

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

DISEÑO FOTOVOLTAICO PARA EDIFICACIONES
CONSIDERANDO APRENDIZAJE DE MÁQUINAS,
CONECTADO A UN DASHBOARD DE MONITOREO EN
TIEMPO REAL

PROYECTO INTEGRADOR

Previo la obtención del Título de:

Ingeniero en electricidad

Presentado por:

Homero Rodrigo Ojeda Guevara

Kimberlyn Stephanie Solano Villao

GUAYAQUIL - ECUADOR

Año: 2022

DEDICATORIA

El presente proyecto lo dedico a: mi fuerte, generoso y amoroso padre Rodrigo Ojeda, maestro de vida y amigo, a mi cariñosa y amorosa madre Kelly Guevara, a mi querido hermano Lucas y a mi pequeño amado sobrino Johan, a mi familia. También a mis amigos, pero especialmente: Jocelyne, Stephanie, Franklin, Gabriel, Víctor, Azael, Edwin; en quienes siempre encontré las virtudes de lo que es ser un buen amigo y me sentí cobijado por su cariño.

Homero Ojeda

Este proyecto va dirigido a mis amados padres Luis y Angela por siempre creer en mí, a mis hermanos Luis Xavier y Christian por enseñarme el valor de la perseverancia y dedicación, a mi abuelita Piedad que, aunque no esté presente físicamente, estoy segura que desde el cielo me acompaña, a toda mi familia, a mi amigo Homero, y demás amigos que me dieron su apoyo incondicional durante toda esta etapa.

Stephanie Solano

AGRADECIMIENTOS

Me gustaría extender mi agradecimiento a quienes hicieron materialmente posible este proyecto: Mis amados padres, y a mi querida amiga y compañera de tesis Stephanie Solano, ellos son quienes merecen mi reconocimiento, por su amor y paciencia. Profundamente agradecido también con: PhD. Miguel Torres, PhD José Córdova y M.sc. Ángel Recalde, espero que los resultados del trabajo justifiquen la confianza depositada.

Homero Ojeda

Estoy profundamente agradecida con Dios, mis queridos padres, hermanos y a mi amigo y compañero de tesis Homero, así como también con Matthew y César quienes aportaron con los valores experimentales del proyecto. De igual forma, con PhD. Miguel Torres y M.sc. Ángel Recalde por su buena disposición a la hora de revisar el trabajo.

Stephanie Solano

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, nos corresponde conforme al reglamento de propiedad intelectual de la institución; *Homero Rodrigo Ojeda Guevara y Kimberlyn Stephanie Solano Villao* damos nuestro consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”

Homero Rodrigo Ojeda
Guevara

Kimberlyn Stephanie
Solano Villao

EVALUADORES

A handwritten signature in blue ink, appearing to read 'Miguel Angel Torres', is positioned above a dotted line. The signature is stylized and cursive.

.....
Nombre del Profesor

DR. ANGEL ANDRES RECALDE

.....
Nombre del Profesor

Dr. MIGUEL ANGEL TORRES

RESUMEN

Durante bastante tiempo el gas natural y los derivados del petróleo han sido las principales fuentes de energía en todo el mundo, sin embargo, estos recursos además de ser agotables también generan grandes emisiones de gases contaminantes que causan el efecto invernadero, los cuales son perjudiciales para el medio ambiente. Por lo tanto, a través de los años se han realizado búsquedas de energías alternativas que permitan disminuir recursos energéticos y que sean amigables con el medio ambiente. En base a lo expuesto, el presente proyecto de titulación tiene como objetivo proponer un diseño mejorado de la planta solar ubicada en el edificio 6A de la Escuela Superior Politécnica del Litoral, utilizando técnicas *machine learning* para el tratamiento de los parámetros meteorológicos de entrada. El diseño incluye una conexión a un dashboard que muestra los resultados relacionados a los datos de energía que posee el edificio, previo a ello fue necesario realizar la evaluación del estado actual de dicha planta solar. La metodología buscó condensar la información relacionada a la energía solar, los algoritmos de clasificación de datos de aprendizaje automático utilizados en este trabajo tales como *k-means* y *Gaussian Mixture Models*. Paralelo a esto, se realizó cálculos de la potencia y temperatura del módulo fotovoltaico para la comparación de las técnicas descritas anteriormente.

Finalmente, se presentó el estado actual de la instalación, y una propuesta de mejora en el diseño existente, así como la respectiva discusión de los hallazgos y conclusiones del trabajo.

Palabras claves: Energía Solar, K-means, Gaussian Mixture Models, paneles fotovoltaicos.

ABSTRACT

For a long time, natural gas and petroleum derivatives have been the main sources of energy around the world. However, these resources, in addition to being exhaustible, also generate large emissions of polluting gases that cause the greenhouse effect, which is harmful to the environment. Therefore, through the years many researchers have been indagating for alternatives that allow energy resources to be reduced and also friendly to the environment.

Based on the above, this degree project aims to propose an improved design for the solar plant located on building 6A of the Escuela Superior Politécnica del Litoral, using machine learning techniques for the treatment of input meteorological parameters. The design includes a connection to a dashboard that shows the results related to the energy data that the building has, prior to this it was necessary to evaluate the current state of the solar plant.

The methodology sought to condense the information related to solar energy, the automatic learning data classification algorithms used in this work such as k-means and Gaussian Mixture Models. Parallel to this, calculations of the power and temperature of the photovoltaic module were carried out to compare the techniques described above.

Finally, the current state of the installation was presented, and a proposal to improve the existing design, as well as the respective discussion of the findings and conclusions of the work.

Keywords: Solar energy, K-means, Gaussian Mixture Models, photovoltaic panels.

ÍNDICE GENERAL

RESUMEN.....	I
<i>ABSTRACT</i>	II
ÍNDICE GENERAL.....	III
ÍNDICE DE FIGURAS.....	V
ÍNDICE DE TABLAS.....	VI
ABREVIATURAS.....	VIII
SIMBOLOGÍA.....	IX
CAPÍTULO 1.....	1
1. Introducción.....	1
1.1 Descripción del problema.....	2
1.2 Justificación del problema.....	3
1.3 Objetivos.....	4
1.3.1 Objetivo General.....	4
1.3.2 Objetivos Específicos.....	4
1.4 Marco teórico.....	5
1.4.1 Energía Solar.....	5
1.4.2 Irradiancia solar.....	5
1.4.3 Radiación solar.....	6
1.4.4 Sistema Fotovoltaico.....	6
1.4.5 Inversor.....	7
1.4.6 Bases de datos meteorológicos.....	8
1.4.7 <i>K-Means</i>	9
1.4.8 <i>Gaussian Mixture Models (GMM)</i>	10
1.4.9 Cálculo de Potencia y Energía de una instalación fotovoltaica.....	11
1.4.10 Scikit-Learn.....	13

CAPÍTULO 2.....	14
2. Metodología	14
2.1 Radiación Solar y Factor de planta.....	14
2.1.1 Aplicación de K-means.....	15
2.1.2 Aplicación de GMM.	17
2.1.3 Cálculo de potencia DC del PV	19
2.1.4 Estimación de la producción de energía y del consumo.	21
2.1.5 Inversor y selecciónSelección del número de PV.....	22
2.1.6 Métricas utilizadas para evaluación del sistema y comparación de las técnicas de ML	24
2.1.7 Monitoreo en tiempo real y Dashboard	26
2.2. Caso de estudio	28
2.2.1. Condiciones climáticas.....	29
2.1.8 Cargas del edificio.....	29
2.2.2. Datos meteorológicos	31
2.2.3. Solcast	32
2.2.4. <i>Adquisición de datos</i>	32
CAPÍTULO 3.....	38
3. Resultados y análisis	38
3.1 Recurso Solar	38
3.2 Resultados K-means.....	41
3.3 Resultados GMM	43
3.4 Comparación de los 3 enfoques (k-means, GMM, media) en la estimación de potencia.	44
3.5 Dimensionamiento de la Instalación fotovoltaica utilizando el mejor enfoque	

3.6	Dashboard.....	48
3.6.1	Segmentación de mercado	49
3.7	Discusión Final.....	50
CAPÍTULO 4.....		52
4.	conclusiones y recomendaciones	52
4.1	CONCLUSIONES	52
4.2	Recomendaciones.....	53
5.	Bibliografía	55
6.	APÉNDICES	61
6.1	Código K-means	61
6.2	GMM.....	63
6.3	Mediciones del Inversor Solis-1P10K-4G-US	65

ÍNDICE DE FIGURAS

Figura 1.1	pseudocódigo del algoritmo k-means (Alpaydm, 2010)	9
Figura 2.1	Diagrama de flujo de K-means para el agrupamiento de Ta, GHI, Vf fuente: autores.....	16
Figura 2.2	Diagrama de flujo de la aplicación del modelo de GMM. Fuente: autores	19
Figura 2.3	Esquema de una planta solar conectado a la red. Fuente: [54].....	22
Figura 2.4	Conexión realizada para el monitoreo en tiempo real.....	27
Figura 2.5	Implementación utilizando un Arduino.....	27
Figura 2.6	Localización geográfica del caso de estudio. Fuente: Autores	28
Figura 2.7	Edificio 6A (Rectorado). Fuente: Autores.....	29
Figura 2.8	División del consumo energético de Rectorado.....	31
Figura 2.9	Ubicación caso de estudio	33
Figura 2.10	Implementación de la metodología	34

Figura 2.11 Curva de corriente vs voltaje de los paneles existentes en el edificio 6A. Fuente: Autores	35
Figura 2.12 Datos de los paneles BSM280P-60. Fuente: Autores.....	35
Figura 2.13 Datos del inversor Solis 1P10K-4G-US	36
Figura 3.1 Distribución de datos de temperatura del aire en verano.....	38
Figura 3.2 Distribución de datos de velocidad de viento en verano. Fuente: Autores	40
Figura 3.3 Distribución de datos de GHI en invierno. Fuente: Autores	40
Figura 3.4 Distribución de datos de velocidad de viento en invierno. Fuente: Autores	41
Figura 3.5 Conglomerados obtenidos mediante k-means.....	42
Figura 3.6 Resultados de la aplicación de BIC	43
Figura 3.7 Dashboard	48
Figura 3.8 Número de árboles a plantar presentado en el dashboard	49

ÍNDICE DE TABLAS

Tabla 1.1 Base de datos sobre el territorio global [23].....	8
Tabla 1.2 Algunas Bases de Datos satelitales [23]	9
Tabla 2.1 Datos a mostrar en el Dashboard	28
Tabla 2.2 Equipos del edificio 6 ^a [60].....	30
Tabla 2.3 - Capacidad instalada por las cargas enfriamiento del edificio [60]	30
Tabla 2.4 – Datos obtenidos de solcast	33
Tabla 3.1 Resultados valor más alto k-means	42
Tabla 3.2 Resultados valor de más baja radiación k-means.....	42
Tabla 3.3 Resultados valor más alto GMM	44
Tabla 3.4 Resultados valor más bajo GMM	44
Tabla 3.5 Coeficientes de la ecuación de temperatura del módulo.....	44
Tabla 3.6 Potencia y temperatura del módulo PV de día soleado	45
Tabla 3.7 Potencia y temperatura del módulo PV de día nublado	45
Tabla 3.8 Error relativo en la estimación de la potencia de salida (día soleado)	45

Tabla 3.9 Error relativo en la estimación de la potencia de salida (día nublado)	45
Tabla 3.10 Índices de evaluación utilizados en la evaluación de la instalación solar (día soleado)	46
Tabla 3.11 Diseño existente vs diseño propuesto.....	47
Tabla 6.1 Acceso a las mediciones del inversor	65

ABREVIATURAS

IPCC	Grupo Intergubernamental de Expertos sobre el Cambio Climático
ESPOL	Escuela Superior Politécnica del Litoral
ML	<i>Machine Learning</i>
CC	Corriente Continua
AC	Corriente Alterna
PV	Módulo fotovoltaico
AM	Masa de aire
MPP	Punto de máxima potencia
NASA	La Administración Nacional de Aeronáutica y el Espacio
GMM	Modelo de mezcla Gaussiana
EM	Estimación máxima
MAP	Estimación de un modelo previo
Taire	Temperatura del aire
GHI	Irradiancia Horizontal Global
Vf	Velocidad de viento
BIC	Criterio de Información Bayesiana
WRDC	World Radiation Data Center
NIWE	Solar Radiation Resource Assessment
NSRD	National Solar Radiation Database
DWD	Deutscher Wetterdienst
BSRN	Baseline Surface Radiation Network
IDMP	International Daylight Measurement Program
GEBA	Global Energy Balance Archive
LED	Light-emitting diode

SIMBOLOGÍA

°C	Grados centígrados
kW	Kilovatio
TW	Terawatt
He	Helio
H	Hidrógeno
M	Metro
W	Watt
s	Segundo
km	Kilómetro
BTU	Unidad térmica británica

CAPÍTULO 1

1. INTRODUCCIÓN

El Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC) en su último informe dejó constancia de como las acciones de las personas han provocado un aumento del calentamiento global, y estimó este incremento en 1,0 °C, teniendo como referencia niveles que le denomina preindustriales (1851-1901) con un margen de probabilidad de 0,8 a 1,2°C [1]. Dentro de las consecuencias que genera esto según los expertos se tiene: aumento del nivel del mar, calentamiento del mar, disminución de las capas de hielo en la Antártida; tal como aparece citado en [2].

Existe consenso de que el CO₂ emitido de las actividades que realiza el ser humano impactan en el calentamiento global, aunque existe discusión respecto a en qué medida afecta, no obstante los países constantemente están desarrollando políticas públicas con el fin de mitigar las emisiones de CO₂, uno de los campos de mayor interés ha sido la transición energética hacia fuentes de energía renovables, esto ha ocurrido en mayor proporción en Europa, no tanto así en América, pero en virtud de la situación que se atravesó a raíz de la pandemia, las tendencias a invertir en este tipo de fuentes de energía experimente un alza.

En el contexto antes relatado, Ecuador no se ha quedado atrás, en el año 2021 a través de la resolución ministerial Nro. MERNNR-VEER-2021-0008, realizó modificaciones en la planificación del sistema eléctrico y se proyectó a captar hasta 2200\$ millones en inversión en estas fuentes de energía no contaminante, en un plazo de 10 años [3], esto se inscribe en lo acordado por el país en el “Acuerdo de París” de reducir las emisiones de los gases contaminantes que causan el calentamiento global. La proyección que hace el gobierno es que al año 2025 se puedan reducir un valor promedio de 6 millones de toneladas de CO₂ [4]. La consecución de estos objetivos no solo queda reducido a los organismos gubernamentales, resulta indispensable la participación de la academia, en el caso de la Escuela Superior Politécnica del Litoral (ESPOL) a partir del año 2018, implementó un programa denominado Sostenibilidad, con el cual busca implementar acciones que contribuyan a generar procesos que permitan un mejor aprovechamiento de los recursos dentro del campus, dentro de este programa en el año 2019 se logró la instalación de una planta solar de 50 KW [5], dicha planta cuenta con 6 inversores a los cuales hay conectados 30 paneles en cada uno no obstante, dicha instalación no cuenta

con un sistema de monitoreo completo, las mediciones se las obtiene del inversor de forma manual y por lo tanto no existe un registro pormenorizado de la cantidad de energía que está ahorrando a la red pública, cabe acotar de igual forma que dicha planta está interconectada a la red dado que no cubre toda la demanda del edificio de 6A, que es la denominación en el plano de la instalación civil donde está ubicado.

El presente trabajo de titulación tuvo como finalidad, proponer un diseño mejorado de la instalación solar antes mencionada, utilizando dos técnicas de *Machine Learning(ML)*, conectado a un dashboard donde se presentó resultados de algunos parámetros que normalmente sirven para la evaluación de estas plantas solares, partiendo de esta idea se formuló de forma teórica 2 distintas pantallas a mostrar en función del cliente.

1.1 Descripción del problema

La variabilidad de los parámetros meteorológicos, es una de las consideraciones realmente importante a la hora de aprovechar el recurso solar, ya que tiene relación con el valor de potencia máxima (PM) de cada módulo solar. Dichos módulos vienen con este parámetro PM especificados por la empresa que los fabrica, los ingenieros que se encargan del diseño suelen usar este valor como referencia para establecer comparaciones entre las distintas tecnologías [6]. La dificultad principal de lo mencionado en el párrafo anterior es que dichos valores ofrecidos por los fabricantes suelen estar alejados del funcionamiento real de los paneles, y esto se debe a la variabilidad de la irradiancia, temperatura y la masa de aire. [7]

Por otro lado la implementación de la tecnología fotovoltaica tiene una incidencia en la autosuficiencia energética de las obras civiles donde se consume electricidad, porque la energía generada por la misma vivienda sirve para su autoconsumo, esto permite mitigar problemas vinculados al impacto del ser humano en el ambiente y también ayuda en la economía, particularmente con la reducción de CO₂ [8], sin embargo para una mayor penetración de energías limpias, es indispensable mejorar la eficiencia de dichas instalaciones, esto no solo se reduce a la implementación de tecnologías renovables sino también a la gestión de la demanda, en esto último entra la promoción de patrones de ahorro de energía [9].

En este contexto, la planta solar de 50 kW ubicada en la Escuela Superior Politécnica del Litoral, Campus Prosperina, en el edificio de 6A, adolece un control y evaluación de

parámetros de autoconsumo y autosuficiencia energética, ni hay un monitoreo en tiempo real de la generación fotovoltaica, dicho control es realizado por un técnico de manera empírica, ya que este recoge los datos que brinda el inversor, para ello no utiliza una metodología documentada, ni tampoco existe una pantalla que le permita al usuario que no es técnico conocer cuanta energía está ahorrando, o cuan eficiente están siendo los paneles, ni tampoco la conversión de toneladas de CO₂ que se dejan de emitir con el uso de los módulos. Adicional a esto la instalación se la efectuó exclusivamente siguiendo lo establecido por el fabricante, y tal como se abordó al inicio de este apartado estos valores varían dependiendo de las condiciones ambientales del lugar. En otras palabras, no se tiene conocimiento de cuan eficiente está siendo la instalación, ni tampoco se propuso un diseño que mitigue el impacto de la variabilidad del recurso meteorológico, y finalmente tampoco hay una gestión de la demanda debido a que los usuarios desconocen cuanta energía están ahorrando y en qué hora se está generando mayor potencia en la planta

1.2 Justificación del problema

En año 2015, Ecuador junto 194 países más se juntaron para el acuerdo de París, el mismo que fue finalmente suscrito por el país en el año 2016 [10], aquí se comprometió a reducir las emisiones de CO₂, dentro de este marco, la transición energética tiene un rol principal, bien es cierto que en esta región la matriz energética está principalmente conformada por energía renovables con el 72,58%, de lo cual la fuente fotovoltaica, que es de interés de este trabajo, representa solo el 0.18% de ese valor [11]. Como se puede observar en las cifras antes mencionadas, el mayor recurso es hídrico, no obstante, la lejanía de las centrales donde se aprovecha este recurso representa una serie de dificultades para el sistema de potencia, principalmente por las pérdidas de potencia en el sistema de transmisión y todo los costos asociados a la construcción de las estructuras mecánicas para el traslado de la electricidad desde lugares tan lejanos. [12].

Así mismo Ecuador está ubicado en una región donde tiene posibilidades de aprovechar mejor este recurso, de hecho el valor de la irradiación oscila alrededor de la media diaria de 4.1 kWh/m² [13], sin embargo la penetración de las fuentes de recurso solar sigue siendo bajo, una forma de contribuir a una mayor utilización, es la propuestas de diseño que mejoren el dimensionamiento y tengan en cuenta las métricas relacionadas con el

autoconsumo y autoeficiencia, de tal forma que resulte atractiva tanto al usuario como alguna empresa interesada en la comercialización del servicio.

En este sentido resulta pertinente, partir de una instalación ya hecha, como es el caso de este proyecto, y mejorar la implementación ya realizada, teniendo en cuenta la variabilidad del recurso meteorológico. La utilización de *machine learning* para el tratamiento de la variabilidad de la irradiancia, temperatura y velocidad de aire, implica un mejor abordaje de este problema, y es de mucho interés dado que se puede identificar patrones que permitan un modelado no tan alejado del funcionamiento real. Por otro lado, la propuesta da un *dashboard* permite gestionar mejor una parte incontrolable, como es la demanda, esto implica un cambio de filosofía donde el usuario participa de forma consciente en la transición de recursos de fuentes fósiles a recursos renovables no convencionales, además da apertura a estudios de otros campos sean estos: sociología o ciencias económicas.

Estudios previos evaluando la planta solar del edificio 6A de la ESPOL, no existen documentalmente, esto significa que el presente trabajo tiene mayor pertinencia no solo a nivel de propuesta, sino también a dejar documentada la evaluación realizada del funcionamiento de la instalación solar, de los datos monitoreados por el inversor en tiempo real. Al ampliar a una propuesta de comercialización 2 tipos de clientes, a los cuales se les mostrará resultados acordes a unas determinadas características, el beneficio se extendería más allá de la ESPOL.

1.3 Objetivos

1.3.1 Objetivo General

Proponer un diseño mejorado de la planta solar ubicada en la ESPOL, edificio de rectorado 6A, utilizando técnicas *machine learning* para el tratamiento de los parámetros meteorológicos de entrada, conectado a un *dashboard* donde se muestren los resultados relacionados con la energía ahorrada

1.3.2 Objetivos Específicos

- Evaluar el autoconsumo y la autoeficiencia de la planta solar utilizando datos en tiempo real.

- Determinar patrones de agrupamiento de los parámetros meteorológicos: Irradiancia, temperatura y velocidad del viento.
- Calcular los valores de energía ahorrada y toneladas de CO2 no emitidas a la atmósfera, que se mostrarán en el *dashboard*.
- Proponer 2 tipos de pantallas a mostrar en *dashboard*, en función de 2 diferentes tipos de cliente.
- Comparar las 2 técnicas de ML utilizando métricas con los valores reales medidos en la planta solar del edificio 6 A de la ESPOL.

1.4 Marco teórico

En la presente sección se detallan algunos fundamentos teóricos relacionados con el presente trabajo de titulación.

1.4.1 Energía Solar

Es la principal fuente de energía del planeta y a partir de esta se genera todos los demás tipos de energía renovable, la energía del sol que entra por la atmósfera cuadruplica el requerimiento actual de energía en el globo terráqueo la cual es 15 TW [14]. La fuente es el sol, aquí ocurre una serie de interacciones nucleares las cuales causan la conversión del elemento químico He en H, esta energía es transportada mediante ondas electromagnética, en física esto se le denomina fotones, visto de manera macro, la energía vendría a constituir un gran colector de energía proveniente del sol, las formas que los seres vivos aprovechan este recurso son diversas, algunas pueden ser de forma directa como la energía fotovoltaica mientras que también existen formas indirectas en que se utiliza el recurso [15].

1.4.2 Irradiancia solar

La irradiancia solar sirve para predecir el rendimiento de un sistema de conversión solar, y está también relacionado con la estabilidad del funcionamiento en la integración con la red eléctrica, además permite estimar la cantidad de energía eléctrica que genera la tecnología utilizada [16]. Hoy en día los procedimientos de calificación energética o predicción del rendimiento deben tener en consideración cuanta irradiancia está efectivamente disponible para los módulos y sistemas fotovoltaicos ya que es de vital

importancia saber cuál es el impacto energético anual y mensual de la irradiancia espectral variable debido a que esto afecta al rendimiento de los módulos fotovoltaicos (PV).

1.4.3 Radiación solar

La radiación solar se la puede entender como una densidad de potencia instantánea, en la noche tiene un valor de $0 \frac{kW}{m^2}$ mientras que en el día oscila alrededor de $1 \frac{kW}{m^2}$, el instrumento que se utiliza para tomar mediciones de forma intermitente a lo largo del día y saber los valores de radiación se denomina piranómetro, aunque también se la concibe como la energía que transmite el sol [17]. Una dificultad siempre presente en el estudio de la radiación es su naturaleza estadística, lo que ha conllevado la utilización de diferentes modelos computacionales robustos. Hay que tener en cuenta que el impacto de la radiación es amplio, porque en la oferta incide en la generación de electricidad y en la demanda afecta en los patrones culturales de consumo, esto ocurre porque dependiendo de la radiación y la temperatura, los usuarios utilizan refrigeración y calefacción, para una discusión más detallada de los modelos sofisticados utilizados para predecir la radiación revisar [18].

1.4.4 Sistema Fotovoltaico.

Son dispositivos que emplean principalmente materiales semiconductores con la finalidad de realizar la conversión de energía solar en energía eléctrica, y el principio físico bajo el cual operan es el efecto fotoeléctrico, los diseños han ido evolucionando en el tiempo, actualmente existen principalmente de inclinación fija o con seguidores del sol, de igual manera la integración puede ser de forma centralizada o distribuida, dependiendo de lo antes mencionado, la cantidad de radiación tendrá un impacto diferente, razón por la cual los técnicos que evalúan esta tecnología prestan principal atención en el histórico de energía solar para un estudio más profundo revisar [16].

El fenómeno fotovoltaico ocurre porque los fotones presentes en las ondas electromagnéticas transmiten una cantidad de paquetes de energía, la cual puede ser modelada por [19]:

$$Ef = h * \frac{c}{\lambda} \quad (1.1)$$

Donde h es la constante de Planck cuyo valor es $6,63 \times 10^{-34}$, y c la velocidad de la luz, y λ es la longitud de onda en metros. Las celdas solares pueden ser representadas mediante sus parámetros eléctricos de voltaje y corriente como salida, y parámetros de entrada meteorológicos como irradiancia y temperatura, los fabricantes normalmente proporcionan la información en la placa, dichos valores son obtenidos en condiciones de prueba estándar, la cual establece que el valor de irradiancia es de 1000 W/m^2 , el valor de temperatura ambiente de: 25°C y de masa de aire (AM) de 1.5 [20]. El conjunto de celdas fotovoltaicas conectadas y encapsuladas es lo que se comercializa y se le denomina panel o módulo, esto genera corriente continua, por lo que necesita un inversor para convertir a corriente alterna [21].

1.4.5 Inversor

Es un componente de electrónica de potencia que convierte la corriente continua en corriente alterna, una de las denominaciones que se utiliza para abreviar es DC-AC, además toda planta solar, tiene una etapa posterior al inversor donde se hace el ajuste de la frecuencia y de igual forma el voltaje, para sincronizarlo con la red [22], existen varias configuraciones en los cuales se los utiliza [21]:

- Sistemas conectados a la red
- Sistemas autónomos con baterías recargables
- Sistema de bombeo sin acumuladores

El inversor es uno de los elementos mas importantes de una instalación solar, en el caso de los sistemas conectados a la red determina el nivel de tensión del lado CC, aquí existe la interconexión, que no es otra cosa que conectar en paralelo a la red, algunos vienen aquí con un control inteligente que sirve para seguimiento del punto de máxima potencia (MPP), con el fin de ajustar a las condiciones más optimas de trabajo, los inversores modernos poseen un registrador de datos de la planta.

En el caso de los inversores para sistemas autónomos, tienen que poseer un valor de eficiencia alto, y la salida debe tener bajas fluctuaciones de tal forma que tenga similitud con la onda sinusoidal pero también existen dispositivos que la señal de salida es una onda rectangular.

1.4.6 Bases de datos meteorológicos

Las bases de datos donde se recopila información de los datos meteorológicos de una determinada área, son fundamentales para el estudio del recurso solar y el dimensionamiento de las instalaciones fotovoltaicas [22]. No son homogéneas ya que varían en los métodos utilizados para obtener las mediciones, las resoluciones temporales y espaciales, pero existen varias estaciones encargadas del monitoreo del recurso expandidas a lo largo del planeta. A continuación, se resume en la Tabla 1.1 de las estaciones alrededor del globo que realizan estudios del recurso solar. Así mismo en la Tabla 1.2 se muestran las bases de datos satelitales, esta información utiliza especificaciones iguales a las de la NASA, la cual también se incluye dentro de esta tabla.

Tabla 1.1 Base de datos sobre el territorio global [23]

Nombre	Área	Resolución temporal
<i>World Radiation Data Center (WRDC)</i>	Mundial	Diario, valores desigualmente distribuidos
<i>Solar Radiation Resource Assessment (NIWE)</i>	India	Datos de 1 minuto
<i>National Solar Radiation Database (NSRD)</i>	EE. UU	Promedios por hora
<i>Deutscher Wetterdienst (DWD)</i>	Alemania	Promedios por hora
<i>Baseline Surface Radiation Network (BSRN)</i>	Mundial	Intervalos de 1-5 min
<i>Global Energy Balance Archive (GEBA)</i>	Mundial	Promedios mensuales
<i>International Daylight Measurement Program (IDMP)</i>	Mundial	Intervalos de 1-5 min

Tabla 1.2 Algunas Bases de Datos satelitales [23]

Nombre	Área	Resolución temporal	Parámetros
3TIER	Europa, Oceanía, África	Cada hora	G, D, DNI
EnMetsol	África, Europa	15min/ 1hora	G, D, DNI, espectro
HelioClim-1	África, Europa	Diariamente	G, DNI
NASA SSE reléase 6.0	Mundial	Cada 3 horas	G, D, DNI, nubes
Satel-Light	Europa	30 min	G, B, D, iluminancia
SolarGis	Mundial	15 min	G, D, DNI
Solemi	África, Europa, Asia	Cada hora	G, DNI

1.4.7 K-Means

Es un algoritmo bastante usado para agrupar, la filosofía del algoritmo es encontrar centros conglomerados en cuyo agrupamiento estén representados ciertas regiones de interés de un conjunto de datos, en otras palabras, lo que hace es asignar a cada elemento del conjunto de datos, a un centro más cercano para luego establecer a este punto como la media de los elementos que le han sido asignados [24]. Un pseudocódigo del algoritmo k-medias se puede visualizar en la Figura 1.1.

```

Initialize  $\mathbf{m}_i, i = 1, \dots, k$ , for example, to k random  $x^t$ 
Repeat
  For all  $x^t \in X$ 
    
$$b_i^t \leftarrow \begin{cases} 1 & \text{if } \|x^t - \mathbf{m}_i\| = \min_j \|x^t - \mathbf{m}_j\| \\ 0 & \text{otherwise} \end{cases}$$

  For all  $\mathbf{m}_i, i = 1, \dots, k$ 
    
$$\mathbf{m}_i \leftarrow \left( \sum_t b_i^t x^t \right) / \left( \sum_t b_i^t \right)$$

  Until  $\mathbf{m}_i$  converge

```

Figura 1.1 pseudocódigo del algoritmo k-means [25]

Los vectores que mejor minimizan el error de b_i^t , teniendo en cuenta que esto depende del valor de m_i , lo cual no existe posibilidad de resolver de forma analítica, esta es la razón por la cual se utiliza iteraciones, donde primero se inicializa el valor de m_i aleatoriamente, con este valor lo siguiente que se hace según el pseudocódigo de la Figura 1.1 es calcular el valor de b_i^t , para todo vector x^t , aquí existen 2 posibilidades si este valor es 1, se puede entender que x^t pertenece al grupo de m_i [25].

1.4.8 Gaussian Mixture Models (GMM)

Un modelo de mezcla Gaussiana (GMM) es un algoritmo de probabilidad paramétrica que está representado como una suma ponderada de las densidades de los componentes gaussianos. Los parámetros GMM se estiman con los datos de entrenamiento por medio de un algoritmo iterativo de estimación máxima (EM) o (MAP) la estimación de un modelo previo [26]. Este modelo normalmente es utilizado en el aprendizaje automático o reconocimiento de imágenes ya que su tarea es distribuir y crear grupos dependiendo de los parámetros propios de cada grupo. Para obtener la suma ponderada de M densidades gaussianas se utiliza la ecuación (1.2) mostrada a continuación.

$$p(x | \lambda) = \sum_{i=1}^M w_i g(x | \mu_i, \Sigma_i) \quad (1.2)$$

Donde x es un vector de datos de valor continuo y w_i , M , son las densidades gaussianas de los componentes.

El método EM es un método general para encontrar la estimación de máxima verosimilitud de los parámetros de una distribución subyacente a partir de un conjunto de datos dado cuando los datos están incompletos o faltan valores. [27]. Por otro lado, el método de estimación de parámetros de máxima verosimilitud es proceso que consiste en calcular las estimaciones de las estadísticas de los datos de entrenamiento y combinarlas con las estadísticas antiguas utilizando coeficientes de mezcla dependiente de los datos.

1.4.9 Cálculo de Potencia y Energía de una instalación fotovoltaica

Es necesario conocer la potencia en condiciones estándar de un módulo fotovoltaico para determinar la energía que este producirá, su viabilidad económica y su funcionamiento. Para calcular la energía que genera una instalación fotovoltaica se la realiza de manera indirecta a la red o en términos energéticos.

El primer método consiste en calcular la potencia proporcionada del sistema como primer paso, para luego calcular la energía integrando la potencia en lapsos de tiempo diarios, mensuales o anuales. El segundo método se basa en modelos de circuitos equivalentes que poseen diodos en el que se calcula la potencia máxima a partir de una curva I-V, esta curva la puede proveer el fabricante o se la puede obtener de manera experimental por medio de redes neuronales [28]. Este método proporciona valores de potencia muy precisos, sin embargo, necesitan de un entrenamiento preliminar de la red y un conjunto de datos de prueba de entrada. El tercer método consiste en calcular la potencia máxima del módulo fotovoltaico de manera directa utilizando ecuaciones con parámetros meteorológicos y características eléctricas brindadas por el fabricante.

Existen otros procesos como el método [29] que consiste en calcular la potencia por medio de los parámetros eléctricos suministrados por el fabricante como se muestra en la ecuación (1.3)

$$P_M = P_M^{STC} * \frac{G}{G_{STC}} * [1 + \gamma * (T_M - T_M^{STC})] \quad (1.3)$$

El método de Taylor mostrado en la ecuación (1.4) calcula la potencia basándose en datos como la temperatura ambiente y velocidad del viento, tomando en consideración valores de irradiancia mayores a 500 W/m² [30]. Es necesario mencionar que los parámetros como a_1 , a_2 , a_3 , a_4 y a_5 , son valores obtenidos de las mediciones en estado condiciones estandar (STC).

$$P_M = a_1 + (a_2 + a_3 * T_A) * G + (a_4 + a_5 * V_v) * G^2 \quad (1.4)$$

Por otro lado, el método de cálculo de potencia máxima que usa un modelo de regresión no lineal en función de la irradiancia, así mismo como el enfoque de Taylor, sólo se emplea para valores de irradiancia mayores a 500 W/m² y en días claros, y utiliza una predicción de la potencia por exceso como se observa en la ecuación (1.5). [31]

$$P_M = b_1 * G + b_2 * G^2 + b_3 * G * \ln G \quad (1.5)$$

El método PVUSA [32] mostrado en la ecuación (1.6), consiste en obtener la potencia de salida del módulo fotovoltaico en función de la temperatura ambiente y la velocidad local del viento, tomando en consideración que no se puede utilizar para valores de irradiancia mayores a 400 W/m², y a su vez, debe estar a una altura de 10 metros para evitar desviaciones causadas por el viento ya que esto provoca que la temperatura del módulo disminuya

$$P_M = G * (c_1 + c_2 * G + c_3 * T_A + c_4 * V_v) \quad (1.6)$$

El método [33] mostrado en la ecuación (1.7), permite calcular la potencia a través de un método simple de dependencia lineal con la ecuación mostrada a continuación.

$$P_M = \frac{G}{G_{STC}} * P_M^{STC} \quad (1.7)$$

Existen otras formas como el de [34], en el que plantea calcular la potencia en un lugar determinado por medio de la irradiancia global y la temperatura, sin embargo, este enfoque necesita una base de datos tomadas desde un laboratorio con el fin de elaborar una superficie de rendimiento. El método de [35] es complejo de utilizar pero permite calcular la potencia a través de la relación que existe entre las variables meteorológicas y parámetros eléctricos. El método de [36] ayuda a obtener la potencia máxima en condiciones estándares y utiliza un polinomio de segundo grado con las constantes del modelo calculadas a través de regresión polinomial como se muestra en la ecuación (1.8)

$$P = A + B * T_M * G + C * G + D * G \quad (1.8)$$

Para el cálculo de energía, existen métodos directos que utilizan parámetros basados en modelos de regresión lineal o en el comportamiento del módulo fotovoltaico para realizar la predicción de manera más rápida. Uno de ellos es el método de [37] mostrado en la ecuación (1.9) el cual permite calcular la energía total diaria producida en un módulo por medio de la temperatura ambiente y la irradiación solar.

$$E = d_1 * H + d_2 * H * T_{A,max}^{-2} + d_3 * T_{A,max} \quad (1.9)$$

1.4.10 Scikit-Learn

Scikit-learn es una biblioteca que se desarrolló en el 2007 por David Cournepeau y es muy utilizada para el aprendizaje automático en Python debido a que posee una gran cantidad de herramientas para el modelado estadístico o análisis predictivo de datos. Esta biblioteca es de código abierto ya que admite el aprendizaje supervisado y no supervisado, y también utiliza otras librerías externas como *numpy*, *scipy* y *matplotlib*, que se pueden integrar fácilmente al flujo de trabajo.

Scikit-learn posee métodos algorítmicos de tomas de decisiones como:

- Clasificación: Identifica y categoriza datos basados en patrones
- Regresión: Predice o proyecta valores de datos basados en la media promedio de los datos existentes y planificados, es decir, permite saber cómo se comporta una variable en relación en otras variables de cuales puede depender.
- Agrupación: Agrupación automática de datos similares en conjunto de datos.
- Procesamiento: Normalización o transformación del conjunto de datos.
- Reducción de dimensionalidad: Extrae los datos más importantes a considerar.
- Selección del modelo: Compara y valida los modelos a utilizar.

Antes de la ejecución de las líneas de código, hay que seleccionar la base de datos con la cual se va a trabajar, para ello el lector puede revisar la información descrita en la sección 1.4.6, y escoger acorde a lugar geográfico la que resulte conveniente, para el caso de estudio como se explicó en la introducción de este capítulo, se utilizó la información brindada por Solcast [38].

CAPÍTULO 2

2. METODOLOGÍA

El presente capítulo describe todos los pasos que fueron esenciales para la elaboración del diseño fotovoltaico propuesto. La metodología adoptada se puede desglosar en las siguientes secciones: I) Estudio de la radiación solar y determinación del factor de planta, II) Cálculo de la Potencia de salida, considerando el efecto que tiene la irradiancia, temperatura y velocidad del viento. III) Estimación de la producción de potencia IV) Selección de los inversores y arreglo de los paneles fotovoltaicos, V) Métricas utilizadas para evaluación del sistema y comparación de las técnicas de ML, VI) Monitoreo en tiempo real y Dashboard. VII) Caso de estudio

El caso de estudio para este trabajo de titulación consistió en la aplicación de la metodología propuesta en el edificio 6A de la ESPOL, y comenzó con la búsqueda de datos meteorológicos a través de la información proporcionada por la empresa global de modelado de datos Solcast [38] en base a la latitud y longitud de su ubicación. No obstante, una descripción más detallada se brindará en la sección final de esta capítulo, también es importante acotar que el tipo de instalación que se abordó fue sin baterías y conectado a la red.

2.1 Radiación Solar y Factor de planta

El estudio de la radiación solar para las áreas donde se desea realizar una instalación de fotovoltaica es indispensable tanto en lo técnico como en los aspectos relacionados con la economía, la principal dificultad es la escasez de información de mediciones directas en determinadas zonas, y los ingenieros a cargo de los estudios a menudo recurren a modelos de estimación donde utilizan datos del tiempo en el lugar geográfico donde desean desarrollar el proyecto, no obstante estos datos acarrearán un error debido a los instrumentos utilizados para la medición de los mismos, y dicho error aumenta con la técnica utilizada para la estimación de un promedio en de la irradiancia solar [39].

Razón por la cual se adoptó la implementación de 2 técnicas de ML para la estimación, aunque estas herramientas habían sido aplicadas para el estudio del recurso solar, a menudo estos trabajos oscilaron en el agrupamiento de emplazamientos [40]; [41]; [42], para encontrar funciones de densidad de probabilidades [43] o relacionados con la predicción de radiación o de energía por horas [44], en cambio aquí lo que se propuso

fue determinar el valor más representativo de un histórico de años, los parámetros de entrada para los dos algoritmos utilizados fueron 3: Temperatura(Taire), Irradiancia Horizontal Global(GHI), Velocidad del viento a 10 metros de altura(Vf).

Los modelos de aprendizaje de máquinas fueron implementados utilizando el lenguaje de programación python® y su biblioteca integral *Scikit-learn*, se optó por esta herramienta debido a su flexibilidad, opciones de personalización y eficiencia tal como se reporta en experiencias previas [45], a continuación se realiza una breve descripción del uso de la herramienta

2.1.1 Aplicación de K-means

El propósito de haber utilizado este enfoque, fue el de tratar de agrupar las muestras en n grupos de igual varianza, para ello la herramienta realizó una minimización de la inercia, cuya definición, se puede observar en la ecuación (2.1)

$$In = \sum_{i=0}^n \|x_i - \mu\|^2 \quad (2.1)$$

Donde n representa el número de grupos o *clusters*, bajo este mismo criterio el algoritmo seleccionó el número de centroides que minimizan la inercia, sin embargo, las falencias de este criterio son: suposición de convexidad en los grupos con los datos que se está trabajando y el problema de la métrica no normalizada que constituye este criterio tal como se encuentra reportado en [46], el metodo se le conoce como *Elbow* que no es otra cosa que realizar iteraciones para minimizar la inercia con el fin de determinar el valor de n. También ayudó en la selección de n, la realización de la investigación previa de las condiciones climatológicas del lugar, y la revisión reportes de los datos meteorológicos que brindó atlas solares y los institutos nacionales de meteorología, de tal forma que la selección del número de clústeres fue consecuente con la realidad.

El diagrama de flujo bajo el cual operó el algoritmo se puede mirar en la Figura 2.1, las librerías que se importaron fueron *numpy*, *pandas*, *matplotlib*, *seaborn*, *sklearn*, tanto la primera como la segunda librería sirvieron para trabajar con vectores y matrices, mientras que *matplotlib* sirvió para realizar las gráficas y la última librería fue la utilizada para la resolución del problema. El escalamiento de datos se empleó para evitar problemas con las escalas, ya que, al ser cada eje de unidades de distinta naturaleza, es preferible

trabajar con valores escalados, aquí se utilizó la función *MinMaxScaler* () la cual se encargó de normalizar los datos en un rango entre 0 y 1.

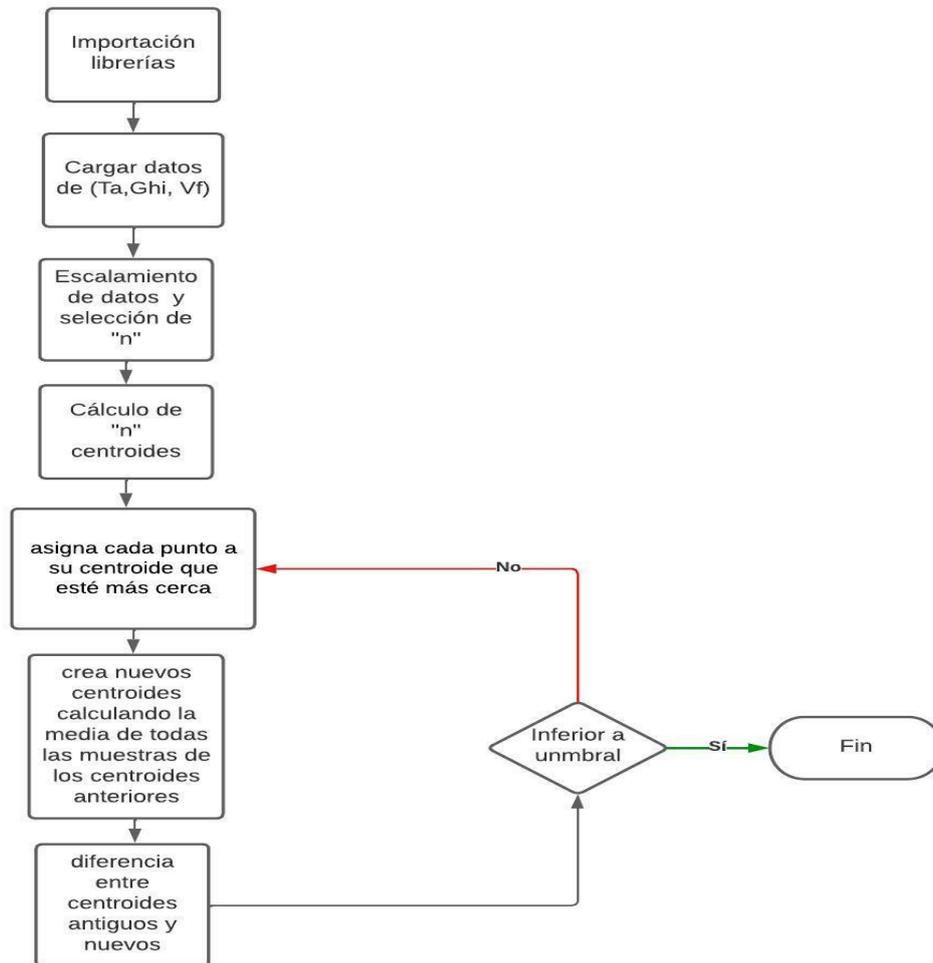


Figura 2.1 Diagrama de flujo de *k-means* para el agrupamiento de Ta, GHI, Vf. Fuente: autores

Los resultados que se obtengan, constituirán los valores representativos de los datos, la discusión de como contextualizar esta información se detallará en el siguiente capítulo. Conocido los valores de Ta, GHI y Vf, el dato que requirió atención fue el de la irradiancia, se debió tener en cuenta que había que volver a sus unidades originales para ello se necesitó emplear la ecuación (2.2).

$$x = x_{esc} * (x_{max} - x_{min}) + x_{min} \quad (2.2)$$

Donde:

x : valor en sus unidades correspondientes

x_{esc} : valor normalizado en el rango de (0,1)

x_{max} : valor máximo del conjunto de datos con el que se trabajó

x_{min} : valor mínimo del conjunto de datos con el que se trabajó

Las unidades de la irradiancia a las que se volvió fueron w/m^2 , para calcular la irradiancia se requirió utilizar la ecuación (2.3).

$$Irradiación = \int_{h_0}^{h_1} (Ghi)dh \left[\frac{Wh}{m^2} \right] \quad (2.3)$$

Dependiendo de la ecuación que se utilice para modelar la salida de potencia de un panel solar, se utilizará el valor de GHI, o de irradiancia. Por otro lado, el intervalo que se consideró prudente utilizar fue el de la hora pico solar.

Por último, para la determinación del factor de la planta, se adoptó la expresión formulada expresada en [40], tal como se puede ver en la (2.4).

$$FC = \frac{P_{fv}(di)(h_1 - h_0)}{365(a_{pv})(b_{pv})Irradiación} * 100 \quad (2.4)$$

Donde:

a_{pv} : Ancho del módulo fotovoltaico en metros

b_{pv} : Largo del módulo fotovoltaico en metros.

di : días que se realizó el análisis.

P_{fv} : Potencia del panel

h_0 : valor inicial de la hora solar pico.

h_1 : valor final de la hora solar pico.

2.1.2 Aplicación de GMM.

Como ya se mencionó anteriormente, el desarrollo de los modelos se lo realizó utilizando el lenguaje de programación Python, para este caso se invocó el paquete *sklearn.mixture* el cual permitió implementar el modelo. A diferencia de *k-means*, la clasificación realizada

por este enfoque, fue “suave”, ya que proporcionó la probabilidad de que un punto sea de un determinado grupo, es decir aquí tuvo en cuenta superposiciones de grupos, por lo que fue de naturaleza probabilística. Antes de implementar el algoritmo de GMM, fue necesario determinar la cantidad de mezclas gaussianas, en vez de los “n” centroides, pero a la vez fue necesario la determinación de la matriz de covarianza a utilizar, para determinar esto se utilizó el Criterio de Información Bayesiana (BIC), lo que trató de encontrar fue la máxima probabilidad posterior o máxima verisimilitud marginal. La explicación detrás de este criterio BIC, se escapa de los objetivos del presente trabajo, pero para una discusión más amplia del tema revisar el artículo [47]. Para determinar la solución del modelo de mezclas gaussianas, se utilizó el comando `GaussianMixture()` donde los argumentos que colocaron fueron:

- Números de mezclas gaussianas.
- Tipo de matriz de covarianza(completa, diagonal, esférica)
- El valor de `random_state()` que garantizó que los resultados sean reproducibles)
- Inicialización con k-means

El diagrama de flujo del modelo, se puede observar en la Figura 2.3, de igual forma lo primero que se realizó fue la importación de las librerías y el escalamiento de datos, lo particularmente nuevo aquí fue la utilización del algoritmo de Esperanza-Maximización(EM), que fue de naturaleza iterativa, con la esperanza estimó el valor de la variable que no se podía observar que estaría detrás y que relacionaría los parámetros estudiados, y con el paso de maximización actualizó los parámetros del modelo: peso, media, matriz de covarianza. Para efectuar lo descrito anteriormente, el programa internamente lo que hizo fue inicializar los índices de las medias a los términos de las covarianzas y los coeficientes de mezclas, y evaluar el valor inicial en la función del logaritmo de la probabilidad, luego realizó el calculo de la variable latente para el enésimo gaussiano que se haya ingresado dividido para la sumatoria de todas las distribuciones gaussianas, luego actualizó los valores de la media, covarianza, y peso, con el paso de maximización, considerando la variable latente, así hasta cumplir el umbral ingresado en el algoritmo, lo descrito anteriormente se encuentra en el diagrama de flujo de la Figura 2.2:

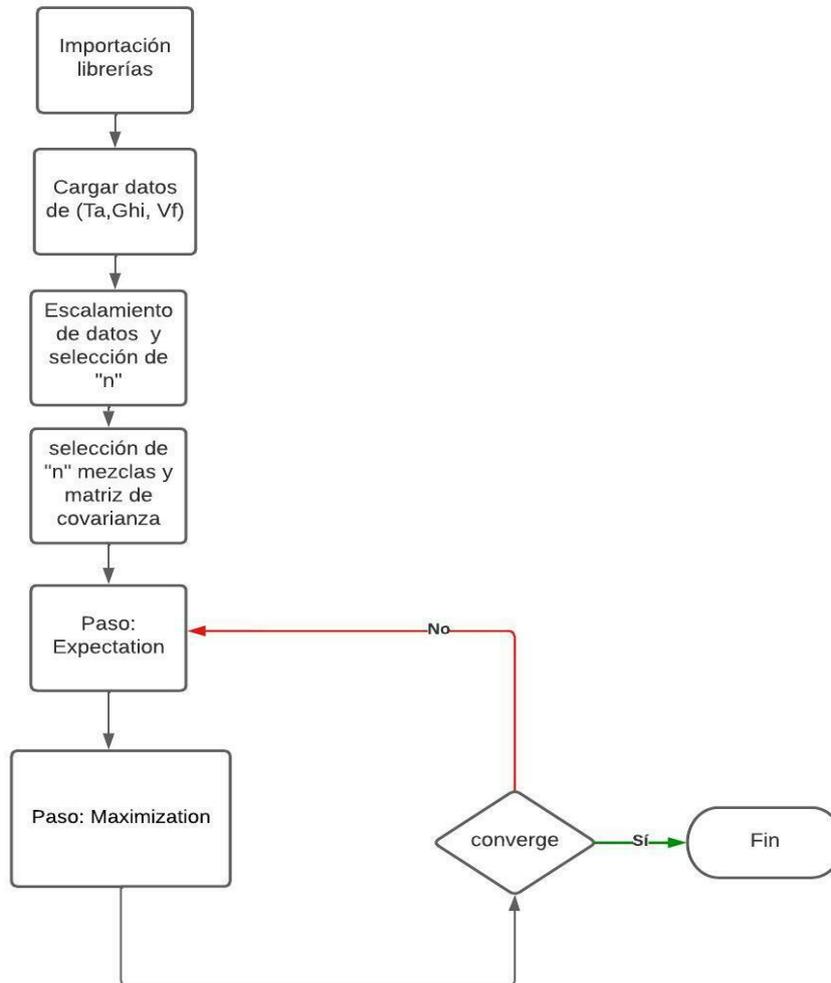


Figura 2.2 Diagrama de flujo de la aplicación del modelo de GMM. Fuente: autores

Una vez obtenido los resultados del modelo, se procedió de igual manera a calcular el factor de planta, y para ello se empleó la ecuación (2.4). descrita en la sección anterior

2.1.3 Cálculo de potencia DC del PV

Una vez obtenidos los valores representativos de los tres parámetros meteorológicos, lo siguiente fue considerarlos en la determinación de la potencia de salida DC del módulo PV. En la sección 1.4.9 se efectuó una revisión de algunas metodologías, no obstante, se optó por lo reportado en [48], donde se consideró las pérdidas en los sistemas fotovoltaicos, producto de el cableado, el inversor, y la ecuación finalmente utilizada fue la ecuación (2.5) que originalmente fue publicada en [49].

(2.5)

$$P_{dc} = P_{pv} * \frac{Ghi}{Ghi_{STC}} \left(1 + r * (T_{cell} - T_{cell}^{STC}) \right) \left(a_1 + a_2 \frac{Ghi}{Ghi_{STC}} + a_3 \ln \left(\frac{Ghi}{Ghi_{STC}} \right) \right) Per_{DC}$$

Donde:

P_{pv} : Potencia del módulo (visto en la placa)

Ghi : Irradiancia global horizontal en W/m² (Obtenida mediante las técnicas de ML)

Ghi_{STC} : Irradiancia bajo STC (1000 W/m²)

r : coeficiente de variación de potencia debido a la temperatura de la celda(%/°C)

T_{cell} : Temperatura de operación de la celda(°C).

T_{cell}^{STC} : Temperatura de la celda bajo STC

a_1, a_2, a_3 : Parámetros que describen la relación entre irradiancia y la eficiencia.

Per_{DC} : factor de perdidas DC.

Los parametros a_1 , a_2 y a_3 , fueron ajustados siguiendo las indicaciones encontradas en [49], donde recomendaban realizar el ajuste de la curva utilizando la eficiencia a distintos puntos de valores de irradiancia: 1000, 600 y 200. No todos los *datasheet* poseen directamente estos valores, pero una explicación más amplia se puede revisar en el caso de estudio, y como utilizar la curva corriente y voltaje para ajustar estos valores.

Si bien la expresión antes descrita permitió modelar el efecto de la irradiancia y la temperatura, aún no se había considerado el efecto del viento, para ello se utilizó lo informador por [50] donde se evidenció que las ecuaciones presentadas en [31] funcionaban con precisión en los modelos de silicio(Si) policristalino. Si se desea trabajar con silicio monocristalino, el lector puede revisar el mismo artículo donde se abordan para otro tipo de tecnologías

$$T_{cell} = T_a + \frac{Ghi}{Ghi_{NOCT}} (T_{NOCT} - T_{a,NOCT}) * \frac{f_{w,NOCT}}{f_w(v)} * \left[1 - \frac{n_{STC}}{\tau * \alpha} (1 - \beta_{STC} T_{STC}) \right] \quad (2.6)$$

Donde:

T_a : Temperatura ambiente (25 °C)

T_{NOCT} : Temperatura que viene indicada en la placa del panel (normalmente 45°C).

Ghi_{NOCT} : Irradiancia de 800 W/m².

$T_{a,NOCT}$: (20 °C)

$f_{w,NOCT}$: velocidad del viento 1 [m/s].

$f_w(v)$: coeficiente de convección del viento.

η_{STC} : eficiencia del módulo (%)

T_{STC} : 25°C

β_{STC} : coeficiente de temperatura de potencia máxima (0.45)

$\tau\alpha$: puede ser asumida como 0.9 según [31]

De la (2.6) quedó pendiente conocer el coeficiente de convección del viento, mismo que fue modelado con la ecuación (2.7) y (2.8), y la relación descrita en la ecuación (2.9) detallada también en [51].

$$f_w = 8.91 + 2.00 v \quad (2.7)$$

$$f_w = 5.7 + 2.8 w \quad (2.8)$$

$$w = 0.68v - 0.5 \quad (2.9)$$

Donde:

v : velocidad del viento 10m por arriba del terreno.

w : velocidad del viento cercana al panel.

2.1.4 Estimación de la producción de energía y del consumo.

Previo a la determinar la energía producida en un intervalo de tiempo, es necesario determinar la potencia entregada en AC, ya que la ecuación (2.5) modeló la potencia de salida DC. Para esto así mismo se adoptó lo explicado en [52] y [48], de donde se tomó las ecuaciones (2.10).

$$P_{AC} = P_{DC}\eta_i C_{AC} \quad (2.10)$$

η_i : Eficiencia del inversor

C_{AC} : Coeficiente que agrupa pérdidas AC.

Y de la cual se deriva la ecuación (2.11) que sirvió para determinar la energía en un periodo de tiempo.

$$E_{AC} = \int_0^t P_{AC} dt \quad (2.11)$$

Una vez conocido el consumo de energía mensual, se utilizó la ecuación estipulada por la ARCONEL 003/018 [53], para la capacidad nominal instalada, esta expresión se puede observar en la ecuación (2.12)

$$Cap. nominal = \frac{\sum_1^{12} Em_i (kWh)}{FP * 8760} (kW) \quad (2.12)$$

Donde el FP fue encontrado en la sección 2.1 y la Em_i es la energía mensual facturada por mes.

2.1.5 Inversor y selección Selección del número de PV

Como ya se estipuló en la introducción del presente capítulo, la metodología adoptada está enfocada en instalaciones conectadas a la red, para que el lector pueda darse una idea del esquema básico de este tipo de conexión, puede observar la Figura 2.3.

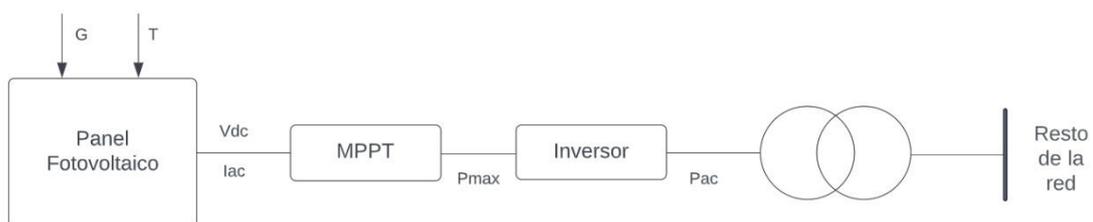


Figura 2.3 Esquema de una planta solar conectado a la red. Fuente: [54]

Aunque a menudo el bloque MPPT, Inversor vienen incorporados en un solo dispositivo. Una expresión utilizada que se sugirió para la elección del inversor fue la ecuación (2.13) y (2.14, basada en su rendimiento tal como se informó en [55] .

$$\eta_{total} = \eta_{MPPT} * \eta_{inv} \quad (2.13)$$

$$\eta_{inv} = \frac{P_{AC}}{P_{DC}} = \frac{P_{AC}}{I_{max} * V_{max}} \quad (2.14)$$

Donde:

I_{Imax} : Corriente en el punto de máxima potencia

V_{Imax} : Voltaje en el punto máxima potencia

Aspectos económicos también son necesarios tener en cuenta a la hora de seleccionar el inversor, así como condiciones climatológicas y técnicas, pero dicho contenido se escapa de los objetivos de este trabajo, sin embargo, para una revisión más profunda y rigurosa, revisar [56] así como también [57].

En referencia con la cantidad de paneles a instalar por cada inversor, la ecuación se que utilizó fue la (2.15, referenciada en la literatura de [19].

$$N_{PT} = \frac{P_{IDCmax}}{P_{pv}} \quad (2.15)$$

P_{IDC} : Potencia máxima del inversor

P_{pv} : Potencia pico del PV.

Mientras que para conocer el número de paneles en serie se utilizó la ecuación (2.16 como la ecuación (2.17 para determinar el número de ramas en paralelo.

$$N_{pseries} = \frac{V_{max-inversor}}{V_{OCm}} \quad (2.16)$$

$V_{max-inversor}$: Voltaje máximo en la placa del inversor

V_{OCm} : Voltaje de circuito abierto máximo del PV

$$N_{ramas} = \frac{N_{PT}}{N_{pseries}} \quad (2.17)$$

N_{PT} : Número de paneles totales

$N_{pseries}$: Número de paneles en serie

Se contrastó finalmente si el producto de la cantidad de paneles en serie multiplicado por la cantidad en paralelo, no sobrepasó el número de paneles totales.

2.1.6 Métricas utilizadas para evaluación del sistema y comparación de las técnicas de ML

Las siguientes métricas fueron extraídas de [8] y de igual forma del estudio realizado para el Distrito metropolitano de Quito [39].

$$Cons = \sum_{k=1}^n C^{k \text{ diario}} * Tm \quad (2.18)$$

C^{diario} : Proyección del consumo diario.

Tm : resolución de tiempo que se consume la carga.

n : cantidad de mediciones

$$E_{diaria} = \sum_{k=1}^n P_{pv}^k * tm \quad (2.19)$$

P_{pv}^k : Medición de la potencia entregada en la k medición.

tm : resolución del tiempo expresado en horas.

n : cantidad de mediciones

La (2.18 se utilizó para estimar el consumo diario, mientras que la ecuación (2.19 fue empleada para la determinación de la energía entregada por la instalación fotovoltaica. Por otro lado, para estimar la potencia autoconsumida se usó la (2.20 mientras que para conocer cuanta potencia se inyecta a la red se hizo uso de la (2.21).

$$P_{\text{autoconsumida}} = \begin{cases} C_{\text{cons}} & P_{Tpv} > C_{\text{cons}} \\ P_{Tpv} & \text{lo opuesto} \end{cases} \quad (2.20)$$

Donde:

P_{Tpv} : Potencia total entregada por la instalación fotovoltaica.

$$P_{\text{importada}} = C_{\text{cons}} - P_{Tpv} \quad (2.21)$$

Respecto a la potencia importada, esta es 0 si la potencia entregada por la instalación fotovoltaica es mayor que la carga diaria consumida. La ecuación (2.22) fue empleada para la estimación de horas de producción del sistema.

$$Y_{\text{prod}} = \frac{E_{\text{entre}}}{P_{\text{nomPV}}} \quad (2.22)$$

E_{entre} : Energía entregada al sistema

P_{nomPV} : Potencia nominal de la instalación

Finalmente, para la comparación de las técnicas de Machine learning se utilizó el índice descrito por la ecuación (2.23), que se denomina rendimiento de referencia, y la métrica del error relativo definido por la ecuación (2.24).

$$Y_{\text{Irra}} = \frac{I_{\text{cal}}}{I_{\text{ref}}} \quad (2.23)$$

Donde:

I_{cal} : Irradiación calculada

I_{ref} : Irradiación de referencia.

$$E_r = \left| \frac{P_{predicido} - P_{medido}}{P_{medido}} \right| \quad (2.24)$$

$P_{predicido}$: Predicciones de potencia de un módulo.

$P_{medidos}$: Valores medidos de potencia.

Finalmente se utilizó lo que establece la IEC 61724-1 [58], para la evaluación de la calidad del sistema fotovoltaico, que relaciona las ecuaciones: 2.22 y 2.23. Esto se encuentra descrito en la ecuación 2.25 y se denomina *Performance Ratio* (PR).

$$PR = \frac{Y_{prod}}{Y_{Irra} * C_k} \quad (2.25)$$

Donde:

Y_{prod} : Estimación de horas de producción del sistema

Y_{Irra} : Rendimiento de referencia

C_k : Índice de corrección

El índice de corrección se determinó utilizando la expresión descrita en la ecuación 2.26, para una discusión más rigurosa sobre estas métricas, revisar lo publicado en [59].

$$C_k = 1 + \gamma * (T_{cell} - 25^{\circ}C) \quad (2.26)$$

Donde:

γ : Coeficiente de variación de temperatura especificado en la placa.

T_{cell} : Temperatura del módulo

2.1.7 Monitoreo en tiempo real y Dashboard

Esta sección se desarrolló de forma interdisciplinaria el área de Ingeniería en Telemática, no obstante, se brinda una descripción general del enfoque al cual el trabajo se acogió en este punto. En la primera fase se realizó la prueba de campo para verificar el acceso a los datos del dispositivo, se observó si existía un registro de datos inalámbricos. Luego se optó por la conexión, esto se puede visualizar en la Figura 2.4.

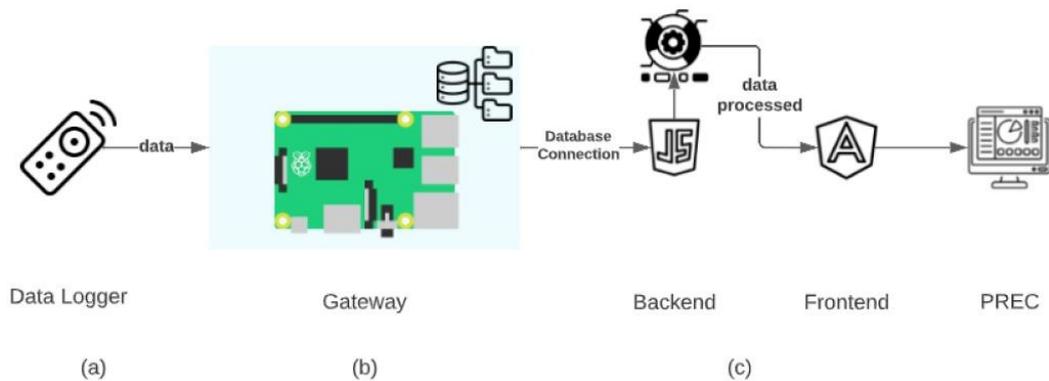


Figura 2.4 Conexión realizada para el monitoreo en tiempo real. Fuente: autores

No obstante, si la conexión antes mencionada presenta dificultades, otra posibilidad fue realizar una prueba de campo mediante la conexión de un Arduino “uno” en la interfaz RS485, utilizando el módulo RS385 a TTL. Una vez conectado, se empleó una computadora conectada al Arduino “uno”, para ello fue necesario un cable USB tipo A, de tal forma que permita abrir el monitor serial en el programa Arduino IDE, y poder verificar si la información fue enviada por el inversor, todo esto se puede observar en la Figura 2.5.

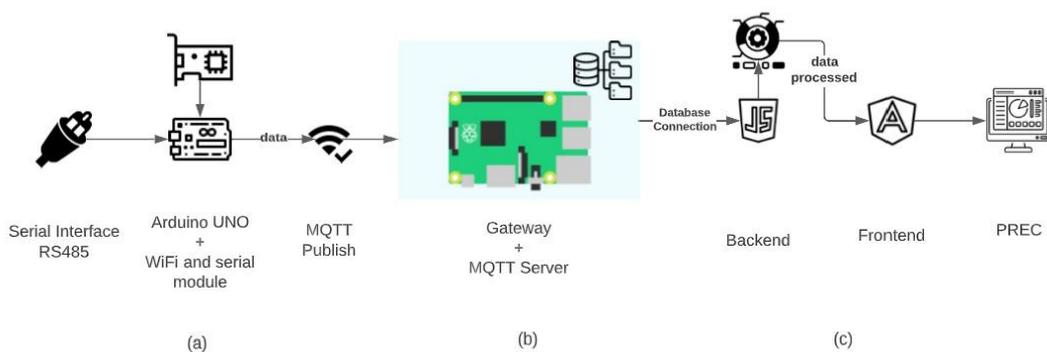


Figura 2.5 Implementación utilizando un Arduino. Fuente: autores

En relación con el *Dashboard* y los resultados a presentar, simplemente consistió en representar los parámetros que se extrajeron del monitoreo detallado en los párrafos anteriores, los nombres que se debieron tener en cuenta, están referidos en la Tabla 2.1.

Tabla 2.1 Datos a mostrar en el Dashboard

Parámetros	Unidades
Energía diaria	kWh
Energía mensual	kWh
Potencia inyectada	kW
kg de CO2.	kg

2.2. Caso de estudio

El proyecto se lo realizó en el edificio 6A de rectorado ubicado en la ESPOL en el km 30.5 de la vía perimetral, Guayaquil, Ecuador, como se muestra en la **Figura 2.6**.

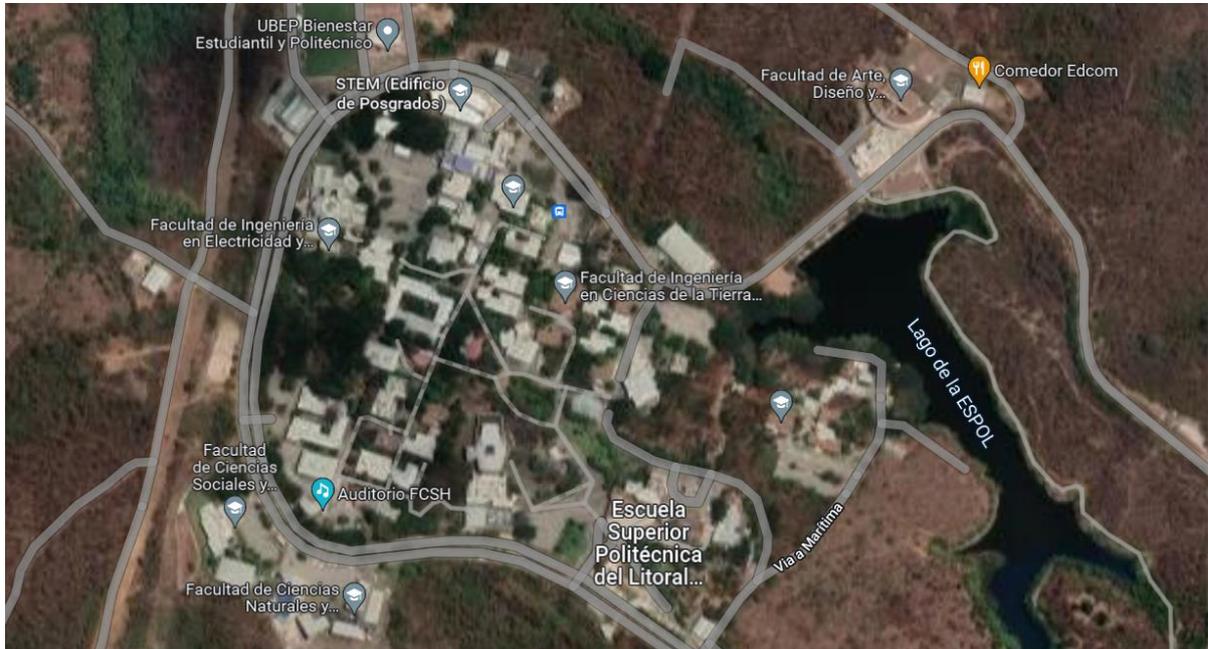


Figura 2.6 Localización geográfica del caso de estudio. Fuente: autores



Figura 2.7 Edificio 6A (Rectorado). Fuente: autores

El edificio de 6A está compuesto por tres pisos y un sótano. Tiene un área de 5136 m² aproximadamente y está distribuido entre salas, oficinas, baños y auditorios.

2.2.1. Condiciones climáticas

Debido a que el edificio se encuentra en una región donde existen condiciones climáticas calientes y húmedas, la temperatura del aire varía entre 18.8 °C y 35.7 °C en épocas de verano e invierno. La irradiancia total recibida sobre una superficie horizontal tiene un promedio mensual de 185.88 W/m² y 169.26 W/m² en las estaciones del año descritas anteriormente y un promedio de la componente horizontal de la irradiancia difusa de 120.77 W/m² durante todo el año.

2.1.8 Cargas del edificio

El edificio posee cargas climáticas, cargas de iluminación y equipos eléctricos, sin embargo, debido a su localización y los diversos factores climáticos, el edificio dispone de cargas de climatización constituidas por aires acondicionados techados y aires individuales.

La información de cargas del edificio se lo obtuvo de [60] en donde se detalla la clasificación de la capacidad instalada de todos los equipos que tiene la infraestructura del edificio 6A. En la Tabla 2.2 se pueden observar dichos datos.

Tabla 2.2 Equipos del edificio 6ª [60]

DESCRIPCIÓN		
ILUMINACIÓN	Fluorescente	Sistema de ahorro de energía (LED)
EQUIPOS	Máquinas de impresión	Fotocopiadoras, impresoras
	Electrodomésticos	Cafeteras, dispensadores de agua, televisión, microondas, refrigeradora, secadora de mano
	Computadora	
	Otros	Estructura de cables, Wifi, equipos de audio, micrófono, proyector, trituradora de papel

Es necesario mencionar que el edificio utiliza sistemas de calefacción, ventilación y aire acondicionado (HVAC)), y unidades divididas para climatizar cada uno de los pisos. A continuación, en la Tabla 2.3 se presenta un resumen de las divisiones de las cargas del sistema de aire acondicionado tomado de [60].

Tabla 2.3 - Capacidad instalada por las cargas enfriamiento del edificio [60]

Distribución	Total (BTU/h)
Sótano	606000
Planta baja	1092000
Piso 1	504000
Piso 2	330000
Total	2532000

El informe también indica que el consumo energético anual es de 465903 kWh, en donde la mayor parte del consumo lo ocupa el sistema de climatización como se puede muestra en la Figura 2.8.

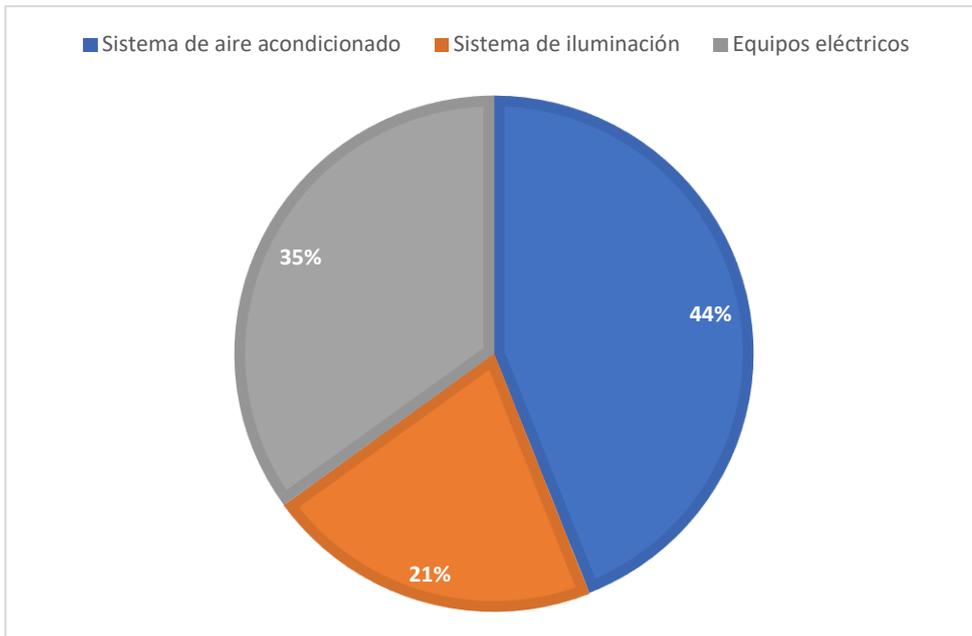


Figura 2.8 División del consumo energético de Rectorado. Fuente: autores

Lo mencionado en esta sección constituyó los parámetros de entrada en lo referente a consumo energético, en la sección 2.1.4 de la metodología implementada se hizo referencia a que era necesario no solo la estimación de producción de los paneles sino también del consumo, no obstante, estos datos presentados aquí son proyecciones que fueron realizadas en el estudio mencionado al inicio de este rótulo, para una profundización del modelo utilizado revisar el artículo [60].

2.2.2. Datos meteorológicos

En la presente sección se realizó un estudio de los parámetros meteorológicos que fueron necesarios para el diseño de una instalación fotovoltaica. Existen algunos modelos de sistemas satelitales propuestos que ayudan a predecir datos de la radiación solar, temperatura, precipitación, duración de insolación, humedad o velocidad de viento dependiendo de las estaciones del año y fecha requerida que son comúnmente usados en softwares como Pvsyst, SAM o HomerPro para el diseño de sistemas solares tal como se abordó en la sección 1.4.7. Estas bases de datos brindan información histórica de 10 años o más, sin embargo, poseen variaciones dependiendo de la calidad de datos y método de medición ya que la naturaleza del instrumento empleado puede ser terrestres o satelitales. En la siguiente sección se presenta la base de datos seleccionada.

2.2.3. Solcast

El estudio de datos se lo realizó a través de la compañía global de pronóstico solar Solcast, la cual permite la adquisición de datos de irradiación solar y la vez ofrece maneras diferentes de acceder a los datos ya sea por radiación, en vivo o pronosticados por medio del (API) mundial, sitios de radiación solar o por *Utility Scale*.

El método Solcast para estimar la irradiancia solar de los satélites meteorológicos geoestacionarios consta de tres pasos. El primero es la detección de la cobertura de nubes y la caracterización de dicha cobertura en términos de su impacto en la radiación solar. En el segundo paso se modela la radiación solar disponible bajo cielos despejados, incluido el tratamiento de aerosoles (polvo, sal, etc.) y el contenido de agua. Y el tercero es la estimación combinada final de la cantidad de radiación solar que llega a la superficie de la Tierra después de pasar a través de las nubes. [38], [61]. A partir de los datos meteorológicos obtenidos de la base antes mencionada como temperatura e irradiación se pudo determinar la eficiencia de los paneles, el dimensionamiento, la energía que estos generan y se evaluó su rendimiento a la vez que se evidenció cual es el efecto que tienen estos parámetros en el diseño de una instalación fotovoltaica.

2.2.4. Adquisición de datos

Como primer paso se especificó geográficamente el lugar en donde se ejecutó el proyecto. La Figura 2.9 muestra la ubicación del edificio de estudio 6A ubicado en la ESPOL en el campus Gustavo Galindo Velasco con una latitud y longitud de -2.147779 y -79.964259 respectivamente.

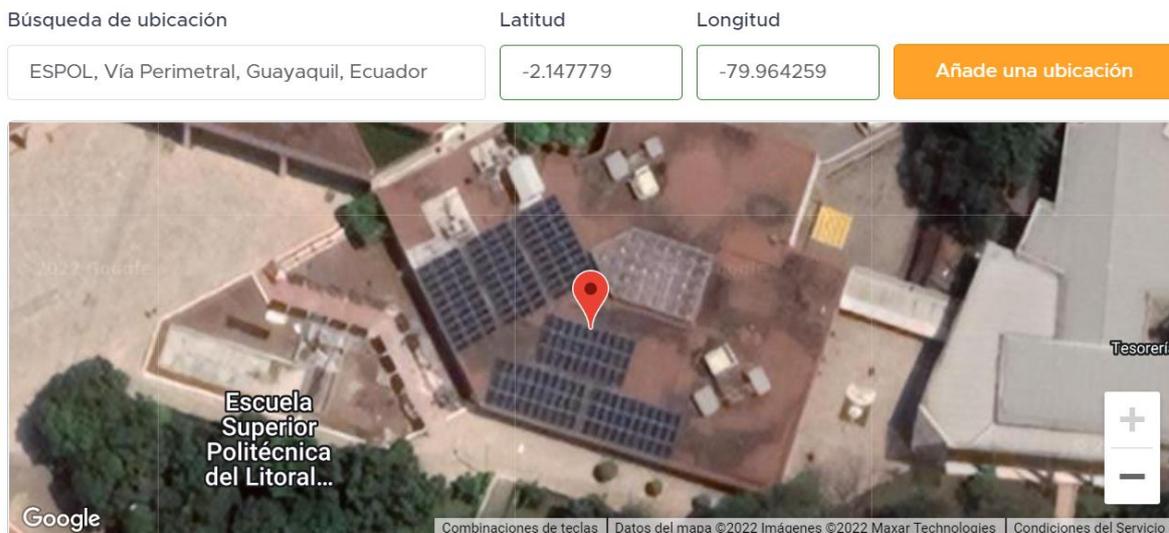


Figura 2.9 Ubicación caso de estudio. Fuente: autores

Los datos obtenidos a través del software se describen en la Tabla 2.4, pero a través del filtrado se seleccionó GHI, Ta y Vf.

Tabla 2.4 – Datos obtenidos de solcast

Parámetros	Significado
<i>Air Temp</i>	La temperatura del aire a 2 metros sobre el suelo
<i>Azimuth</i>	Ángulo de acimut solar (grados).
<i>Cloud Opacity</i>	La atenuación de la luz solar entrante debido a la nube.
<i>Dewpoint</i>	La temperatura del punto de rocío a 2 metros sobre el nivel del suelo
<i>DHI</i>	La componente horizontal de la irradiancia difusa
<i>DNI</i>	Irradiación Normal Directa.
<i>GHI</i>	La irradiancia total recibida sobre una superficie horizontal.
<i>Precipitable Water</i>	Agua precipitable de toda la columna atmosférica
<i>Relative Humidity</i>	La humedad relativa del aire a 2 metros sobre el nivel del suelo
<i>Surface Pressure</i>	La presión del aire a nivel del suelo
<i>Wind Direction</i>	Dirección del viento (grados) a 10 m sobre el nivel del suelo.
<i>Wind Speed</i>	Velocidad del viento (en m/s) a 10 m sobre el nivel del suelo.
<i>Zenith</i>	El ángulo entre una línea perpendicular a la superficie de la tierra y el sol.

2.2.5. Dimensionamiento

La implementación de la metodología descrita en la sección 2.1 se llevó a cabo tal como se puede observar en la Figura 2.10.

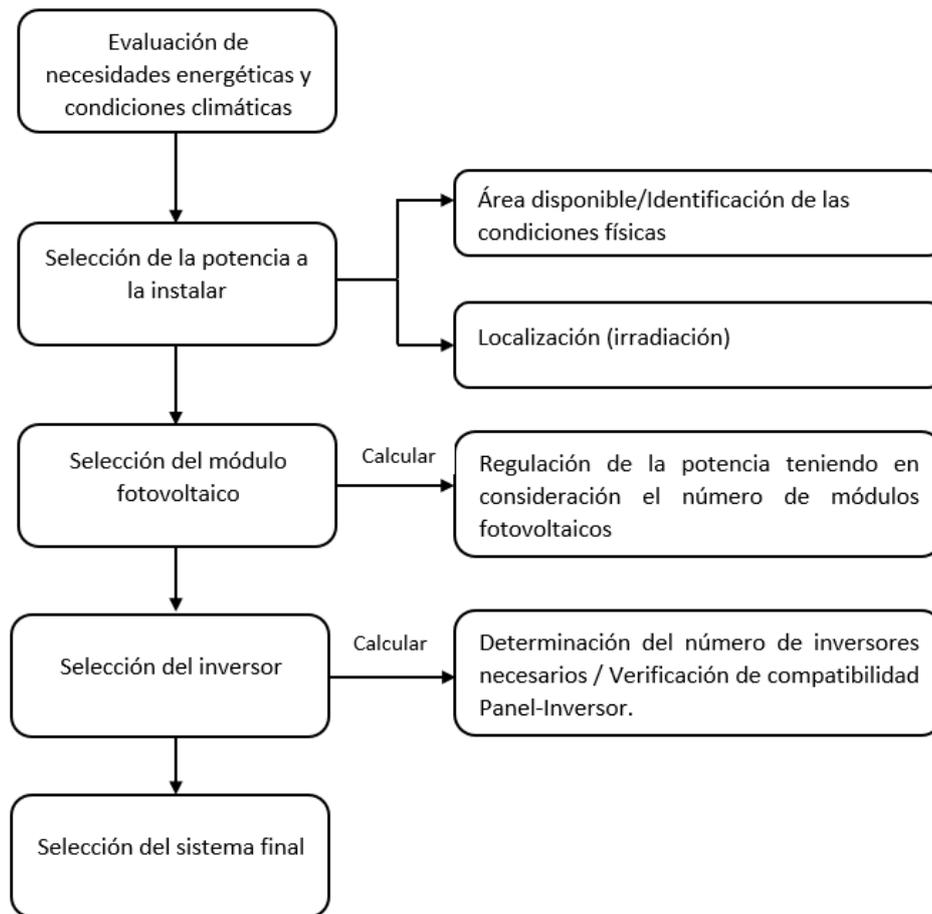


Figura 2.10 Implementación de la metodología

Para calcular la potencia entregada de los paneles se utilizó la ecuación (2.5) y (2.6), pero tal como se ve en la Figura 2.11, pero antes fue necesario extraer los valores representativos de los parámetros GHI, mediante las 2 técnicas de Machine Learning aplicadas en los datos de la sección 2.2.5. Extraídos los valores representativos se determinó lo antes mencionado. Además, se tuvo que en cuenta que, para la expresión de la potencia del módulo, se empleó el valor de GHI más elevado, mientras que la ecuación que modeló la temperatura utilizó el valor menos favorable. Para los coeficientes $a_1, a_2, y a_3$, de la ecuación se tomó como parámetros de entrada en el ajuste de la curva los datos mostrados en la Figura 2.11.

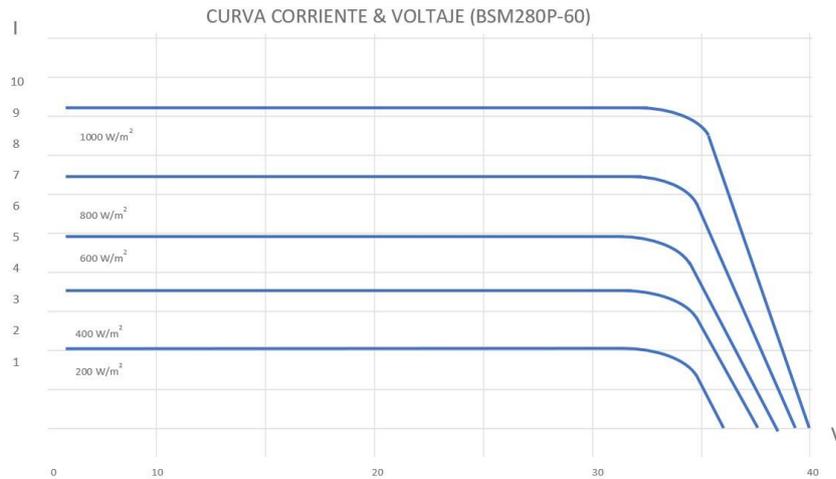


Figura 2.11 Curva de corriente vs voltaje de los paneles existentes en el edificio 6A.

Fuente: autores

Así mismo fue de bastante relevancia los valores mostrados en la placa de paneles, los mismos que se encuentran referidos en la Figura 2.12, fueron utilizados en las ecuaciones antes mencionadas.



Figura 2.12 Datos de los paneles BSM280P-60. Fuente: autores

Las características del inversor también representaron parámetros de entrada para el dimensionamiento de la instalación solar, dichos datos se pueden evidenciar en la

Figura 2.13, de aquí se requirió los valores del: número de entradas por MPP, el voltaje máximo que soporta el dispositivo, el voltaje mínimo con el que se enciende.



Figura 2.13 Datos del inversor Solis 1P10K-4G-US. Fuente: autores

Con los datos referidos en el párrafo anterior, se determinó el número de paneles en serie y los números de *string*. Para ello fue necesario considerar 2 restricciones, con la finalidad de que los inversores que se fueran a utilizar, ante condiciones adversas, logren su encendido y no excedan su valor máximo de tensión que pueden soportar, dichas limitaciones están expresadas en las ecuaciones (2.27 y (2.28).

$$V_{oc}(T_{min}) = V_{oc}(25^{\circ}C) + (\beta * dT_1) < V_{DCmax}^i \quad (2.27)$$

Donde:

$V_{oc}(T_{min})$: Voltaje de circuito abierto en función de la temperatura mínima

T_{min} : Temperatura mínima del módulo.

β : -0.330 %/°C

dT_1 : Diferencia de temperatura entre T_{min} -25.

V_{DCmax}^i : Voltaje máximo DC del inversor.

$$V_{mpp}(T_{max}) = V_{mpp}(25^{\circ}C) + (\beta * dT_2) > V_{DCmin}i \quad (2.28)$$

Donde:

$V_{mpp}(T_{max})$: Voltaje en el punto de máxima potencia en el valor de temperatura máximo.

$V_{mpp}(25^{\circ}C)$: Voltaje en el punto de máxima potencia a 25 °C

β : -0.330 %/°C

dT_2 : diferencia de temperatura entre el valor máximo del módulo y 25°C

CAPÍTULO 3

3. RESULTADOS Y ANÁLISIS

3.1 Recurso Solar

Se utilizó el histograma de frecuencias para la representación gráfica del conjunto de datos obtenidos en Solcast con el fin de visualizar la variabilidad de datos e identificar la tendencia central de los mismos. Estos datos se dividieron en las temporadas de frío y calor del año 2021, comenzando con los meses de diciembre a mayo y luego de junio a noviembre.

La Figura 3.1 muestra los datos obtenidos de la temperatura del aire versus la frecuencia de la estación de verano, donde se observó que el pico de datos ocurre en torno a los 21.9 °C y la dispersión de datos es a partir de los 18.9 °C a los 35.9 °C. La Figura 3.2 describe los datos de la irradiación horizontal global en donde se presentó que el pico de datos de GHI se encuentra en 33.5 W/m² y la Figura 3.3 presenta los datos que se obtuvieron de velocidad de viento versus frecuencia, el cual indica que el pico de datos ocurre alrededor de 1.45 m/s y la dispersión de datos es entre 0.45 m/s y 4.45 m/s.

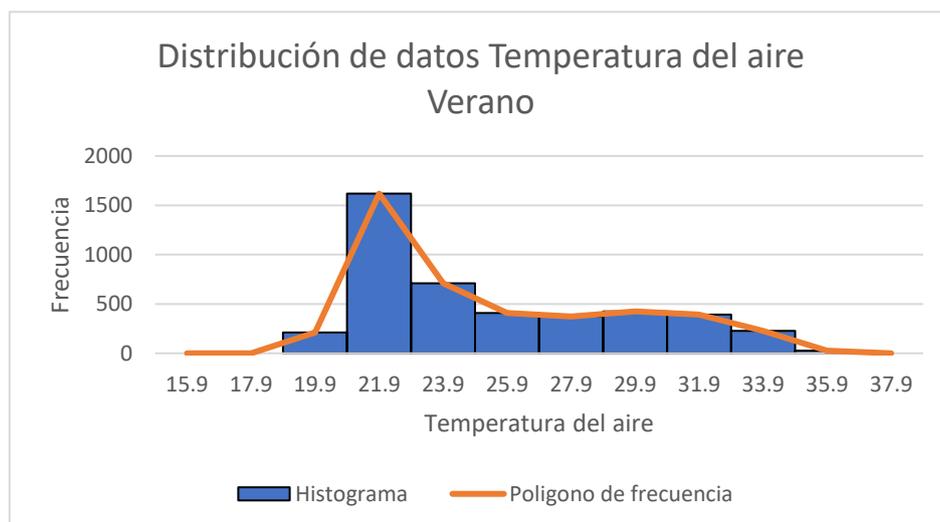


Figura 3.1 Distribución de datos de temperatura del aire en verano. Fuente: autores

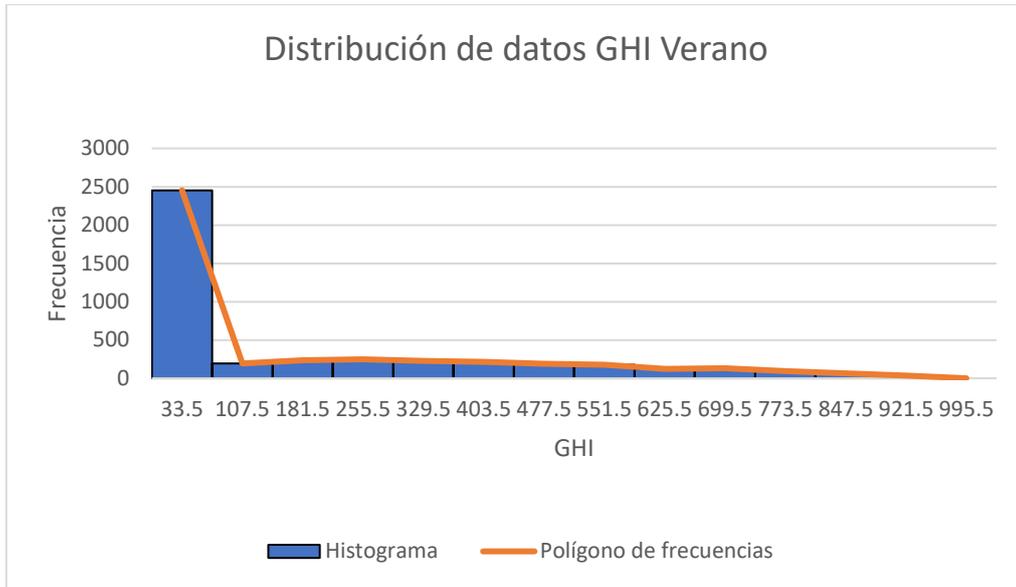


Figura 3.2 Distribución de datos GHI en verano. Fuente: autores

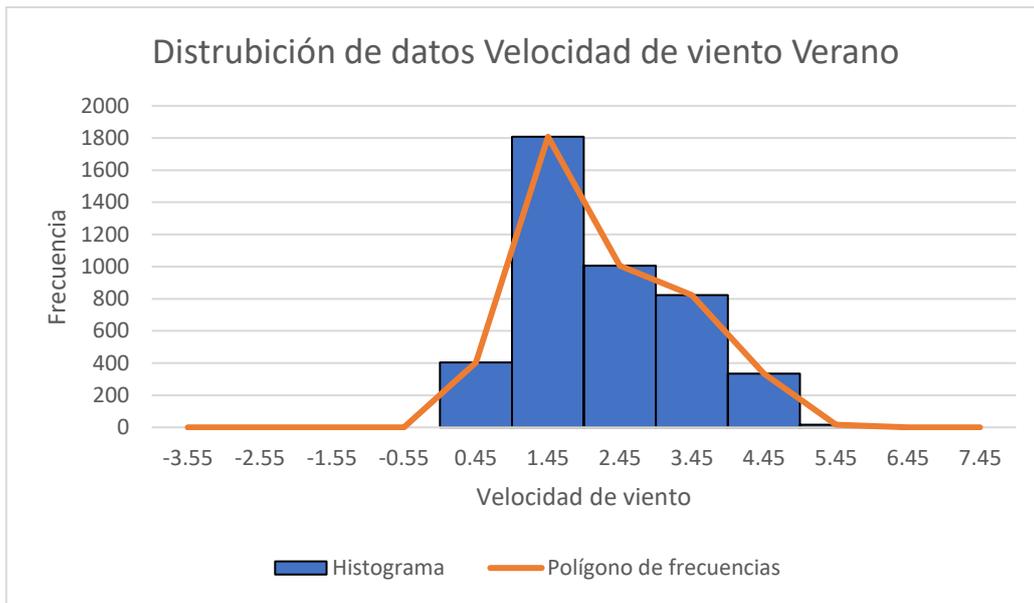


Figura 3.3 Distribución de datos de velocidad de viento en verano. Fuente: autores

La Figura 3.4 presenta la distribución de datos de temperatura del aire versus la frecuencia, pero para la estación de invierno. Se observa que la dispersión de datos ocurre desde los 20.85 °C hasta los 30.85 °C y el pico de datos es en torno de los 22.85 °C. En la Figura 3.5 ocurre una situación similar a la Figura 3.2 en donde los datos se encuentran en 33.5 W/m² y, por último, en la Figura 3.6, se muestran los datos de

velocidad de viento donde el pico de datos es alrededor de 1.5 m/s y su distribución es el rango de 0.7 m/s a 4.7 m/s.

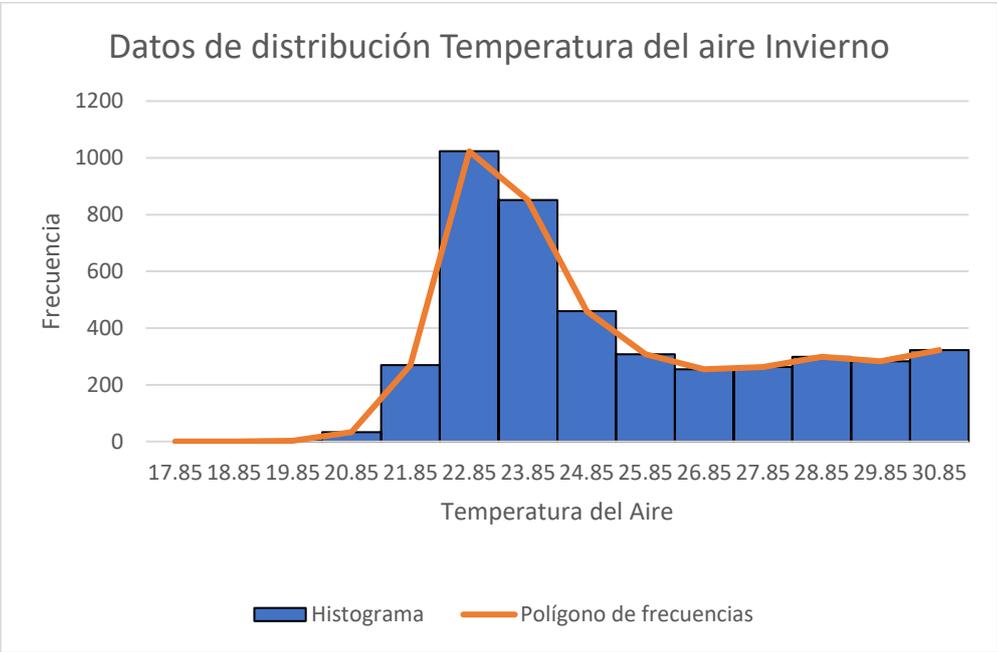


Figura 3.2 Distribución de datos de velocidad de viento en verano. Fuente: autores

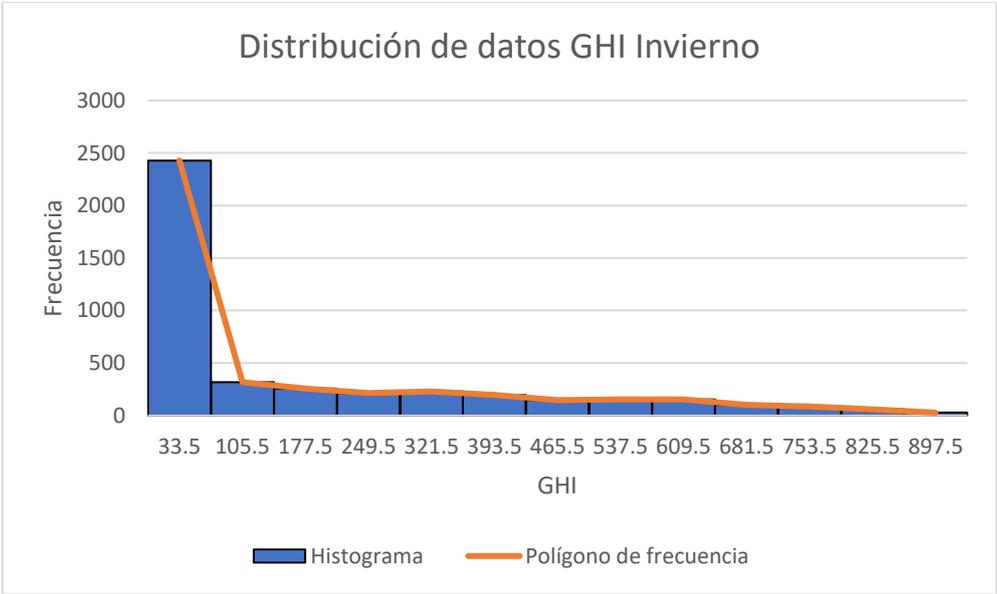


Figura 3.3 Distribución de datos de GHI en invierno. Fuente: autores

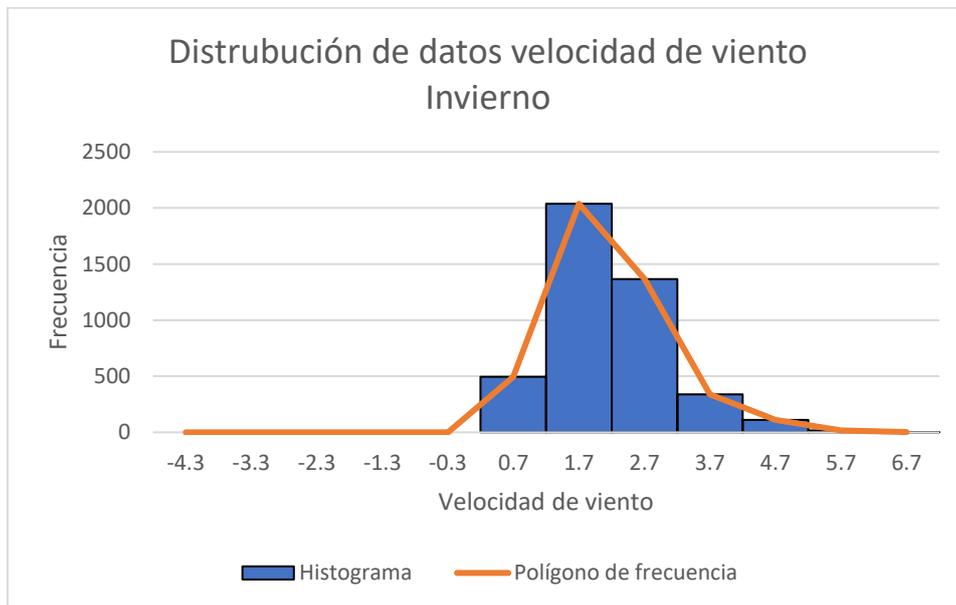


Figura 3.4 Distribución de datos de velocidad de viento en invierno. Fuente: autores

3.2 Resultados K-means

Los hallazgos producto de la implementación de la técnica de ML, k-means, se encuentran resumidos en la Tabla 3.1, dichos valores están escalados en el rango de 0 a 1 en la columna x_{esc} , mientras que la etiqueta x_{real} , representa el valor real representativo obtenido de cada grupo o *clusters*, esto se denomina centroides, 2 de los valores fueron utilizados para la determinación de la ecuación (2.5) y (2.6), en el caso de la potencia de salida DC, se utilizó el valor de GHI más alto, mientras que en el modelamiento de la temperatura de la celda se usó el valor de GHI intermedio que se puede observar en Tabla 3.2, bajo el criterio de que dicha expresión tiene como entrada el escenario menos favorable. Aunque para la consideración de la velocidad del viento, se tuvo en cuenta solo el valor más alto, esto debido a que era una cantidad pequeña y su efecto no fue tan notorio.

Tabla 3.1 Resultados valor más alto k-means

K means				
Valor más alto				
	xesc	xmin	xmax	xreal
Ta	0,48874078	16,2	37,6	26,65905269
GHI	0,54553132	0	962	524,8011298
Vf	0,33154288	0	6,7	2,221337296

Tabla 3.2 Resultados valor de más baja radiación k-means

K means				
Valor de más baja radiación				
	xesc	xmin	xmax	xreal
Ta	0,40646848	16,2	37,6	24,89842547
GHI	0,11940762	0	962	114,8701304
Vf	0,43976514	0	6,7	2,946426438

Una mejor representación se puede visualizar en la gráfica 3D, de la Figura 3.1, donde se muestran los aglomerados arrojados por el algoritmo.

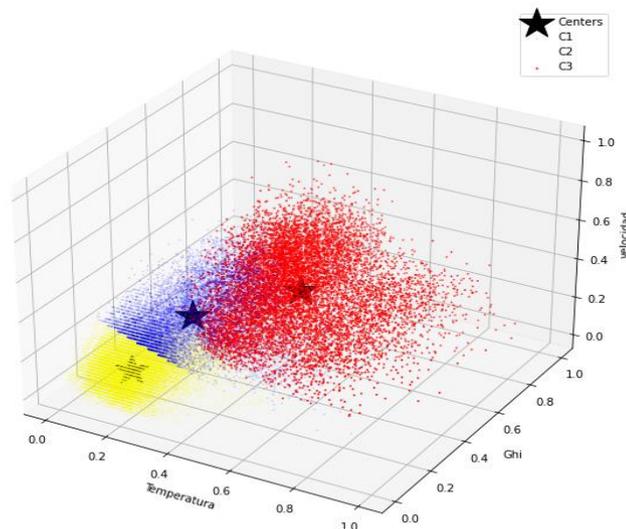


Figura 3.5 Conglomerados obtenidos mediante k-means. Fuente: autores

3.3 Resultados GMM

Como se estipuló en la metodología, uno de los parámetros a determinar previo a la ejecución del algoritmo, era que tipo matriz de covarianza tenía que establecerse como parámetro de entrada de tal forma que en las distribuciones asociadas a los datos, también muestre la matriz de covarianza por variable, para ello se aplicó el criterio BIC, cuyo resultado puede visualizarse en la Figura 3.6, donde se puede ver que la matriz a calcular tenía que ser completa, mientras que el número de mezclas, debería ser 3.

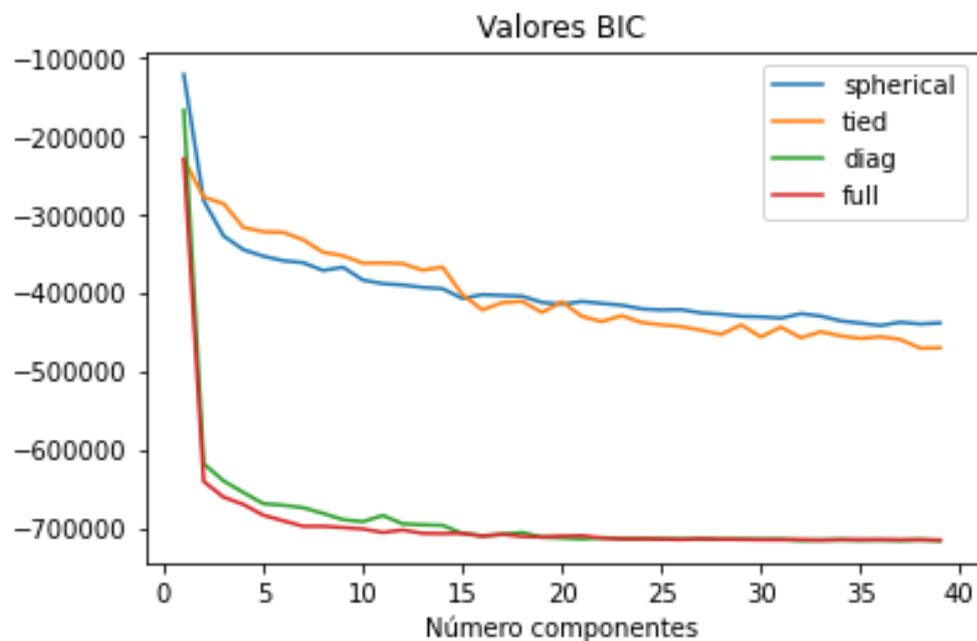


Figura 3.6 Resultados de la aplicación de BIC. Fuente: autores

Una vez aplicado el algoritmo de GMM, los resultados encontrados se pueden observar en la Tabla 3.3 y Tabla 3.4. Hay que tener en cuenta que este modelo realiza asignaciones “suaves”, a diferencia de k-means, es decir aquí presupone que puede existir superposición de mezclas gaussianas, en otras palabras, un elemento puede pertenecer a 2 mezclas, pero con un peso distinto, siendo el peso el factor a tener en cuenta a la hora de saber a cuál mezcla tiene más probabilidad de pertenecer dicho dato, además de obtener como representativo la media de cada una de las distribuciones, también se puede analizar la matriz de covarianza.

Tabla 3.3 Resultados valor más alto GMM

GMM				
	xesc	xmin	xmax	xreal
Ta	0,49640931	16,2	37,6	26,82315923
GHI	0,541511822	0	962	520,9343731
Vf	0,31174008	0	6,7	2,088658536

Tabla 3.4 Resultados valor más bajo GMM

GMM				
Valor de más baja radiación				
	xesc	xmin	xmax	xreal
Ta	0,352554383	16,2	37,6	23,7446638
GHI	0,157096145	0	962	151,1264915
Vf	0,303618204	0	6,7	2,034241967

3.4 Comparación de los 3 enfoques (k-means, GMM, media) en la estimación de potencia.

Antes de calcular la potencia y la temperatura de la celda, fue necesario encontrar los coeficientes a_1 , a_2 y a_3 de la ecuación 2.5, los valores hallados se encuentran referidos en la Tabla 3.5, estos fueron realizados a partir del ajuste de curvas de la Figura 2.11.

Tabla 3.5 Coeficientes de la ecuación de temperatura del módulo

Parámetros	Valor
a_1	1.968059
a_2	15.24194
a_3	0.8676617

Así mismo en la Tabla 3.6 se puede visualizar los resultados del modelamiento de la potencia del módulo y la temperatura del mismo, considerando GHI, Ta, vf, utilizando 3 enfoques, cabe aclarar que en este cuadro se muestra los valores que se obtuvieron, para un día soleado, mientras que en la Tabla 3.7 se puede revisar los valores obtenidos para un día nublado.

Tabla 3.6 Potencia y temperatura del módulo PV de día soleado

Enfoque	Potencia PV [W]	Temperatura PV[°C]
<i>k-means</i>	231.94	32.98
GMM	275.8	35.14
Media	218.71	43.56
Valor medido	262.6	-

Tabla 3.7 Potencia y temperatura del módulo PV de día nublado

Enfoque	Potencia PV [W]	Temperatura PV[°C]
<i>k-means</i>	7.68	30.60
GMM	17.47	32.40
Media	7.37	30.61
Valor medido	25.2	-

Es importante enfatizar, que el valor medido de la potencia PV en la hora pico, se lo encontró, efectuando una división entre el valor que brindó el inversor, para el número de paneles existentes que tiene dicho elemento, el cual fue 30, en otras palabras, la potencia que se obtuvo del inversor fue 7880 W, y el día de extracción de este dato fue el 27 de agosto del 2022, día soleado. En la Tabla 3.8 se puede observar el error de cada uno de los enfoques para este día, y así mismo en la Tabla 3.9 se puede observar para el día nublado que fue el 23 de agosto del 2022, la potencia que mostró el inversor en esta fecha fue de 756 W.

Tabla 3.8 Error relativo en la estimación de la potencia de salida (día soleado)

Enfoque	Error [%]
<i>k-means</i>	11.69
<i>GMM</i>	5
Media	16.734772

Tabla 3.9 Error relativo en la estimación de la potencia de salida (día nublado)

Enfoque	Error [%]
<i>k-means</i>	69.52
<i>GMM</i>	30.67
Media	70.75

El enfoque de GMM, fue el que mejores resultados brindó, por ende, se utilizó ese enfoque para el cálculo del número de *strings*, dicha información se detallará con más precisión en la sección 3.5 luego de mostrar los valores encontrados de las métricas planteadas en la sección 2.1.6.

El factor de planta encontrado, utilizando el GHI parametrizado mediante la técnica de GMM fue de 0.3308, y la capacidad nominal máxima a instalar para la carga de climatización fue de 70.83 kW, en relación con el factor de planta este valor está dentro de lo que se reporta normalmente en instalaciones solares, que es entre el 25 y 50%, para poder afirmar esto se utilizó lo informado por *National Renewable Energy Laboratory (NREL)* en el estudio [62], así como lo detallado en su página oficial de estadísticas sobre la línea de base de tecnología anual [63]

En la Tabla 3.10, se puede visualizar los valores de algunos índices utilizados para la evaluación de la planta solar de 50 kW del edificio 6A, se debe tener en cuenta que tal como se dijo en la introducción de este capítulo, la carga considerada fue la de climatización (aires acondicionados específicamente)

Tabla 3.10 Índices de evaluación utilizados en la evaluación de la instalación solar (día soleado)

Energía diaria producida	Potencia autoconsumida	Potencia importada para cubrir 50% de la carga	Y_{prod}	Y_{irra}	PR
204 kWh	47.28 kW	22.92 kW	0.698	0.3411	0.587

Si el pliego tarifario, establece que el valor a pagar por kWh es de 0.1709 \$ para consumos entre 1001 y 1500 kWh, el dinero que se ahorra diariamente la ESPOL, por la planta solar existente es de 34\$, lo que constituye mensualmente 1045\$, no obstante, esta estimación está sujeta a la variabilidad del recurso solar, bien es cierto que se eligió en enfoque de GMM, pero incluso este método, en días nublados, conlleva un error elevado, en consecuencia, en días donde las horas de mayor irradiación son pocas, estos valores pueden sufrir grandes cambios, por lo que resulta oportuno enfatizar, que para un día nublado, la energía diaria producida es de 71.4 kWh, por lo que el dinero que se ahorraría en estos días sería de 12\$.

3.5 Dimensionamiento de la Instalación fotovoltaica utilizando el mejor enfoque

Tal como se mostró en la sección 3.4, el enfoque con mejores resultados fue el de GMM, lo primero que se tuvo en cuenta fue la cantidad de puntos de máxima potencia (Mpp) que tenía, y el número de entradas por Mpp por inversores, la ventaja que ofrece cada uno de estos, es que puedes cambiar la configuración, acorde a las necesidades, es decir si se requiere que un inversor tenga conectados 3 *strings* de 9 paneles en serie en 2 de ellos, mientras que en el tercero se desea 10 celdas en serie, el *string* diferente debe colocarse en otro Mpp. Teniendo eso claro, la propuesta de este trabajo se detalla a continuación.

Debido a que el voltaje de circuito abierto en condiciones estándar de los paneles es de 38.85 V y el voltaje máximo DC del inversor es de 600 V, en consecuencia, el número máximo de paneles en serie que se puede colocar es de 15, mientras que si se tiene en cuenta que el voltaje de encendido de este mismo elemento es de 120 V, el número mínimo es de módulos a instalar es de 4, de igual manera para determinar el número total de celdas solares que puede tener un inversor, se consideró la potencia máxima del mismo, que era de 11500 W, y la potencia de salida de cada módulo, extraída del modelo GMM.

En virtud de lo expuesto en el párrafo anterior, en la Tabla 3.11, se muestran los datos del diseño existente, así como la propuesta del presente trabajo, así mismo se menciona el área en m^2 que se utilizaría de la disponible en la terraza del edificio en mención.

Tabla 3.11 Diseño existente vs diseño propuesto

Valores del diseño existente				
<i>strings</i>	Paneles en serie	Potencia de 1 inversor	Potencia total	Área utilizada
3	10	8272.35	49634	698.58 m^2
Valores propuestos				
<i>strings</i>	Paneles en serie	Potencia de 1 inversor	Potencia total	Área utilizada
5	8	11029	66.178	743.18 m^2

Resulta pertinente matizar algunas cuestiones: El área disponible en la terraza del edificio 6A, es de 1409 m^2 , y además el modelo existente solo cubre el 35% de la carga de los aire acondicionados, mientras que esta propuesta pretende satisfacer el 50% de

dicho consumo, así mismo, lo que se buscó fue aprovechar la cuarta entrada no utilizada en los Mpp de los 6 inversores, y cambiar la cantidad de paneles por inversor con la finalidad de no comprar otros inversores, en otras palabras se agregó 10 módulos por inversor, por lo que en se tendrían 240 celdas en total(existentes +añadidos). Se deja disponible la utilización de tanto las otras entradas como de los Mpp, para diseños donde se consideren otras variables, hay que tener en cuenta que, por cada Mpp, equivale como tener un inversor distinto, por lo que queda a criterio del ingeniero utilizarlos como considere necesario. Finalmente, también dentro del diseño propuesto, se enfatiza la importancia de la implementación de un *Dashboard*, esto se abordará con más detalles en la siguiente sección.

3.6 Dashboard.

En relación con los resultados monitoreados, utilizados para la propuesta del Dashboard, en la sección 6 de apéndices, se encuentran disponibles estos valores, no se los incluyó en la presente sección, porque dicho trabajo se encuentra en fase de producción por parte del grupo interdisciplinar.

El dashboard presentado en la Figura 3.7 es el bosquejo final que permite visualizar al usuario los resultados extraídos del inversor tiempo real, así como también ayuda a tener información acerca de la energía del último mes, y valores voltaje y corriente del sistema.

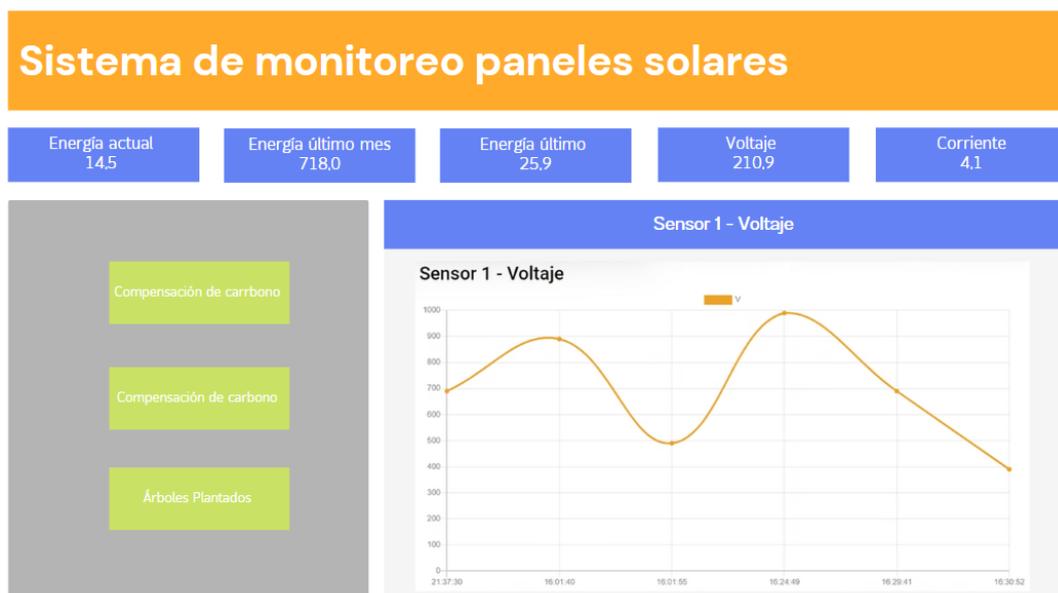


Figura 3.7 Dashboard. Fuente: autores

Otros de los datos presentados en el Dashboard, es la información acerca de la cantidad de árboles que se deben plantar para la absorción de dióxido de carbono en un

determinado lugar equivalente en ahorro que representaría la implementación de la tecnología propuesta PV y el cálculo de las emisiones de CO2 reducidas durante el año.

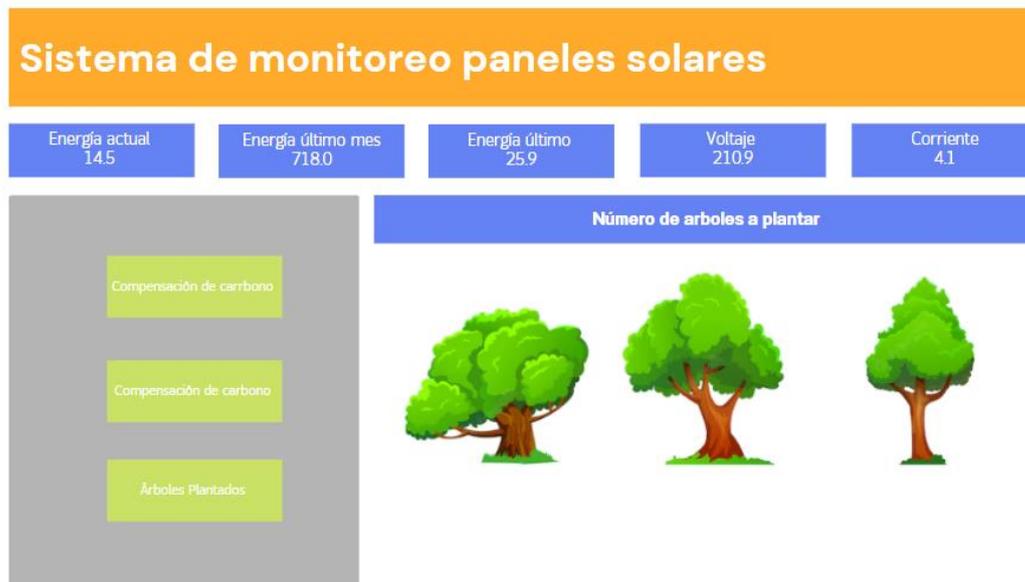


Figura 3.8 Número de árboles a plantar presentado en el dashboard. Fuente: autores

3.6.1 Segmentación de mercado

Para el enfoque de la audiencia se segmentó varios grupos en específico, con el fin de saber con exactitud cuales son los clientes potenciales que requieren del servicio.

La segmentación aplicada en el presente trabajo fue la segmentación psicográfica debido a que permitió separar a un grupo de personas considerando detalles sobre sus intereses y estilos de vida. También se consideró la segmentación por industria ya que ayudó a conocer qué empresas brindan el mismo servicio y qué lo diferencia con el nuestro. Por lo tanto, los clientes potenciales son:

- Clientes comerciales
- Clientes residenciales
- Ambientalistas o ecologistas
- Personas apasionadas de la tecnología

Tanto los clientes comerciales y residenciales requieren disminuir los gastos de energía mensual y llevar un monitoreo de la información y del estado actual del sistema de la

manera más rápida y sin necesidad de realizar bastantes informes. Los ambientalistas y ecologistas son aquellos que desean reducir su huella de carbono y optan por utilizar tecnologías amigables con el medio ambiente y, por último, lo aficionados a la tecnología escogen este tipo de sistemas por innovación y ahorrar su tarifa en los pagos de energía eléctrica.

3.7 Discusión Final.

Uno de los aspectos llamativos del presente trabajo y que puede resultar de interés para el lector, es de porque se decidió satisfacer solo el 50% del consumo, y no se optó por el 100% o al menos un porcentaje mayor, ante la tentativa de pensar que esto está justificado en criterios económicos, cabe matizar que no es así, su fundamento está en el factor de planta obtenido, el cual fue presentado en la sección 3.5, y esto fue de gran relevancia debido a que afecta la ecuación que propone la regulación 003/18 de la Arconel, para la capacidad máxima a instalar, en consecuencia la potencia a instalar no puede sobrepasar 70.88 kW, y esto depende del recurso solar disponible en dicha región, por lo que si se colocaban más inversores, no se cumplía este requerimiento, y si se observa el 50% de la carga de los aires acondicionados representa un valor aproximado a 70 kW.

Lo descrito en el párrafo anterior permite al lector atisbar que cualquier mejora que se desee implementar necesariamente tiene que estar centrada en un mejor aprovechamiento del GHI, teniendo en cuenta que la técnica de GMM brindó una aproximación con un error bajo, lo cual representa una ventaja para planteamientos más profundos de problemas de optimización o de mejoras tecnológicas. De igual forma los índices de producción, de referencia y *Performance Ratio* confirma lo observado en el párrafo anterior, dichos valores deberían estar comprendidos entre 0.8 y 1, esto no ocurrió, para contrastar ello referirse a las Tabla 3.10 y Tabla 3.11, trabajos similares no se han realizado para dicha instalación, por lo que esto permite enfatizar la importancia de adoptar algunos aspectos esenciales de la metodología aquí recogida para implementaciones similares en otros edificios.

Recapitulando, lo expuesto aquí confirma que un modelo basado en minería de datos brinda mayor precisión a la hora de realizar este tipo de instalaciones, donde la variabilidad del recurso incide en la producción de energía, quedan muchos aspectos

que a evaluar, pero se ha tratado de abordar los puntos esencialmente importantes reportados en el examen de literatura efectuado, de seguro esto está sujeto a cambios, pero es bastante seguro que la línea aquí planteada permitirá mejorar muchos aspectos, teniendo cuenta las deficiencias que presentan los cálculos tradicionales y como no responden a las exigencias que plantea el reto aprovechar al máximo el recurso solar, con la finalidad de mitigar los impactos del cambio climático cuyas consecuencias hoy en día resultan cada vez más evidentes.

CAPÍTULO 4

4. CONCLUSIONES Y RECOMENDACIONES

4.1 CONCLUSIONES

En función del trabajo desarrollado, del análisis y la discusión realizada en sus respectivos apartados, sobre la evaluación y propuesta diseño de la instalación fotovoltaica del edificio 6A de la ESPOL, se concluye que:

- Se determinó los valores representativos de 3 parámetros meteorológicos, mediante el empleo 2 técnicas de ML: k-means y GMM. Dichos enfoques permitieron el estudio del recurso solar disponible, así como también obtener los parámetros de entrada para la evaluación de las ecuaciones de temperatura de la celda y la salida de potencia DC del módulo PV, el contraste con los valores reales medidos extraídos sirvió para realizar la comparación con el enfoque tradicional.
- La utilización del método tradicional, mediante el cálculo de la media, de los valores de los 10 años evidenció tener un error bastante considerable, en comparación con las técnicas de ML antes mencionadas, especialmente frente al enfoque de GMM, esta afirmación está fundamentada en los errores reportados para la estimación de la salida de la potencia DC de un módulo fotovoltaico, por lo que se puede argumentar que las técnicas de inteligencia artificial, muestran mejores beneficios en términos de precisión a la hora de obtener un diagnóstico en índices de evaluación, en los diseños fotovoltaicos como son: horas de producción, factor de planta, capacidad nominal a instalar, *performance ratio*.
- La capacidad nominal máxima a instalar en el edificio 6 A, manteniendo las características tanto de los inversores como de los módulos PV fue de 70.83 kW, debido al factor de planta reportado con el valor representativo de GHI, el cual fue de 0.33, esto mostró la importancia de aprovechar mejor el recurso solar, e implementar mejoras más profundas con la finalidad de conseguir mejorar el factor de planta, de tal forma que garantice subir la capacidad nominal máxima a instalar.
- Debido a la limitación reportada de la restricción que plantea la Arconel como capacidad máxima a instalar, la propuesta a implementar, sin la necesidad de

comprar más inversores, fue la instalación de 60 módulos PV adicionales, garantizando así el aprovechamiento de las otras entradas de los Mpp de cada inversor, y logrando así cubrir el 50% del consumo de los aires acondicionados.

- En el trabajo conjunto realizado con el grupo interdisciplinar, se dejó en fase de producción la plataforma de monitoreo de la instalación fotovoltaica, sin embargo, la conexión de manera remota para la extracción de mediciones reales garantizó la evaluación de la planta solar. Finalmente se dejó estipulado la propuesta de dos pantallas de Dashboard, para clientes de tipo: residencial y comercial, los valores presentados en dichas pantallas fueron valores reales obtenidos de la extracción de datos antes mencionada.

4.2 Recomendaciones.

La línea argumentativa de este trabajo de titulación ha evidenciado: la dificultad y las exigencias que se presentan a la hora de abordar el diseño de este tipo de instalaciones. A la luz de las fuentes bibliográficas actualizadas disponibles, y sin caer en una abstracción que dificulte la comprensión de las implicaciones prácticas que estas dificultades tienen en el día a día del ingeniero eléctrico, se buscó consolidar una guía práctica, bien es cierto que los hallazgos permitieron inferir conclusiones de interés y en función de estas observaciones, se recomienda lo siguiente.

- La precisión de la información que presenta Solcast fue un aspecto no discutido en este trabajo, y teniendo en cuenta que esta base de datos es satelital global, la obtención de mediciones de parámetros meteorológicos mediante aparatos remotos podría garantizar una mejor aproximación a los resultados, o garantizar una discusión más profunda sobre el efecto de la variabilidad del recurso solar en la ciudad de Guayaquil.
- Se sugiere para implementaciones futuras de este tipo de instalaciones, la consideración del factor de planta aquí reportado, y como tecnologías más sofisticadas podrían garantizar aumentar este valor, de tal forma que permita satisfacer un consumo más elevado, y se considera utilizar la misma tecnología, tener en cuenta un modelo de optimización no centrado en minimizar el costo de la instalación sino en maximizar el factor de planta y la capacidad nominal máxima a instalar.

- En función de una de las conclusiones a las cuales se llegó en este trabajo, resulta oportuno sugerir la integración en alguna interfaz gráfica los pasos aquí planteados para que cualquier ingeniero pueda adoptar esta metodología, es necesario reparar, que lo observado hasta ahora y validado con valores reales, permitieron obtener un diagnóstico de mayor precisión, y teniendo en cuenta que existieron 2 etapas, al menos la etapa de evaluación mostró resultados que pueden resultar de bastante interés a las exigencias que se plantea a la hora de abordar estos diseños.

5. BIBLIOGRAFÍA

- [1] IPCC, «Calentamiento global de 1,5 °C,» IPCC, Ginebra, 2018.
- [2] S. VijayaVenkataRaman, S. Iniyar y R. Goic, «A review of climate change, mitigation and adaptation,» *Renewable and Sustainable Energy Reviews*, pp. 878-897, 2012.
- [3] CELEC EP, «Ecuador actualiza su Plan Maestro de Electricidad para impulsar inversiones en Energía Renovables no Convencionales por cerca de USD 2.200 millones,» 2021. [En línea]. Available: <https://www.celec.gob.ec/gensur/index.php/553-ecuador-actualiza-su-plan-maestro-de-electricidad-para-impulsar-inversiones-en-energias-renovables-no-convencionales-por-cerca-de-usd-2-200-millones>.
- [4] Gobierno del Encuentro, «Ministerio de Energía y Minas:Noticias: Ecuador reafirma su compromiso de cumplir con el “Acuerdo de París”, en el marco de la Conferencia de las Naciones Unidas sobre el Cambio Climático,» 14 septiembre 2021. [En línea]. Available: <https://www.recursosyenergia.gob.ec/ecuador-reafirma-su-compromiso-de-cumplir-con-el-acuerdo-de-paris-en-el-marco-de-la-conferencia-de-las-naciones-unidas-sobre-el-cambio-climatico/>.
- [5] ESPOL, «espol:noticias:Representantes de Mexichmen Ecuador recorrieron planta fotovoltaica de Espol,» 12 Junio 2019. [En línea]. Available: <https://www.espol.edu.ec/noticias/representantes-de-mexichem-ecuador-recorrieron-planta-fotovoltaica-de-espol>.
- [6] T. Huld, E. Dunlop, H. Georg Beyer y R. Gottschalg, «Data sets for energy rating of photovoltaic modules,» *Solar Energy*, pp. 267-279, 2013.
- [7] C. Cañete Torralvo, «Modelos para la carecterización eléctrica de módulos fotovoltaicos en condiciones de sol real,» Universidad de Málaga, Málaga, 2015.
- [8] A. C. Ayala Gilardón, «Modelos para la evaluación y optimización de instalaciones fotovoltaicas de autoconsumo,» Universidad de Málaga, Málaga, 2019.
- [9] A. Ansuategi y J. Delgado, «Green Energy and Efficiency: An Economic Perspective,» Springer, Switzerland, 2015.

- [10] Gobierno del Encuentro, «Gobierno del Encuentro: Ministerio del Ambiente, Agua y Transición Ecológica: Comunicamos: noticias: Ecuador suscribe Acuerdo de París sobre cambio climático,» 26 Julio 2016. [En línea]. Available: <https://www.ambiente.gob.ec/ecuador-suscribe-acuerdo-de-paris-sobre-cambio-climatico/>.
- [11] Ministerio de Energía y Recursos Naturales No Renovables, Plan Maestro de Electricidad, Quito: Ministerio de Energía y Recursos Naturales No Renovables, 2018.
- [12] J. Wirfs-Brock, «Inside Energy: Lost in Transmission: How Much Electricity Disappears Between A Power Plant And Your Plug?,» 6 Noviembre 2015. [En línea]. Available: <http://insideenergy.org/2015/11/06/lost-in-transmission-how-much-electricity-disappears-between-a-power-plant-and-your-plug/>.
- [13] G. F. Velasco y E. Cabrera, «Generación solar fotovoltaica dentro del esquema de generación distribuida para la provincia de Imbabura.,» *V Seminario Euro Latinoamericano de Sistemas de Ingeniería (FIEE)*, 2009.
- [14] G. H. Bauer, Photovoltaic Solar Energy Conversion, Berlin: Springer, 2015.
- [15] R. Foster, M. Ghassemi y A. Cota, SOLAR ENERGY: Renewable Energy and the Environment, Cleveland: CRC Press, 2010.
- [16] T. Stoffel, Solar Resources and Forecasting Group, National Renewable Energy Laboratory, Elsevier, 2013.
- [17] S. Sumathi, L. Ashok Kumar y P. Surekha, Solar PV and Wind Energy Conversion Systems: An Introduction to Theory, Modeling with MATLAB/SIMULINK , and the Role of Soft Computing Techniques, Switzerland: Springer, 2015.
- [18] A. Mcevoy, T. Markvart y L. Castañer, Practical Handbook of photovoltaics: Fundamentals and Applications, Kidlington: Elsevier, 2012.
- [19] R. Zilles, W. Negrão, M. Barros y S. Ferreira de Oliveira , Sistemas fotovoltaicos conectados à rede elétrica, São Paulo: oficina de textos, 2009.
- [20] S. Odeh, «Analysis of the Performance Indicators of the PV Power System,» *Journal of Power and Energy Engineering*, vol. 6, pp. 59-75, 2018.
- [21] A. Luque y S. Hegedus, Handbook of Photovoltaic Science and Engineering, Chichester: WILEY, 2003.

- [22] R. M. Rocha Val de Ovelha, «Projeto, Dimensionamento e Instalação de Solução Fotovoltaica numa moradia offgrid,» Ciências ULisboa, Lisboa, 2017.
- [23] G. Kwinjo, T. Munetsiwa y R. Arora, «A Review of the Existing Solar Irradiation Databases,» *IJESC*, pp. 15306-15309, 2017.
- [24] A. C. Müller y S. Guido, *Introduction to Machine Learning with Python: A Guide For Data Scientists*, Sebastapol: O'Reilly, 2016.
- [25] E. Alpaydm, *Introduction to Machine Learning*, London: The MIT Press, 2010.
- [26] D. Reynolds, «Gaussian Mixture Models,» 2022.
- [27] B. P. y. R. J. R. Singh, «Representación estadística de las cargas del sistema de distribución mediante el modelo de mezcla gaussiana,» Febrero 2010. [En línea]. Available: <https://ieeexplore.ieee.org/abstract/document/5298967>.
- [28] C. L. D.-B. L. T. D. C. P. y. O. H. Poissant Y, «Simple test methods for evaluating the energy,» Gothenburg, Sweden, 2003.
- [29] C. Osterwald, «Translation of device performance measurements to reference,» USA, 1985, pp. 269-279.
- [30] R. W. Taylor, «System and module rating: Advertised versus actual capability,» USA, 1985, pp. 335-344.
- [31] E. Skoplaki, A. Boudouvis y J. Palyvos, «A simple correlation for the operating temperature of photovoltaic modules of arbitrary mounting,» *Solar Energy Materials and Solar Cells*, vol. 92, 2008.
- [32] B. K. Farmer, *Model Technical Specifications*, 1995.
- [33] E. K. M. D. y. M. L. Kroposki B, «A comparison of photovoltaic module performance evaluation methodologies for energy ratings.,» 1994, pp. 858-862.
- [34] D. B. y. D. E. Anderson J, «Energy rating of photovoltaic modules,» 2000 , pp. 2087-2091.
- [35] B. W. E. y. K. J. King D. L, *Photovoltaic array performance model.*, 2004 .
- [36] A. M. A. y. D. P. D. Massi Pavan, de *Solar Energy*, 2011, pp. 1128-1136.
- [37] E. y. v. D. E. Meyer, «Development of energy model based on total daily irradiation and maximum ambient temperature.,» de *Renewable Energy* , 2000, pp. 37-47.

- [38] SOLCAST, 2022. [En línea]. Available: <https://solcast.com/historical-and-tmy/inputs-and-algorithms/>.
- [39] J. Constante Segura y E. Palacios Chacón, «El RECURSO SOLAR PARA GENERACIÓN DE ENERGÍA ELÉCTRICA: Análisis para el Distrito Metropolitano de Quito,» Universidad Politécnica Salesiana, Cuenca, 2014.
- [40] H. Ramírez Murillo, C. A. Torres Pinzón y E. F. Forero García, «Estimación del potencial fotovoltaico mediante minería de datos en cuatro ciudades de Colombia,» *TecnoLógicas*, vol. 22, nº 46, pp. 77-97, 2019.
- [41] V. Heidrich-Meisner y R. Wimmer-Schweingruber, «Solar Wind Classification Via K-Means Clustering Algorithm,» *Physics, Computer Science*, 2018.
- [42] J. D. Salinas-González, A. García-Hernández, D. Riveros-Rosas, G. Moreno-Chávez, L. F. Zarzalejo, J. Alonso-Montesinos, C. E. Galván-Tejada, A. Mauricio-González y A. E. González-Cabrera, «Multivariate Analysis for Solar Resource Assessment Using Unsupervised Learning on Images from the GOES-13 Satellite,» *Remote Sens*, pp. 1-15, 2022.
- [43] M. Wahbah, T. EL-Fouly y B. Zahawi, «Gaussian Mixture Model for Estimating Solar Irradiance Probability Density,» de *2020 IEEE Electric Power and Energy Conference(EPEC)*, 2020.
- [44] K. Benmouiza y A. Cheknane, «Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models,» *Energy Conversion and Management*, pp. 561-569, 2013.
- [45] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller y O. Grisel, «API design for machine learning software: experiences from the scikit-learn project,» *Languages for Data Mining and Machine Learning*, pp. 108-122, 2013.
- [46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion y O. Grisel, «Scikit-learn: Machine Learning in Python,» *Microtome*, 2011, pp. 537-539.
- [47] A. Mehrjou, R. Hosseini y B. Nadjar Araabi, «Improved Bayesian Information criterion for mixture model selection,» *Pattern Recognition Letters*, pp. 22-27, 2016.
- [48] B. Cédric, C. Housmans, J. Leloux y M. Journée, «Solar irradiation from the energy production of residential PV systems,» *Renewable Energy*, vol. 125, pp. 306-318, 2018.

- [49] J. Carrillo, J. Muñoz, A. Makibar, A. Luna y L. Navarte, «SISIFO: THE OPEN-SOURCE SIMULATION TOOL OF PV SYSTEMS DEVELOPED IN PVCROPS,» *Instituto de Energía Solar-Universidad Politécnica de Madrid*, pp. 1-4, 2015.
- [50] C. Schwingshackl, M. Petitta, J. Wagner, G. Belluardo, D. Moser, M. Castelli, M. Zebisch y A. Tetzlaff, «Wind effect on PV module temperature: Analysis of different techniques for an accurate estimation,» *Energy Procedy*, vol. 40, pp. 77-86, 2013.
- [51] D. Loveday y A. Taki, «Convective heat transfer coefficients at a plane surface on a full-scale building facade,» *International Journal of Heat and Mass Transfer*, vol. 39, nº 8, pp. 1729-1742, 1996.
- [52] M. Fuentes, G. Nofuentes, J. Aguilera, D. Talavera y M. Castro, «Application and validation of algebraic methods to predict the behaviour of crystalline silicon PV modules in Mediterranean climates,» *Solar Energy*, vol. 81, nº 11, pp. 1396-1408, 2007.
- [53] Arconel, «Resolución 053/18,» Quito, 2019.
- [54] R. Castro, *Energias Renováveis e Produção Descentralizada - Introdução à Energia Fotovoltaica*, vol. 2008, 2002, pp. 1-48.
- [55] E. da Silva João, «Metodologia de dimensionamento para instalações fotovoltaicas em autoconsumo,» Técnico Lisboa, 2016.
- [56] H. Ramos López y R. Luna Puente, «Diseño de un sistema fotovoltaico integrado a la red para el área de estacionamiento de la Universidad Tecnológica de Salamanca,» Centro de Investigación en Materiales Avanzados, S.C, Chihuahua, 2014.
- [57] M. Georg Kratzenberg, E. Martins Deschamps, L. Nascimento, R. Rüter y H. Helmut Zürn, «Optimal Photovoltaic Inverter Sizing Considering Different Climate Conditions and Energy Prices,» *Energy Procedia*, pp. 226-234, 2014.
- [58] UNE: Normalización Española, «UNE-EN-IEC 61724-1:2022,» [En línea]. Available: <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0068724>. [Último acceso: 21 agosto 2022].
- [59] V. Castro Landínez, «Evaluación del desempeño del sistema solar fotovoltaico implementado en el Edificio Santo Domingo de la Universidad de Los Andes mediante un modelo computacional,» Universidad de Los Andes, Bogotá, 2020.

- [60] R. H.-L. J. M. K. D. a. G. S. J. Litardo, «Estimating energy consumption and conservation measures for ESPOL Campus main building model using EnergyPlus,» 2019.
- [61] C. Gueymard, «High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – Validation with a benchmark dataset,» de *Solar Energy* , 2008, pp. 272-285.
- [62] B. Sigrin, P. Sullivan, E. Ibanez y R. Margolis, «Representation of the Solar Capacity Value in the ReEDS Capacity Expansion Model,» de *4th IEEE Photovoltaic Specialists Conference*, Denver, 2014.
- [63] National Renewable Energy Laboratory(NREL), «NREL: Annual Technology Baseline,» 2022. [En línea]. Available: <https://atb.nrel.gov/>. [Último acceso: 27 Agosto 2022].
- [64] W. Zheng, I. Koprinska y M. Rana, «Clustering Based Methods for Solar Power Forecasting,» de *International Joint Conference on Neural Networks (IJCNN)*, Vancouver, 2016.
- [65] E. y. P. J. Skoplaki, «On the temperature dependence of photovoltaic,» 2009, pp. 614-624.
- [66] F. Martínez-Moreno, E. Lorenzo, L. Narvarte, R. Moretón, N. Tyutyundzhiev y I. De La Parra, «TECHNICAL SPECIFICATIONS AND QUALITY CONTROL PROCEDURES FOR REDUCING THE UNCERTAINTY IN PV INSTALLATIONS: RESULTS OF THE FP7 PROJECT PVCROPS,» de *31st European Photovoltaic Solar Energy Conference and Exhibition*, 2015.
- [67] D. Dirnberger, B. Müller y C. Reise, «On the uncertainty of energetic impact on the yield of different PV technologies,» *Solar Energy*, pp. 82-96, 2015.
- [68] T. Díaz Corcobado y G. Carmona Rubio, *Instalaciones solares fotovoltaicas*, McGraw-Hill Interamericana de España S.L, 2010.
- [69] «Ley de eficiencia energética N° 449,» 2019.
- [70] Solar Bluesun, julio 2019. [En línea]. Available: <https://www.almaden.com.au/wp-content/uploads/2019/07/Bluesun-Solar-Panel-Datasheet.pdf>. [Último acceso: 8 agosto 2022].

6. APÉNDICES

6.1 Código K-means

```
reset-f
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
#import warnings
#warnings.filterwarnings('ignore')
import os
dataset = pd.read_csv('10añoscompletos.csv')
dataset.shape
dataset.head

X= dataset
from sklearn.preprocessing import LabelEncoder
X.info()
cols = X.columns
from sklearn.preprocessing import MinMaxScaler
ms = MinMaxScaler()
X = ms.fit_transform(X)
X = pd.DataFrame(X, columns=[cols])
X.head()
from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=3, random_state=0)
kmeans.fit(X)
kmeans.cluster_centers_
kmeans.inertia_
labels = kmeans.labels_

from sklearn.cluster import KMeans
cs = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters = i, init = 'k-means++', max_iter = 300, n_init = 10, random_state = 0)
    kmeans.fit(X)
    cs.append(kmeans.inertia_)
plt.plot(range(1, 11), cs)
plt.title('The Elbow Method')
plt.xlabel('Number of clusters')
```

```

plt.ylabel('CS')
plt.show()

from sklearn.cluster import KMeans

kmeans = KMeans(n_clusters=2,random_state=0)

kmeans.fit(X)

label = kmeans.fit_predict(X)

labels = kmeans.labels_
y_clusters = kmeans.fit_predict(X)

filtered_label0 = dataset[label == 0]
print(label)
##-----
kmeans = KMeans(n_clusters=2)           # Number of clusters == 3
kmeans = kmeans.fit(X)                 # Fitting the input data
labels = kmeans.predict(X)             # Getting the cluster labels
centroids = kmeans.cluster_centers_    # Centroid values
print("Centroids are:", centroids)

import matplotlib.pyplot as plt
from sklearn.datasets import make_swiss_roll
from mpl_toolkits.mplot3d import Axes3D
from sklearn.cluster import KMeans

fig = plt.figure(figsize=(10,10))
ax = fig.gca(projection='3d')

x = np.array(labels==0)
y = np.array(labels==1)
z = np.array(labels==2)

#ax.scatter(centroids[:,0],centroids[:,1],centroids[:,2],c="black",s=1000,marker= "*",label="Centers",alpha=1)
ax.scatter(centroids[:,0],centroids[:,1],c="black",s=1000,marker= "*",label="Centers",alpha=1)

ax.scatter(X.values[x,0],X.values[x,1],X.values[x,2],c="blue",s=10,label="C1",alpha=0.1)
ax.scatter(X.values[y,0],X.values[y,1],X.values[y,2],c="yellow",s=10,label="C2",alpha=0.1)
#ax.scatter(X.values[z,0],X.values[z,1],X.values[z,2],c="red",s=10,label="C3")
ax.legend()
ax.set_xlabel('Temperatura')
ax.set_ylabel('Ghi')
ax.set_zlabel('velocidad')
print("Centroids are:", centroids)

centroids = kmeans.cluster_centers_    # Centroid values
etiquetas = kmeans.labels_

```

6.2 GMM

```
reset-f
import numpy as np
import pandas as pd
#from matplotlib import loadmat
from sklearn.datasets import make_blobs
import matplotlib.pyplot as plt
from matplotlib import style
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import warnings
from sklearn.mixture import GaussianMixture
from sklearn.pipeline import make_pipeline
from sklearn.preprocessing import StandardScaler
Xg=pd.read_csv("10añoscompletos.csv")
cols = Xg.columns
from sklearn.preprocessing import MinMaxScaler
ms = MinMaxScaler()
Xg = ms.fit_transform(Xg)
Xg = pd.DataFrame(Xg, columns=[cols])
Xg.head()
Xg.columns = ["col_" + str(i) for i in Xg.columns]
modelo_gmm = GaussianMixture(
    n_components = 3,
    covariance_type = 'full',
    random_state = 123
)

modelo_gmm.fit(X=Xg)

## Gráfico de la covarianza
fig, ax = plt.subplots(figsize=(6, 3.84))

n_components = range(1, 40)
covariance_types = ['spherical', 'tied', 'diag', 'full']

for covariance_type in covariance_types:
    valores_bic = []

    for i in n_components:
        modelo = GaussianMixture(n_components=i, covariance_type=covariance_type)
        modelo = modelo.fit(Xg)
        valores_bic.append(modelo.bic(Xg))

    ax.plot(n_components, valores_bic, label=covariance_type)

ax.set_title("Valores BIC")
ax.set_xlabel("Número componentes")
ax.legend();

modelo_gmm = GaussianMixture(
    n_components = 3,
    covariance_type = 'full',
    random_state = 123
)
modelo_gmm.fit(X=Xg)

### predicción de anomalías
#log_probabilidad_predicha = modelo_gmm.score_samples(X=Xg)
#log_probabilidad_predicha
#fig, ax = plt.subplots(nrows=1, ncols=1, figsize=(7, 3.5))
#sns.distplot(
#    log_probabilidad_predicha,
#    hist = False,
#    rug = True,
#    color = 'blue',
#    kde_kws = {'shade': True, 'linewidth': 1},
#    ax = ax
#)

#ax.set_title('Distribución predicciones')
#ax.set_xlabel('Logaritmo densidad de probabilidad');
modelo_gmm.means_
modelo_gmm.covariances_
probabilidades = modelo_gmm.predict_proba(Xg)
probabilidades
# Clasificación (asignación a la componente de mayor probabilidad)
# =====
# Cada fila es una observación y cada columna la probabilidad de pertenecer a
```

```

# cada una de las componentes.
clasificacion = modelo_gmm.predict(Xg)
clasificacion

#####

# Representación gráfica
# =====
# Código obtenido de:
#
fig, ax = plt.subplots(1, 1, figsize=(6, 3.84))
ax.scatter(
    x = Xg.values[:, 0],
    y = Xg.values[:, 1],
    z = Xg.values[:,2]),
    c = 'blue',
    marker = 'o',
    edgecolor = 'red',
)
ax.set_title('Datos simulados');

import numpy as np
import pandas as pd
from sklearn.datasets import make_blobs

# Gráficos
# =====
import matplotlib.pyplot as plt
import matplotlib as pl
from matplotlib.patches import Ellipse
from matplotlib import style
style.use('ggplot') or plt.style.use('ggplot')

# Preprocesado y modelado
# =====
from sklearn.mixture import GaussianMixture
from sklearn.preprocessing import scale
from sklearn.metrics import silhouette_score

# Configuración warnings
# =====
import warnings
warnings.filterwarnings('ignore')

def make_ellipses(gmm, ax):
    for n in range(gmm.n_components):
        if gmm.covariance_type == 'full':
            covariances = gmm.covariances_[n]
        elif gmm.covariance_type == 'tied':
            covariances = gmm.covariances_
        elif gmm.covariance_type == 'diag':
            covariances = np.diag(gmm.covariances_[n])
        elif gmm.covariance_type == 'spherical':
            covariances = np.eye(gmm.means_.shape[1]) * gmm.covariances_[n]
        v, w = np.linalg.eigh(covariances)
        u = w[0] / np.linalg.norm(w[0])
        angle = np.arctan2(u[1], u[0])
        angle = 180 * angle / np.pi # convert to degrees
        v = 2. * np.sqrt(2.) * np.sqrt(v)

        for i in range(1,3):
            ell = plt.patches.Ellipse(gmm.means_[n], i*v[0], i*v[1],
                                     180 + angle, color="blue")
            ell.set_clip_box(ax.bbox)
            ell.set_alpha(0.1)
            ax.add_artist(ell)

fig, axs = plt.subplots(1, 2, figsize=(12, 3.84))

# Distribución de probabilidad de cada componente
for i in np.unique(clasificacion):
    axs[0].scatter(
        x = Xg.values[clasificacion == i, 0],
        y = Xg.values[clasificacion == i, 1],
        c = plt.rcParams['axes.prop_cycle'].by_key()['color'][i],
        marker = 'o',
        edgecolor = 'black',
        label= f"Componente {i}"
    )

```

```

make_ellipses(modelo_gmm, ax = axs[0])
axs[0].set_title('Distribución de prob. de cada componente')
axs[0].legend()

# Distribución de probabilidad del modelo completo
xs = np.linspace(min(Xg.values[:, 0]), max(Xg.values[:, 0]), 1000)
ys = np.linspace(min(Xg.values[:, 1]), max(Xg.values[:, 1]), 1000)
xx, yy = np.meshgrid(xs, ys)
scores = modelo_gmm.score_samples(np.c_[xx.ravel(), yy.ravel()], )
axs[1].scatter(Xg.values[:, 0], Xg.values[:, 1], s=5, alpha=.6, c=plt.cm.tab10(clasificacion))
scores = np.exp(scores) # Las probabilidades están en log
axs[1].contour(
    xx, yy, scores.reshape(xx.shape),
    levels=np.percentile(scores, np.linspace(0, 100, 10))[1:-1]
)
axs[1].set_title('Distribución de prob. del modelo completo');

```

6.3 Mediciones del Inversor Solis-1P10K-4G-US

Dado que la cantidad de mediciones fue de 3744, por inversor, se adjunta un acceso a los 3 archivos.csv, en caso de requerir los datos por parte del lector, tenga en cuenta que los 6 inversores fueron de iguales características, por lo que solo se extrajo los datos de 3 de ellos, sin embargo, la información necesaria de la marca, placa y datasheet del tipo de inversor existente se puede revisar la Figura 2.13 así como la referencia [64].

Tabla 6.1 Acceso a las mediciones del inversor

Mediciones	Link:
Datos de los tres inversores	https://drive.google.com/drive/folders/122_BTKmMi-T6agDJlwkp0PJ_9EAENCIO?usp=sharing