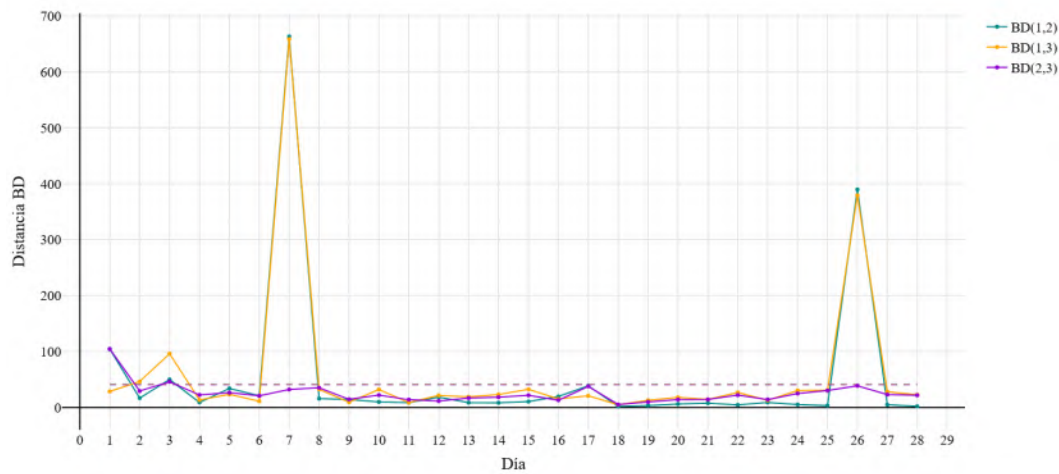


Figura 6.5: Distancias entre los diagramas asociados a la PCC y la PCA durante el mes de Febrero al simular pérdida por clima los días 7 y 21.

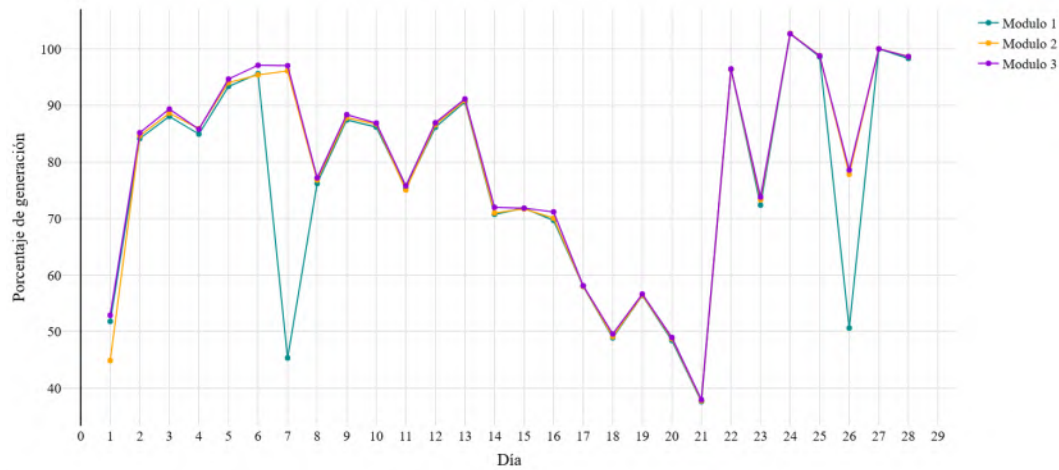


Figura 6.4: Cambio en el porcentaje de la topología al simular pérdida por clima en los días 7 y 21.

Se consideraron fallas asociadas al inversor o a la red eléctrica, que afectan la conversión de la energía generada. Estas se representaron mediante una disminución en la potencia alterna, simulando fallos de conversión, desconexiones o errores en el inversor. La simulación de fallas en los inversores fue realizada de manera similar, se aplicaron valores aleatorios en la potencia convertida para los días 7 y 26, lo cual provoca cambios en las distancias obtenidas, Figura 6.6. El modelo detectó éstas fallas sin mayor complicación. Sin embargo, ya que las fallas eran provenientes de la potencia en corriente continua con fallas, éstas fueron señaladas con la etiqueta de un -1 en la Tabla 6.6.

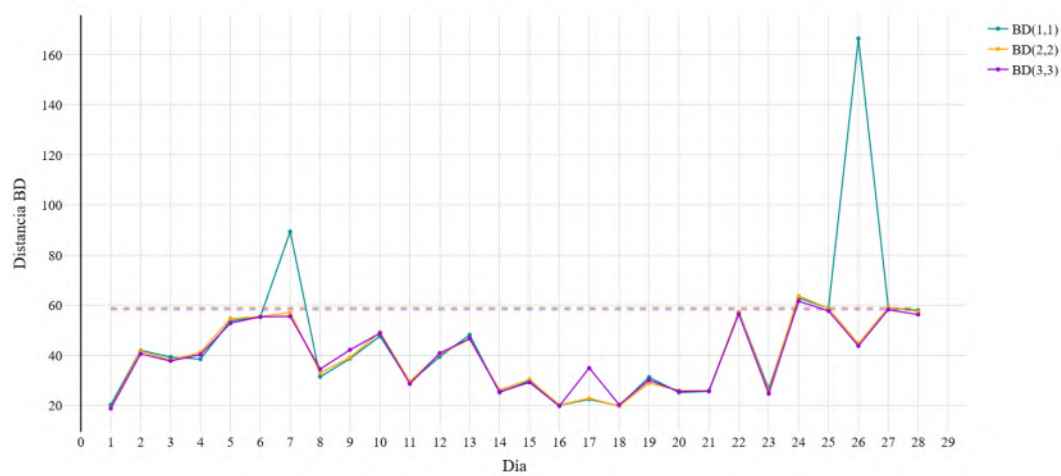


Figura 6.6: Distancias entre los diagramas asociados a la PCC y la PCA durante el mes de Febrero.

Tabla 6.5: Fallas del módulo 1 al reducir su rendimiento de manera aleatoria.

D	M ₁	M ₂	M ₃
1	0.5	0	0.5
2	0.5	0.5	0.5
3	0	0	0
4	0.5	0.5	0.5
5	0.5	0.5	0.5
6	1	1	1
7	0	0.5	0.5
8	0.5	0.5	0.5
9	0.5	0.5	0.5
10	0.5	0.5	0.5
11	0.5	0.5	0.5
12	0.5	0.5	0.5
13	0.5	0.5	0.5
14	0.5	0.5	0.5

D	M ₁	M ₂	M ₃
15	0.5	0.5	0.5
16	0.5	0.5	0.5
17	0.5	0.5	0.5
18	0.5	0.5	0.5
19	0.5	0.5	0.5
20	0.5	0.5	0.5
21	0.5	0.5	0.5
22	0.5	0.5	0.5
23	0.5	0.5	0.5
24	1	1	1
25	1	1	1
26	0	0.5	0.5
27	1	1	1
28	1	1	1

Tabla 6.6: Fallas en inversores al reducir el rendimiento del módulo 1 de manera aleatoria.

D	I ₀	I ₁	I ₂
1	1	1	1
2	1	1	1
3	-1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	-1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1

D	I ₀	I ₁	I ₂
15	1	1	1
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	0	1	1
25	1	1	1
26	-1	1	1
27	1	1	1
28	1	1	1

Falla por daño o desconexión

La simulación de un módulo defectuoso o desconectado se realizó en los días 6,7,8,13 y 14. Para realizar un análisis se propone que el módulo dejó de reportar información o se reporta con valor

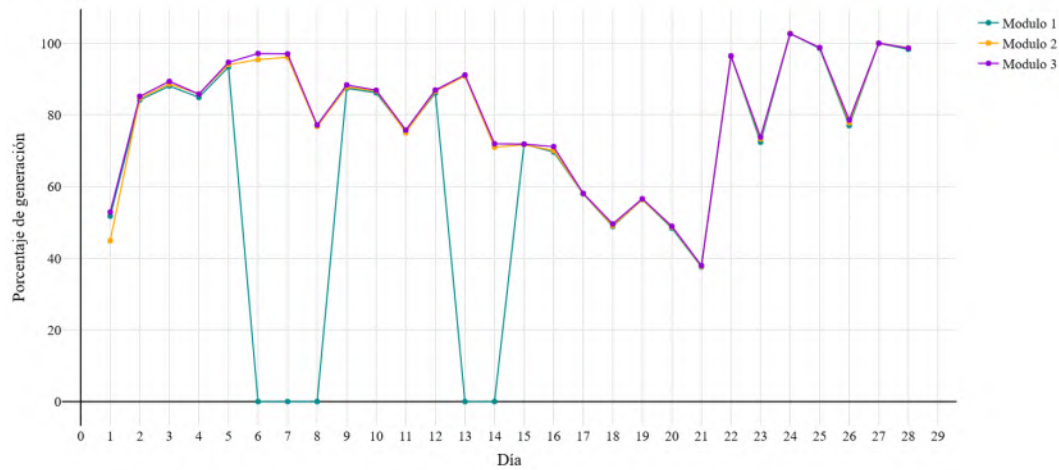


Figura 6.7: Cambios en el porcentaje de generación suponiendo la pérdida o des conexión de un módulo.

de 0 en todo momento. El resultado de este comportamiento es un porcentaje de generación del 0%, Figura 6.7. Al igual que en la simulación de falla anterior, las distancias entre los módulos que dependen del módulo afectado crecen en gran medida, Figura 6.8. Por otro lado, la distancia entre los diagramas de PCC y PCA reportan una distancia de 0, Figura 6.9, este valor es asignado para evitar problemas de procesamiento entre los datos, aunque no existe una distancia debido a la falta de información.

Las anomalías observadas en las Figuras mencionadas son reportadas en las Tablas 6.7 y 6.8. Estas secciones nos mostraron gráficamente la manera en la que una falla del sistema afecta a las topologías. En secciones posteriores el análisis se basa en la revisión de las tablas.

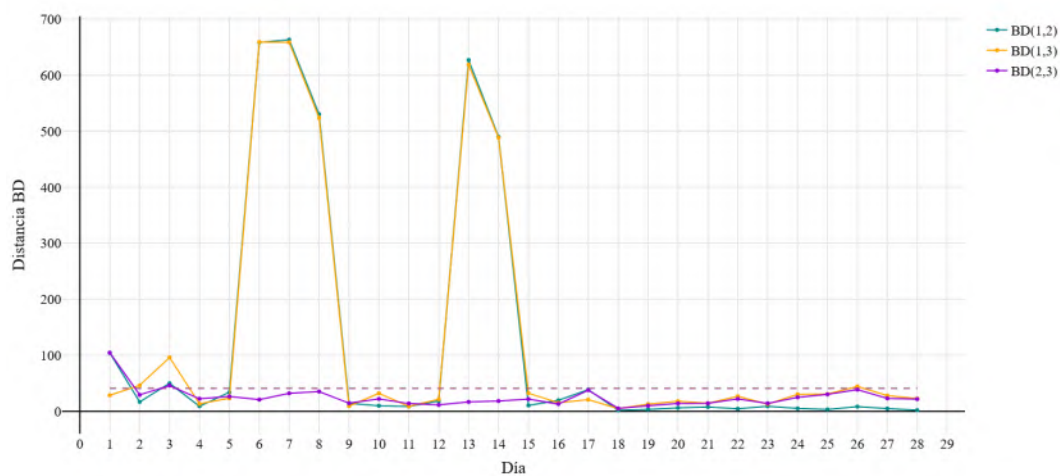


Figura 6.8: Distancias entre los diagramas asociados a la PCC y la PCA simulando la pérdida de un módulo, los días 6,7,8,13 y 14.

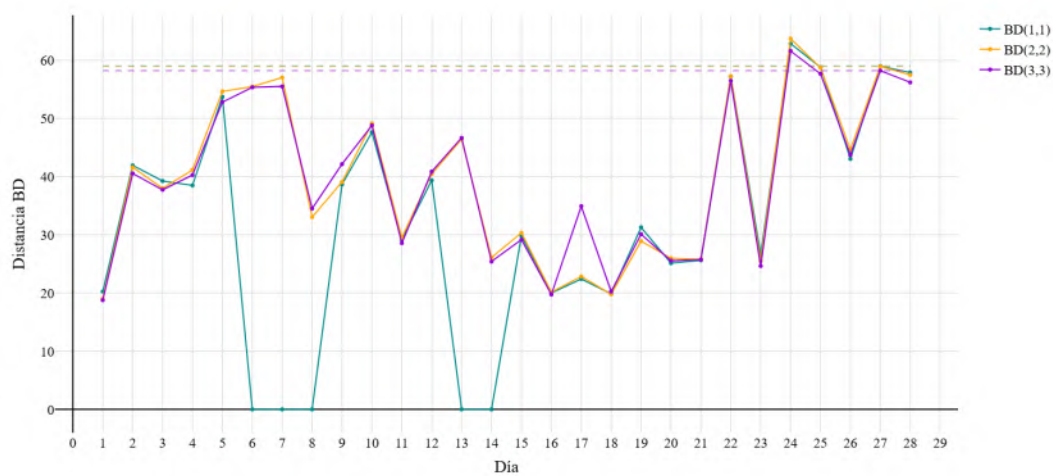


Figura 6.9: Distancias entre los diagramas asociados a la PCC y la PCA al simular la pérdida o des conexión de un módulo.

Tabla 6.7: Fallas por simulación de pérdida en el módulo 1.

D	M_1	M_2	M_3	D	M_1	M_2	M_3
1	0.5	0	0.5	15	0.5	0.5	0.5
2	0.5	0.5	0.5	16	0.5	0.5	0.5
3	0	0	0	17	0.5	0.5	0.5
4	0.5	0.5	0.5	18	0.5	0.5	0.5
5	0.5	0.5	0.5	19	0.5	0.5	0.5
6	0	1	1	20	0.5	0.5	0.5
7	0	0.5	0.5	21	0.5	0.5	0.5
8	0	0.5	0.5	22	0.5	0.5	0.5
9	0.5	0.5	0.5	23	0.5	0.5	0.5
10	0.5	0.5	0.5	24	1	1	1
11	0.5	0.5	0.5	25	1	1	1
12	0.5	0.5	0.5	26	0.5	0.5	0.5
13	0	0.5	0.5	27	1	1	1
14	0	0.5	0.5	28	1	1	1

Tabla 6.8: Falla en inversores por simulación de pérdida en un módulo.

D	I_0	I_1	I_2	D	I_0	I_1	I_2
1	1	1	1	15	1	1	1
2	1	1	1	16	1	1	1
3	-1	1	1	17	1	1	1
4	1	1	1	18	1	1	1
5	1	1	1	19	1	1	1
6	-1	1	1	20	1	1	1
7	-1	1	1	21	1	1	1
8	-1	1	1	22	1	1	1
9	1	1	1	23	1	1	1
10	1	1	1	24	0	1	1
11	1	1	1	25	1	1	1
12	1	1	1	26	1	1	1
13	-1	1	1	27	1	1	1
14	-1	1	1	28	1	1	1

6.2.1. Simulación considerando diodos *bypass*

Para representar este comportamiento, se simularon fallas parciales en las que uno o más paneles de una cadena quedaban inactivos, reduciendo proporcionalmente la potencia total. Por ejemplo, una falla en dos de los doce paneles de una cadena se reflejó como una disminución del $\frac{2}{12}$ en la potencia continua generada por dicha cadena. Este enfoque permite capturar de forma más realista el efecto de fallas parciales, evitando sobreestimar su impacto y representando adecuadamente la respuesta del sistema ante este tipo de eventos. Se realizó una serie de fallas en módulos para determinar cuál es la cantidad mínima de paneles dañados para que el modelo los detecte como una falla. El análisis se divide en los casos donde existen paneles dañados, pero el inversor funciona de manera correcta y donde no existe panel dañado, pero el inversor presenta fallas. El caso donde no existe panel dañado se presenta en la Sección 6.1.

Falla por pérdida de paneles

Las simulaciones generadas presentaron una efectividad del método desde la eliminación de un panel en el módulo 2. La Tabla 6.9 muestra los resultados obtenidos, los únicos dos valores cambiantes son resaltados en color amarillo, estos corresponden a los días 11 y 17 del mes. La Tabla 6.10 no modifica sus valores, debido a que la potencia en los inversores se mantuvo en relación a la potencia de los módulos.

Las Tablas 6.11 y 6.12 corresponden a los resultados de eliminar dos paneles en el módulo 2, los días 8 y 18 de mes.

Tabla 6.9: Fallas en módulos por simulación de pérdida de un panel.

D	M_1	M_2	M_3
1	0.5	0	0.5
2	0.5	0.5	0.5
3	0	0	0
4	0.5	0.5	0.5
5	0.5	0.5	0.5
6	1	1	1
7	0.5	0.5	0.5
8	0.5	0.5	0.5
9	0.5	0.5	0.5
10	0.5	0.5	0.5
11	0.5	0	0.5
12	0.5	0.5	0.5
13	0.5	0.5	0.5
14	0.5	0.5	0.5

D	M_1	M_2	M_3
15	0.5	0.5	0.5
16	0.5	0.5	0.5
17	0.5	0	0.5
18	0.5	0.5	0.5
19	0.5	0.5	0.5
20	0.5	0.5	0.5
21	0.5	0.5	0.5
22	0.5	0.5	0.5
23	0.5	0.5	0.5
24	1	1	1
25	1	1	1
26	0.5	0.5	0.5
27	1	1	1
28	1	1	1

Tabla 6.10: Fallas en inversores por pérdida de un panel.

D	I_1	I_2	I_3
1	1	1	1
2	1	1	1
3	-1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1

D	I_1	I_2	I_3
15	1	1	1
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	0	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1

Tabla 6.11: Fallas en módulos por pérdida de dos paneles.

D	M_1	M_2	M_3	D	M_1	M_2	M_3
1	0.5	0	0.5	15	0.5	0.5	0.5
2	0.5	0.5	0.5	16	0.5	0.5	0.5
3	0	0	0	17	0.5	0.5	0.5
4	0.5	0.5	0.5	18	0.5	0	0.5
5	0.5	0.5	0.5	19	0.5	0.5	0.5
6	1	1	1	20	0.5	0.5	0.5
7	0.5	0.5	0.5	21	0.5	0.5	0.5
8	0.5	0	0.5	22	0.5	0.5	0.5
9	0.5	0.5	0.5	23	0.5	0.5	0.5
10	0.5	0.5	0.5	24	1	1	1
11	0.5	0.5	0.5	25	1	1	1
12	0.5	0.5	0.5	26	0.5	0.5	0.5
13	0.5	0.5	0.5	27	1	1	1
14	0.5	0.5	0.5	28	1	1	1

Tabla 6.12: Fallas en inversores por pérdida de dos paneles.

D	I_1	I_2	I_3	D	I_1	I_2	I_3
1	1	1	1	15	1	1	1
2	1	1	1	16	1	1	1
3	-1	1	1	17	1	1	1
4	1	1	1	18	1	1	1
5	1	1	1	19	1	1	1
6	1	1	1	20	1	1	1
7	1	1	1	21	1	1	1
8	1	1	1	22	1	1	1
9	1	1	1	23	1	1	1
10	1	1	1	24	0	1	1
11	1	1	1	25	1	1	1
12	1	1	1	26	1	1	1
13	1	1	1	27	1	1	1
14	1	1	1	28	1	1	1

Fallas en potencia alterna

Existen dos casos posibles en la Potencia en corriente alterna, esta puede verse reducida por factores eléctricos o físicos, o puede perderse por completo el componente. El primer caso fue analizado en la Figura 6.6. A continuación se muestran las tablas correspondientes a la pérdida del inversor 2 los días 8 y 18, se asume que se hizo el cambio de componente en el momento de su detección.

Tabla 6.13: Fallas en módulos por pérdida de un inversor.

D	M_1	M_2	M_3
1	0.5	0	0.5
2	0.5	0.5	0.5
3	0	0	0
4	0.5	0.5	0.5
5	0.5	0.5	0.5
6	1	1	1
7	0.5	0.5	0.5
8	0.5	0.5	0.5
9	0.5	0.5	0.5
10	0.5	0.5	0.5
11	0.5	0.5	0.5
12	0.5	0.5	0.5
13	0.5	0.5	0.5
14	0.5	0.5	0.5

D	M_1	M_2	M_3
15	0.5	0.5	0.5
16	0.5	0.5	0.5
17	0.5	0.5	0.5
18	0.5	0.5	0.5
19	0.5	0.5	0.5
20	0.5	0.5	0.5
21	0.5	0.5	0.5
22	0.5	0.5	0.5
23	0.5	0.5	0.5
24	1	1	1
25	1	1	1
26	0.5	0.5	0.5
27	1	1	1
28	1	1	1

Tabla 6.14: Fallas en inversores por pérdida de un inversor.

D	I_1	I_2	I_3
1	1	1	1
2	1	1	1
3	-1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	0	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1

D	I_1	I_2	I_3
15	1	1	1
16	1	1	1
17	1	1	1
18	1	0	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	0	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1

El análisis de estas simulaciones permitió observar el comportamiento del modelo ante distintas condiciones de falla, tanto internas como externas al arreglo FV, y evaluar su capacidad para identificar anomalías topológicas en las representaciones generadas.

6.3. Discusión

El modelo desarrollado revela que, aunque el sistema FV está expuesto a cambios por condiciones ambientales y presenta periodos con menor rendimiento en algunos módulos, cuenta con componentes que, en general, aseguran una generación de potencia adecuada. Este comportamiento resalta la importancia de contar con una supervisión continua y un análisis de los datos, que permita distinguir entre variaciones normales de operación y verdaderas anomalías que requieran atención. En este contexto, el uso del TDA se presenta como una herramienta capaz de diagnosticar fallas. A lo largo del estudio, demostró su capacidad para captar patrones sutiles en la geometría general de los datos. En particular, el análisis de diagramas de persistencia permitió identificar cambios significativos en la forma de las curvas de potencia cuando se introdujeron fallas simuladas, como la pérdida de generación en ciertos módulos o la desconexión parcial del sistema, lo cual comprueba la hipótesis planteada al comienzo del estudio. Durante los experimentos, se observó que, bajo condiciones normales, los datos presentan una estructura topológica regular, mientras que en presencia de fallas, esta estructura se modifica de forma notoria. Los ciclos que aparecen en los diagramas y que representan características importantes de la forma de los datos tienden a ser más persistentes en escenarios estables y a dividirse cuando existe alguna anomalía. Esta diferencia

fue cuantificada mediante la distancia cuello de botella, la cual permitió establecer umbrales de detección y evaluar la magnitud de los cambios de manera sistemática.

Durante el análisis se identificaron **fallas de comunicación** en el sistema de almacenamiento de datos, provocando intermitencias en el registro. La Tabla 6.15 resume los meses afectados. Cabe destacar que durante este periodo no se cuenta con un mecanismo automático para verificar la conexión con la nube y verificar el almacenamiento.

Tabla 6.15: Meses y cantidad de días almacenados. La última fila indica si se detectó falla en la comunicación (valor 0) o no (valor 1).

Mes	08/2021	09/2021	08/2024	09/2024	10/2024	02/2025	04/2025	05/2025
Días	25	30	24	30	31	28	14	12
Falla	0	0	1	1	1	1	0	0

Una de las principales ventajas del TDA es que no depende de modelos físicos específicos ni de supuestos previos sobre el comportamiento del sistema. A diferencia de los métodos tradicionales que requieren conocer en detalle las características eléctricas o estructurales del sistema FV, el TDA opera directamente sobre los datos disponibles, lo que lo hace adaptable a diferentes configuraciones. Además, mostró ser menos sensible al ruido, lo que resulta especialmente útil en entornos donde los sensores presentan incertidumbre o los datos son incompletos. No obstante, el enfoque también presenta ciertas limitaciones. Si bien es posible detectar la presencia de una falla, no permite localizar con precisión el componente afectado. Además, la interpretación de los resultados, particularmente los diagramas de persistencia, requiere conocimientos especializados que pueden dificultar su adopción por parte de usuarios sin formación técnica. Esto sugiere la necesidad de desarrollar herramientas de visualización más intuitivas o procesos de capacitación que faciliten su integración en sistemas de monitoreo operativos.

Finalmente, aunque los resultados obtenidos con datos simulados y casos de estudio fueron prometedores, logrando una efectividad cercana al 95 % en la detección de fallas, es necesario validar esta metodología con datos reales de sistemas con fallas. En estos escenarios, factores como la variabilidad ambiental, el desgaste desigual de los componentes o errores de medición podrían influir en los resultados, por lo que se recomienda continuar con estudios adicionales que refuercen su aplicabilidad práctica. En conjunto, los resultados alcanzados reafirman el valor del TDA como una herramienta novedosa para el análisis del comportamiento de sistemas FV. Su capacidad para detectar patrones en los datos ofrece una nueva perspectiva para el monitoreo y diagnóstico, complementando las metodologías tradicionales.

Comparación con otras metodologías

En el Capítulo 3 se describieron diversas metodologías empleadas para el diagnóstico de fallas en sistemas FV, incluyendo enfoques basados en inteligencia artificial, modelos eléctricos y técnicas estadísticas. A continuación se presenta, de manera breve, una comparación cualitativa del enfoque basado en TDA con algunas de las metodologías más representativas. Debido a la arquitectura del sistema analizado, no fue posible realizar una cuantificación directa del desempeño en términos de métricas de detección. No obstante, la revisión de la literatura ofrece suficientes elementos para establecer una comparación cualitativa que aporte una visión sobre su aplicabilidad, robustez y capacidad de generalización.

Las redes neuronales artificiales son una de las herramientas más empleadas en el diagnóstico de fallas, principalmente por su capacidad para modelar relaciones no lineales complejas entre variables y capturar de manera automática interacciones entre predictores. Además, pueden desarrollarse con distintos algoritmos de entrenamiento y requieren un formalismo matemático menos estricto en comparación con otros enfoques [62]. Sin embargo, su principal desventaja radica en que operan como una “caja negra”, lo que dificulta la interpretación causal de los resultados. A esto se suma la alta demanda de recursos computacionales, la tendencia al sobreajuste y la dificultad de aplicar estos modelos directamente en entornos de campo, lo cual limita su uso práctico sin una infraestructura tecnológica adecuada. Por otro lado, las pruebas basadas en la curva I–V ofrecen ventajas notables en términos de rapidez y detalle en el diagnóstico. Estas permiten obtener resultados precisos sin necesidad de poner en línea el inversor, lo que reduce tiempos de prueba y ofrece una línea base detallada del rendimiento del sistema fotovoltaico [63], [64]. Asimismo, la gestión de datos resultante es eficiente y facilita un análisis a nivel de módulo. No obstante, presentan limitaciones importantes al no ser aplicables como técnica de monitorización en tiempo real; la técnica depende de mediciones puntuales en un instante específico; y requiere equipos especializados, lo que puede elevar los costos de implementación. En cuanto a los modelos eléctricos equivalentes, destacan por su simplicidad y bajo costo computacional. Su estructura facilita la implementación en laboratorio y permite obtener resultados básicos de diagnóstico sin necesidad de una infraestructura compleja [65]. A pesar de estas ventajas, su precisión es reducida, muestran un desempeño deficiente bajo condiciones de sombreado parcial y, en general, no logran representar de forma fiel la totalidad de la curva I–V. Esto los convierte en herramientas útiles en escenarios controlados, pero con limitaciones considerables en aplicaciones reales y dinámicas. Finalmente, el TDA se presenta como un enfoque novedoso y prometedor. Su principal fortaleza radica en la independencia de modelos físicos, lo que le permite adaptarse a datos provenientes de contextos diversos sin requerir un conocimiento exhaustivo del sistema, como lo hacen los modelos basados en redes neuronales. Además, es particularmente robusto frente al ruido y ofrece la posibilidad de identificar cambios sutiles en la forma de los datos. Sin embargo, también enfrenta desafíos; su interpretación requiere una formación especializada, carece de clasificadores automáticos que identifiquen una falla, y al aumentar en gran medida la cantidad de variables analizadas puede causar un elevado costo computacional.

En resumen, el TDA se perfila como una alternativa complementaria a las metodologías tradicionales, especialmente en contextos donde los datos son limitados, el conocimiento del sistema es parcial o se requiere una interpretabilidad del comportamiento más profunda. Aunque aún se encuentra en una etapa de consolidación, su potencial para enriquecer el diagnóstico de fallas en sistemas fotovoltaicos es evidente y lo coloca como una herramienta con gran proyección hacia el futuro.

Capítulo 7

Conclusiones y trabajo futuro

Conclusiones

El presente estudio demuestra la eficacia del TDA como una herramienta robusta e innovadora para el diagnóstico de fallas en sistemas fotovoltaicos. Su capacidad para extraer información estructural directamente de la forma de los datos permite identificar patrones de funcionamiento y detectar comportamientos anómalos que podrían pasar desapercibidos mediante métodos tradicionales. Uno de los hallazgos más relevantes es que la **presencia, cantidad y persistencia de los 1-hoyos en los diagramas de persistencia** se correlacionan con el rendimiento del sistema. Una menor persistencia implica una reducción en la generación de energía, mientras que un incremento del ruido y su alejamiento de la diagonal se asocian con degradación operativa. Estas características permiten construir indicadores cuantitativos de deterioro.

La **métrica de cuello de botella** se posiciona como una herramienta clave para comparar la topología de distintos módulos y establecer umbrales precisos para la detección automática de fallas. Este enfoque permite distinguir entre pérdidas energéticas debidas a condiciones ambientales y aquellas provocadas por fallas internas. El modelo desarrollado que integra el preprocesamiento de datos, el cálculo de diagramas de persistencia, el uso de métricas topológicas y la definición de umbrales constituye un marco integral y escalable para el diagnóstico automático. La validación realizada con datos reales de diversos periodos, como febrero de 2025, confirma su aplicabilidad operativa y su capacidad para detectar comportamientos atípicos en módulos individuales. Asimismo, el análisis de la **potencia generada a lo largo del día** mediante TDA, representada como nubes de puntos y caracterizada con diagramas de persistencia, permitió distinguir días con funcionamiento óptimo de aquellos con pérdidas energéticas. Bajo condiciones ideales, el conjunto de datos presenta un ciclo bien definido con alta persistencia; en cambio, las fallas se manifiestan como ciclos adicionales o menos persistentes, reflejando una menor eficiencia.

En conjunto, los resultados obtenidos no solo validan la viabilidad del TDA como una técnica no invasiva y sensible para el diagnóstico de sistemas fotovoltaicos, sino que también permiten avanzar hacia **modelos automáticos de detección de fallas** basados en la estructura topológica de los datos. La metodología propuesta alcanzó una **efectividad del 95 %** en la detección de fallas dentro del conjunto de datos existente junto a las simulaciones realizadas, lo cual respalda su uso como una alternativa confiable para el monitoreo inteligente y la gestión preventiva de sistemas energéticos renovables. Aunque el sistema analizado es de pequeña escala, presenta características representativas de instalaciones mayores. El modelo fue diseñado con criterios de escalabilidad y puede implementarse en sistemas compuestos por miles de paneles. Además, el registro de varia-

bles clave por parte del inversor permite realizar comparaciones entre la potencia generada y la entregada a la red, facilitando la localización de fallas en distintas etapas del proceso de conversión energética.

Trabajo futuro

El trabajo realizado abre múltiples líneas de investigación y desarrollo orientadas a potenciar el diagnóstico de fallas en sistemas fotovoltaicos mediante el uso de tecnologías emergentes y enfoques computacionales avanzados. El primer trabajo futuro, y quizá uno de los más alcanzables, consiste en la **implementación del modelo en una granja solar de tamaño considerable**. Esto permitiría validar el enfoque en un entorno real, evaluar su desempeño frente a condiciones operativas variables y recolectar datos que retroalimenten y mejoren la precisión del sistema. Este paso sería fundamental para escalar la solución y convertirla en una herramienta útil para la industria energética.

Uno de los próximos pasos clave consiste en avanzar del diagnóstico a la **localización e identificación precisa de las fallas** detectadas. Para lograrlo, se plantea el desarrollo de técnicas de *aislamiento automático de fallas*, que permitan ubicar los módulos o componentes específicos afectados dentro del sistema. La implementación del modelo en **plataformas embebidas o en la nube** permitiría su integración en sistemas de monitoreo en tiempo real, ampliando su aplicabilidad a instalaciones fotovoltaicas a gran escala. Esta integración facilitaría la validación del modelo en condiciones operativas diversas y permitiría evaluar su desempeño frente a diferentes arquitecturas de hardware y volúmenes de datos. Se propone también la **incorporación de información multimodal** en el análisis topológico. El uso de datos meteorológicos (irradiancia, temperatura ambiente, velocidad del viento), variables internas del sistema (temperatura de operación, corriente, voltaje), e incluso imágenes térmicas o visuales, enriquecería el contexto del diagnóstico y podría mejorar significativamente la discriminación entre pérdidas debidas a factores ambientales y fallas internas.

Otra línea de trabajo consiste en **explorar modelos de aprendizaje automático**, como redes neuronales profundas, para la clasificación automática de fallas a partir de representaciones vectoriales derivadas de los diagramas de persistencia. Esto permitiría automatizar la interpretación de los resultados topológicos y lograr una **clasificación robusta y en tiempo real** de los distintos tipos de anomalías. Se considera también prioritario **evaluar y optimizar los tiempos de procesamiento** del modelo, especialmente en escenarios de gran escala. Esto permitirá determinar si los diagramas de persistencia son la representación más eficiente, o si es preferible una transformación previa hacia vectores persistentes u otras codificaciones que reduzcan la carga computacional sin pérdida de precisión diagnóstica. Además, se recomienda investigar mecanismos para la **calibración dinámica de los umbrales de detección**. Actualmente, los umbrales son fijos; sin embargo, su ajuste automático en función de variables como la temperatura, la irradiancia o el historial del sistema podría mejorar la sensibilidad y especificidad del diagnóstico, minimizando falsos positivos y negativos.

Resulta relevante el desarrollar la capacidad del modelo para **detectar fallas complejas o combinadas**, que involucren simultáneamente distintos componentes del sistema (como módulos, inversores o cableado). Esto requerirá análisis topológicos más complejos que consideren interacciones entre múltiples variables y regiones del sistema. Para validar la generalización del enfoque, es indispensable realizar **pruebas exhaustivas con conjuntos de datos más amplios y diver-**

sos, abarcando periodos prolongados y condiciones ambientales variadas. Esto permitirá evaluar la robustez del modelo y su adaptabilidad a distintos contextos geográficos y tecnológicos.

Finalmente, se plantea el desarrollo de una **interfaz gráfica de usuario intuitiva**, que permita visualizar en tiempo real los resultados del análisis topológico, la clasificación de fallas y el estado operativo del sistema. Esta herramienta facilitaría la interpretación por parte de técnicos y operadores, promoviendo una gestión más eficiente y preventiva del mantenimiento. Estas líneas de trabajo permitirán consolidar el Análisis Topológico de Datos como una herramienta fundamental para el monitoreo inteligente, el diagnóstico automático y la operación optimizada de sistemas fotovoltaicos basados en energías renovables.

Apéndice A

Códigos implementados, Python

Función que interpola los puntos necesarios

```
def interpolar(datos, puntos_intermedios=7):
    interpolados = []
    # Normalizamos los datos
    datos = datos/100
    n = 70
    for i in range(len(datos) - 1):
        punto_inicial = datos[i]
        punto_final = datos[i + 1]
        for t in np.linspace(0, 1, puntos_intermedios + 2):
            punto_intermedio = punto_inicial + t * (punto_final - punto_inicial)
            interpolados.append(punto_intermedio)
    x = np.arange(0, n, n / len(interpolados))
    finales = []
    for index in range(len(interpolados)):
        data_final.append([x[index], interpolados[index]])
        data_final.append([x[index], 0])
    return data_final
```

Función de carga de la base de datos

```
def carga(dia):
    archivo = f"mes/{dia}.csv"
    df = pd.read_csv(archivo, delimiter=';', skiprows=5)
    Pac1 = np.array(interpolar(df.iloc[:, 12].astype(float)))
    Vpv1 = df.iloc[:, 17].astype(float)
    Ipv1 = df.iloc[:, 10].astype(float)
    Ppv1 = np.array(interpolar(np.multiply(Vpv1, Ipv1)))
    Pac2 = interpolar(df.iloc[:, 29].astype(float))
    Vpv2 = df.iloc[:, 34].astype(float)
    Ipv2 = df.iloc[:, 27].astype(float)
    Ppv2 = np.array(interpolar(np.multiply(Vpv2, Ipv2)))
    Pac3 = interpolar(df.iloc[:, 46].astype(float))
    Vpv3 = df.iloc[:, 51].astype(float)
    Ipv3 = df.iloc[:, 44].astype(float)
    Ppv3 = np.array(interpolar(np.multiply(Vpv3, Ipv3)))
    return Pac1, Ppv1, Pac2, Ppv2, Pac3, Ppv3
```

Función que extrae los 1-hoyos

```
def diagrama(datos):
    if np.array(datos).any():
        ac = gd.AlphaComplex(points=datos)
        st = ac.create_simplex_tree()
        diag = st.compute_persistence()
        #Recuperamos los valores originales de la informacion
        p1 = np.sqrt(st.persistence_intervals_in_dimension(1)) * 100
    else:
        p1 = np.array([[0, 0]])
    return p1
```

Función encargada de la distancia y el porcentaje entre los módulos y el comportamiento ideal.

```
def Distancia(modulo1, modulo2, modulo3, Ideal1, Ideal2, Ideal3):
    distancia1 = persim.bottleneck(Ideal1, modulo1, matching=False)
    distancia2 = persim.bottleneck(Ideal2, modulo2, matching=False)
    distancia3 = persim.bottleneck(Ideal3, modulo3, matching=False)
    max1 = np.amax(Ideal1[:, 1] - Ideal1[:, 0])
    max2 = np.amax(Ideal2[:, 1] - Ideal2[:, 0])
    max3 = np.amax(Ideal3[:, 1] - Ideal3[:, 0])
    percent1 = (np.amax(modulo1[:, 1] - modulo1[:, 0]) / max1) * 100
    percent2 = (np.amax(modulo2[:, 1] - modulo2[:, 0]) / max2) * 100
    percent3 = (np.amax(modulo3[:, 1] - modulo3[:, 0]) / max3) * 100
    return [distancia1, distancia2, distancia3, porcentaje1, porcentaje2,
            porcentaje3]
```

Función que compara los datos obtenidos de los módulos por pares.

```
def compara(diagrama1, diagrama2, diagrama3):
    dB12 = persim.bottleneck(diagrama1, diagrama2, matching=False)
    dB13 = persim.bottleneck(diagrama1, diagrama3, matching=False)
    dB23 = persim.bottleneck(diagrama2, diagrama3, matching=False)
    return [dB12, dB13, dB23]
```

Función que compara los comportamientos de los inversores con los módulos.

```
def comparaInversor(diagPpv1, diagPpv2, diagPpv3, diagPac1, diagPac2, diagPac3):
    :
    dB12 = persim.bottleneck(diagPpv1, diagPac1, matching=False)
    dB13 = persim.bottleneck(diagPpv2, diagPac2, matching=False)
    dB23 = persim.bottleneck(diagPpv3, diagPac3, matching=False)
    return [dB12, dB13, dB23]
```

Detección de fallas

```
# Inicializaci n de listas para almacenar resultados
fallasPpv, fallasPac = [], []
distanciasPac, distanciasPpv = [], []
distIdealesPpv, distIdealesPac = [], []

for i in range(1, 29): #Ajustar a la cantidad de dias
    try: #Si no existe problema de comunicacion
        Pac1, Ppv1, Pac2, Ppv2, Pac3, Ppv3 = carga(i)

        # Diagramas de persistencia de las potencias
        diam1Ppv = diagrama(Ppv1)
```

```

diam2Ppv = diagrama(Ppv2)
diam3Ppv = diagrama(Ppv3)

diam1Pac = diagrama(Pac1)
diam2Pac = diagrama(Pac2)
diam3Pac = diagrama(Pac3)

# Comparaciones con referencias ideales
distIdealesPpv.append(
    Distancia(diam1Ppv, diam2Ppv, diam3Ppv,
              Idealm1Ppv, Idealm2Ppv, Idealm3Ppv) )
distIdealesPac.append(
    Distancia(diam1Pac, diam2Pac, diam3Pac,
              Idealm1Pac, Idealm2Pac, Idealm3Pac) )

# Comparaciones entre modulos
distanciasPpv.append(compara(diam1Ppv, diam2Ppv, diam3Ppv))
distanciasPac.append(comparaInversor(
    diam1Ppv, diam2Ppv, diam3Ppv,
    diam1Pac, diam2Pac, diam3Pac))

# Inicializar deteccion de fallas
fallasPpv.append([1, 1, 1])

# Modulo 1
if distanciasPpv[i - 1][0] > umbralPpv[0] and distanciasPpv[i - 1][1] >
    umbralPpv[0]:
    if distIdealesPpv[i - 1][0] > umbralPpv[0]:
        fallasPpv[i - 1][0] = 0
    else:
        fallasPpv[i - 1][0] = 0.5
# El proceso para los dem s m dulos es an logo

# Evaluaci n de fallas en PAC
fallasPac.append([1, 1, 1])

for j in range(3):
    if fallasPpv[i - 1][j] == 1:
        if distanciasPac[i - 1][j] > 1.03 * UmbralIdeal[j]:
            fallasPac[i - 1][j] = 0
        elif distanciasPac[i - 1][j] > UmbralIdeal[j]:
            fallasPac[i - 1][j] = 0.5
    elif fallasPpv[i - 1][j] == 0:
        fallasPac[i - 1][j] = -1

except Exception as e: #En caso de existir falla asignamos valores por
    default
    fallasPpv.append([-1, -1, -1])
    fallasPac.append([-1, -1, -1])
    distanciasPpv.append([np.nan, np.nan, np.nan])
    distanciasPac.append([np.nan, np.nan, np.nan])
    distIdealesPpv.append([np.nan]*6)

```


Apéndice B

Meses analizados

Agosto 2021

Durante el mes de agosto del 2021, se generó un problema de conexión con la base de datos entre los días 15 y 20.

Tabla B.1: Fallas del sistema en agosto de 2021, Potencia en continua

D	M_1	M_2	M_3
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	-1	-1	-1
16	-1	-1	-1

D	M_1	M_2	M_3
17	-1	-1	-1
18	-1	-1	-1
19	-1	-1	-1
20	-1	-1	-1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	1	1	1
31	1	1	1

Tabla B.2: Fallas del sistema en agosto de 2021, Potencia en alterna

D	I_1	I_2	I_3
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	-1	-1	-1
16	-1	-1	-1

D	I_1	I_2	I_3
17	-1	-1	-1
18	-1	-1	-1
19	-1	-1	-1
20	-1	-1	-1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	1	1	1
31	1	1	1

Septiembre 2021

El mes de Septiembre del 2021 se contó con una pérdida en la conexión el día 22. A lo largo de todo el resto de mes, se reportaron pérdidas en la generación por factores climáticos, pero los inversores funcionaron de manera correcta.

Tabla B.3: Fallas del sistema en Septiembre de 2021, Potencia en continua

D	M_1	M_2	M_3
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	1	1	1

D	M_1	M_2	M_3
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	-1	-1	-1
23	1	1	1
24	1	1	1
25	1	1	1
26	0	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	1	1	1

Tabla B.4: Fallas del sistema en septiembre de 2021, Potencia en alterna

D	I_1	I_2	I_3
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	1	1	1

D	M_1	M_2	M_3
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	-1	-1	-1
23	1	1	1
24	1	1	1
25	1	1	1
26	-1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	1	1	1

Septiembre 2024

Tabla B.5: Fallas del sistema en septiembre de 2024, Potencia en continua

D	M_1	M_2	M_3	D	M_1	M_2	M_3
1	1	1	1	16	1	1	1
2	1	1	1	17	1	1	1
3	1	1	1	18	1	1	1
4	1	1	1	19	1	1	1
5	1	1	1	20	1	1	1
6	1	1	1	21	1	1	1
7	1	1	1	22	1	1	1
8	1	1	1	23	1	1	1
9	1	1	1	24	1	1	1
10	1	1	1	25	1	1	1
11	1	1	1	26	1	1	1
12	1	1	1	27	1	1	1
13	1	1	1	28	1	1	1
14	1	1	1	29	1	1	1
15	1	1	1	30	1	1	1

Tabla B.6: Fallas del sistema en septiembre de 2024, Potencia en alterna

D	I_1	I_2	I_3	D	I_1	I_2	I_3
1	1	1	1	16	1	1	1
2	1	1	1	17	1	1	1
3	1	1	1	18	1	1	1
4	1	1	1	19	1	1	1
5	1	1	1	20	1	1	1
6	1	1	1	21	1	1	1
7	1	1	1	22	1	1	1
8	1	1	1	23	1	1	1
9	1	1	1	24	1	1	1
10	1	1	1	25	1	1	1
11	1	1	1	26	1	1	1
12	1	1	1	27	1	1	1
13	1	1	1	28	1	1	1
14	1	1	1	29	1	1	1
15	1	1	1	30	1	1	1

Octubre 2024

Tabla B.7: Fallas del sistema en octubre de 2024, Potencia en continua

D	M_1	M_2	M_3
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	1	1	1
16	1	1	1

D	M_1	M_2	M_3
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	0
28	1	1	1
29	1	1	1
30	1	1	1
31	1	1	1

Tabla B.8: Fallas del sistema en Octubre de 2024, Potencia en alterna

D	I_1	I_2	I_3
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	0	1	1
11	0	1	1
12	0	1	1
13	1	1	1
14	1	1	1
15	1	1	1
16	1	1	1

D	I_1	I_2	I_3
17	0	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	0	1	1
25	0	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	1	1	1
31	1	1	1

Abril 2025

Tabla B.9: Fallas del sistema en abril de 2025,
Potencia en continua

D	M_1	M_2	M_3
1	-1	-1	-1
2	-1	-1	-1
3	-1	-1	-1
4	-1	-1	-1
5	-1	-1	-1
6	-1	-1	-1
7	-1	-1	-1
8	-1	-1	-1
9	-1	-1	-1
10	-1	-1	-1
11	-1	-1	-1
12	-1	-1	-1
13	-1	-1	-1
14	-1	-1	-1
15	-1	-1	-1

D	M_1	M_2	M_3
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	-1	-1	-1

Tabla B.10: Fallas del sistema en abril de 2025,
Potencia en alterna

D	I_1	I_2	I_3
1	-1	-1	-1
2	-1	-1	-1
3	-1	-1	-1
4	-1	-1	-1
5	-1	-1	-1
6	-1	-1	-1
7	-1	-1	-1
8	-1	-1	-1
9	-1	-1	-1
10	-1	-1	-1
11	-1	-1	-1
12	-1	-1	-1
13	-1	-1	-1
14	-1	-1	-1
15	-1	-1	-1

D	I_1	I_2	I_3
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	-1	-1	-1

Mayo 2025

Tabla B.11: Fallas del sistema en mayo de 2025, Potencia en continua

D	M_1	M_2	M_3
1	-1	-1	-1
2	-1	-1	-1
3	-1	-1	-1
4	-1	-1	-1
5	-1	-1	-1
6	-1	-1	-1
7	-1	-1	-1
8	-1	-1	-1
9	-1	-1	-1
10	-1	-1	-1
11	-1	-1	-1
12	-1	-1	-1
13	-1	-1	-1
14	-1	-1	-1
15	-1	-1	-1
16	-1	-1	-1

D	M_1	M_2	M_3
17	-1	-1	-1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	-1	-1	-1
31	-1	-1	-1

Tabla B.12: Fallas del sistema en mayo de 2025, Potencia en alterna

D	I_1	I_2	I_3
1	-1	-1	-1
2	-1	-1	-1
3	-1	-1	-1
4	-1	-1	-1
5	-1	-1	-1
6	-1	-1	-1
7	-1	-1	-1
8	-1	-1	-1
9	-1	-1	-1
10	-1	-1	-1
11	-1	-1	-1
12	-1	-1	-1
13	-1	-1	-1
14	-1	-1	-1
15	-1	-1	-1
16	-1	-1	-1

D	I_1	I_2	I_3
17	-1	-1	-1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	-1	-1	-1
31	-1	-1	-1

Bibliografía

- [1] Z. Gao, C. Cecati y S. X. Ding, “A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part I: Fault Diagnosis With Model-Based and Signal-Based Approaches,” *IEEE Transactions on Industrial Electronics*, vol. 62, n.º 6, págs. 3757-3767, 2015. DOI: 10.1109/TIE.2015.2417501.
- [2] C. Edwards y C. P. Tan, “Robust fault detection using sliding mode observers,” en *Advances in variable structure and sliding mode control*, Springer, 2006, págs. 293-312.
- [3] “Distributed bearing fault diagnosis based on vibration analysis,” *Mechanical Systems and Signal Processing*, vol. 66-67, págs. 521-532, 2016, ISSN: 0888-3270. DOI: <https://doi.org/10.1016/j.ymssp.2015.06.007>. dirección: <https://www.sciencedirect.com/science/article/pii/S0888327015002939>.
- [4] B. Wang et al., “Automatic Fault Diagnosis of Infrared Insulator Images Based on Image Instance Segmentation and Temperature Analysis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, n.º 8, págs. 5345-5355, 2020. DOI: 10.1109/TIM.2020.2965635.
- [5] R. Isermann, “Model-based fault-detection and diagnosis – status and applications,” *Annual Reviews in Control*, vol. 29, n.º 1, págs. 71-85, 2005, ISSN: 1367-5788. DOI: <https://doi.org/10.1016/j.arcontrol.2004.12.002>. dirección: <https://www.sciencedirect.com/science/article/pii/S1367578805000052>.
- [6] H. Ballesteros-Moncada, E. J. Herrera-López y J. Anzurez-Marín, “Fuzzy model-based observers for fault detection in CSTR,” *ISA transactions*, vol. 59, págs. 325-333, 2015.
- [7] D. Rotondo, A. Cristofaro, T. A. Johansen, F. Nejjari y V. Puig, “Robust fault and icing diagnosis in unmanned aerial vehicles using LPV interval observers,” *International Journal of Robust and Nonlinear Control*, vol. 29, n.º 16, págs. 5456-5480, 2019.
- [8] E. Bernardi y E. J. Adam, “Observer-based fault detection and diagnosis strategy for industrial processes,” *Journal of the Franklin Institute*, vol. 357, n.º 14, págs. 10 054-10 081, 2020.
- [9] S. Firth, K. Lomas y S. Rees, “A simple model of PV system performance and its use in fault detection,” *Solar Energy*, vol. 84, n.º 4, págs. 624-635, abr. de 2010.
- [10] P. Jieyang et al., “A systematic review of data-driven approaches to fault diagnosis and early warning,” *Journal of Intelligent Manufacturing*, vol. 34, n.º 8, págs. 3277-3304, 2023.
- [11] B. Guo, B. Yang, S. Wang, W. Shi, F. Yang y D. Wang, “Machine learning-based fault diagnosis and classification of three-phase transmission lines with RFE and domain knowledge,” *Electric Power Systems Research*, vol. 247, pag. 111 777, 2025.

- [12] M. Alrifaei et al., “Hybrid deep learning model for fault detection and classification of grid-connected photovoltaic system,” *IEEE Access*, vol. 10, págs. 13 852-13 869, 2022.
- [13] H. Yousif y Z. Al-Milaji, “Fault detection from PV images using hybrid deep learning model,” *Solar Energy*, vol. 267, pág. 112 207, 2024.
- [14] G. G. A. El-wahhab, T. A. M. Abdelwahab, Y. K. O. Taha Osman, M. F. A. Abuhussein, A. E. M. Fodah y K. A. M. Ali, “Performance evaluation of solar panels under different dust accumulation conditions using thermography: focusing on PV temperature variation,” *Materials for Renewable and Sustainable Energy*, vol. 12, n.º 3, págs. 247-255, 2023.
- [15] H. Bakır, “Detection of Faults in Photovoltaic Modules of SPPS in Turkey; Infrared thermographic diagnosis and recommendations,” *Journal of Electrical Engineering & Technology*, vol. 18, n.º 3, págs. 1945-1957, 2023.
- [16] M. Alsafasfeh, I. Abdel-Qader, B. Bazuin, Q. Alsafasfeh y W. Su, “Unsupervised fault detection and analysis for large photovoltaic systems using drones and machine vision,” *Energies*, vol. 11, n.º 9, pág. 2252, 2018.
- [17] S. Fadhel et al., “PV shading fault detection and classification based on IV curve using principal component analysis: Application to isolated PV system,” *Solar Energy*, vol. 179, págs. 1-10, 2019.
- [18] S. Sivagamasundari, M. S. Rayudu et al., “IoT based solar panel fault and maintenance detection using decision tree with light gradient boosting,” *Measurement: Sensors*, vol. 27, pág. 100 726, 2023.
- [19] S. Jaybhaye, V. Sirvi, S. Srivastava, V. Loya, V. Gujarathi y M. Jaybhaye, “Classification and Early Detection of Solar Panel Faults with Deep Neural Network Using Aerial and Electroluminescence Images,” *Journal of Failure Analysis and Prevention*, págs. 1-13, 2024.
- [20] K. A. K. Niazi, W. Akhtar, H. A. Khan, Y. Yang y S. Athar, “Hotspot diagnosis for solar photovoltaic modules using a Naive Bayes classifier,” *Solar Energy*, vol. 190, págs. 34-43, 2019.
- [21] H. R. Rocha et al., “Tracking Defective Panel on Photovoltaic Strings with Non-Intrusive Monitoring and Deep Learning,” *Journal of Control, Automation and Electrical Systems*, págs. 1-14, 2024.
- [22] L. Li, Z. Wang y T. Zhang, “Gbh-yolov5: Ghost convolution with bottleneckcsp and tiny target prediction head incorporating yolov5 for pv panel defect detection,” *Electronics*, vol. 12, n.º 3, pág. 561, 2023.
- [23] A. Et-taleby, Y. Chaibi, B. Elkari, M. Benslimane y Z. Chalh, “Detection and Classification of Faults in PV Systems Based on Thermal Imaging and Fuzzy Logic Algorithm,” en *International Conference on Electrical Systems & Automation*, Springer, 2023, págs. 91-95.
- [24] M. Dhimish, V. Holmes, B. Mehrdadi y M. Dales, “Diagnostic method for photovoltaic systems based on six layer detection algorithm,” *Electric Power Systems Research*, vol. 151, págs. 26-39, 2017.
- [25] A. A. Malek, M. A. Alias, F. A. Razak, M. S. M. Noorani, R. Mahmud y N. F. S. Zulkepli, “Persistent Homology-Based Machine Learning Method for Filtering and Classifying Mammographic Microcalcification Images in Early Cancer Detection,” *Cancers*, vol. 15, n.º 9, pág. 2606, 2023.

- [26] A. Maurya et al., “Hybrid Topological Data Analysis and Deep Learning for Basal Cell Carcinoma Diagnosis,” *Journal of Imaging Informatics in Medicine*, vol. 37, n.º 1, págs. 92-106, 2024.
- [27] W. Guo y A. G. Banerjee, “Identification of key features using topological data analysis for accurate prediction of manufacturing system outputs,” *Journal of Manufacturing Systems*, vol. 43, págs. 225-234, 2017.
- [28] Y. Wang, Y. Li, Y. Song, D. Xu y W. Zheng, “A Novel Fault Diagnosis Method Based on Topological Data Analysis,” en *2021 CAA Symposium on Fault Detection, Supervision, and Safety for Technical Processes (SAFEPROCESS)*, IEEE, 2021, págs. 1-6.
- [29] B. Wang, C. Lin, H. Inoue y M. Kanemaru, “Induction motor eccentricity fault detection and quantification using topological data analysis,” *IEEE Access*, 2024.
- [30] C. S. Solanki, *Solar photovoltaics: fundamentals, technologies and applications*. Phi learning pvt. Ltd., 2015.
- [31] V. Quaschnig, *Understanding renewable energy systems*. Routledge, 2014.
- [32] P. Rappaport, “The photovoltaic effect and its utilization,” *Solar Energy*, vol. 3, n.º 4, págs. 8-18, 1959, ISSN: 0038-092X. DOI: [https://doi.org/10.1016/0038-092X\(59\)90002-7](https://doi.org/10.1016/0038-092X(59)90002-7). dirección: <https://www.sciencedirect.com/science/article/pii/0038092X59900027>.
- [33] M. V. Damhare, B. Butey y S. Moharil, “Solar photovoltaic technology: A review of different types of solar cells and its future trends,” en *Journal of Physics: Conference Series*, IOP Publishing, vol. 1913, 2021, pág. 012053.
- [34] T. S. Gutiérrez Sepúlveda, “Análisis comparativo entre inversores y microinversores para la generación de energía solar fotovoltaica,” 2023. dirección: <https://hdl.handle.net/2445/194049>.
- [35] D. Spiers, “Chapter IIB-2 - Batteries in PV Systems,” en *Practical Handbook of Photovoltaics (Second Edition)*, A. McEvoy, T. Markvart y L. Castañer, eds., Second Edition, Boston: Academic Press, 2012, págs. 721-776, ISBN: 978-0-12-385934-1. DOI: <https://doi.org/10.1016/B978-0-12-385934-1.00022-2>. dirección: <https://www.sciencedirect.com/science/article/pii/B9780123859341000222>.
- [36] A. Hali e Y. Khelifi, “New translation method to STC of photovoltaic module characteristics: A comparison of conventional approaches and proposal of a novel method,” *Microelectronics Reliability*, vol. 154, pág. 115324, 2024, ISSN: 0026-2714. DOI: <https://doi.org/10.1016/j.microrel.2024.115324>. dirección: <https://www.sciencedirect.com/science/article/pii/S0026271424000040>.
- [37] I. Jamil, H. Lucheng, S. Habib, M. Aurangzeb, A. Ali y E. M. Ahmed, “Performance Ratio Analysis Based on Energy Production for Large-Scale Solar Plant,” *IEEE Access*, vol. 10, págs. 5715-5735, 2022. DOI: 10.1109/ACCESS.2022.3141755.
- [38] E. G. Novoa Guamán, “Manual de operacion y mantenimiento preventivo y correctivo para parques fotovoltaicos en el Ecuador,” B.S. thesis, Quito, EPN, 2015., 2015.
- [39] I. Rena, “Renewable power generation costs in 2022,” *International Renewable Energy Agency, Abu Dhabi*, 2023.
- [40] J. Y. Kim, J.-W. Lee, H. S. Jung, H. Shin y N.-G. Park, “High-efficiency perovskite solar cells,” *Chemical reviews*, vol. 120, n.º 15, págs. 7867-7918, 2020.

- [41] A. K. Shukla, K. Sudhakar y P. Baredar, “A comprehensive review on design of building integrated photovoltaic system,” *Energy and Buildings*, vol. 128, págs. 99-110, 2016, ISSN: 0378-7788. DOI: <https://doi.org/10.1016/j.enbuild.2016.06.077>. dirección: <https://www.sciencedirect.com/science/article/pii/S0378778816305667>.
- [42] M. Bragard, N. Soltan, S. Thomas y R. W. De Doncker, “The Balance of Renewable Sources and User Demands in Grids: Power Electronics for Modular Battery Energy Storage Systems,” *IEEE Transactions on Power Electronics*, vol. 25, n.º 12, págs. 3049-3056, 2010. DOI: 10.1109/TPEL.2010.2085455.
- [43] J. J. Moreno Escobar, O. Morales Matamoros, R. Tejeida Padilla, I. Lina Reyes y H. Quintana Espinosa, “A Comprehensive Review on Smart Grids: Challenges and Opportunities,” *Sensors*, vol. 21, n.º 21, 2021, ISSN: 1424-8220. DOI: 10.3390/s21216978. dirección: <https://www.mdpi.com/1424-8220/21/21/6978>.
- [44] S. E. SUNCORE. “Prueba de curva IV para paneles solares: ¿Qué son?” Consultado el 10 de septiembre de 2025. dirección: <https://suncore.com.mx/paneles-solares/prueba-de-curva-iv-para-paneles-solares-que-son/#:~:text=Componentes%5C%20de%5C%20la%5C%20curva%5C%20IV,que%5C%20el%5C%20panel%5C%20puede%5C%20entregar.>
- [45] T. Pei y X. Hao, “A Fault Detection Method for Photovoltaic Systems Based on Voltage and Current Observation and Evaluation,” *Energies*, vol. 12, n.º 9, 2019, ISSN: 1996-1073. DOI: 10.3390/en12091712. dirección: <https://www.mdpi.com/1996-1073/12/9/1712>.
- [46] M. Q. Baig, H. A. Khan y S. M. Ahsan, “Evaluation of solar module equivalent models under real operating conditions—A review,” *Journal of Renewable and Sustainable Energy*, vol. 12, n.º 1, 2020.
- [47] S. K. Firth, K. J. Lomas y S. J. Rees, “A simple model of PV system performance and its use in fault detection,” *Solar energy*, vol. 84, n.º 4, págs. 624-635, 2010.
- [48] E. Skoplaki y J. A. Palyvos, “On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations,” *Solar energy*, vol. 83, n.º 5, págs. 614-624, 2009.
- [49] I. M. Moreno-Garcia et al., “Real-time monitoring system for a utility-scale photovoltaic power plant,” *Sensors*, vol. 16, n.º 6, pág. 770, 2016.
- [50] A. Mellit y S. Kalogirou, “Assessment of machine learning and ensemble methods for fault diagnosis of photovoltaic systems,” *Renewable Energy*, vol. 184, págs. 1074-1090, 2022.
- [51] A. Phinikarides, G. Makrides, B. Zinsser, M. Schubert y G. E. Georgiou, “Analysis of photovoltaic system performance time series: Seasonality and performance loss,” *Renewable Energy*, vol. 77, págs. 51-63, 2015.
- [52] A. Michail et al., “A comprehensive review of unmanned aerial vehicle-based approaches to support photovoltaic plant diagnosis,” *Heliyon*, vol. 10, n.º 1, 2024.
- [53] G. Carlsson y M. Vejdemo-Johansson, *Topological data analysis with applications*. Cambridge University Press, 2021.
- [54] Ž. Virk, *Introduction to Persistent Homology*. Založba UL FRI, 2022.
- [55] N. A. Baas, G. E. Carlsson, G. Quick, M. Szymik y M. Thauale, *Topological data analysis*. Springer, 2020.

- [56] V. Pascucci, X. Tricoche, H. Hagen y J. Tierny, *Topological Methods in Data Analysis and Visualization: Theory, Algorithms, and Applications*. Springer Science & Business Media, 2010.
- [57] D. Attali, A. Lieutier y D. Salinas, “Vietoris-rips complexes also provide topologically correct reconstructions of sampled shapes,” en *Proceedings of the Twenty-Seventh Annual Symposium on Computational Geometry*, ép. SoCG '11, Paris, France: Association for Computing Machinery, 2011, págs. 491-500, ISBN: 9781450306829. DOI: 10.1145/1998196.1998276. dirección: <https://doi.org/10.1145/1998196.1998276>.
- [58] S. Gaukstad, “The Bottleneck distance and the Interleaving distance,” 2021.
- [59] C. Maria, J.-D. Boissonnat, M. Glisse y M. Yvinec, “The GUDHI Library: Simplicial Complexes and Persistent Homology,” en *Mathematical Software – ICMS 2014*, H. Hong y C. Yap, eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, págs. 167-174.
- [60] G. contributors, *Tutorials for Topological Data Analysis with the Gudhi Library*, GitHub repository, MIT License; 328 commits; fecha de acceso: 12 septiembre 2025, 2025. dirección: <https://github.com/GUDHI/TDA-tutorial>.
- [61] CONAHCYT, *Plataforma Nacional de Energía, Ambiente y Sociedad*, Revisado el 12-07-2025, 2022. dirección: <https://energia.conacyt.mx/planeas/electricidad/sistema-electrico-nacional>.
- [62] C. Dumitru y V. Maria, “Advantages and Disadvantages of Using Neural Networks for Predictions,” *Ovidius University Annals, Series Economic Sciences*, vol. 13, n.º 1, 2013.
- [63] Elegrow, *I-V Curve Analysis for Photovoltaic System*, Revisado el 2-08-2025, 2023. dirección: <https://elegrow.com/i-v-curve-analysis-for-photovoltaic-system/>.
- [64] M. M. Rahman, I. Khan y K. Alameh, “Potential measurement techniques for photovoltaic module failure diagnosis: A review,” *Renewable and Sustainable Energy Reviews*, vol. 151, pág. 111 532, 2021, ISSN: 1364-0321. DOI: <https://doi.org/10.1016/j.rser.2021.111532>. dirección: <https://www.sciencedirect.com/science/article/pii/S1364032121008108>.
- [65] S. R. Fahim, H. M. Hasanien, R. A. Turkey, S. H. E. A. Aleem y M. Calasan, “A Comprehensive Review of Photovoltaic Modules Models and Algorithms Used in Parameter Extraction,” *Energies*, vol. 15, n.º 23, 2022, ISSN: 1996-1073. DOI: 10.3390/en15238941. dirección: <https://www.mdpi.com/1996-1073/15/23/8941>.

Brayan Hernández Calvillo

El Problema del Diagnóstico de Fallas en Sistemas Fotovoltaicos basado en Análisis Topológico de Datos..pdf

 Universidad Michoacana de San Nicolás de Hidalgo

Detalles del documento

Identificador de la entrega

trn:oid:::3117:502404161

Fecha de entrega

23 sep 2025, 7:40 a.m. GMT-6

Fecha de descarga

23 sep 2025, 7:51 a.m. GMT-6

Nombre del archivo

El Problema del Diagnóstico de Fallas en Sistemas Fotovoltaicos basado en Análisis Topológico d....pdf

Tamaño del archivo

8.0 MB

103 páginas

33.639 palabras

171.289 caracteres

Formato de Declaración de Originalidad y Uso de Inteligencia Artificial

Coordinación General de Estudios de Posgrado
Universidad Michoacana de San Nicolás de Hidalgo



A quien corresponda,

Por este medio, quien abajo firma, bajo protesta de decir verdad, declara lo siguiente:

- Que presenta para revisión de originalidad el manuscrito cuyos detalles se especifican abajo.
- Que todas las fuentes consultadas para la elaboración del manuscrito están debidamente identificadas dentro del cuerpo del texto, e incluidas en la lista de referencias.
- Que, en caso de haber usado un sistema de inteligencia artificial, en cualquier etapa del desarrollo de su trabajo, lo ha especificado en la tabla que se encuentra en este documento.
- Que conoce la normativa de la Universidad Michoacana de San Nicolás de Hidalgo, en particular los Incisos IX y XII del artículo 85, y los artículos 88 y 101 del Estatuto Universitario de la UMSNH, además del transitorio tercero del Reglamento General para los Estudios de Posgrado de la UMSNH.

Datos del manuscrito que se presenta a revisión

Programa educativo	Maestría en Ciencias de la Ingeniería Eléctrica	
Título del trabajo	El Problema del Diagnóstico de Fallas en Sistemas Fotovoltaicos basado en Análisis Topológico de Datos	
	Nombre	Correo electrónico
Autor/es	Brayan Hernández Calvillo	brayan.hernandez@umich.mx
Director	Juan Anzurez Marín	juan.anzurez@umich.mx
Codirector	Juan Ahtziri González Lemus	ahtziri.lemus@umich.mx
Coordinador del programa	Norberto García Barriga	norberto.garcia@umich.mx

Uso de Inteligencia Artificial

Rubro	Uso (sí/no)	Descripción
Asistencia en la redacción	No	Ninguna
Traducción al español	No	Ninguna
Traducción a otra lengua	No	Ninguna

Formato de Declaración de Originalidad y Uso de Inteligencia Artificial

Coordinación General de Estudios de Posgrado
Universidad Michoacana de San Nicolás de Hidalgo



Uso de Inteligencia Artificial		
Rubro	Uso (sí/no)	Descripción
Revisión y corrección de estilo	Sí	Ayuda en la mejora de la redacción para que el texto sea formal y coherente, corrigiendo el estilo, la fluidez y la estructura
Análisis de datos	No	Ninguna
Búsqueda y organización de información	No	Ninguna
Formateo de las referencias bibliográficas	No	Ninguna
Generación de contenido multimedia	Si	Las imágenes del Capítulo 2 fueron generadas con IA para evitar situaciones de copyright
Otro	No	Ninguna

Datos del solicitante	
Nombre y firma	Brayan Hernández Calvillo
Lugar y fecha	Morelia, Michoacán a 19 de septiembre del 2025