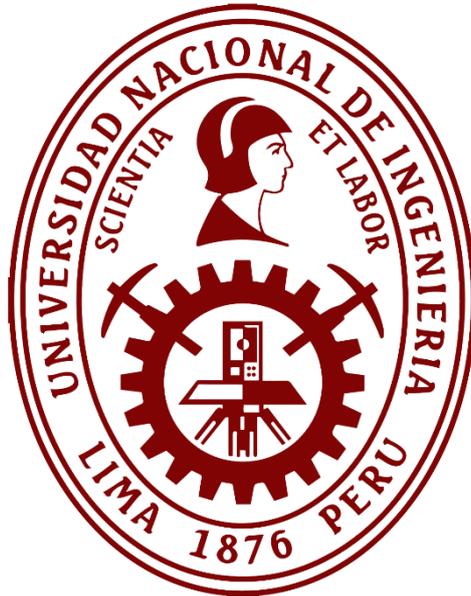


**UNIVERSIDAD NACIONAL DE INGENIERÍA**  
**FACULTAD DE INGENIERÍA MECÁNICA**



**TESIS**

**“OPTIMIZACIÓN DE UN SISTEMA FOTOVOLTAICO HÍBRIDO Y  
LA PREDICCIÓN DE LA DEMANDA ENERGÉTICA Y  
VARIABLES CLIMÁTICAS UTILIZANDO LA INTELIGENCIA  
ARTIFICIAL”**

PARA OBTENER EL GRADO ACADÉMICO DE DOCTOR EN CIENCIAS CON  
MENCIÓN EN ENERGÉTICA

ELABORADO POR:

**MSc. JHON JHONATHAN PEÑALVA SANCHEZ**

ASESOR:

**Dr. JOSÉ CARLOS DANIEL ALVAREZ MERINO**

**LIMA, PERÚ**

**2024**

### **DEDICATORIA:**

La presente tesis doctoral está dedicada a mis abuelos, padres y hermanos y por su amor, motivación y apoyo que me brindaron durante todo mi camino académico y en el desarrollo de esta investigación.

En memoria de mi abuela, Epifania Chuquichampi Melo, cuyo apoyo incondicional y amor han sido una constante inspiración en mi vida y en mi formación académica. Dedico esta tesis doctoral como tributo a su memoria.

### **AGRADECIMIENTO:**

Al Dr. José Carlos Daniel Álvarez Merino, le expreso mi agradecimiento sincero por su motivación y orientación constante durante todo el desarrollo de esta investigación. Así mismo, un agradecimiento al Centro de Energía Renovable (CER) de la Universidad Nacional de Ingeniería por las capacitaciones brindadas durante el desarrollo de la investigación.

# ÍNDICE

ÍNDICE DE FIGURA .....	7
ÍNDICE DE TABLAS.....	10
RESUMEN .....	12
ABSTRACT .....	13
INTRODUCCIÓN.....	14
CAPÍTULO I – PLANTEAMIENTO DEL PROBLEMA.....	16
1.1 Justificación .....	16
1.2 Planteamiento del problema.....	17
1.2.1 Planteamiento del problema general .....	17
1.2.2 Planteamiento del problema específico.....	17
1.3 Objetivo general y específico.....	18
1.3.1 Objetivo general .....	18
1.3.2 Objetivo específico.....	18
1.4 Hipótesis .....	19
1.5 Motivación .....	19
CAPÍTULO II – MARCO TEÓRICO.....	21
2.1 Antecedentes bibliográfico .....	21
2.1.1 Pronóstico del recurso renovable y estrategia de operación de los sistemas renovables híbridos .....	21
2.1.1 Optimización en el dimensionamiento y operación del sistema híbrido fotovoltaico .....	27
2.2 Las energías renovables y la inteligencia artificial .....	35
2.2.1 Las energías renovables en el mundo.....	35
2.2.2 Estado de las energías renovables en el Perú .....	39
2.2.3 La inteligencia artificial y las energías renovables .....	40
2.3 Bases conceptuales .....	43
2.3.1 Recurso renovable y modelos de pronóstico.....	43
2.3.2 Sistema de generación híbridos.....	44
CAPÍTULO III – METODOLOGÍA.....	47
3.1 Tipo de investigación.....	47
3.2 Nivel de investigación .....	47
3.3 Diseño de investigación .....	47
3.4 Fuentes de recolección de datos.....	51
3.5 Población y muestra.....	52

3.6 Técnicas de procesamiento de datos .....	52
3.6.1 Desarrollo de modelos de pronóstico de series de tiempo .....	52
3.6.2 Modelamiento de los componentes del sistema fotovoltaica híbrido .....	62
3.6.3 Programa de optimización para el sistema de generación híbrido: HOMER	67
3.6.4 Función objetivo de la optimización y sus restricciones.....	72
3.6.5 Diseño y modelamiento del sistema fotovoltaico híbrido.....	76
CAPÍTULO IV - RESULTADOS.....	92
4.1 Análisis de la base de datos .....	92
4.2 Modelos predictivos de las series de tiempo.....	94
4.2.1 Modelo predictivo SARIMA.....	94
4.2.2 Modelo predictivo ANN.....	98
4.2.3 Modelo LSTM.....	104
4.2.4. Análisis comparativo de los rendimientos de los modelos SARIMA, ANN y LSTM .....	112
4.3 Modelamiento y simulación del sistema de generación fotovoltaica mediante el programa HOMER.....	113
4.3.1 Diseño el sistema de generación .....	113
4.3.2 Datos climáticos y consumo de energía del lugar en estudio.....	114
4.3.3 Parámetros técnicos de los componentes del sistema basados en sus tecnologías y costos .....	116
4.3.4 Optimización del sistema de generación.....	118
4.3.5 Capacidades y potencias óptimas de los componentes del sistema .....	120
4.3.6 Análisis tecno-económico y producción eléctrica del sistema.....	123
4.3.5 Operación e impacto medio ambiental de los sistemas.....	126
4.4 Optimización del sistema híbrido fotovoltaico .....	129
4.4.1. Área de estudio y diseño del sistema .....	129
4.4.3 Diseño del algoritmo de optimización y gestión energética del sistema fotovoltaico híbrido .....	130
DISCUSIONES .....	138
CONCLUSIONES.....	146
RECOMENDACIONES Y TRABAJOS FUTUROS .....	148
CONTRIBUCIONES CIENTÍFICAS Y PARTICIPACIÓN EN LA COMUNIDAD ACADÉMICA .....	149
REFERENCIA BIBLIOGRÁFICA.....	150
ANEXO .....	165
Anexo 1: Consumo energético de la comunidad rural.....	165

Anexo 2: Ficha técnica de los componentes del sistema .....	166
Anexo 3: Código en Python del Modelo SARIMA.....	167
Anexo 4: Código en Python del Modelo ANN.....	173
Anexo 5: Código en Python del Modelo LSTM.....	180
Anexo 6: Código en Python de la operatividad del sistema y el algoritmo de optimización Evolución Diferencial .....	187
Anexo 7: Artículo científico .....	205

## ÍNDICE DE FIGURA

<b>Figura 1</b> Generadores de energía renovable en base a su potencia instalada (GW) desde el año 2008 hasta 2018.....	37
<b>Figura 2</b> Potencia instalada acumulada en GW para los generadores de energía renovable desde el año 2000 hasta 2022 y sus proyecciones. ....	37
<b>Figura 3</b> Comparación de LCOE de diferentes tecnologías renovables entre el año 2009 y 2023.....	39
<b>Figura 4</b> Evolución de la potencia instalada en los años de 2008 a 2018 para las centrales de generación renovable en el Perú.....	40
<b>Figura 5</b> Esquema de un sistema híbrido conformado por paneles fotovoltaico, inversores, baterías, generador diésel y red eléctrica.....	46
<b>Figura 6</b> Sistema de generación híbrido afectado directa o indirectamente por factores externos .....	46
<b>Figura 7</b> Los tres procesos de investigación realizados en la tesis doctoral.....	50
<b>Figura 8</b> Pre procesamiento y procesamiento de los datos para cada uno de los modelos .....	54
<b>Figura 9</b> Estructura y procesos del multilayer perceptron (MLP).....	57
<b>Figura 10</b> Estructura de la celda de memoria LSTM .....	58
<b>Figura 11</b> Diagrama de flujo de los procesos realizados por el programa HOMER.....	71
<b>Figura 12</b> Algoritmo de optimización Evolución Diferencial.....	86
<b>Figura 13</b> Diagrama de flujo del ingreso de la base de datos y la operatividad del sistema considerando el costo presente neto como función objetivo.....	87
<b>Figura 14</b> Diagrama de flujo de los programas que gestiona la descarga y carga del banco de baterías .....	88
<b>Figura 15</b> Serie de tiempo de la radiación solar global y la potencia de consumo del lugar de estudio desconectada de la red eléctrica.....	93
<b>Figura 16</b> Función de Autocorrelación y Parcial Autocorrelación para la radiación solar y potencia de consumo.....	95
<b>Figura 17</b> Predicción del modelo SARIMA sobre la data de prueba de la radiación solar .....	97
<b>Figura 18</b> Visualización de la data de prueba y las predicciones del modelo SARIMA con sus diferentes configuraciones en la potencia de consumo .....	98

<b>Figura 19</b>	Métricas de error para diferentes arquitecturas con SGD, RMSprop, Adam Optimizer y funciones de activación (Relu, Tanh y Sigmoid) para diferentes casos con y sin neurona dropout. ....	100
<b>Figura 20</b>	Curva de entrenamiento del modelo ANN: (a) Entrenamiento de la data para la radiación solar. (b) Entrenamiento de la data para potencia de consumo	103
<b>Figura 21</b>	Visualización de las predicciones de las diferentes configuraciones de los modelos ANN sobre la data de prueba de la radiación solar .....	103
<b>Figura 22</b>	Visualización de las predicciones de los modelos ANN sobre la data de prueba de la potencia de consumo .....	104
<b>Figura 23</b>	Curva de entrenamiento de las diferentes configuraciones del modelo LSTM para la radiación sola, potencia de consumo y temperatura ambiente .....	106
<b>Figura 24</b>	Predicciones de las configuraciones del modelo LSTM con mayores rendimientos sobre la radiación, potencia de consumo y temperatura ambiente .....	107
<b>Figura 25</b>	Predicción de valores futuros de la configuración con mayor rendimiento en las series de tiempo .....	108
<b>Figura 26</b>	Curva de entrenamiento de todas las configuraciones del LSTM para las series de tiempo de la Biblioteca Universitaria.....	109
<b>Figura 27</b>	Predicción de las 3 mejores configuraciones del LSTM para las series de tiempo de la biblioteca universitaria .....	110
<b>Figura 28</b>	Predicción a futuro de la configuración del LSTM con mayor rendimiento para las series de tiempo de la Biblioteca Universitaria .....	111
<b>Figura 29</b>	Diseño del Sistema fotovoltaico híbrido para la comunidad desconectada de la red en estudio .....	114
<b>Figura 30</b>	Perfil de potencia de consumo horaria y radiación solar mensual de la comunidad San Juan de Tarucani.....	115
<b>Figura 31</b>	Potencia de salida de los componentes del Sistema A durante el año .....	122
<b>Figura 32</b>	Potencia de salida de los componentes del Sistema B a lo largo del año ...	123
<b>Figura 33</b>	Flujo de caja del Sistema A (Fotovoltaico, controlador, baterías, inversor y generador diésel) obtenido del programa HOMER .....	124
<b>Figura 34</b>	Flujo de potencia del SHF para cubrir la potencia de consumo.....	128
<b>Figura 35</b>	Convergencia del CPN, COE y el número de baterías y fotovoltaicos .....	133
<b>Figura 36</b>	Operación horaria del sistema de generación .....	134

<b>Figura 37</b> Costo Capital, reemplazo, O6M y salvage obtenidos de la simulación y operación del sistema.....	135
<b>Figura 38</b> Potencias de descarga y carga de las baterías, exceso y déficit del sistema	137

## ÍNDICE DE TABLAS

<b>Tabla 1</b>	Detalles del lugar de estudio y su consumo de energía .....	93
<b>Tabla 2</b>	Diferentes combinaciones del modelo SARIMA con sus respectivos errores métricos .....	96
<b>Tabla 3</b>	Hiperparámetros y sus valores para diferentes arquitecturas del modelo de red neuronal.....	99
<b>Tabla 4</b>	Diferentes configuraciones del modelo ANN para la radiación solar y la potencia de consumo con bajos errores predictivos.....	102
<b>Tabla 5</b>	Hiperparámetros y sus valores para cada configuración del modelo LSTM..	105
<b>Tabla 6</b>	Resultado de la evaluación de las distintas configuraciones del modelo LSTM sobre las series de tiempo estudiadas .....	107
<b>Tabla 7</b>	Resultado de la evaluación de todas las configuraciones del LSTM para las series de tiempo de la biblioteca universitaria.....	110
<b>Tabla 8</b>	Comparación de las configuraciones con mayor rendimiento del modelo LSTM para las distintas series de tiempo en la comunidad rural .....	112
<b>Tabla 9</b>	Comparación de las configuraciones con mayor rendimiento del modelo LSTM para las distintas series de tiempo en la Biblioteca Universitaria .....	113
<b>Tabla 10</b>	Energía y potencia de consumo promedio del lugar de estudio .....	115
<b>Tabla 11</b>	Radiación solar (kWh/m <sup>2</sup> /d) y temperatura ambiente (°C) en la comunidad	116
<b>Tabla 12</b>	Parámetros técnico-económicos y los componentes del SFH .....	117
<b>Tabla 13</b>	Resultado de la optimización del tipo categorizada y general acorde al Costo de Energía (COE) y el Costo Presente Neto (CPN) o NPC .....	119
<b>Tabla 14</b>	Resultados de la optimización del sistema mediante el programa HOMER	121
<b>Tabla 15</b>	Resumen de los costos de cada uno de los componentes para los sistemas A y B .....	124
<b>Tabla 16</b>	Producción de energía eléctrica y consumo anual, exceso de electricidad y fracción renovable .....	126
<b>Tabla 17</b>	Emisión de gases contaminantes para ambos sistemas A y B.....	129
<b>Tabla 18</b>	Parámetros tecno-económicos y parámetros del optimizador Evolución Diferencial.....	131
<b>Tabla 19</b>	Costos de cada uno de los componentes para el cálculo del CPN.....	135
<b>Tabla 20</b>	Resultados obtenidos de las variables temporales en estudio y comparadas con otros resultados.....	140

<b>Tabla 21</b> Diferentes sistemas fotovoltaicos híbridos obtenido en esta investigación y comparación con resultados de otros autores .....	145
---	-----

## RESUMEN

Los sistemas fotovoltaicos híbridos están teniendo gran atención en el campo de la energía renovable, debido a que son una solución energética confiable y sostenible para áreas remotas. Sin embargo, los principales desafíos que enfrenta este sistema son la variabilidad de las condiciones climática que afectan en la generación de energía y el dimensionamiento inadecuado que la hace menos eficiente y rentable. En base a lo mencionado, el primer objetivo de esta tesis doctoral se centra en evaluar diferentes modelos predictivos para estimar la demanda energética y condiciones climáticas (radiación solar y temperatura ambiente) con el fin de determinar un modelo predictivo con alta precisión y confiabilidad. Por otro lado, el segundo objetivo tiene como propósito diseñar un algoritmo que simule y optimice la operatividad dinámica de un sistema fotovoltaico híbrido, el cual determine de manera óptima las cantidades de sus componentes con la finalidad de minimizar el costo presente neto durante la vida del proyecto. Los resultados de la evaluación y comparación de modelos predictivos (SARIMA, ANN y LSTM) sobre la información del lugar en estudio, en forma de series temporales, han indicado que el modelo LSTM tiene un notable potencial predictivo con valores de MSE y RSME próximos a cero y un valor de  $R^2$  por encima de 0.97, con respecto a los otros modelos. Por otro lado, la simulación del sistema fotovoltaica híbrido apoyado por batería y diésel, mediante el programa HOMER, ha determinado las capacidades óptimas de sus componentes y obtenido mejores resultados del Costo Presente Neto (CPN), costo de la electricidad y la confiabilidad del sistema, que un sistema fotovoltaico apoyado únicamente con baterías. Así mismo, el desarrollo e implementación de un algoritmo evolutivo en la operatividad del sistema fotovoltaico híbrido (Fotovoltaico, Baterías) ha permitido determinar las cantidades óptimas de módulo fotovoltaico y baterías con un mínimo valor del CPN, costo de electricidad y un valor de probabilidad de pérdida de energía de cero, indicando la confiabilidad del sistema. Se ha identificado, en estos dos últimos estudios, que el tiempo de vida esperado de las baterías en un sistema fotovoltaico híbrido es un factor importante que hace que el costo de reemplazo aumente o disminuya generando cambios en el CPN del proyecto.

**Palabras claves:** Predicción, algoritmo inteligente de Optimización, LSTM, ANN, SARIMA, HOMER, Evolución diferencial, Energía renovable, sistema fotovoltaico híbrido, Radiación solar, Demanda energética.

## ABSTRACT

Hybrid photovoltaic systems are getting a lot of attention in the renewable energy field because they are a reliable and sustainable energy solution for remote areas. However, the main challenges faced by this system are the variability of climatic conditions that affect power generation and the inadequate sizing that makes it less efficient and cost-effective. Based on the above, the first objective of this doctoral thesis focuses on evaluating different predictive models to estimate energy demand and climatic conditions (solar radiation and ambient temperature) in order to determine a predictive model with high accuracy and reliability. On the other hand, the second objective is to design an algorithm that simulates and optimizes the dynamic operation of a hybrid photovoltaic system, which optimally determines the quantities of its components in order to minimize the net present cost during the life of the project. The results of the evaluation and comparison of predictive models (SARIMA, ANN and LSTM) on the information of the site under study, in the form of time series, have indicated that the LSTM model has a remarkable predictive potential with MSE and RSME values close to zero and an  $R^2$  value above 0.97, with respect to the other models. On the other hand, the simulation of the hybrid photovoltaic system supported by battery and diesel, using the HOMER program, has determined the optimal capacities of its components and obtained better results of Net Present Cost (NPC), cost of electricity and reliability of the system, than a photovoltaic system supported only by batteries. Likewise, the development and implementation of an evolutionary algorithm in the operation of the hybrid photovoltaic system (PV, batteries) has allowed determining the optimal quantities of PV module and batteries with a minimum NPC value, electricity cost and a zero energy loss probability value, indicating the reliability of the system. It has been identified, in these last two studies, that the expected lifetime of the batteries in a hybrid PV system is an important factor that makes the replacement cost increase or decrease generating changes in the CPN of the project.

**Keywords:** Prediction, intelligent Optimization algorithm, LSTM, ANN, SARIMA, HOMER, Differential evolution, Renewable energy, hybrid photovoltaic system, Solar resource, Electricity consumption.

## INTRODUCCIÓN

Los combustibles fósiles son recursos tradicionales para la generación de energía, sin embargo, estos recursos tienen impactos negativos en el medio ambiente porque produce gases contaminantes que afectan la atmósfera. Actualmente, muchos gobiernos están implementando regulaciones para mitigar estos impactos, de modo que las industrias y los investigadores presten más atención a los recursos energéticos renovables para la producción de energía.

El objetivo de generar electricidad a partir de recursos renovables es satisfacer la creciente demanda mundial de electricidad a través de fuentes de energía sostenibles. Esto permite una reducción en el uso de fuentes de energía no sostenibles, un menor impacto en el medio ambiente y un aprovechamiento de recursos locales (solar, eólico y otros) para cubrir la demanda energética de localidades con y sin conexión a red (Nasir *et al.*, 2016)(Nasir *et al.*, 2018). Por otro lado, los sistemas de generación de energía renovable (eólica y solar) y la demanda energética del sitio presentan problemas a la hora de estimar su generación y demanda energética. Al no tener una estimación correcta, estos generan varios problemas técnicos y económicos en el buen funcionamiento del sistema de generación al tornarse una generación insegura e inestable para cubrir la demanda energética.

Por el lado de la generación fotovoltaica, la radiación solar es variable durante todo el año debido a los cambios de estación, horas del día, periodos de lluvia y nubes pasajeras, las cuales provocan discontinuidades en la producción de energía fotovoltaica, haciéndola poco confiable (Teixeira *et al.*, 2015). De manera similar, la demanda energética de un sitio se ve afectada por diversas condiciones a lo largo del tiempo, tales como: factores meteorológicos, la economía, el comportamiento humano y la demografía (Ma y Ma, 2018).

Esta variabilidad en la demanda energética provoca una sobreestimación o subestimación que confunde a los planificadores del sistema de generación de energía, desperdiciando recursos y costos operativos, provocando fallas y escasez de generación (Sadeghi *et al.*, 2009)(Qiu, 2013). Por lo tanto, una previsión precisa de la radiación solar y la demanda energética en una ubicación determinada son importantes porque reducen

el efecto de la variabilidad y la incertidumbre de la radiación solar, permitiendo una gestión óptima de la energía y una mejor planificación en la instalación de generadores con energías renovables con el fin de lograr operaciones confiables, seguras y estables para satisfacer la demanda energética (Singh *et al.*, 2019)(Raza y Khosravi, 2015).

Por otro lado, la combinación de fuentes de energía renovables y otras fuentes de generación (como los sistemas de almacenamiento de energía y el diésel convencional) son denominados sistemas de generación de energía híbrida. Este sistema aporta electricidad de forma estable y confiable al suministro eléctrico, debido a que este sistema combina las generaciones de energía de manera eficiente (Gan *et al.*, 2015).

SEH ofrece muchas ventajas, entre ellas: reducir el Costo de la Energía (COE), aumentar la proporción de recursos energéticos renovables, reducir la emisión de gases de efecto invernadero (NO<sub>2</sub>, CO<sub>2</sub> entre otros) y permitir el acceso a la electricidad en áreas remotas y rurales. Estas ventajas están en línea con tres de los ejes centrales del Desarrollo Sostenible (DS), que son la protección del planeta, a las personas y la prosperidad (Mohammed *et al.*, 2014)(Ghenai y Bettayeb, 2017). SEH es una opción apropiada que permitirá promover el desarrollo local en áreas donde la conexión a la red es costosa o en áreas rurales donde existen problemas de conexión a la red de transmisión eléctrica, debido a barreras geográficas y condiciones climáticas (Jung *et al.*, 2019).

La tesis es organizada de la siguiente forma: **El capítulo I** presenta la justificación, planteamiento del problema, objetivo general y específico, hipótesis y motivación de la investigación. Los antecedentes, la energía renovable y la inteligencia artificial, y las bases conceptuales son descritos en **el capítulo II**. **En el capítulo III** presenta la metodología de investigación donde se describe el tipo y diseño de investigación, fuentes de recolección y técnicas de procesamiento de datos. Los resultados obtenidos de los modelos predictivos de la serie de tiempo, optimización y simulación del sistema de generación híbrido son mostrados en **el capítulo IV**. **Discusión de resultados y conclusiones. Recomendaciones y trabajos futuros.**

# CAPÍTULO I – PLANTEAMIENTO DEL PROBLEMA

## 1.1 Justificación

La energía solar fotovoltaica presenta muchos beneficios, lo cual la convierte una opción importante en la generación de energía eléctrica. Esta energía aprovecha la radiación solar del lugar que es inagotable, generando ahorro en el costo de la energía eléctrica a largo plazo y es amigable con el medio ambiente. Sin embargo, esta generación presenta dificultades relacionadas a la variabilidad de la radiación solar lo que genera poca confiabilidad del sistema ante la demanda energética de lugares desconectado de la red, así mismo, esta intermitencia relacionada a su recurso genera inestabilidad en la red eléctrica. Por lo que muchas investigaciones están buscando soluciones ante estos desafíos con el fin de mejorar su eficiencia energética y su confiabilidad.

Existe un gran interés en la comunidad científica en abordar los desafíos en la generación fotovoltaica, las cuales están relacionado a su variabilidad e intermitencia de su recurso solar y su confiabilidad ante el consumo eléctrico. En base a ello, esta investigación pretende abordar esta problemática utilizando las herramientas de inteligencia artificial que permita brindar soluciones ante estos desafíos. Por lo que se propone: 1) Identificar modelos inteligente predictivos con alto rendimiento para predecir la variabilidad de la radiación solar, la temperatura ambiente y el consumo de energía de un lugar de estudio. 2) Analizar la operatividad dinámica de un sistema fotovoltaico híbrido, apoyado con batería y diésel, mediante la simulación y optimización inteligente que permita determinar el óptimo rendimiento y la confiabilidad del sistema.

Los hallazgos encontrados en esta investigación tendrán aportes significativos en el campo de la energía solar fotovoltaica. Estos resultados permitirán un estudio del recurso del lugar y la confiabilidad del sistema considerando costos mínimos de inversión y una gestión eficiente de la energía generada. Esta investigación dará lugar a los profesionales del diseño e implementación del sistema energéticos tomar mejores decisiones en base a la previsión del recurso del lugar y la operatividad del sistema. Además, los resultados encontrados en base a la aplicación de la Inteligencia artificial promoverán el uso de esta herramienta en los desafíos que presenta esta generación.

## **1.2 Planteamiento del problema**

### **1.2.1 Planteamiento del problema general**

¿Cuál es el mejor modelo de predicción de series temporales con mayor rendimiento en precisión y confiabilidad para pronosticar la demanda energética de un lugar y variables climáticas (radiación solar y temperatura ambiente) relevantes para el diseño y simulación de un sistema de generación híbrido?

¿Cómo se puede simular y optimizar un sistema de generación híbrido a partir de la demanda energética de una localidad y variables climáticas (radiación solar y temperatura ambiente) para determinar el dimensionamiento óptimo de las componentes del sistema, considerando el mínimo costo presente neto, confiabilidad y otras restricciones del sistema?

### **1.2.2 Planteamiento del problema específico**

- ¿Es posible diseñar, implementar y evaluar modelos predictivos y comparar su precisión y confiabilidad a través de sus métricas de error (MSE, RMSE,  $R^2$ ), utilizando las series de tiempo horarias de la demanda energética de un lugar de estudio y variables climáticas como la radiación solar y la temperatura ambiente?
- ¿Se podrá diseñar y simular la operación dinámica de un sistema de generación híbrido (fotovoltaico-batería-diésel) para cubrir la demanda energética de una comunidad alejada de la red, a través del programa HOMER, y encontrar las óptimas capacidades de los componentes y su operación integrada con el fin de minimizar el costo presente neto sujeto a ciertas restricciones mínimas de eficiencia y fiabilidad del sistema durante toda la vida del proyecto?
- ¿Es posible diseñar e implementar un algoritmo inteligente en la operación dinámica de un sistema de generación híbrido (fotovoltaico y baterías) para encontrar el óptimo número de módulos fotovoltaicos y baterías, considerando un mínimo costo de electricidad (COE), costo presente neto y la confiabilidad del sistema?

## **1.3 Objetivo general y específico**

### **1.3.1 Objetivo general**

El objetivo principal de esta investigación está conformado por:

- Desarrollar diferentes modelos predictivos para estimar las series temporales de la demanda energética y condiciones climáticas (radiación solar y temperatura ambiente) y compararlos con el fin de determinar un modelo predictivo con alta precisión y confiabilidad en la calidad de las predicciones.
- Diseñar un algoritmo que simule y optimice la operatividad dinámica de un sistema de generación fotovoltaica, apoyada con diésel y baterías, el cual optimice las cantidades de sus componentes sujeto a restricciones de capacidad y técnicas del sistema para minimizar el costo presente neto durante la vida del proyecto.

### **1.3.2 Objetivo específico**

El objetivo principal será obtenido determinando los siguientes objetivos específicos:

- 1) Diseñar, implementar y evaluar modelos predictivos y comparar sus rendimientos en precisión y confiabilidad utilizando las series de tiempo horarias de la radiación solar, la temperatura ambiente y la demanda energética de un lugar de estudio.
- 2) Diseñar y simular la operación dinámica de un sistema de generación híbrido (fotovoltaico-batería-diésel) para cubrir la demanda energética de una comunidad alejada de la red a través del programa HOMER y encontrar las óptimas capacidades de los componentes que conforman el sistema, minimizando el costo presente neto del sistema durante toda la vida del proyecto.
- 3) Diseñar e implementar un algoritmo inteligente en la operación dinámica de un sistema de generación híbrido (fotovoltaico y baterías) para encontrar el óptimo número de módulos fotovoltaicos y baterías, considerando un mínimo costo de electricidad (COE), costo presente neto durante el proyecto y la confiabilidad del sistema.

## **1.4 Hipótesis**

Si se diseña, implementa y evalúa modelos predictivos y se comparan su precisión y confiabilidad a través de sus métricas de error (MSE, RMSE,  $R^2$ ), utilizando las series de tiempo horarias de la demanda energética de un lugar de estudio y variables climáticas como la temperatura ambiente y la radiación solar, se obtendrá un modelo inteligente con la capacidad de predecir con precisión las series de tiempo mencionadas. Esto permitirá una mejor planificación y gestión de la energía entre la generación fotovoltaica y la demanda energética, proporcionando información valiosa en la toma de decisión relacionada con la inversión del sistema de generación fotovoltaica híbrido.

El diseño e implementación de un algoritmo inteligente que simule y optimice la operación dinámica de un sistema de generación híbrido con restricciones de capacidad, características técnicas y costos de los componentes, resultará en un dimensionamiento óptimo de las componentes del sistema con valores mínimos de costo presente neto, costo de electricidad generada y confiabilidad. Esto permitirá una reducción de costos, una mejora en la eficiencia energética y un aumento en la confiabilidad del sistema, logrando obtener beneficios significativos en término tecno-económico, medio ambiental y autonomía.

## **1.5 Motivación**

Esta investigación doctoral surge de la necesidad de enfrentar problemas actuales en el campo de la generación de energía en áreas remotas y aisladas sin acceso a la red eléctrica. Las comunidades en estas áreas enfrentan a dificultades para acceder a servicios eléctricos asequibles y confiables, siendo dependientes de fuentes de energía no sostenibles y costosas, como el diésel.

La implementación de sistemas de generación híbridos en áreas remotas es una solución prometedora para abordar estas problemáticas, debido a que estos sistemas aprovechan el recurso disponible en el lugar y combinan las tecnologías de generación, como energía fotovoltaica, baterías y generadores diésel. Estos sistemas son más eficientes energéticamente, confiables, autónomos y a costos bajos. Sin embargo, para que la efectividad de este sistema sea óptima, es de suma importancia contar con un

modelo predictivo preciso y un algoritmo de optimización que permitan el diseño y operación de los componentes del sistema de manera rentable y eficiente.

Esta investigación doctoral tiene como propósito fundamental desarrollar y comparar diferentes modelos predictivos que permitan estimar de forma precisa las series temporales de la demanda energética y las condiciones climáticas, como la temperatura ambiente y la radiación solar. Así mismo, se plantea como objetivo diseñar un algoritmo inteligente que optimice la operación dinámica de un sistema de generación híbrido, considerando restricciones de capacidad, gestión de la energía generada y parámetros técnicos de los componentes del sistema con el objetivo de minimizar el costo presente neto durante toda la vida del proyecto.

Al finalizar los objetivos propuestos en esta investigación, se espera contribuir significativamente al campo de la generación de energía renovable y aislada, ofreciendo una metodología y herramientas que proporcionen una mejora en la confiabilidad, rentabilidad y eficiencia en los sistemas de generación híbridos. Esta tesis doctoral permitirá una influencia directa en el nivel de bienestar de las comunidades alejadas de la red eléctrica, al permitirles acceso a una energía más asequible, sostenible y limpia.

## CAPÍTULO II – MARCO TEÓRICO

### 2.1 Antecedentes bibliográfico

En la literatura científica existen diversos estudios relacionados a la predicción de los recursos naturales con aplicaciones energéticas y optimización de sistemas energéticos. Así mismo, una revisión de la literatura científica es realizada en esta sección donde se aborda estudios relacionado a la predicción y la optimización de sistemas generación eléctrica.

#### 2.1.1 Pronóstico del recurso renovable y estrategia de operación de los sistemas renovables híbridos

**Paoli *et al.* (2010)** presentan una aplicación de las Redes Neuronales Artificiales (ANN, siglas en inglés) del tipo Perceptrón Multicapa (MLP, siglas en inglés) para la predicción diaria de la radiación solar. El método de ventana deslizante (Sliding-Window, en inglés) fue utilizada en el procesamiento de la data de radiación solar para generar los datos de entrada del modelo MLP. Los resultados del MLP son prometedores, con un Error Cuadrático Medio Normalizado (nRMSE, siglas en inglés) de alrededor del 21% y un Error Cuadrático Medio (RMSE, siglas en inglés) de alrededor de 3.59 MJ/m<sup>2</sup>. Estos resultados fueron mejores a otros modelos como el Autorregresivo Integrado de Promedio Móvil (ARIMA, siglas en inglés), Inferencia Bayesiana, cadenas de Markov y los k-Vecinos más Cercanos (k-Nearest-Neighbors, en inglés).

**Hong *et al.* (2011)** proponen un modelo de regresión lineal múltiple ingenuo como referencia para la predicción del consumo eléctrico a corto plazo de una empresa de servicios públicos en Estados Unidos. El modelo propuesto está conformado por variables cuantitativas (tendencia y temperatura ambiente en horas), variables de clase (horas, días y meses) y efectos de interacción o cruzados. La predicción del modelo para los días siguientes, en etapas de tiempo horarias, tuvo como resultado un Error Porcentual Absoluto Medio (MAPE, siglas en inglés) de 5%, el cual resulto ser más menor que otros modelos.

**Voyant *et al.* (2013)** proponen una técnica original para modelar las series temporales de la radiación solar basada en la combinación de una Red Neuronal Artificial

(ANN, siglas en inglés) y un modelo Auto-Regressivo y de Media Móvil (ARMA, siglas en inglés). Los modelos ANN y ARMA son utilizados en forma independiente y conjunta para modelar la radiación solar en las estaciones de invierno, primavera, verano y otoño. Los resultados indican que los modelos ANN y ARMA tienen valores de nRMSE alrededor de 15% a 17% para las diferentes estaciones que otros modelos. La combinación de ANN y ARMA permite una mejora en los valores de nRMSE, alcanzando un valor mínimo de 11.5%.

**Pillai *et al.* (2014)** investigan la viabilidad de utilizar datos de carga y clima disponibles públicamente para generar perfiles de carga sintéticos, utilizando Redes Neuronales Artificiales con algoritmo de entrenamiento del tipo Regulación Bayesiano. El modelo propuesto se entrena primero, utilizando datos meteorológicos y datos de perfiles de carga de regiones vecinas como entrada y salida del modelo. Los resultados indican que los perfiles de carga sintéticamente generados para diferentes lugares se asemejan estrechamente a los reales, teniendo valores aproximados de RMSE, MAE y MAPE de 0.03, 0.02 y 1.95, respectivamente.

Así mismo, **Huang *et al.* (2014)** presentan un método híbrido basado en el clima, recopilada del Buró Central de Clima de Taiwán, que combina Mapas Auto-Organizativos (SOM, siglas en inglés) y redes de Cuantización Vectorial de Aprendizaje (LVQ, siglas en inglés), Regresión de Vectores de Soporte (SVR, siglas en inglés) y el método de Inferencia Difusa para la predicción horaria con un día de anticipación de la producción de energía fotovoltaica. El método presenta tres etapas (clasificación, entrenamiento y predicción), donde la primera etapa consiste en recopilar y clasificar datos históricos de la producción de energía fotovoltaica a través de SOM y LVQ y la segunda etapa se basa en entrenar los conjuntos de datos de entrada/salida de temperatura ambiente, probabilidad de precipitación e irradiación solar de horas utilizando el SVR. Los resultados de la predicción fueron evaluados por los criterios de MSE y RMSE, tomando valores de aproximadamente 3% y 350W, respectivamente. Estos valores resultaron ser menores que los obtenidos por métodos simples de SVR y ANN tradicionales.

**Sinha y Chandel (2015)** estudian sistemas híbridos basados en módulos fotovoltaicos y micro eólica en 12 ubicaciones del estado occidental del Himalaya,

Himachal Pradesh, en la India. La metodología propuesta se basa en dimensionar sistemas de generación híbrido, a través del programa HOMER, utilizando datos proporcionados por la NASA, datos pronosticados de Redes Neuronales Artificiales y datos medidos o estimado. Los datos estimados del viento y la radiación solar, tuvieron valores de RMSE alrededor de 1.6 kWh/m<sup>2</sup> y 1 m/s, respectivamente, los cuales indican que las estimaciones del modelo se acercan a los datos medidos. La optimización del sistema de generación híbrido para una demanda energética promedio diario de energía de 5.2 kWh/día en Hamirpur tiene una configuración óptima de una turbina micro eólica de 5 kWp, un convertidor de 2 kW, 10 baterías y un sistema fotovoltaico de 8 kWp, mientras que para las otras ubicaciones se obtienen una turbina micro eólica de 5 kWp, un convertidor de 2 kW, 10 baterías y sistemas fotovoltaicos de 2 a 9 kWp.

**Azimi et al. (2016)** proponen un pronóstico híbrido para la radiación solar, utilizando un algoritmo de Transformación basado en K-medias (TB K-means, en inglés) para aumentar la precisión del pronóstico a diferentes horizontes temporales (1 hora, 2 horas, ..., 48 horas por delante). El agrupamiento propuesto se combina con un análisis de series temporales, un nuevo algoritmo de selección de clústeres y una Red Neuronal de Perceptrón Multicapa (MLPNN, siglas en inglés) para desarrollar el método híbrido del pronóstico de la radiación solar. El rendimiento del agrupamiento TB K-medias propuesto se evalúa utilizando varios conjuntos de datos diferentes y se compara con diferentes variantes del algoritmo de K-medias. Los resultados de la comparación directa con otros modelos de pronóstico bien establecidos demuestran el rendimiento superior del método híbrido propuesto.

**Pombeiro et al. (2017)** examinan la adecuación de los modelos no lineales de diseño sencillo para predecir el consumo de energía en un entorno real de un edificio inteligente, utilizando variables predictoras simples como la hora del día, las condiciones meteorológicas y la ocupación estimada a partir del tráfico de WiFi. Los modelos desarrollados de Sistemas Difusos y Redes Neuronales obtienen un mayor rendimiento y precisión considerablemente mejores que los modelos de Regresión Lineal, teniendo una Varianza de la Explicación (VAF, siglas en inglés) del 79.3% y 75.7% y Error Absoluto Medio (MAE, por sus siglas en inglés) de 5.9 y 7.1 kWh, respectivamente, en comparación con un VAF del 36.6% y un MAE de 12.8 kWh para los modelos de Regresión Lineal.

**Lu y Chang (2018)** presenta un modelo predictivo del tipo Red Neuronal de Funciones de Base Radial (RBFNN, siglas en inglés) con un método de desacople para predecir la energía fotovoltaica para el día siguiente. Este modelo es comparado con los modelos del tipo Promedio Móvil Integrado Autorregresivo, Red Neuronal de propagación hacia atrás (BPNN, siglas en inglés) y Red Neuronal de Función de Base Radial a través de las métricas MAPE y RMSE. Como trabajo futuro, los autores proponen aplicar el modelo propuesto a diversas condiciones climáticas que permitan el pronóstico de la energía solar para el día siguiente.

Por otro lado, **Fouilloy et al. (2018)** estudiaron once modelos predictivos del tipo estadístico y aprendizaje automático para el previsionamiento de la radiación solar horaria, con horizontes temporales de 1 a 6 horas, para tres localidades con variabilidad meteorológica baja, media y alta. La predicción de los modelos de Regresión de Soporte Vectorial (SVR, siglas en inglés), Proceso Gaussiano (GP, siglas en inglés), Bosque Aleatorio (RF, siglas en inglés), Árbol de Regresión (RT, siglas en inglés), Perceptrón Multicapa (MLP, siglas en inglés), Media Móvil Autorregresivo (ARMA, siglas en inglés), Árbol de Regresión Potenciado (BRT, siglas en inglés), entre otros, fueron comparados a través del MAE, nRMSE y Puntuación de Habilidad. Los resultados determinaron que los modelos ARMA y MLP fueron más eficiente para una débil variabilidad de la radiación solar, mientras que los modelos ARMA, RT y otros modelos fueron más eficientes para una variabilidad media. Los modelos RF y RT predijeron de manera más precisa la alta variabilidad de la radiación.

**Alsharif y Younes (2019)** estudiaron los cambios de la radiación solar en los últimos 37 años en la ciudad de Seúl. El modelo predictivo llamado Sistema de Inferencia Neuro-Difuso Adaptativo (ANFIS, siglas en inglés) fue utilizado en este estudio, el cual tiene la capacidad de predecir la radiación solar utilizando entradas de series temporales caóticas. El resultado del modelo proporcionó una predicción solar mensual relativamente buena con un nRMSE del 0.22%, RMSE del 55.4, un coeficiente de determinación ( $R^2$ ) del 0.8 y una Distribución Normal del Residuo Estandarizado con un valor de  $p=0.222$ .

**Lou et al. (2020)** proponen un método híbrido de Media Móvil Integrada Autorregresiva Estacional (SARIMA, siglas en inglés) y Regresión de Vector de Soporte (SVR, siglas en inglés) para el pronóstico horario de la potencia de consumo. Este modelo

híbrido usa al modelo SARIMA para ajustar la parte lineal de la serie y al modelo SVR para corregir la desviación. En la comparación, el modelo propuesto presenta un valor de MAPE y MSE de  $7.193E-5$  y  $6.843\%$ , respectivamente, el cual es mucho más preciso que otros modelos.

Así mismo, **Sorkun et al. (2020)** realizaron la estimación de la radiación solar con el modelo de Memoria a Corto y Largo Plazo (LSTM, siglas en inglés) donde consideran el uso de datos multivariantes como la humedad, la temperatura ambiente y la nebulosidad. La predicción del modelo determina un valor de nRMSE de 0.1592 a través de la optimización de sus hiperparámetros, siendo este valor el menor que otros modelos.

**Aslam et al. (2021)** desarrollan un modelo de aprendizaje profundo conformado por un mecanismo de atención en dos etapas sobre la Memoria de Corto y Largo Plazo (LSTM) para pronosticar la energía fotovoltaica para el día siguiente. Los óptimos hiperparámetros del modelo fueron obtenidos a través del algoritmo de optimización Bayesiana y varias características de entrada para las 21 instalaciones fotovoltaicas en diferentes ubicaciones geográficas de Alemania. Además, los datos ingresados al modelo fueron la temperatura ambiente, la radiación solar, la humedad, las nevadas, el albedo, entre otros. El modelo propuesto supera al modelo de Atención LSTM, Red Neuronal Convolutiva-LSTM y modelo de Ensemble en términos de precisión, obteniendo una puntuación de la Habilidad de Pronóstico de 0.4813, RMSE de 0.0638 y MAE de 0.0324.

**Medina y Cárdenas (2022)** proponen una metodología en el dimensionamiento de un sistema de generación híbrido (fotovoltaico, batería, aerogenerador y diésel) a una comunidad rural en la costa del Pacífico de México, el cual incluye un modelo predictivo del tipo LSTM para predecir condiciones climáticas, tales como la velocidad del viento, temperatura ambiente y la radiación solar. El dimensionamiento del sistema se basó en la Programación Lineal, el cual es un algoritmo de optimización que minimiza el costo anualizado del sistema, el factor renovable y probabilidad de carga de pérdida. La precisión de la predicción tuvo valores de RMSE de 0.668, 170.033 y 2.057 para la velocidad del viento, la radiación solar y temperatura ambiente, respectivamente. La optimización del sistema obtuvo un valor de costo anualizado de \$4 526.536 y una probabilidad de pérdida de carga de 0.06008, el cual representa la confiabilidad del sistema.

Por otro lado, **Kim et al. (2022)** evaluaron varios modelos para predecir la radiación solar en 18 estaciones en Corea, los cuales utilizan diferentes variables meteorológicas (temperatura ambiente, humedad, nubosidad y radiación solar) como datos de entrada. El resultado del modelo Autorregresivo Integrado de Promedio Móvil con Variables Externas (ARIMAX, siglas en inglés), Árbol de Decisión con Refuerzo de Gradiente Extremo (XGBoost, , por sus siglas en inglés), Memoria a Corto y Largo Plazo (LSTM, siglas en inglés), Unidad Recurrente con Compuertas (GRU, siglas en inglés), Red Neuronal Recurrente Simple (RNN, siglas en inglés), Árbol de Decisión con Categoría Mejorada (CatBoost, siglas en inglés) tienen valores de MAE entre 0.108 a 0.705 para la precisión de la radiación solar en todas las estaciones. El MAE del modelo SARIMA tiene los valores más alto, mientras que los otros modelos tienen valores más bajos.

**Sumorek y Idzkowski (2023)** realizaron un análisis predictivo de los datos históricos de la producción energética de plantas fotovoltaica con y sin seguimiento solar en el noreste de Polonia. Los modelos de pronóstico basados en Redes Neuronales de Perceptrón Multicapa (MLP-ANN, siglas en inglés) y técnicas Aprendizaje Automático tradicionales (ML, siglas en inglés) no usaron la radiación solar como característica de entrada. El modelo MLP-ANN del tipo regresión tuvo mejor resultado que los modelos clasificadores MLP-ANN a través de las métricas de error. El valor de  $R^2$  es de alrededor de 0.6 para el modelo MLP-ANN del tipo regresión.

En la revisión literaria existen diversos modelos de predicción para determinar el comportamiento preciso de las variables climática, las cuales son de gran importancia en el diseño de un sistema fotovoltaico híbrido. La predicción de la temperatura ambiente, la radiación solar, viento, humedad y la potencia de consumo de un lugar de forma precisa a través de las métricas de error (MAE, MAPE, MSE, RMSE, nRMSE) han sido determinado por diversos modelos predictivos, como ARMA, ARIMAX, SARIMA, XGBosst, CatBoost, MLP, LSTM, RNN, SVR, ANFIS, entre otros. En esta investigación, los modelos SARIMA, MLP y LSTM han sido considerados en la predicción de series temporal de la demanda energética, la temperatura ambiente y la radiación solar de un lugar de estudio con el fin de obtener el mejor modelo predictivo con mayor precisión y confiabilidad.

### **2.1.1 Optimización en el dimensionamiento y operación del sistema híbrido fotovoltaico**

**Ohsawa *et al.* (1993)** utilizaron un modelo de inteligencia artificial del tipo Red Neuronal Artificial (en inglés Artificial Neural Network-ANN) y programación dinámica (en inglés Dynamic Programation-DP) con la finalidad de controlar la operación del sistema híbrido de generación fotovoltaica y el generador diésel. La información de la radiación solar y la demanda energética son ingresadas a los modelos para determinar la operatividad del sistema. Esta operatividad es luego entrenada por una red neuronal, donde la radiación solar, demanda energética y baterías son ingresadas a la entrada del modelo. En la salida del modelo, la potencia del generador diésel y el tiempo son especificados. Los resultados indican que el modelo ANN tiene una mejor descripción de la operatividad del sistema que el modelo DP.

**Ashari y Nayar (1999)** presentan la operatividad de un sistema de energía híbrido solar (sistema fotovoltaico, diésel y baterías), considerando un “Punto de ajuste o set point” y valores óptimos de los puntos de ajuste para el arranque y parada del generador diésel con el fin de minimizar los costos generales del sistema. El programa informático desarrollado por esta metodología permite una estrategia de despacho, una predicción del rendimiento energético a largo plazo y el costo de inversión del sistema durante su vida útil. Los resultados indican que el consumo de combustible del sistema presenta el mayor costo total de operación que incurre el sistema a lo largo de su vida útil

**Park *et al.* (2001)** utilizaron un sistema de generación fotovoltaica para cubrir la demanda energética de un barco. En la simulación del sistema se considera la variación de la potencia fotovoltaica como un factor importante de estudio. El objetivo de la investigación es controlar la capacidad de almacenamiento de la batería hacia un valor mínimo, mientras que el generador diésel mantiene constante su salida a pesar de la fluctuación de la potencia fotovoltaica. Los resultados de la simulación fueron validados con datos experimentales, durante el mes de Julio del 1998. Estos resultados también determinaron la tasa de utilización del recurso de energía fotovoltaica y el consumo de combustible para diferentes capacidades de almacenamiento de la batería. Además, los autores plantean como trabajo futuro que el método propuesto permite predecir el tamaño óptimo de la capacidad de la batería y del sistema fotovoltaico en términos de costos,

efectos ambientales y limitaciones para las embarcaciones, y sugieren que se estudie el control de operación del sistema.

**Hove y Tazvinga (2012)** han realizado la implementación y aplicación de un programa de simulación, basado en un modelo de optimización, que permite analizar el dimensionamiento de las componentes, el análisis económico y el rendimiento energético del sistema de energía híbrido (Fotovoltaico, Baterías y Diésel). El modelo desarrollado determina diversas curvas de dimensionamiento que definen el tamaño de las componentes del sistema, asegurando una confiabilidad a través de la Probabilidad de Pérdida de Carga deseada (LLP, siglas en inglés). En la investigación resaltan que la vida útil del generador diésel y la eficiencia del combustible tiene una dependencia del grado y la frecuencia de carga, mientras que el tiempo de la vida útil de la batería tiene dependencia de la velocidad y profundidad de descarga de los ciclos. Los resultados demuestran que la elección de la estrategia de despacho del generador diésel es otro factor de gran importancia que influye en el rendimiento energético y el costo del sistema. Como trabajo futuro, los autores concluyen que estudios experimentales en base a su simulación sería un importante aporte para continuar sus estudios.

**Huneke et al. (2012)** desarrollan un modelamiento general de un sistema híbrido fuera de la red en un grupo de aldeas en India y Titumate, en Colombia, para reflejar las condiciones reales del sistema híbrido (Fotovoltaico-Generador diésel y Baterías del tipo plomo-acido) con el fin de lograr la optimización económica y ecológica del sistema de energía. En su metodología utiliza métodos de Programación Lineal en el entorno de Sistema de Modelado Algebraico General (GAMS, siglas en inglés), en el cual encuentran una configuración óptima del sistema de suministro de energía eléctrica siguiendo restricciones características, como el clima por hora y los datos de la demanda energética. Concluyen que la combinación de energía renovable en combinación con dispositivos de almacenamiento eléctrico ayuda a disminuir el costo de la energía en comparación con los grupos electrógenos de diésel independientes.

**Hsu y Kang (2014)** analizan dos regímenes de despacho, (i) una estrategia de Despacho de carga de Ciclo Completo y (ii) una estrategia de Despacho de Carga Parcial (PSOC, siglas en inglés) con puntos de ajuste del Estado de Carga (SOC, siglas en inglés) correspondiente del sistema de alimentación de la batería y el generador diésel. Luego,

comparan en término de Gastos Operativos del Sistema (OPEX, siglas en inglés) y Costo Presente Neto (NPC, siglas en inglés). El programa de simulación se establece en función de los cálculos de eficiencia del sistema y su régimen de carga de la batería, donde los resultados demuestran que: 1) El ciclo poco profundo puede traer un tiempo de funcionamiento largo del generador diésel y un alto consumo de combustible, mientras que el ciclo profundo está a favor de reducir el tiempo de funcionamiento del generador diésel y el consumo de combustible. 2) El ciclo superficial está a favor de la duración de la batería bajo el régimen (i), mientras que el ciclo profundo está a favor de la duración de la batería bajo el régimen (ii). 3) La Profundidad de Carga (DOD, siglas en inglés) están presentes para ambos regímenes (i) y (ii). 4) La descarga del régimen (ii) conduce al NPC más bajo. Los resultados concluyen que deben proporcionarse ciclos de DOD precisos para determinar la falla de las baterías, permitiendo un beneficio en el diseño de las estrategias de despacho del sistema.

Así mismo, **Bortolini et al. (2014)** presentan el diseño de un sistema híbrido (Baterías y Fotovoltaico) conectada a la red, en los nuevos edificios de la escuela de arquitectura ubicados en la universidad de Bologna. El modelo realizado para la operación del sistema tiene como fin determinar la potencia nominal del sistema fotovoltaico, la capacidad del sistema de banco de baterías y el mínimo valor del Costo Nivelado de Electricidad (LCOE, siglas en inglés). El valor de LCOE es hallado mediante un algoritmo de Control de Flujo de Potencia orientado a cubrir la demanda energética, donde considera varios escenarios de prueba que varían la potencia y la capacidad nominal del sistema. El resultado de la simulación tuvo un LCOE igual a € 0.151/kWh para los edificios, con una reducción del costo de la energía del 24.5% en comparación con el precio de la electricidad de la red. Como trabajos futuros plantean que se puede realizar otros escenarios para comparar sus rendimientos bajo varias restricciones y diferentes características de ubicación. Además, se plantea el diseño del modelo multiobjetivo que considere el desempeño económico y medio ambiental con el objetivo de encontrar una óptima configuración del sistema que minimice el LCOE y los indicadores de impacto ambiental.

**Bortolini et al. (2015)** realizan un modelo de optimización multiobjetivo para un sistema fotovoltaico-batería y generador diésel aislado de la red. El objetivo fue identificar la capacidad del sistema de baterías, la potencia nominal de la planta

fotovoltaico y la configuración técnica capaz de reducir la Huella de Energía de Carbono (CFOE) y el Costo Nivelado de Electricidad (LCOE). El modelo tiene en cuenta la información del lugar (la demanda energética por horas, los perfiles de radiación solar y temperatura ambiente) para el cálculo del rendimiento de la planta fotovoltaica por hora, la energía requerida del generador, los procesos de carga y descarga de la batería. Los resultados destacan la viabilidad económica, técnica y ambiental del sistema con un costo de combustible relativamente bajo.

**Scioletti *et al.* (2017)** presentan un modelo de optimización que determina el diseño y la estrategia despacho del sistema híbrido (diésel-fotovoltaico y tres opciones de tecnologías de baterías) para cubrir el perfil de demanda energética de un año de bases militares con diferentes zonas climáticas. La optimización del sistema híbrido usa un modelo de Entero Mixto (Mixed-Integer, en inglés) y determina una reducción del consumo de combustible en un 30% en comparación con usar un solo generador. Además, las soluciones indican una estrategia de diseño y despacho del sistema que permite la carga y descarga de la batería cuando la demanda energética es baja y alta, respectivamente, con el fin de evitar que los generadores funcionen con una potencia alta. El beneficio implícito de esta estrategia es que los generadores mantienen niveles superiores al 30% de su potencia nominal y el tiempo de uso en un 96%. En lugar de minimizar costos, su modelo podría incorporar fácilmente objetivos como minimizar el impacto ambiental o el volumen total de las tecnologías adquiridas.

**Ahmed *et al.* (2017)** proponen tres sistemas de energía renovable que permitan alimentar de forma óptima un sistema de bombeo de riego y la casa de un agricultor en dos lugares diferentes en el Sinaí, Egipto. Los tres sistemas de energía propuestos son: un sistema eólico, sistema fotovoltaico independiente y un sistema híbrido fotovoltaico-eólico. El programa de simulación HOGA (Hybrid Optimization by Genetic Algorithms) basado en Algoritmos Genéticos (GA, siglas en inglés) es usado para el dimensionamiento, optimización y evaluación económica de los tres sistemas. Las configuraciones óptimas obtenidas en la simulación determinaron que el sistema híbrido fotovoltaico-eólico y el sistema fotovoltaico independiente fueron más rentables para las ubicaciones de Abu-Rudies y El-Arish, respectivamente.

**Charfi et al. (2018)** proponen un nuevo algoritmo que permita controlar la energía total producida de un sistema de energía híbrido (fotovoltaico, banco de baterías y generador diésel) con la finalidad de cubrir la demanda energética de una casa conectada a la red. En esta investigación proponen un algoritmo que calcula de manera recurrente la profundidad de descarga de la batería y el consumo de combustible del generador diésel, los cuales son usados para pronosticar la potencia de salida del sistema fotovoltaico cada dos minutos hacia delante. Los cálculos del algoritmo consideran criterios que garantizan el uso máximo de la energía renovable generada, el tiempo mínimo de operación del generador diésel y la confiabilidad del sistema híbrido para cubrir demanda energética durante todo el día. Los autores concluyen que el algoritmo propuesto es eficiente para la gestión de energía que garantiza un tiempo operativo mínimo de generación diésel y un suministro continuo de la demanda energética.

**Muthukumar y Balamurugan (2019)** proponen un sistema de generación de energía eólica fotovoltaica optimizada que busca maximizar la energía generada del sistema con el apoyo de un modelo de Red Neuronal y un algoritmo de Colonia de Abejas. La optimización del sistema híbrido (fotovoltaico y eólico) se ha realizado en el programa SIMULINK con diferentes patrones de la velocidad del viento y la radiación solar. El sistema incorpora un mecanismo de conmutación entre las dos fuentes de energía y la salida del convertidor con el fin de sintonizar la generación de potencia máxima con ayuda del modelo neuronal, lo que garantiza la estabilidad del sistema. Los autores indican como trabajo futuro que el sistema híbrido debe incluir una mayor cantidad de recursos energéticos que sean de naturaleza renovable y usar el algoritmo de Colonia de Abeja en la simulación.

**Kurukuru y Mohamed (2019)** presentan un modelo práctico y preciso del sistema de celdas de combustible-fotovoltaico, el cual ha sido simulado en el programa SIMULINK con la finalidad de mejorar la eficiencia del sistema. Este sistema simulado utiliza un controlador inteligente del tipo Lógica Difusa (Fuzzy, en inglés). La investigación muestra que el modelo propuesto puede mejorar el estado medio de carga de la batería en un 20% en comparación con otros modelos y que un control predictivo permite obtener mejoras aumentando del 20 al 30% en el uso de energía fotovoltaica y una reducción en el uso de la celda de combustible en un 25%. Estos resultados muestran una solución óptima sobre la administración de la energía de diferentes sistemas.

**Shezan (2019)** realizan un diseño integral del sistema de generación de energía híbrido con módulos fotovoltaica, diésel y baterías para 37 viviendas familiares en áreas de ecoturismo de Malasia desconectadas de la red. Los resultados de la optimización del sistema propuesto son económicamente más viables que otras combinaciones en el programa HOMER, debido a su Costo Nivelado de Energía (LCOE) se acerca a \$ 0.895/kWh y el Costo Presente Neto (NPC) de \$ 158,206. Este mismo sistema ha sido simulado por el programa de Herramienta de Sistema Fotovoltaico (PVSYST), el cual tuvo un COE de \$ 0.885/kWh y NPC de \$ 157 907.

**Kharrich et al. (2020)** han propuesto una versión mejorada del Optimizador Bonobo (BO, siglas en inglés) basada en un método cuasi-oposicional para encontrar el mejor diseño de un sistema híbrido en Asuán, Egipto. Este sistema está comprendido por paneles fotovoltaicos, turbinas eólicas, baterías y generadores diésel. Los resultados muestran que el algoritmo de optimización propuesto ha determinado mínimos valores del Costo Presente Neto (NPC, siglas en inglés) y Probabilidad de Perdida de Suministro de Energía (LPSP, siglas en inglés) en comparación a los tradicionales BO tradicional y otras técnicas de optimización denominadas Algoritmo de Campo Eléctrico Artificial (AEFA, siglas en inglés), Optimización de Harris Hawks (HHO, siglas en inglés) y Optimización de Malezas Invasivas (IWO, siglas en inglés).

Por otro lado, **Masih y Verma (2020)** realizaron el análisis de optimización y confiabilidad del sistema de energía renovable (fotovoltaico, turbina eólica, baterías) para abastecer la demanda energética y carga de riego de una granja ubicada en un área remota. La optimización del sistema a través del Algoritmo de Optimización de Saltamontes (GOA, siglas en inglés) determinó valores mínimos del costo presente neto (NPC), Costo Operativo (OC), Probabilidad de Pérdida de Carga (LOLP) y Costo Nivelado de Energía (LCOE), encontrando un dimensionamiento óptimo de paneles fotovoltaicos, la turbina eólica y el sistema de energía de la batería. El sistema híbrido de energía fotovoltaica, eólica y de batería tienen un NPC de \$ 26.450, LCOE de \$ 0.502 /kWh y LOLP de 0.0000103.

**Sultan et al. (2021)** principalmente se centran en el diseño óptimo de un sistema híbrido de energía renovable, conformado por turbina eólica, paneles fotovoltaicos y pila de combustible con un tanque a gas de hidrógeno para almacenar la energía conectado de

la red y fuera de la red en la región de Ataka, Golfo de Suez-Egipto. Los tamaños óptimos de los componentes del sistema de generación híbrido propuesto se logran utilizando un novedoso algoritmo denominado Optimización Mejorada de Ecosistemas Artificiales (IAEO, siglas en inglés), el cual proporciona un valor mínimo del Costo de la Energía, Costo Presente Neto, exceso de energía y el índice de confiabilidad presentado por la Probabilidad de Pérdida de Suministro de Energía de \$0.42957/kWh, \$10 437 193, 434 630 kWh y 0.02449, respectivamente, para el sistema aislado.

Por otro lado, **Rinaldi et al. (2021)** determinaron la configuración óptima de un sistemas de generación renovable, conformada por componentes únicos (fotovoltaico, eólico y diésel) y híbridos, para cubrir la demanda energética de comunidades remotas peruanas (Campo Serio, el Potrero y Silicucho) con diferentes zonas climáticas del Perú. El programa HOMER fue usado para determinar la configuración y el tamaño óptimo del sistema con menor Coste de la Energía (COE), Costo Presente Neto (NPC) y emisiones de CO<sub>2</sub> para diferentes escenarios. Los resultados obtenidos han revelado que el sistema híbrido fotovoltaico-eólico-diésel es la configuración más viable económicamente para todas las comunidades investigadas con un NPC de \$ 183.851 (COE: \$ 0.460/kWh) para el Potrero, \$ 227.335 (COE: \$ 0.478/kWh) para Campo Serio y \$ 146.583 (COE: \$ 0.504/kWh) para Silicucho.

**Mahmoud et al. (2022)** estudiaron métodos de optimización recientes llamados Algoritmos de Enjambre de Salpas (SSA), el Optimizador del Lobo Gris (GWO) y el Optimizador mejorado del Lobo Gris (IGWO) para encontrar la configuración óptima de un sistema de energía híbrido. Este sistema está comprendido por turbinas eólicas con un sistema de almacenamiento de baterías, paneles fotovoltaicos y generadores diésel de respaldo. La simulación del sistema es realizada en el lenguaje de programación MATLAB y mediante la optimización del sistema se busca minimizar el Coste de la Energía (COE) y la Probabilidad de Pérdida de Suministro de Energía del sistema (LPSP), bajo ciertas restricciones operativas y de confiabilidad. Los resultados revelaron el número óptimo de fotovoltaicos, turbinas eólicas, la capacidad del generador diésel y del banco de baterías. Además, el IGWO mostró el mejor rendimiento en comparación con el SSA y GWO.

Así mismo, **Anand et al. (2022)** estudiaron el dimensionamiento de diferentes configuraciones de un sistema de energía renovable híbrido para cubrir la demanda energética de 533 familias de la comunidad de Sonipat, Haryana, en la India. El sistema está configurado por un sistema fotovoltaico, biomasa, biogás acompañado con baterías, el cual su dimensionamiento es determinado por el Optimizador de Harris Hawk centrado en la Búsqueda Exploratoria Aleatoria (hHHO-RES, siglas en inglés) donde la función objetivo es el Costo Presente Neto (NPC) bajo un conjunto de restricciones, como los límites de las componentes del sistema y la confiabilidad. La configuración del sistema conformado por un sistema fotovoltaico y biomasa con baterías tuvo valores de NPC de \$ 781 000 y COE de \$ 0.115/kWh, siendo estos valores los más bajos que los valores del Algoritmo de Optimización de Búsqueda de Armonía (HS, siglas en inglés) y del Enjambre de Partículas (PSO, siglas en inglés).

**Arif et al. (2023)** realizan la simulación y optimización de un sistema fotovoltaico apoyado con baterías para abastecer la demanda energética de un área rural de forma rentable y confiable, denominada Energía Neta Cero (NZE). La radiación solar de la ubicación seleccionada es de 6.16 kWh/m<sup>2</sup>/día, mientras que la carga eléctrica estimada para toda la aldea es de 64.259 kWh. El programa HOMER estudió la viabilidad técnica y económica del sistema a través del Costo Neto Presente (NPC) y Costo de Electricidad (COE), alcanzado un valor mínimo de \$0.511 M y de 2.26 \$/unidad, respectivamente. El análisis de sensibilidad de ciertos parámetros (precio de los paneles solares, el precio de las baterías, la radiación solar, las variaciones en la carga eléctrica y las tasas de descuento) redujeron la inversión de capital inicial de \$0.434 M, permitiendo obtener un sistema a NZE y rentable.

En la literatura científica, los sistemas de generación híbridos con diferentes configuraciones de tecnologías (fotovoltaico, baterías, diésel, turbinas eólicas, biomasa, celda combustible) han sido estudiados en base a la optimización de diferentes algoritmos (Búsqueda de Armonía, Enjambre de Partículas, Harris Hawk centrado en la Búsqueda Exploratoria Aleatoria, Enjambre de Salpas, el Optimizador del Lobo Gris, Ecosistemas Artificiales, Saltamontes, Colonia de Abeja). La optimización en estos sistemas tiene como fin encontrar el menor Costo Presente Neto y Costo de la Energía e impacto medio ambiental, bajo ciertos criterios de confiabilidad del sistema. Así mismo, la operatividad y dimensionamiento de estos sistemas de generación también han sido estudiado por

programa comerciales, como HOMER, PVSYST y HOGA, dando resultados de diferentes indicadores tecnoeconómicos y ambientales.

Por lo tanto, en esta investigación tendrá un enfoque en la optimización y operatividad de un sistema de generación híbrido mediante la simulación en el programa HOMER y una implementación de un algoritmo que simule y optimice el sistema a través de un optimizador del tipo evolutivo, llamado Evolución Diferencial, que permita el dimensionamiento óptimo del sistema que garantice la confiabilidad.

## **2.2 Las energías renovables y la inteligencia artificial**

### **2.2.1 Las energías renovables en el mundo**

Las consecuencias ambientales generadas por el uso de recursos fósiles en la generación eléctrica impactan en gran medida a diversos países, generando aumento de la temperatura ambiente, inundaciones y cambios climáticos extremos cada vez con mayor frecuencia. Este panorama representa un peligro a la vida de muchas personas y la sostenibilidad del planeta. En base a ello, se ha llegado a acuerdos internaciones, como el Acuerdo de Paris y el Protocolo de Kioto, cuyo propósito es controlar y solucionar el problema del impacto ambiental y la sostenibilidad del planeta.

Por otro lado, las Naciones Unidas a través de una agenda global, llamada Los Objetivos de Desarrollo Sostenible (ODS), tiene como objetivo combatir el cambio climático, la erradicación de la pobreza, la igualdad de género, la energía asequible y no contaminante, entre otros (Naciones Unidas-CEPAL, 2019). Estos acuerdos y agenda global están alineadas en promover la sostenibilidad ambiental y abordar el tema del cambio climático. En ese sentido, diversos países han empezado a disminuir sus emisiones de gases que generan impacto negativo al ecosistema y buscan impulsar su transición energética en base a fuente de energía limpias, como las energías renovables.

#### **2.2.1.1 La potencia instalada de generadoras tipo RER en el mundo**

Según lo reportado por la Agencia Internacional de Energía Renovable (*The International Renewable Energy Agency-IRENA*) (IRENA, 2020b), las capacidades

totales de generación renovable han evolucionado a lo largo del 2008 hasta el 2018, teniendo un notable incremento que logran alcanzar un valor de 2 351 GW de capacidad instalada en el año 2018 a nivel mundial y que representa el 100% respecto a todas las renovables (Schmerler *et al.*, 2019).

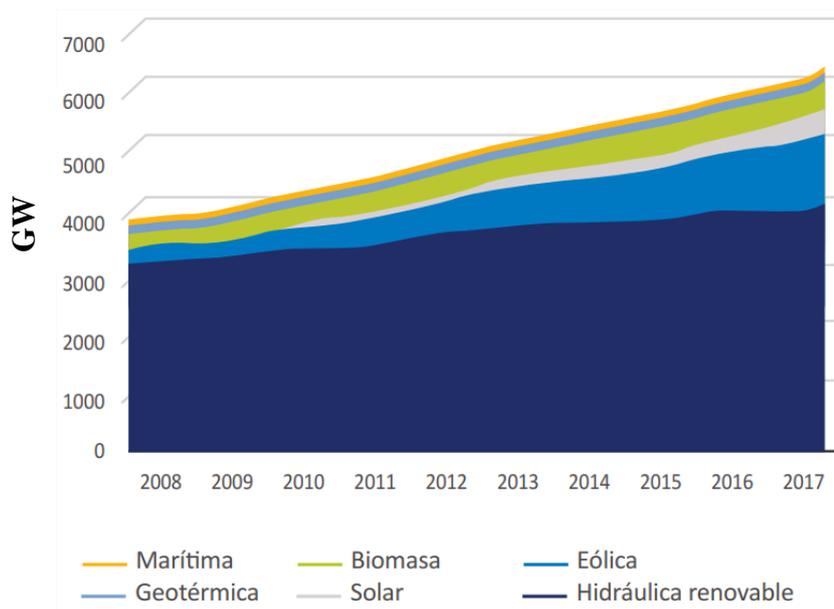
Dentro del incremento de la capacidad de generación renovable en el año 2018: La energía solar presenta una potencia instalada de 564 GW y la energía eólica con una potencia instalada de 480 GW, los cuales representan un 24 % y 26% de todas las energías renovables, respectivamente. Con respecto a las otras energías renovable, la energía marítima, biomasa y geotérmica representaron una potencia instalada de 500 MW, 121 GW y 13GW, respectivamente, que representan el 0.4%, 5% y 4% de las energías renovables (Schmerler *et al.*, 2019). Estos aumentos de la potencia instalada son mostrados en la Figura 1.

Por otro lado, en el 2022, las energías renovables tuvieron un aumento en la capacidad total, alcanzando 3 371.793 GW según los informes técnicos de IRENA (2023) y IRENA (2022). En el mismo año, la energía fotovoltaica y eólica incrementaron en un 191 GW y 75 GW. Así mismo, la capacidad instalada de biomasa y geotérmica aumentaron en 7.6 GW y 181 MW, respectivamente.

Así mismo, en el 2023, la potencia instalada total alcanza 4 137.3 GW según el reporte del Seguimiento del Proceso de Energía Renovable (en inglés, *Renewable Energy Progress Tracker-REPT*) (IEA, 2024). En este reporte, las instalaciones hidroeléctricas y bioenergía tuvieron 1 411.3 GW y 170.1 GW, respectivamente. Mientras que las instalaciones fotovoltaicas a escala de servicios públicos y distribuidos alcanzaron 1 550 GW, el cual es mayor a las instalaciones eólicas, onshore y offshore, que tienen un total de 1 005.0 GW. Estas informaciones son mostradas en la Figura 2.

**Figura 1**

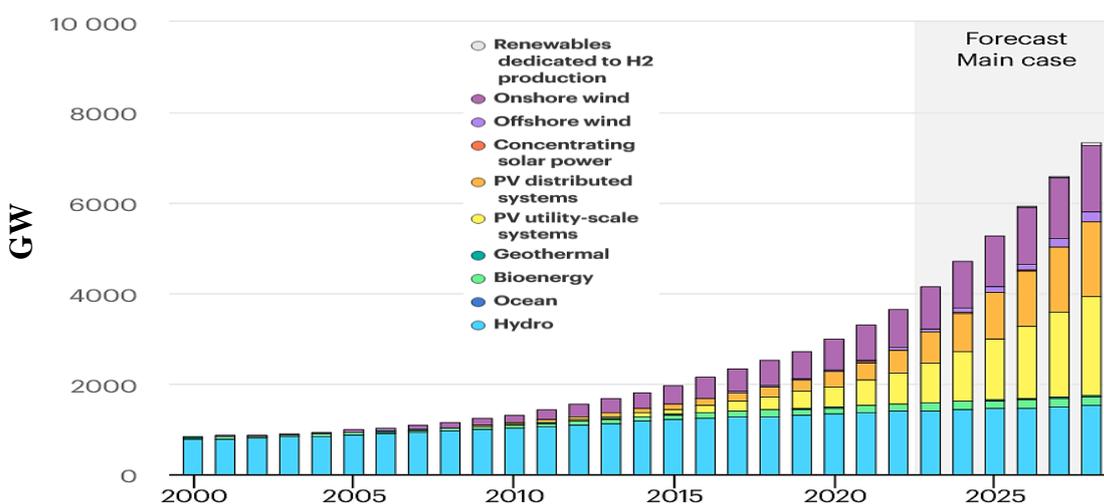
Generadores de energía renovable en base a su potencia instalada (GW) desde el año 2008 hasta 2018.



Nota: Imagen tomada de Schmerler *et al.* (2019) y visualizados por Gerencia de Políticas y Análisis Económico (GPAE)-Osiderming e IRENA.

**Figura 2**

Potencia instalada acumulada en GW para los generadores de energía renovable desde el año 2000 hasta 2022 y sus proyecciones.



Nota: Imagen tomada de Energía Renovable Progreso Tracker -IEA (IEA, 2024).

### 2.2.1.2 Evolución del costo nivelado de energía para las renovables

El costo de la energía de un proyecto en toda su vida útil es determinado por El Costo Nivelado de la Energía (en inglés Level Cost Of Energy-LCOE), en unidades de \$/kWh, donde se incluye la tasa de descuento, los costos de operación y mantenimiento (O&M), costo de inversión, tiempo de vida útil y otros costos en las cuales se incurre (Colla *et al.*, 2020) (Comello *et al.*, 2017). Este es un indicador que permite identificar y comparar diferentes tecnologías en base a su competitividad económica (Colla *et al.*, 2020)(Comello *et al.*, 2017).

En un análisis del LCOE de las tecnologías renovables entre los años 2010 y 2020 realizado en el resumen técnico de IRENA (IRENA, 2020a), se determina lo siguiente:

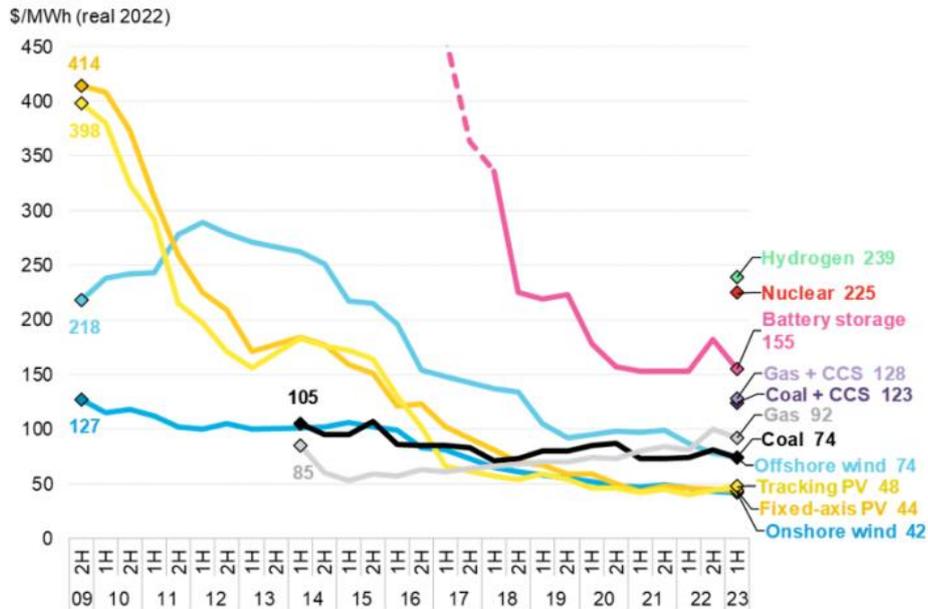
- En las tecnologías solar fotovoltaica se identificó que su LCOE promedio tiene un descenso de \$ 0.381/kWh hasta \$ 0.057/kWh, específicamente para proyectos de servicio público, en los años 2010 a 2020.
- En ese mismo año, las tecnologías eólicas terrestres disminuyeron de 0.089 USD/kWh hasta \$ 0.039/kWh. La reducción de ambas tecnologías solar y eólico se deben a su disminución de sus costos de instalación y caída de costos de los módulos fotovoltaico y turbinas a lo largo de esos años. Por otro lado, el LCOE promedio de la tecnología de biomasa se mantuvo a \$ 0.076/kWh.

Por otro lado, según el reporte de Bloomberg NEF (BloombergNEF, 2023) en el 2023, la tecnología solar fotovoltaica con seguimiento solar y eje fijo han alcanzado valores bajos de LCOE de \$ 0.048/kWh y \$ 0.044/kWh, respectivamente. Así mismo, la tecnología de turbinas marítimas y terrestres han llegado a tener un LCOE de \$ 0.074/kWh y \$ 0.042/kWh, siendo este último cercano a la tecnología fotovoltaica. Esta información es mostrada en la Figura 3.

Es notable un gran descenso del LCOE para las distintas tecnologías renovables (turbinas eólicas, fotovoltaica, baterías, hidrógeno, nuclear), incluso alcanzando valores por debajo de las tecnologías que utilizan carbón y gas. Esto es un indicador de que las tecnologías renovables están siendo más competitivas a lo largo de los años con

tecnologías tradicionales, generando una reducción en los costos de operación y mantenimiento y los costos de inversión.

**Figura 3**  
Comparación de LCOE de diferentes tecnologías renovables entre el año 2009 y 2023



Nota: Imagen tomada de la página web Bloomberg NEF (BloombergNEF, 2023)

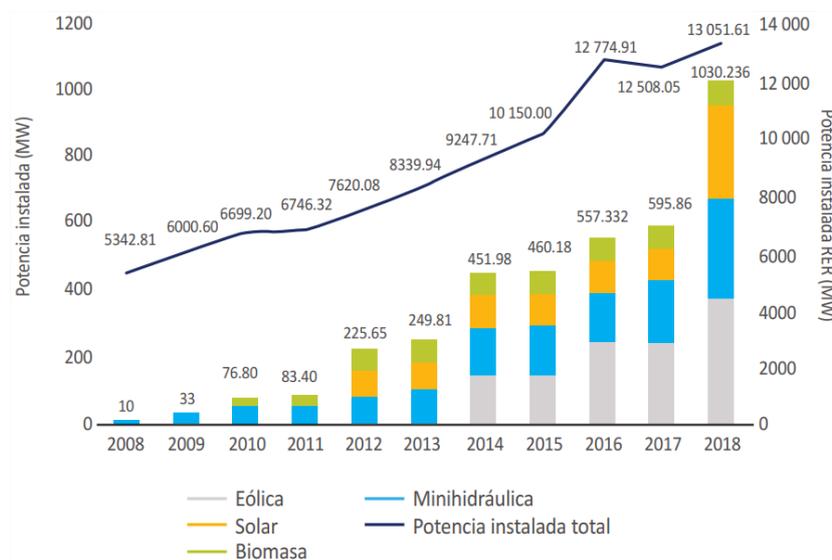
## 2.2.2 Estado de las energías renovables en el Perú

### 2.2.2.1 Potencia instalada de energía renovable en el Perú

Entre el año 2008 hasta 2018, Schmerler *et al.* (2019) indican que: La potencia instalada de las centrales de energía eléctrica renovable ha incrementado desde 10 MW hasta 1 030.2 MW, obteniendo una participación de 7.9% del total de potencia instalada en la industria eléctrica. Así mismo, es notable identificar que, a partir del 2012, la potencia de energía solar inicia y tiende al incremento e incluso alcanzando a la potencia instalada minihidráulica. Por lo que el incremento de estas potencias instaladas, tanto solar y minihidráulica, son debido al ingreso de nuevas centrales lo que haría su incremento en 196.9% y 58.8%, respectivamente, en el 2018 con respecto al 2017. La potencia instalada de la energía renovable a lo largo de los años es mostrada en la Figura 4.

Por otro lado, según el reporte estadístico del sector eléctrico (Asociación Peruana de Energía Renovable, 2023) para Julio del 2023, a través de la base de datos del COES (COES, 2023), indica que: La potencia instalada total de energía renovable (Eólica, Solar, Hidroeléctrica y Termosolar) ha alcanzado un total de 13 560.5 MW. Así mismo, la energía solar y eólica tienen una potencia instalada de 282.3 MW y 412 MW, respectivamente, las cuales representan en suma el 7 % de a capacidad instalada de energía renovable.

**Figura 4**  
Evolución de la potencia instalada en los años de 2008 a 2018 para las centrales de generación renovable en el Perú



Nota: Imagen tomada de Schmerler *et al.* (2019) y elaborada por Gerencia de Políticas y Análisis Económico (GPAE) de Osinergmin y COES (2019).

### 2.2.3 La inteligencia artificial y las energías renovables

#### 2.2.3.1 Las implicancias de la IA en la industria energética

La Inteligencia Artificial (IA) ha tomado gran importancia en los años más recientes en distintas áreas de la investigación, tanto en el desarrollo e implementación de modelos y las aplicaciones en la ingeniería e industrias. Actualmente, la IA ha atraído la atención de diversas industrias, como: El transporte, la salud y medicina, las finanzas, los servicios al cliente y la producción manufacturera, agricultura, comercio electrónico, energía y entre otros (Wamba *et al.*, 2020)(Yin *et al.*, 2021) (US-European Commission, 2022); permitiendo a las empresas un gran avance en la automatización de su operatividad

de forma eficiente y mejorando sus tomas de decisiones a través de modelos de IA que aprenden de gran cantidad de datos históricos.

La IA en el sector energético ha tenido un significativo impacto, permitiendo mejoras en la gestión de recursos y eficiencia, entre ellos se tienen (Feng *et al.*, 2021):

- La optimización en la generación de energía (Zahraee *et al.*, 2016). En este enfoque, el modelo de IA permite el ajuste de forma óptima entre el consumo energético y las distintas tecnologías de generación, permitiendo minimizar la producción y distribución de la energía de forma eficiente. Por lo que la aplicación de la IA genera en el sistema de generación una disminución de costos en la operatividad y mantenimiento, reducción del impacto ambiental y reducción de pérdidas de energía generando un ahorro energético.
- El pronóstico de las condiciones climática, generación de la energía, control y gestión de la demanda energética (Zhang, 2021). La IA, a través de algoritmos inteligentes como máquinas de aprendizaje o aprendizaje profundo, permite una mejora de la generación y distribución de la energía eléctrica a las empresas de manera más exacta con la predicción y previsionamiento en tiempo real o futuro del consumo de energía, velocidad del viento, temperatura ambiente, la radiación solar y entre otros. A través de estas informaciones, las empresas generadoras de energía presentan una mejor planificación, un buen ajuste entre su producción y consumo de energía.
- Mantenimiento predictivo en sistemas energéticos (Zhao *et al.*, 2019). La IA realiza un monitoreo y evaluación a tiempo real de todos los parámetros técnicos de un equipo y parámetros climáticos, a través de sensores especializados. En base a esto, los modelos de IA analizan los datos obtenidos e identifican patrones de fallos en el sistema, permitiendo así realizar acciones preventivas que reducen el costo de mantenimiento y evitando pérdidas de rendimiento.
- Nuevos materiales para el desarrollo de tecnologías energéticas (Liu *et al.*, 2021). La IA está siendo constantemente utilizada para el desarrollo de diversos materiales con propiedades sobre saliente en diversas tecnológicas como la

fotovoltaica, eólica, baterías, productor de hidrogeno y supercapacitores. Entre las propiedades, la IA identifica relaciones entre el rendimiento energético y las propiedades de un material, como la capacidad de almacenamiento, conductividad eléctrica, resistencia mecánica y estructura cristalina del material.

En base a lo mencionado, podemos indicar que la IA aplicada en el sector energético permite obtener grandes oportunidades y beneficios tanto en el desarrollo de novedosos materiales e incluso en la disminución de costos de inversión en sistemas de generación energética, aumentando la confiabilidad del sistema y minimizando los impactos ambientales.

### **2.2.3.2 La IA y políticas energéticas en el Perú**

Actualmente, el interés y uso de la inteligencia artificial ha aumentado en el Perú, así como a nivel mundial, con aplicaciones en diversos sectores del país, como educación, política, finanzas y otros (Agencia Andina, 2023).

Al comienzo de Julio del 2023, se publicó en el diario oficial, El Peruano, mediante la Ley N°. 31814: “Ley que promueve el uso de la Inteligencia Artificial (IA) en favor del desarrollo económico y social del país” (El Peruano, 2023). En el cual se menciona sobre la promoción y el impulso en la formación de profesionales, desarrollo y colaboración nacional e internacional y el uso de la IA en el desarrollo del país, así como también, en el fortalecimiento y creación de la infraestructura digital que dé soporte al desarrollo y uso de la inteligencia artificial, dando un uso replicable, transparente y sostenible en base a los lineamientos éticos.

Por otro lado, en el sector energético peruano se ha realizado diversas aplicaciones de la IA:

- El Premio Creatividad Empresarial 2022 fue otorgada a El Organismo Supervisor de la Inversión de Energía y Mina (OSINERGMIN) por su proyecto novedoso denominado “Menos interrupciones del servicio eléctrico para el ciudadano”, en la categoría de Innovación de procesos, organizada por una universidad peruana. Este proyecto utiliza la IA para realizar una supervisión de las desconexiones eléctricas

con la finalidad de predecir interrupciones en el sector eléctrico (OSINERGMIN, 2022).

- Por otro lado, según el Informe de Sostenibilidad Enel 2021 (ENEL, 2021) indica que ha mejorado las precisiones de la disponibilidad de las plantas generadoras eólicas y solares, a través de modelos de inteligencia artificial que brindan pronósticos a corto, mediano y largo plazo del recurso eólico y solar con buenas exactitudes.
- Así mismo, en el reporte ENEL (2021) también menciona que la empresa Enel Green Power Perú realiza inspecciones en base a la termografía a los sistemas de paneles fotovoltaicos, en el cual utiliza la IA para el análisis de imágenes a estos sistemas con el fin de mejorar la detección de fallas.
- De igual importancia, las empresas Enel X y on. Energy han implementado un programa on.Command™ que está basada en inteligencia artificial, el cual predice a tiempo real la demanda energética en su máxima y mínima potencia. En horas pico, donde la demanda de energía es máxima, el sistema de almacenamiento de la empresa Palmosa cubre la demanda máxima y evita los altos costos al utilizar plantas generadoras de energía (ENEL, 2021)(Oficina de Prensa de Enel Perú, 2022). Por lo que este sistema genera una disminución en la facturación eléctrica del usuario final.

## **2.3 Bases conceptuales**

### **2.3.1 Recurso renovable y modelos de pronóstico**

Los recursos energéticos no agotables que son utilizados para la generación de energía renovable presentan desafíos de la variabilidad o fluctuaciones debido a parámetros climático. En base a esta problemática, un campo en constante evolución son los modelos predictivos que permiten determinar el comportamiento de estos recursos renovables (radiación solar, viento, olas y mareas, entre otros) y conocer con exactitud su variabilidad (Islam *et al.*, 2023).

En los últimos años se han propuestos distintos modelos de predicción, con enfoques persistentes, físicos, estadísticos, inteligencia artificial y otros, para conocer la variabilidad de la velocidad del viento y de la radiación solar para su generación de energía (Ssekulima *et al.*, 2016). Entre ellos se tienen:

- El enfoque persistente que determina el valor predicho de una variable  $y_t$  igual al valor de la variable medido en un tiempo anterior  $y_{t-1}$  (Notton y Voyant, 2018). Así mismo, el enfoque físico considera los principios físicos y matemáticos, en base al modelamiento, para predecir parámetros climáticos (Hanifi *et al.*, 2020).
- Por otro lado, el enfoque estadístico está basado en el análisis de series de tiempo de una o varias variables con el fin de determinar una relación entre la variable de interés con otras variables, el cual permite predecirla (Ssekulima *et al.*, 2016). Existen diversos modelos que están comprendidos en este enfoque, entre los más relevantes son: El modelo Media Móvil Integrada Auto-Regresiva (en inglés Auto Regressive Integrated Moving Average-ARIMA), ARIMA estacional (en inglés Seasonal Auto Regressive Integrated Moving Average-SARIMA) y ARIMA con variables exógenas (en inglés Auto Regressive Integrated Moving Average with eXogenous variables-ARIMAX) (Ssekulima *et al.*, 2016)(Hammad *et al.*, 2020).
- Así mismo, en el enfoque de la inteligencia artificial, los modelos predictivos inteligentes presentan una mayor precisión en sus predicciones que les permiten describir el comportamiento de diferentes series de temporales. Entre las técnicas más conocidas: Red Neuronal Artificial (en inglés Artificial Neural Network-ANN), las Máquinas de Vector Soporte (en inglés Support Vector Machine-SVM), Red Neuronal Recurrente (en inglés Recurrent Neural Network-RNN) y otros (H. Wang *et al.*, 2022)(Hammad *et al.*, 2020).

### **2.3.2 Sistema de generación híbridos**

Los sistemas de generación híbrido consisten en integrar distintos sistemas de generación de energía, como sistemas fotovoltaicos, sistema de almacenamiento, sistema de aerogeneradores, diésel y entre otros (Rehman, 2021). Esta combinación de los sistemas, conectado o desconectado de la red, tienen como fin asegurar la

confiabilidad entre el consumo de energía y la generación eléctrica del sistema híbrido, reducir los costos de operación y mantenimiento, disminuir las emisiones de gases contaminantes y aumentar el tiempo de vida de los equipos. Estos tipos de sistemas de generación híbrida están siendo usado para electrificar lugares alejados de la red, los cuales tienen difícil acceso a la red eléctrica.

Además, los sistemas de generación mixta presentan una excelente solución para aumentar la participación de sistemas de generación de energía de fuentes renovables que utilizan recursos naturales, como radiación solar y viento, ya que permite a estos tipos de generación un soporte a través de sistemas, como baterías o generadores diésel, cuando su recurso presenta variabilidad debido a cambios climáticos.

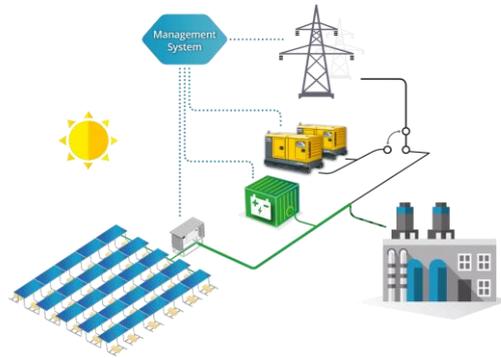
Las configuraciones de un sistema de generación híbrido pueden apoyarse a través de la combinación de sus tecnologías, entre ellas se tienen:

- Sistema fotovoltaico apoyado por un sistema de baterías
- Sistema fotovoltaico y grupos electrógenos, con y sin apoyo de baterías.
- Sistemas fotovoltaicos y sistemas eólicos y electrógenos, con y sin sistemas de baterías.
- Sistema eólico y sistema fotovoltaico con almacenamiento de hidrógeno.

En la Figura 5 se muestra un ejemplo esquemático de la generación híbrida de energía de un sistema fotovoltaico y un sistema de almacenamiento de baterías cubriendo la demanda energética. Si el sistema entra en déficit de generación de energía, el sistema es apoyado por un grupo electrógeno y la red eléctrica.

**Figura 5**

*Esquema de un sistema híbrido conformado por paneles fotovoltaico, inversores, baterías, generador diésel y red eléctrica*



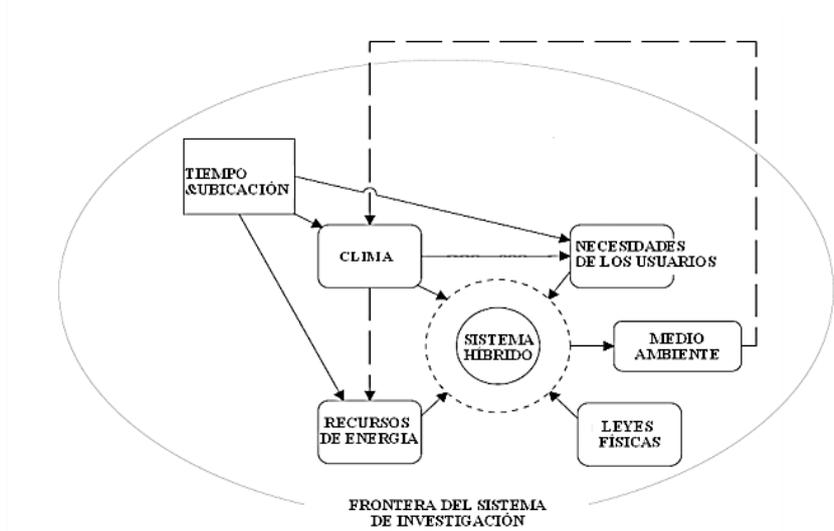
Nota: Imagen tomada de RedDavia Solar Power (2018)

### 2.3.2.1 El sistema híbrido fotovoltaico y sus factores influyentes

Ulleberg (1998) indica en su investigación que existen diversos factores que afectan al sistema de generación híbrido del tipo fotovoltaico. La Figura 6 describe estos factores mencionados por el autor, resaltando que la ubicación y tiempo meteorológico en un lugar son uno de los factores principales que afectan la operatividad eficiente de un sistema de generación, el cual genera cambios en el recurso energético del lugar, clima y necesidades en los usuarios.

**Figura 6**

*Sistema de generación híbrido afectado directa o indirectamente por factores externos*



Nota: Imagen tomada de Cruz (2010) y elaborada en inglés por Ulleberg (1999)

## **CAPÍTULO III – METODOLOGÍA**

### **3.1 Tipo de investigación**

La presente investigación de tesis doctoral tiene un enfoque cuantitativo.

### **3.2 Nivel de investigación**

La tesis presenta un nivel de investigación aplicativo.

### **3.3 Diseño de investigación**

La aplicación de herramientas de inteligencia artificial para predecir información relacionada a un sistema de generación y la optimización del mismo es el enfoque de esta investigación.

Esta tesis tiene los siguientes pasos para llegar a los objetivos propuestos y son esquematizados en la Figura 7:

- 1) Como primer proceso en esta investigación, se hace un estudio relacionado a la predicción de series temporales utilizando modelos predictivos, los cuales siguen los siguientes pasos:
  - La primera etapa es identificar un lugar de estudio donde se adquiere la información climática, como temperatura ambiente y la radiación solar, y la potencia de consumo en series temporales horarias.
  - La segunda etapa se hace una amplia revisión literaria de varios modelos de predicción reportado en artículos científicos, en donde se identifica 3 modelos que serán utilizados en la predicción de las series de tiempo. Estos modelos son comparados con el fin de determinar el modelo con mayor rendimiento.
- 2) El segundo procedimiento en esta investigación, se analiza la operatividad y costos del sistema de generación híbrido (fotovoltaico, baterías y diésel) mediante el programa HOMER.

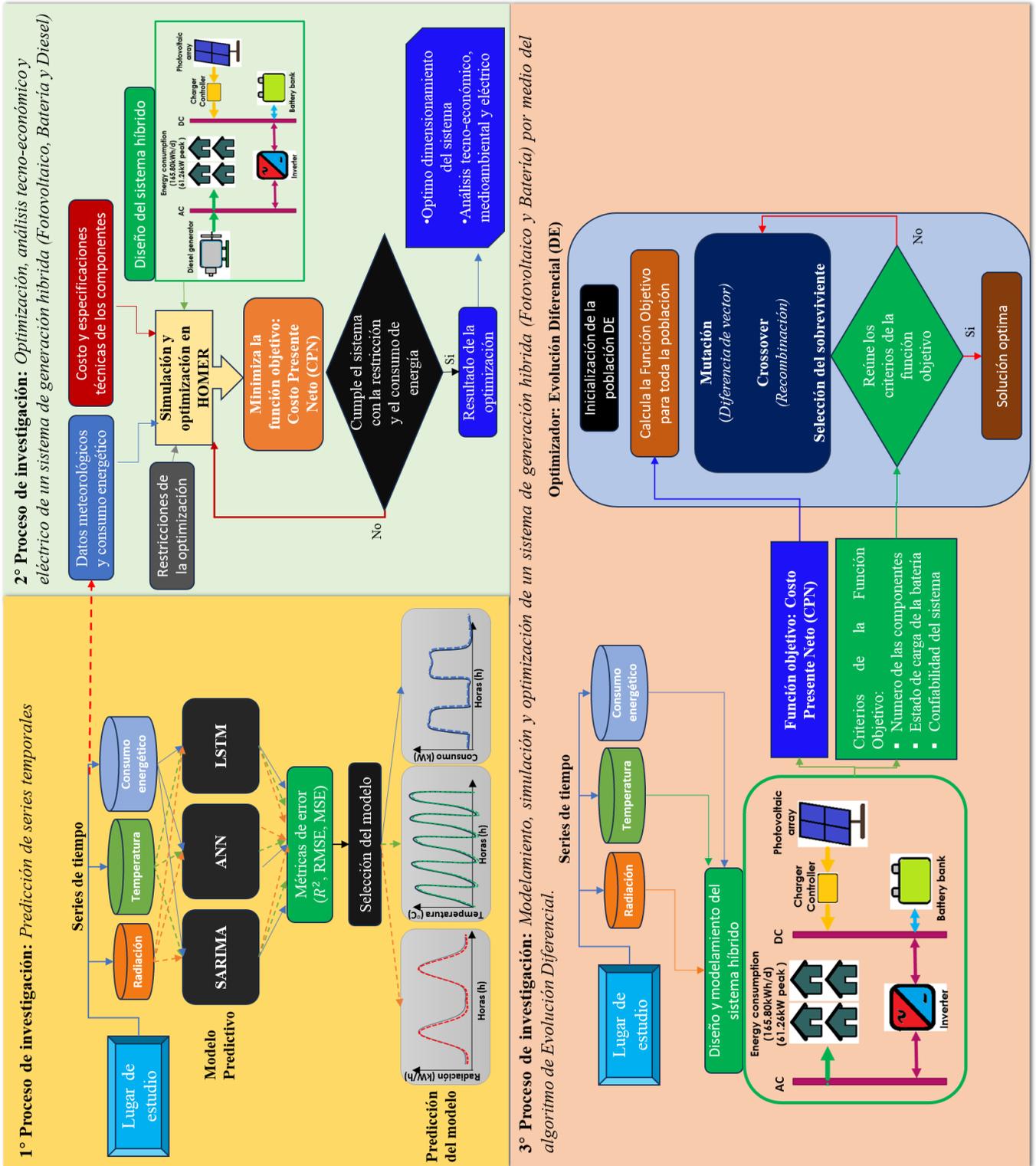
- La primera etapa en este procedimiento se identifica un lugar de estudios y se obtiene la radiación solar, temperatura ambiente y consumo de energía.
  - La segunda etapa se identifica los costos (Capital, operación, tiempo de vida, entre otros) y parámetros técnicos de los componentes del sistema de generación.
  - La tercera etapa se utiliza el programa HOMER para diseñar un sistema de generación, conformada por Fotovoltaico, baterías y diésel, con el fin de optimizar los costos del sistema en base al costo presente neto y la cantidad de componentes del sistema.
  - La cuarta etapa se evalúa el sistema mediante un análisis tecno-económico, basándonos en la producción de energía eléctrica, emisión de contaminantes y otros.
- 3) El tercer procedimiento en esta investigación, se diseña y elabora un sistema de fotovoltaico híbrido en donde se optimiza la cantidad de módulos fotovoltaicos y baterías con el fin de equilibrar la producción eléctrica del sistema y la demanda de energía de un lugar de estudio.
- La primera etapa se diseña y elabora un algoritmo de optimización inteligente del tipo Evolución Diferencial, en lenguaje de programación Python, a fin de optimizar el costo presente neto a través de las variables de cantidad paneles fotovoltaico, baterías, controladores e inversor, estados de la batería y otros parámetros.
  - La segunda etapa se diseña y elabora la operatividad de un sistema de generación fotovoltaico y almacenamiento de baterías, mediante la programación, comprendiendo los siguientes aspectos: la generación fotovoltaica a través de la radiación y temperatura ambiente, estados de descarga y carga de la batería cuando la generación fotovoltaica está en exceso y déficit, el costo presente neto considerando la vida útil de cada

uno de los componentes del sistema, costos totales y unitarios, y parámetros técnicos.

- La tercera etapa se realiza la simulación, donde ingresa la información de la temperatura ambiente, el consumo de energía del lugar y la radiación solar. Esta simulación consiste en realizar la optimización y operación del sistema de generación fotovoltaica y baterías, por medio del algoritmo inteligente y la estrategia de operación para obtener los valores de óptimos del costo presente neto del sistema y las cantidades necesarias de sus componentes.

**Figura 7**

Los tres procesos de investigación realizados en la tesis doctoral



### 3.4 Fuentes de recolección de datos

Los algoritmos predictivos requieren información previa para realizar sus predicciones, por lo que se requiere lo siguiente:

- La información obtenida en esta investigación se basa en datos climático del tipo serie temporales de la temperatura ambiente y la radiación solar. La información mencionada se obtuvo para los años 2020 y fue obtenida de la página web de la NASA.
- Por otro lado, la información del consumo de energía de los lugares en estudio es obtenida de distintas formas: La información del consumo de energía de la comunidad de Tarucani ha sido obtenida en base a un registro de la información basándonos en la potencia de cada dispositivo y el tiempo de uso en horas, mostrado en el Anexo 1. Por otro lado, la potencia de consumo de la Biblioteca de la Universitaria se ha registrado para una semana en etapas de tiempo de horas y convertidas en series temporales a lo largo del año, utilizando el programa HOMER.

Con respecto a la simulación mediante el programa HOMER, la información necesaria fue obtenido de la siguiente forma:

- Los datos para la simulación se requirió la radiación, temperatura ambiente y consumo de energía de un lugar en estudio. Así mismo, la información de cada uno de los componentes del sistema fotovoltaico híbrido (Fotovoltaico, baterías y diésel) se necesitó su costo unitario, costos de operación y mantenimiento (O&M) y costo de reemplazo, entre otros. Esta información técnica fue obtenida desde la página AutoSolar que muestran en sus catálogos. Esta información se observa en el Anexo 2.

Con respecto a la implementación del diseño de un sistema de generación energética híbrido (Fotovoltaico y Baterías), un algoritmo de optimización llamado Diferencial Evolutivo y un algoritmo de gestión energética del sistema han sido elaborados en el lengua de programa Python, en el cual la información del consumo de energía, la radiación solar, temperatura ambiente y los costos de los componentes del

sistema híbrido han sido usados para obtener la óptima gestión de energía del sistema y parámetros económicos como el Costo de Energía y Costo Presente Neto.

### **3.5 Población y muestra**

Esta tesis, la población de estudio está conformada para 200 viviendas, una escuela, una comisaría, un puesto de salud, una iglesia y las iluminaciones de la calle que comprende la comunidad rural.

La muestra se ha considerado la misma cantidad de la población, considerando aspectos representativos de las viviendas e instalaciones.

### **3.6 Técnicas de procesamiento de datos**

#### **3.6.1 Desarrollo de modelos de pronóstico de series de tiempo**

En esta sección se describe el diagrama de flujo de los modelos de predicción, utilizados en esta investigación, y la formulación matemática de los modelos y sus procesos de cálculo. Así mismo, la eficiencia basada en la precisión de cada uno de los modelos es evaluado mediante la métrica de error.

##### **3.6.1.1 Diagrama de flujo de los modelos de pronóstico**

En esta investigación, el modelo SARIMA, ANN y LSTM fueron usados para predecir la información de la potencia de consumo, la radiación solar y la temperatura ambiente para el período de un año con registros de pasos de 1 hora. La adquisición de la información de la temperatura ambiente y radiación solar se obtuvieron de la base de datos de la NASA por medio de las coordenadas del lugar de estudio. Así mismo, la data de la potencia de consumo fue estimada a lo largo del año para el lugar de estudio. Estos registros de datos son series de tiempo univariadas, los cuales fueron pre procesadas y posteriormente procesadas por cada uno de los modelos.

Las metodologías de los modelos son descritas a continuación:

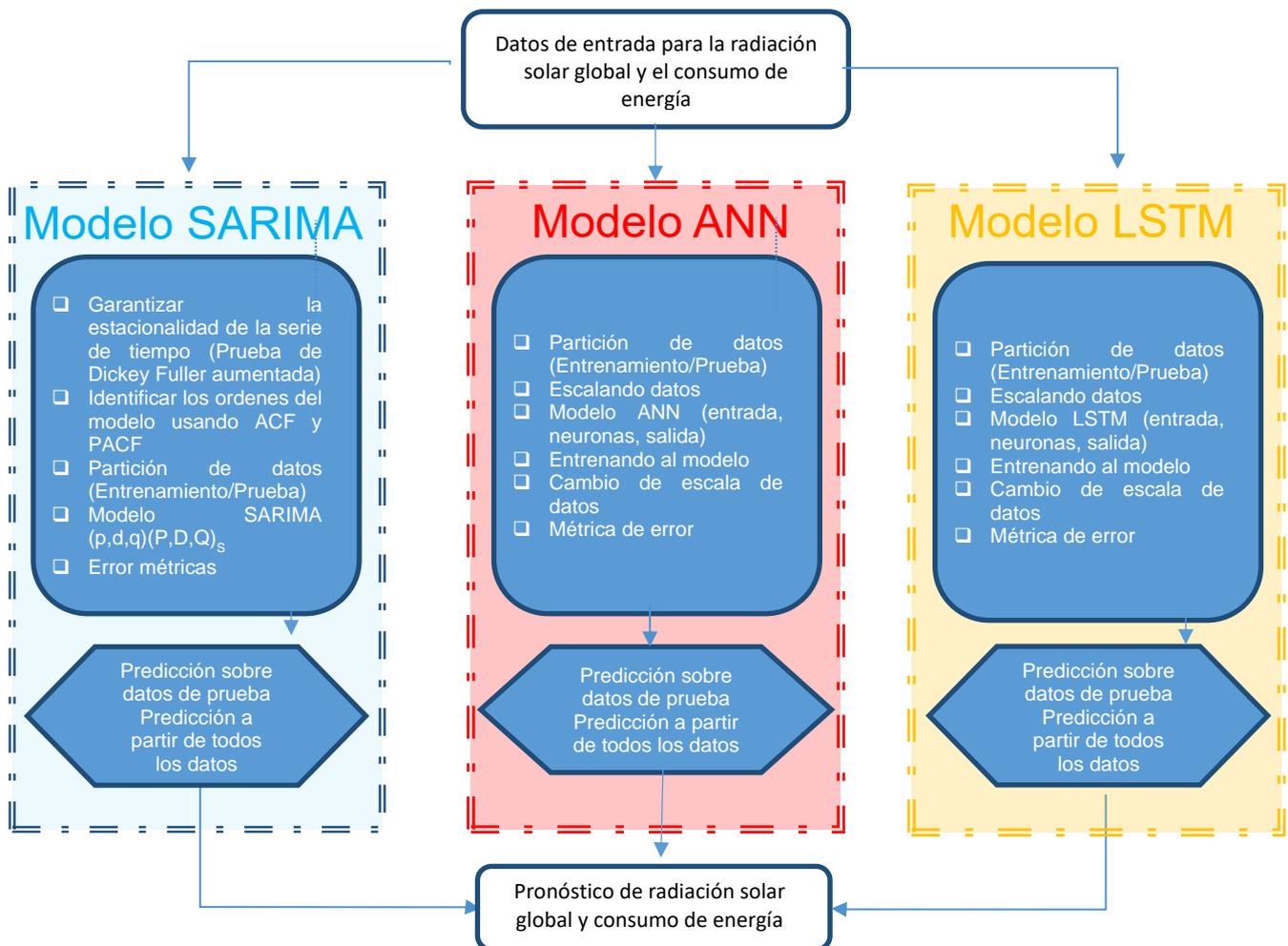
1. El modelo SARIMA consiste en varios procesos para su análisis y evaluación, entre ellos es primero determinar la estacionalidad de las series de tiempo mediante la prueba de Dickey-Fuller Aumentada y luego hacer una división de la data de entrenamiento y prueba, considerando un 10% de todas las series de tiempo para este último. En segundo lugar, se identifican los parámetros u órdenes del modelo SARIMA de la forma  $(p, d, q) (P, D, Q)_m$  analizando las funciones de AutoCorrelacion (ACF) y Parcial ACF (PACF). Finalmente, las pruebas y selección de los modelos para cada serie de tiempo son determinadas mediante las métricas de error MSE, RMSE, MAE y  $R^2$ .
2. De la misma forma, el modelo ANN presenta varias etapas para analizarse y evaluarse: Primero, se empieza con el preprocesamiento de la serie de tiempo, el cual consiste en la limpieza de los datos, la partición de la data en entrenamiento y prueba, el escalamiento o normalización de la data de entrenamiento y la transformación de los datos mediante la técnica de desplazamiento de ventanas (Sliding Window, en inglés) (Vafaeipour *et al.*, 2014). La data de entrenamiento es transformada en pasos de tiempo para generar los input y output que serán interpretados por el modelo ANN. Posteriormente, los modelos son entrenados con la data de entrenamiento y su diagnóstico del aprendizaje se evalúa mediante la curva de entrenamiento. Así mismo, se realiza el rescalamiento de los datos predichos por el modelo y comparados con los datos observables de prueba. Esta comparación es realizada mediante las métricas de error, la cual permite identificar que modelos son más precisos con respecto a los otros.
3. El modelo LSTM es similar al modelo ANN en base a su pre procesamiento y procesamiento de los datos, sin embargo, la diferencia que presenta el LSTM, con respecto al ANN, es que tiene una estructura de celda de memoria que está conformado por compuertas, permitiendo recordar a largo y corto plazo.

Finalmente, las mejores configuraciones de los modelos SARIMA, ANN y LSTM que presenten mayor precisión serán usados para el pronóstico de las series de tiempo para un futuro considerando ciertos intervalos de confianzas. Un resumen de

etapas seguidas para el tratamiento de los datos y los modelos son mostrados en la Figura 8. Los códigos de cada uno de los modelos SARIMA, ANN y LSTM son mostrados en los Anexo 3, Anexo 4 y Anexo 5, respectivamente.

**Figura 8**

*Pre procesamiento y procesamiento de los datos para cada uno de los modelos*



### 3.6.1.2 Modelo predictivo de media móvil integrada autorregresiva estacional (SARIMA)

#### 3.6.1.2.1 Pre-procesamiento e identificación de parámetros

El modelo SARIMA o llamado también estacionario ARIMA es un modelo desarrollado bajo la metodología Box-Jenkins (Tsoku *et al.*, 2017). Este modelo se ha convertido en el más popular modelo para el pronóstico de series de tiempo, su diferencia con respecto al modelo ARIMA es porque incluye el componente estacionario con

patrones horarios, diarios, mensuales y anuales en la serie de tiempo (Hanaa Elgohari, 2018).

El modelo SARIMA está conformado de un modelo auto regresivo (AR), media móvil (MA) quienes presentan componentes no estacionales y estacionales. El modelo está representado de la forma  $(p,d,q)(P,D,Q)_s$  donde sus siete parámetros son enteros y positivos llamados ordenes, siendo representado matemáticamente de la forma (Cui *et al.*, 2016):

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^DY_t = \theta_q(B)\Theta_Q(B^S)e^t \quad (1)$$

donde  $\phi_p(B)$  y  $\theta_q(B)$  son el orden polinomial AR y MA de orden p y q, respectivamente. Similarmente,  $\Phi_P(B^S)$  y  $\Theta_Q(B^S)$  son el orden estacional polinomio AR y MA de orden P y Q. También se tiene  $Y_t$ , el cual es el valor de la serie de tiempo observada y  $e^t$  es el error o ruido blanco en el tiempo t. Los operadores diferenciales representados por  $(1-B)^d$  y  $(1-B^S)^D$  son diferenciales no estacionales y estacionales de orden d y D, respectivamente.

Los órdenes poligonales AR y MA tanto para no estacionales y estacionales son representados de la siguiente forma (Cui *et al.*, 2016)(Khanarsa y Sinapiromsaran, 2020):

$$\phi_p(B) = 1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p \quad (2)$$

$$\Phi_P(B^S) = 1 - \Phi_1B^S - \Phi_2B^{2S} - \dots - \Phi_PB^{Ps} \quad (3)$$

$$\theta_q(B) = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q \quad (4)$$

$$\Theta_Q(B^S) = 1 - \Theta_1B^S - \Theta_2B^{2S} - \dots - \Theta_QB^{Qs} \quad (5)$$

donde S representa la duración del período estacional y B es el operador de retroceso que transforma un  $Y_t$  usando la forma  $BY_t = Y_{t-1}$ .

### 3.6.1.2.2 Modelo de redes neuronales artificiales: Perceptrón multicapa (MLP-ANN)

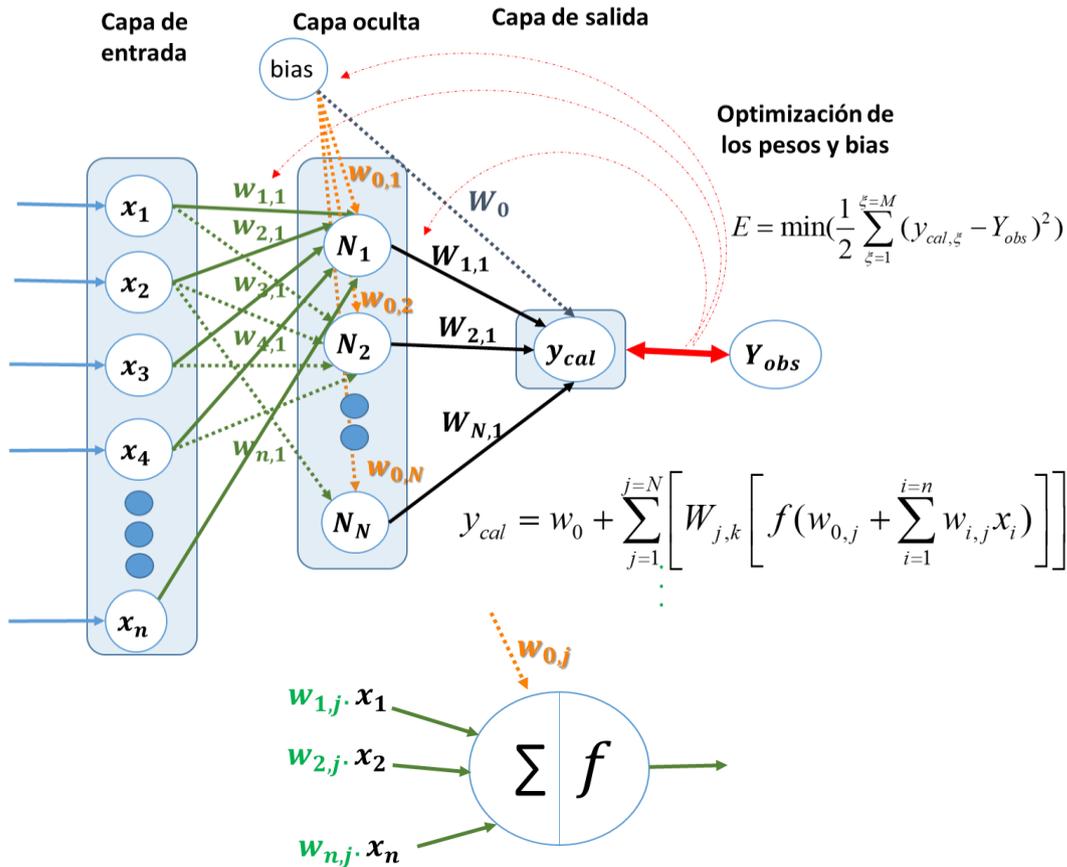
Las redes neuronales artificiales (ANN) es una técnica computacional e inteligente que fue inspirado en el procesamiento de la información por medio de las neuronas y sistema nervioso biológico (Vanneschi y Castelli, 2018). Esta técnica permite resolver problemas difíciles en muchas áreas de la investigación, por ejemplo, para la optimización de sistemas, predicciones o pronósticos de series de tiempo, clasificación de información, reconocimiento de imagen y voz, entre otros (Adebisi *et al.*, 2014)(Abiodun *et al.*, 2019)(Hatami *et al.*, 2018).

Hay variedades de tipos de ANN en la literatura (Durstewitz *et al.*, 2019), sin embargo, el más resaltante es el multilayer perceptron quien es uno de los modelos más usados para predecir series de tiempo. Este modelo está conformado por una capa de salida, capas ocultas y capa de entrada; las cuales están interconectadas por medio de sus pesos. El proceso que realiza una neurona es multiplicar los valores de los parámetros de entradas y cada uno de sus pesos asignados; siendo después sumadas con su sesgo o bias, para posteriormente aplicar la función de activación o transferencia y determinar un valor de salida. La estructura y procesos de una red neuronal es mostrada en la Figura 9 y su representación matemática tiene la forma (Khandelwal *et al.*, 2015)(Tealab *et al.*, 2017):

$$y_k = f_0 \left( \sum_{j=1}^n w_{kj} f_h \left( \sum_{i=1}^m w_{ji} X_i + w_{jb} \right) + w_{kb} \right) \quad (6)$$

donde  $X_i$  es la variable de entrada y  $w_{ji}$  es el peso asignado para conectar de las neuronas  $i$ th y las neuronas  $j$ th que conforman cada capa de entrada y capa oculta, respectivamente. La cantidad de neuronas ocultas es determinada por  $m$ , mientras que el término  $w_{jb}$  es el bias designado para cada  $j$ th neurona oculta. El  $f_h$  representa la función de activación para las neuronas ubicadas en la capa oculta. El término que conecta las neuronas  $j$ th ubicadas en la capa oculta y neuronas  $k$ th ubicadas en la capa de salida es el peso  $w_{kj}$ . Por otro lado, en la capa oculta,  $w_{kb}$  representa el bias para una de las neuronas  $j$ th. El  $f_0$  es la función de activación para las neuronas de salida y  $y_k$  es el valor de salida o valor de predicción  $k$ th.

**Figura 9**  
Estructura y procesos del multilayer perceptron (MLP)



### 3.6.1.2.3 Modelo de memoria a corto y largo plazo

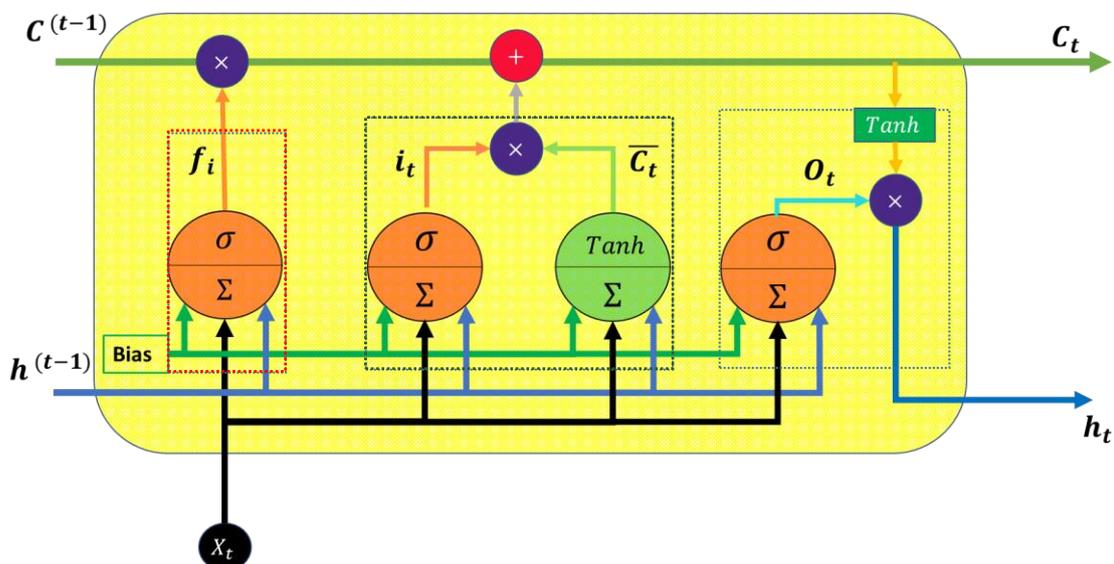
El modelo de memoria a corto y largo plazo (en inglés *Long Short-Term Memory-LSTM*) tiene la estructura mejorada que el modelo de red neuronal recurrente (en inglés Recurrent neural Network-RNN). Este modelo tiene la capacidad de recordar información durante periodos de tiempos cortos y largos, permitiéndole así leer, escribir y borrar información en sus celdas memorias (Sima *et al.*, 2019).

La celda o célula de memoria LSTM está compuesta de compuertas que permiten decidir si la información de la memoria se conserva o ignora durante un largo periodo de tiempo, siendo esta decisión tomada por el modelo a través de un proceso de entrenamiento mediante la designación de pesos y la información ingresada a la celda (Science *et al.*, 2019).

La Figura 10 representa el proceso de funcionamiento de la célula de memoria de la red LSTM. Esta célula está compuesta por funciones del tipo sigmoidea y tangencial, sesgos  $b_n$  y pesos  $W_m$  asignados a cada conexión. Los detalles de la célula de memoria de la red LSTM son descritos a continuación (Do *et al.*, 2020):

- El recuadro de color rojo representa la puerta del olvido, la cual es determinada por la función de activación sigmoid quien toma valores entre 0 a 1 indicando eliminación o retención, respetivamente, del estado actualizado  $C_{t-1}$  de la celda anterior en el tiempo  $t - 1$ . Este estado actualizado también es llamado memoria de largo plazo.
- Así mismo, el recuadro color verde muestra la puerta de entrada quien determina si se almacenan o no los nuevos datos en el estado de la célula permitiendo así tener un nuevo estado de la célula memoria  $C_t$ , también llamado nueva memoria a largo plazo. En esta puerta, el valor que debe actualizarse se determina mediante la función sigmoidea y el vector que debe añadirse al estado de la celda, el cual se genera mediante la función  $Tanh$ .
- Por otro lado, la puerta de salida mostrada en el recuadro azul determina el valor de la salida  $h_t$ , el cual es el resultado de filtrar el valor de  $O_t$  y función  $tanh$  del estado  $C_t$  de la celda.

**Figura 10**  
Estructura de la celda de memoria LSTM



Las formulaciones matemáticas de las puertas internas de la celda de memoria del modelo LSTM son detalladas de la forma:

### 1) Puerta del olvido:

En esta etapa, la celda unidad del LSTM determina que información debe ser eliminada de sus estados de celda anterior  $C^{(t-1)}$ . Esta puerta del olvido es descrita por la función  $f_t$ , el cual tiene valores entre 0 y 1, y su valor de  $f_t$  es determinada por la función activación sigmoid ( $\sigma$ ), la entrada actual  $X_t$ , el bias  $b_f$  y la salida  $h_{t-1}$  de la unidad de memoria anterior ( $t - 1$ ). Este proceso que ocurre en la puerta del olvido es dado por la expresión de la función  $f_t$  dado de la forma (Salman *et al.*, 2018):

$$f_t = \sigma(W_f \otimes X_t + U_f \otimes h_{t-1} + b_f) \quad (7)$$

donde  $W_f$  y  $U_f$  son los pesos asociados con  $X_t$  y  $h_{t-1}$ , respectivamente, mientras que  $b_f$  denota los pesos del bias.

### 2) Puerta de entrada:

En esta etapa, esta puerta toma la decisión sobre la nueva información para ser almacenada en el estado de la celda (también llamada memoria LSTM). Esta puerta consiste de 2 capas: una capa sigmoid y una capa tanh. La capa sigmoid decide cual valor necesita ser actualizado, haciendo que el valor de  $h_{(t-1)}$  pase a través de la función sigmoide y obtener un valor  $i_t$  entre 0 y 1. Por otro lado, la capa tanh genera un vector de valores con nuevos candidatos que serán agregados en la memoria LSTM, donde el valor de  $h_{(t-1)}$  pasa a través de función tanh para obtener un nuevo valor candidato  $\bar{C}_t$ . Las salidas de estas dos capas son expresadas de la forma (Yang *et al.*, 2022):

$$i_t = \sigma(W_{ix} \otimes X_t + U_i \otimes h_{t-1} + b_i) \quad (8)$$

$$\bar{C}_t = \tanh(W_{cx} \otimes X_t + W_{ch} \otimes h_{t-1} + b_c) \quad (9)$$

donde  $i_t$  representa si el valor necesita ser actualizado o no y  $\bar{C}_t$  indica un vector de nuevos valores candidatos que serán agregados en la memoria LSTM.

### 3) Almacenamiento:

A continuación, la combinación de estas dos puertas (la puerta de entrada y puerta del olvido) provee una actualización para la memoria LSTM. Esta actualización consiste en obtener el valor actual  $C_t$  de la memoria mediante la suma de la capa de la puerta de olvido, por medio de la multiplicación del valor  $f_t$  y  $C_{t-1}$ , y la capa de la puerta de entrada, a través de la multiplicación de  $i_t \times \bar{C}_t$  del estado anterior de la célula  $C_{t-1}$ . Esta suma considera qué información debe abandonarse y almacenarse y crea un nuevo estado de la célula  $C_t$ .

Este proceso puede formularse del siguiente modo (Yang *et al.*, 2022):

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \bar{C}_t \quad (10)$$

### 4) Puerta de Salida:

Esta etapa controla la puerta de salida de la unidad LSTM, el cual combina la entrada actual  $X_t$ , la salida de la unidad  $h_{t-1}$  y el valor de la celda  $C_t$ . La salida de la LSTM se basa en actualizar el estado de la celda  $C_t$ . En primer lugar, empleamos la función sigmoide  $\sigma$  para generar un valor  $O_t$  que permite controlar la salida de celda unidad  $h_t$ , llamado también el nuevo estado oculto. A continuación, la función tanh del estado de la celda actualizada  $C_t$  y la salida de la función sigmoidea  $O_t$  se utilizan para generar el estado de la celda  $h_t$ . Por lo tanto, el valor de  $O_t$  y  $h_t$  son calculado por las siguientes ecuaciones (Soares de Araujo, 2020):

$$O_t = \sigma(W_0 \otimes X_t + U_0 \otimes h_{t-1} + b_0) \quad (11)$$

$$h_t = O_t \otimes \text{Tanh}(C_t) \quad (12)$$

donde  $W_0$  y  $U_0$  son los pesos,  $X_t$  es la entrada actual,  $b_0$  es el bias o sesgo,  $O_t$  es la compuerta de salida en el tiempo  $t$ ,  $h_t$  representa el valor de la memoria y  $C_t$  es el valor del estado de la memoria de la celda LSTM en el tiempo  $t$  (Wang *et al.*, 2018).

### 3.6.1.2.4 Métricas de evaluación del error

Para evaluar la precisión, entre los valores observados que conforman la serie de tiempo y los valores predichos obtenido por los modelos, se utilizan las métricas de error. Existen varios criterios que investigadores utilizan para evaluar el rendimiento o error de los modelos mediante las métricas. Estos criterios son definidos: Error Cuadrático Medio Raíz (en inglés *Root Mean Squared Error-RMSE*), Error Absoluto Medio (en inglés *Mean Absolute Error-MAE*), Error Cuadrático Medio (en inglés *Mean Squared Error-MSE*) y el coeficiente de determinación (en inglés *R squared*). MSE y RMSE determinan la desviación entre las observaciones y los valores predictivos.

La diferencia entre los valores predichos y observadas son medidos por MAE; además *R squared* mientras más cercano a 1 indica que el modelo tiene un buen rendimiento. Las expresiones matemáticas de estos criterios son mostradas a continuación (Chi, 2022)(Singh *et al.*, 2019):

$$MSE = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N} \quad (13)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (14)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (15)$$

$$R^2 = -\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} + 1 \quad (16)$$

donde  $N$  es el número de datos,  $y_i$  son los valores observables de la serie de tiempo de la radiación solar, temperatura ambiente y potencia de consumo y  $\bar{y}_i$  su media. Las predicciones de cada modelo para las series de tiempos son  $\hat{y}_i$ .

### 3.6.1.2.5 Intervalo de confianza del modelo

Las predicciones de los modelos presentan grados de incertidumbre que deben considerarse cuando se realizan pronósticos. Esta incertidumbre es cuantificada a través de intervalos de confianza que presentan un límite inferior y un límite superior, donde un valor predicho por el modelo se ubica entre estos límites con una probabilidad determinada (Chatfield y Yar, 1991).

El intervalo de confianza (IC) de una predicción al  $100(1 - \alpha)\%$  es estimado de la siguiente forma (Dos Santos y Franco, 2019):

$$IC_{100(1-\alpha)\%} = [y_i + z_{\alpha/2}RMSE, y_i - z_{\alpha/2}RMSE] \quad (17)$$

donde  $y_i$  es el valor predicho por el modelo,  $\alpha$  es el nivel de significación y  $z_{\alpha/2}$  es el intervalo de confianza de  $100(1 - \alpha)\%$  o el valor de la distribución normal estándar para una probabilidad. Por otro lado, RMSE es el error cuadrático medio obtenido de la ecuación (14).

## 3.6.2 Modelamiento de los componentes del sistema fotovoltaica híbrido

Los modelos matemáticos de cada componente del sistema híbrido fotovoltaico son descritos en esta sección, el cual comprende al sistema fotovoltaico, controlador e inversor, generador diesel y banco de baterías.

### 3.6.2.1 Sistema fotovoltaico

La salida de potencia del sistema fotovoltaico se expresa mediante la siguiente ecuación (18), en donde se muestran los parámetros de la radiación, la temperatura de la

célula fotovoltaica y otros parámetros técnicos del módulo fotovoltaico (Oladigbolu *et al.*, 2019):

$$P_{PVout(t)} = P_{NPV} \left( \frac{G(t)}{G_{ref}} \right) [1 + \alpha_T (T_{Cell(t)} - T_{ref})] \eta_{Mpppt} \eta_{Inv} \quad (18)$$

donde  $P_{NPV}$  (kW) es la capacidad proporcionada por el conjunto fotovoltaico en condiciones estándar o de referencia y  $P_{PVout}$  (kW) es la potencia de salida del sistema fotovoltaico.  $G$  es la radiación solar (W/m) y  $G_{ref}$  es la radiación solar en condiciones de referencia ( $G_{ref}=1000\text{W/m}$ ). El coeficiente de temperatura de la potencia máxima para un tipo de tecnología fotovoltaica, como monocristalina o policristalina, es representada por  $\alpha_T$  (%/°C). La temperatura en condiciones de referencia y la temperatura de la célula fotovoltaica se denotan como  $T_{ref}$  (°C) y  $T_{Cell}$  (°C), respectivamente. Además,  $\eta_{Mpppt}$  y  $\eta_{Inv}$  representan las eficiencias del controlador e inversor, respectivamente.

La variación de temperatura durante el funcionamiento del módulo se expresa mediante la ecuación (Skoplaki *et al.*, 2008):

$$T_{Cell(t)} = T_{Amb(t)} + \left( \frac{NOCT-20^\circ}{800} \right) G(t) \quad (19)$$

donde  $T_{Amb}$  (°C) y  $T_{Cell}$  (°C) son la temperatura ambiente y la temperatura de la célula, respectivamente. La temperatura nominal de funcionamiento de la célula (NOCT) presenta valores entre 44°C y 55°C dependiendo de la ficha técnica del fabricante.

### 3.6.2.2 Modelo de eficiencia del controlador e inversor

La eficiencia del seguidor que controla el punto de máxima potencia (en inglés *Maximum Power Point Tracking-MPPT*) es obtenida por la ecuación (Apeh *et al.*, 2021):

$$\eta_{MPPT\ tracker} = \left( \frac{P_{PV,average}}{P_{PVout}} \right) \times 100\% \quad (20)$$

donde  $P_{PVout}$  (kW) es la potencia de salida y  $P_{PV,average}$  (kW) es la potencia media de salida del sistema fotovoltaico.

El rendimiento del inversor viene determinado entre la potencia de salida del inversor (Corriente Alterna-CA) y la potencia de entrada del inversor (Corriente Continua-CC). La eficiencia del inversor se expresa de la siguiente forma (Shiva y Sudhakar, 2015):

$$\eta_{Inv} = \left( \frac{P_{PV,ac}}{P_{PV,dc}} \right) \times 100\% \quad (21)$$

donde  $P_{PV,ac}$  (kW) representa la potencia de salida y  $P_{PV,dc}$  (kW) es la potencia de entrada del inversor.

### 3.6.2.3 Generador diésel

El generador diésel es un equipo que genera un alto rendimiento y constancia en la generación de energía eléctrica. Este equipo está diseñado para producir la energía eléctrica durante un período más largo de tiempo, en comparación con otro tipo de generadores. Un sistema híbrido al cual se incorpora un generador diésel puede cubrir la carga eléctrica cuando hay déficit en el sistema. La ecuación que describe el consumo de combustible del generador diésel se calcula de la forma (Olatomiwa *et al.*, 2015):

$$F_{Consumo} \left( \frac{L}{h} \right) = C_1 \left( \frac{L}{kWh} \right) P_{Diésel,salida(kW)} + C_2 \left( \frac{L}{kWh} \right) P_{rated(kW)} \quad (22)$$

donde  $F_{Consumo} \left( \frac{L}{h} \right)$  determina la cantidad de combustible en litros por hora (L/h) utilizada para una potencia de salida en operación  $P_{Diésel,salida(kW)}$  y una potencia nominal  $P_{rated(kW)}$  del generador. Donde  $C_1$  y  $C_2$  son la pendiente de la recta del combustible (L/kWh) y la intercepción de la recta de combustible (L/kWh), respectivamente.

Así mismo, la cantidad de combustible anual consumida por el generador puede determinarse mediante la siguiente expresión:

$$F^{anual}_{Consumo} \left( \frac{L}{año} \right) = C_1 \left( \frac{L}{kWh} \right) OH_{Op.h} \left( \frac{h}{año} \right) P_{Diésel,salida\ promedio(kW)} + C_2 \left( \frac{L}{kWh} \right) OH_{Op.h} \left( \frac{h}{año} \right) P_{rated(kW)} \quad (23)$$

donde  $F^{annual}_{Consumo}$  es el consumo de combustible del generador diésel durante todo el año en unidades de litros por año (L/año),  $OH_{Op.h}$  representa las horas de operación del generador por año (h/año) y  $P_{Diésel,salida\ promedio(kW)}$  es la potencia de salida promedio durante el año en (kW).

Por otro lado, la vida esperada del generador diésel es representado por:

$$R_{Vida\ esperada\ Gen.Diesel(año)} = \frac{O_{Tiempo\ de\ vida\ Gen.dies(h)}}{OH_{Op.h}(h/año)} \quad (24)$$

donde  $O_{Tiempo\ de\ vida\ Gen.diesel(h)}$  es el tiempo de vida del diésel en unidades horarias (h) y  $OH_{Op.h}$  representa las horas de operación del generador por año (h/año). Estas dos expresiones determinan la vida esperada del generador diésel  $R_{Vida\ esperada\ Gen.Diesel}$  en años.

#### 3.6.2.4 Banco de baterías

El almacenamiento del exceso de energía producido por el sistema de generación es a través del banco de baterías. Este exceso de energía es suministrado a la demanda energética cuando hay una baja producción del sistema de generación. La capacidad de energía almacenada en la batería (kWh) viene determinada por la expresión (Ramli *et al.*, 2018):

$$E_{Bat} = \frac{(E_L AD)}{\eta_{Inv}\eta_{Bat}DOD} \quad (25)$$

donde  $AD$  es el número de días de autonomía (días) y  $E_L$  (kWh/d) es el consumo medio diario de energía. La profundidad de carga de la batería en porcentaje (%) es representado por  $DOD$ .  $\eta_{Inv}$  y  $\eta_{Bat}$  son las eficiencias del inversor y de la batería en porcentaje (%).

Por otro lado, el estado de carga de la batería (en inglés *Stated of Charge-SOC*) está relacionado entre la capacidad máxima y la capacidad disponible de la batería (Sundén, 2019):

$$SOC(t) = \frac{E_{Bat,available}(t)}{E_{Bat}} \times 100\% \quad (26)$$

donde  $E_{Bat,available}$  (kWh) es la energía disponible en la batería durante un tiempo determinado y  $E_{Bat}$  (kWh) es la capacidad máxima de energía almacenada.

El banco de baterías presenta procesos de descarga y carga de energía durante el funcionamiento u operatividad del sistema. Un proceso del banco de baterías es la carga de energía y ocurre cuando la energía del sistema fotovoltaico es superior que la energía de consumo. Esta energía en exceso es direccionada y almacenada al banco de baterías. Por otro lado, la descarga de energía es el otro proceso del banco de baterías y sucede cuando la energía de consumo es superior al sistema de generación de energía, generando un déficit el cual es cubierto por la energía descargada por las baterías.

La expresión matemática (27) y (28) determinan la energía almacenada en el banco de baterías para un periodo de tiempo  $t$  después de un proceso de carga y descarga (Tazay, 2020)(Diaf *et al.*, 2007):

$$E^{Proc\_carga}_{Bat,Disp(t)} = E_{Bat,Disp(t-1)}(1 - \sigma) + \left( E_{sis(t)} - \frac{E_L(t)}{\eta_{Inv}} \right) \eta_{Bat} \quad (27)$$

$$E^{Proc\_descarga}_{Bat,Disp(t)} = E_{Bat,Disp(t-1)}(1 - \sigma) - \left( \frac{E_L(t)}{\eta_{Inv}} - E_{sis(t)} \right) \quad (28)$$

donde la energía almacenada que queda en el banco de baterías posterior a un proceso de carga y descarga es representada por  $E^{Proc\_carga}_{Bat,Disp(t)}$  y  $E^{Proc\_descarga}_{Bat,Disp(t)}$ , respectivamente.  $E_{Bat,Disp(t-1)}$  representa la energía almacenada en un tiempo anterior  $t - 1$  a los procesos mencionados y  $\sigma$  (%) es la tasa de autodescarga de las baterías. Así mismo,  $\eta_{Inv}$  y  $\eta_{Bat}$  son del inversor y la eficiencia de la batería, respectivamente. Por otro lado,  $E_L$  (kWh) es el consumo de energía y  $E_{sis}$  (kWh) es la energía del sistema generada por los módulos fotovoltaicos.

Por otro lado, un parámetro importante en las baterías o banco de baterías es su tiempo de vida esperada. Este parámetro determina el tiempo en el cual la batería tiene que ser reemplazada y depende del rendimiento energético de la batería. La siguiente

expresión matemática describe el rendimiento energético y la vida esperada de las baterías.

$$E_{Rendi\_Bat(kWh)} = E_{nom,Bat(kWh)} \eta_{roundtrip(\%)} DOD(\%) N_{cycle} \quad (29)$$

$$R_{vida\ esperada\ BB(año)} = \frac{N_{Bat} E_{Rendi\_Bat(kWh)}}{E_{Anual\_Rendi\_Bat(kWh/año)}} \quad (30)$$

donde  $R_{vida\ esperada\ BB}$  es la vida esperada del banco de batería en el tiempo,  $N_{Bat}$  es el número de baterías que comprende el banco de baterías,  $E_{Rendi\_Bat}$  y  $E_{Anual\_Rendi\_Bat}$  representan el rendimiento energético y el rendimiento energético anual de la batería.  $N_{cycle}$ ,  $\eta_{roundtrip}$ ,  $DOD$  y  $E_{nom,Bat(kWh)}$  representan el número de ciclos, la eficiencia de ida y vuelta, la profundidad de descarga y la energía nominal de la batería.

### 3.6.3 Programa de optimización para el sistema de generación híbrido: HOMER

El Optimización híbrida de múltiples recursos energéticos (en inglés *Hybrid Optimization of Multiple Energy Resources-HOMER*) permite realizar simulaciones de sistemas de generación híbrido que determinan el tamaño óptimo de sus componentes mediante un análisis de dimensionamiento y costos. Así mismo, el programa realiza un análisis tecno-económico, sensibilidad y ambiental al sistema de fotovoltaico híbrido, tanto conectados a la red o fuera de la red. HOMER es una herramienta de gran utilidad para diseñar y planear sistemas de generación híbridos.

Los inputs o datos de entrada para iniciar la simulación y el análisis en HOMER incluyen precios y tamaños de las componentes del sistema híbrido propuesto, el perfil de carga o consumo, datos meteorológicos, como radiación y temperatura ambiente, y otras características técnicas. HOMER realiza varias simulaciones hasta encontrar una configuración óptima, la cual se logra obteniendo un mínimo **Costo Presente Neto** (en inglés *Net Present Cost-NPC*) y un mínimo Nivel del Costo de la Energía (en inglés *Cost of Energy-COE*). Finalmente, los resultados obtenidos a través de la configuración óptima del sistema de energía híbrido son detallados como resultados de la optimización donde muestra combinaciones óptimas de los sistemas de generación y sus óptimas capacidades,

un análisis tecno-económico y producción de electricidad e incluso el despacho o flujo de energía y su impacto medio ambiental.

### 3.6.3.1 Criterios e indicadores de evaluación económica

#### 3.6.3.1.1 Costo Presente Neto

El Costo Presente Neto (CPN) o en inglés *Net present Cost (NPC)* representa el costo del ciclo de vida de un sistema de generación híbrido y representa los costos del sistema durante su vida útil al cual incurren, refiriéndose a la suma total de los costos de sus componentes (coste de capital inicial, costo de sustitución, costos anuales de funcionamiento y mantenimiento, costo del combustible y otros) menos los ingresos o valor residual durante la vida del proyecto. El valor del CPN se calcula mediante la siguiente ecuación (Razmjoo *et al.*, 2019):

$$CPN = \frac{C_{Total}}{CRF(i, T_p)} \quad (31)$$

$$CPN = \frac{1}{CRF(i, T_p)} \sum_{k=Comp} (C^{total\ anual}_{capital,k} + C^{total\ anual}_{O\&M,k} + C^{total\ anual}_{rem,k} + C^{total\ anual}_{fuel,k} - C_{salvage,k}) \quad (32)$$

$$\frac{1}{CRF(i, T_p)} = \frac{i(1+i)^{T_p} - 1}{i(1+i)^{T_p}} \cong \sum_{n=1}^{T_p} \frac{1}{(1+i)^n} \quad (33)$$

donde  $CRF(i, T_p)$  es el factor de recapitalización,  $T_p$  es la vida útil del proyecto en número de años y  $i$  (%) es el tipo de descuento real. El coste inicial, el coste de reposición, el coste de operación y mantenimiento (O&M), el coste de recuperación y el coste de combustible son representados por  $C_{capital}$  (\$/kWh),  $C_{rem}$  (\$/kWh),  $C_{O\&M}$  (\$/kWh) y  $C_{salvage}$  (\$/kWh) y  $C_{fuel}$  (\$/kWh), respectivamente, y el término  $k$  representa el número de componentes del sistema.

### 3.6.3.1.2 El costo nivelado de la energía

El costo medio de la energía generada por un sistema en \$/kWh es determinado por el Costo Nivelado de la Energía (en inglés *Levelized Cost Of Energy-COE*). El COE se obtiene dividiendo el costo total del sistema en el año y la generación de energía total anual. La expresión matemática que determina el COE es a través de la siguiente ecuación (Hidalgo *et al.*, 2022):

$$COE = \frac{C_{Total}}{E_{Serverd}} \quad (34)$$

donde  $E_{Serverd}$  (kWh) representa la energía que el sistema genera para cubrir el consumo energético del lugar.  $C_{Total}$  (\$) es el costo total anualizado del sistema.

### 3.6.3.1.3 Fracción de la energía renovable

La fracción renovable está expresada por la división porcentual de la energía generada por fuentes no renovables y renovables. La relación es la siguiente (Sawle *et al.*, 2018):

$$RF(\%) = \left( \frac{E_{renovable}}{E_{renovable} + E_{no-renovable}} \right) \times 100\% \quad (35)$$

donde  $E_{renovable}$  (kWh) y  $E_{no renovable}$  (kWh) son la energía renovable y no renovable generada por el sistema, respectivamente.  $RF(\%)$  es la fracción renovable del sistema en porcentaje.

### 3.6.3.2 Descripción de las etapas del proceso de simulación en el programa en HOMER

El programa HOMER permite diseñar un sistema energético híbrido conectado y aislado de la red para cubrir la demanda energética. Este programa contiene diversos procesos para realizar estudios de sistema de generación y sus análisis tecno-económicas.

La Figura 11 muestra los procedimientos realizados por el programa a través de un diagrama de flujo, donde se consideran las siguientes etapas:

- **Etapas de pre-simulación:** 1) Localización del lugar de estudio y el diseño un sistema energético híbrido. 2) Datos de entrada como datos climáticos (radiación solar, temperatura ambiente), datos de consumo energético de lugar de estudio, costos y especificaciones de los componentes del sistema. 3) Las restricciones de la optimización como la variación de las capacidades de los componentes y los parámetros económicos y técnicos del proyecto.
- **Fase de simulación y optimización:** En la simulación, el programa modela cada uno de los componentes haciendo funcionar el sistema cada hora durante un año y determina la potencia o capacidades de los componentes generadores de energía mediante los datos de entrada y las restricciones asignadas para satisfacer el consumo energético del lugar.

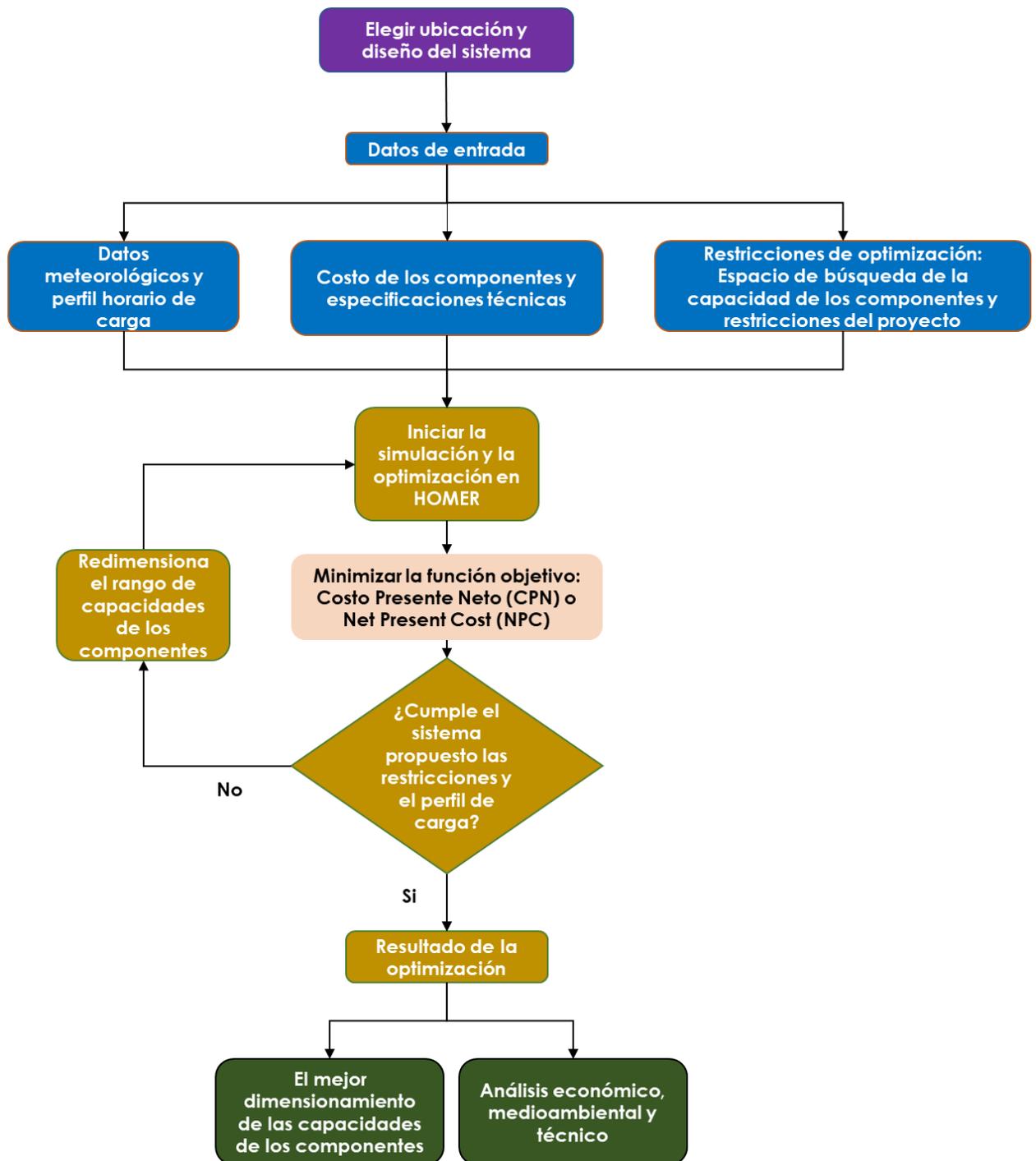
Por otro lado, en esta etapa se inicia el proceso de optimización donde el programa utiliza un algoritmo de tipo espacio de búsqueda que permite encontrar las capacidades óptimas de los componentes del sistema minimizando la función objetivo definida por el Costo Presente Neto (CPN) o *Net Present Cost (NPC)* bajo el cumplimiento de sus restricciones (balance de potencia, restricción de carga, límites de carga y descarga del banco de baterías y las capacidades de los componentes del sistema).

Si el sistema cumple el CPN mínimo satisfaciendo las restricciones y cubriendo el consumo de energía, la optimización finaliza y da como resultado el mejor dimensionamiento de los componentes del sistema y su rendimiento medioambiental y tecno-económico. En caso contrario, se redimensionan los componentes del sistema y se vuelve a simular y optimizar el sistema.

Un punto importante a destacar es que el programa no sólo simula un sistema energético híbrido propuesto, sino que también crea escenarios para diferentes sistemas como: I) Sistema fotovoltaico conectado a un sistema de baterías y generador diésel, controlador de carga e inversor. II) Sistema fotovoltaico conectado a un sistema de baterías, controlador de carga e inversor. III) Sistema fotovoltaico, controlador de carga e inversor.

**Figura 11**

Diagrama de flujo de los procesos realizados por el programa HOMER



### 3.6.4 Función objetivo de la optimización y sus restricciones

La función objetivo a minimizar en la optimización de un sistema de generación es el Costo Presente Neto (CPN) descrito en la ecuación (31). Esta expresión en más detalles es descrita a continuación:

$$CPN = \sum_{n=1}^{T_P} \frac{1}{(1+i)^n} \left( \sum_{k=Comp} C^{total\ anual}_{capital,k} + C^{total\ anual}_{O\&M,k} + C^{total\ anual}_{remp,k} + C^{total\ anual}_{fuel,k} - C_{salvage,k} \right) \quad (36)$$

donde el CPN comprende el costo total anualizado del sistema a través del costo total del capital, costo de mantenimiento y operación, costo de reemplazo, costo de salvamento o residual (en inglés *Salvage*) y el factor de recapitalización (RCF).

Asimismo, el CPN puede quedar expresado en función de sus costos acumulados a lo largo del proyecto considerando su tasa de descuento anual, como se muestra en la siguiente expresión:

$$CPN = \sum_{k=comp} (C^{acumulado\ total}_{capital,k} + C^{acumulado\ total}_{remp,k} + C^{acumulado\ total}_{O\&M,k} + C^{acumulado\ total}_{fuel,k}) - S^{acumulado\ total}_{Salvage,k} \quad (37)$$

donde cada componente k presenta su costo total acumulado  $C^{acumulado\ total}_k$  de capital, reemplazo, O&M, combustible y salvamento, como se describe en la siguiente ecuación:

$$C^{acumulado\ total}_{capital,k} = C^{total\ anual}_{capital,k} \left( \frac{\$}{año} \right) \quad (38)$$

$$C^{acumulado\ total}_{remp,k} = C^{total\ anual}_{remp,k} \left( \frac{\$}{año} \right) \sum_{n=1}^{T_P} \frac{1}{(1+i)^n} (año) \quad (39)$$

$$C^{acumulado\ total}_{O\&M,k} = C^{total\ anual}_{O\&M,k} \left( \frac{\$}{año} \right) \sum_{n=1}^{T_P} \frac{1}{(1+i)^n} (año) \quad (40)$$

$$C_{acumulado\ total}^{fuel,k} = C_{total\ anual}^{fuel,k} (\$/año) \sum_{n=1}^{T_P} \frac{1}{(1+i)^n} (año) \quad (41)$$

$$C_{acumulado\ total}^{salvage,k} = C_{salvage,k} \left( \frac{\$}{año} \right) \sum_{n=1}^{T_P} \frac{1}{(1+i)^n} (año) \quad (42)$$

En esta expresión,  $n$  es el año en el cual se reemplaza cada componente  $k$  y  $T_P$  es el año de vida del proyecto. La tasa de descuento nominal es  $i$  y  $C_{total\ anual}^k$  es su costo total anual de cada componente  $k$  para el capital, reemplazo, O&M, combustible y salvamento. Estos son expresados de la siguiente forma:

$$C_{total\ anual}^{capital,k} (\$) = P_{total,k} (kW) C_{capital\ unitario,k} \left( \frac{\$}{kW} \right) \quad (43)$$

$$C_{total\ anual}^{O\&M,k} \left( \frac{\$}{año} \right) = P_{total,k} (kW) C_{O\&M\ unitario,k} \left( \frac{\$}{kW - año} \right) \quad (44)$$

$$C_{total\ anual}^{remp,k} (\$) = P_{total,k} (kW) C_{remp\ unitario,k} \left( \frac{\$}{kW} \right) \quad (45)$$

$$C_{total\ anual}^{fuel,k} (\$/año) = HO \left( \frac{h}{año} \right) C_{fuel\ unitario,k} \left( \frac{\$}{h} \right) \quad (46)$$

$$C_{salvage,k} \left( \frac{\$}{año} \right) = C_{total\ anual}^{remp,k} \left( \frac{\$}{y} \right) \left( \frac{R_{rem,k}}{R_{com,k}} \right) \quad (47)$$

$$P_{total,k} (kW) = N_k P_{nom,k} (kW) \quad (48)$$

donde  $C_{salvage,k}$  es el costo de salvamento o valor residual de cada componente  $k$ , el cual tiene dependencia con el costo total anual de cada componente  $C_{total\ anual}^{remp,k}$ , los años de remanencia  $R_{rem,k}$  y la vida esperada de cada componente representada por  $R_{com,k}$ .  $N_k$  es el número de fotovoltaico, controlador de carga, batería, inversor y generador diésel y  $P_{nom,k}$  es la potencia nominal de cada uno de los componentes.

El año de remanencia de cada componente  $R_{rem,k}$  es determinado por las siguientes expresiones (Dalton *et al.*, 2009):

$$R_{rep,k} = R_{com,k} INT \left( \frac{R_{proyect}}{R_{com,k}} \right) \quad (49)$$

$$R_{rem,k} = R_{com,k} - (R_{proyect} - R_{rep,k}) \quad (50)$$

donde  $R_{com,k}$  es el tiempo de vida esperada de cada componente k,  $R_{proyect}$  es el año de vida del proyecto que es igual a  $T_p$ , el cual fue mencionado en las ecuaciones anteriores.  $R_{rep}$  es el año de reposición de cada componente k donde el termino *INT* considera el entero de la división.

La función objetivo está bajo ciertas restricciones para encontrar los valores mínimos, ante ello, las siguientes secciones describen las tres restricciones de la optimización.

#### 3.6.4.1 Restricción al Balance de potencia

Las restricciones de la carga y descarga del banco de baterías y el balance de potencia del sistema son descritos por la siguiente expresión (Quispe *et al.*, 2023):

$$P_{PV(t)} \pm P_{Bat(t)} + P_{Diesel(t)} = P_{load(t)} \quad (51)$$

donde  $P_{PV}$  representa la potencia de salida del sistema fotovoltaico y  $P_{Diesel(t)}$  representa la potencia de salida del generador diésel. Así mismo, la potencia del banco de baterías  $P_{Bat(t)}$ , con signos positivo y negativo, indica el modo de carga y descarga para un cierto tiempo t.  $P_{load(t)}$  es la potencia de consumo del lugar.

#### 3.6.4.2 Restricciones variables: Límites superior e inferior

Una de las restricciones para obtener el mínimo CPN y la confiabilidad del sistema de fotovoltaico híbrido propuesta son los límites de las variables definidas por cantidad

de módulos fotovoltaicos, baterías, convertidores, inversores y diésel. Estas restricciones de las variables son descritas a continuación:

$$N_{PV,min} \leq N_{PV} \leq N_{PV,max} \quad (52)$$

$$N_{Bat,min} \leq N_{Bat} \leq N_{Bat,max} \quad (53)$$

$$N_{Conv,min} \leq N_{Conv} \leq N_{Conv,max} \quad (54)$$

$$N_{Conv,min} \leq N_{Conv} \leq N_{Conv,max} \quad (55)$$

$$N_{Diesel,min} \leq N_{Diesel} \leq N_{Diesel,max} \quad (56)$$

donde  $N_{k,min}$  es el número mínimo de k componentes del sistema de generación híbrido y  $N_{k,max}$  es la cantidad máxima.  $N_k$  representa la cantidad óptima de cada componente.

### 3.6.4.3 Restricción del sistema de almacenamiento en batería

Otra restricción está basada en la energía almacenada de las baterías a través de su mínima y máxima capacidad de almacenamiento de energía (Das *et al.*, 2021):

$$E_{Bat}^{min} \leq E_{Bat}^t \leq E_{Bat}^{max} \quad (57)$$

donde la capacidad mínima y máxima del almacenamiento de energía del banco de baterías son representadas por  $E_{Bat}^{min}$  y  $E_{Bat}^{max}$ , respectivamente.

Estas capacidades son determinadas por medio de las siguientes ecuaciones (Das *et al.*, 2021):

$$E_{Bat}^{min} = N_{bat}V_{bat}C_{bat}SOC_{min} \quad (58)$$

$$E_{Bat}^{max} = N_{bat}V_{bat}C_{bat}SOC_{max} \quad (59)$$

donde  $N_{bat}$  representa el número de baterías que componen el banco de baterías.  $C_{bat}$  es la capacidad de una batería y su voltaje  $V_{bat}$ . Así mismo,  $SOC_{min}$  y  $SOC_{max}$  representan el estado de carga mínima y máxima.

### 3.6.5 Diseño y modelamiento del sistema fotovoltaico híbrido

#### 3.6.5.1 Función objetivo CPN y formulación de la optimización

##### 3.6.5.1.1 Costo de ciclo de vida

El Costo de ciclo de Vida (en inglés *Life Cycle Cost-LCC*) representa el costo que incurre el sistema de generación durante su vida útil. Este costo es representado por la siguiente expresión:

$$LCC = Capex(\$) + Opex_{vidatotal}(\$) + C_{reemplazos}^{total}(\$) \quad (60)$$

donde  $Capex$  representa el costo de inversión de cada uno de los componentes. El segundo término representa el costo de mantenimiento y operación  $Opex_{vidatotal}$  que comprende cada uno de los componentes durante la vida útil del sistema y  $C_{reemplazos}^{total}$  representa el costo de reemplazar cada uno de los componentes.

##### 3.6.5.1.2 Costo presente neto

El Costo Presente Neto (CPN), bajo las siglas en inglés NPC, es nuestra función objetivo utilizada en nuestra implementación de un optimizador inteligente para optimizar las cantidades de componentes del sistema de generación (fotovoltaica y baterías). Este CPN está conformada por el Costos de Ciclo de Vida (LCC) y los años de vida útil del sistema, siendo expresada de la siguiente forma:

$$CPN = \sum_{t=0}^{t=20} \frac{LCC}{(1+i)^t} \quad (61)$$

$$CPN = \sum_{t=0}^{t=20} \frac{Capex + Opex_{vidatotal} + C_{reemplazos}^{total} - C_{Salvamento}}{(1+i)^t} \quad (62)$$

$$CPN = Capex + \sum_{t=1}^{t=Tp} \frac{Opex_{vidatotal} + C_{reemplazos}^{total} - C_{Salvamento}}{(1+i)^t} \quad (63)$$

donde el LCC es el costo de ciclo de vida del sistema,  $Opex_{vidatotal}$ ,  $C_{reemplazos}^{total}$  y  $C_{Salvamento}$  representan el costo de las componentes, costos de reemplazo debido a fallos y el costo de ahorro, respectivamente.

El CPN de la ecuación (36) o (63) puede expandirse en su forma más generalizada:

$$\begin{aligned} CPN = P^{fv} & \left( C^{fv} \left[ \frac{\$}{kWp} \right] + \sum_{t=1}^{t=20} \frac{C_{O\&M}^{fv} \left[ \frac{\$}{kWp - \text{año}} \right]}{(1+i)^t} \right) \\ & + E^{baterias} \left( C^{bateria} \left[ \frac{\$}{kWh} \right] + \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{bateria} \left[ \frac{\$}{kWh - \text{año}} \right]}{(1+i)^t} + C_{Reemplazo}^{bateria} \left[ \frac{\$}{kWh} \right] \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{bateria} \left[ \frac{\$}{kWh - \text{año}} \right]}{(1+i)^t} - C_{Reemplazo}^{bateria} \left[ \frac{\$}{kWh - \text{año}} \right] \left( \frac{R_{Rem}^{baterias}}{R_{Com}^{baterias}} \right) \right) \\ & + P^{Inversores} \left( C^{inversor} \left[ \frac{\$}{kW} \right] + \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{inversor} \left[ \frac{\$}{kW - \text{año}} \right]}{(1+i)^t} + C_{Reemplazo}^{inversor} \left[ \frac{\$}{kW} \right] \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{inversor} \left[ \frac{\$}{kW - \text{año}} \right]}{(1+i)^t} - C_{Reemplazo}^{inversor} \left[ \frac{\$}{kW - \text{año}} \right] \left( \frac{R_{Rem}^{inversor}}{R_{Com}^{inversor}} \right) \right) \\ & + P^{Controladores} \left( C^{control} \left[ \frac{\$}{kW} \right] + \sum_{t=1}^{t=20} \frac{C_{O\&M}^{control} \left[ \frac{\$}{kW - \text{año}} \right]}{(1+i)^t} + C_{Reemplazo}^{control} \left[ \frac{\$}{kW} \right] \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{control} \left[ \frac{\$}{kW - \text{año}} \right]}{(1+i)^t} - C_{Reemplazo}^{control} \left[ \frac{\$}{kW - \text{año}} \right] \left( \frac{R_{Rem}^{control}}{R_{Com}^{control}} \right) \right) \end{aligned} \quad (64)$$

donde  $C^k$  es el costo capital por unidad de kW de cada componente k del sistema.  $C_{O\&M}^k$  es el costo de mantenimiento y operación,  $C_{Reemplazo}^k$  es el costo de reemplazo y  $P^k$  es la potencia total de cada uno de los componentes k. La energía del banco de baterías en kWh es representada por  $E^{Baterias}$ .

La potencia total de cada uno de los componentes (fotovoltaico, batería, inversor controlador) del sistema de generación híbrido es dado por las ecuaciones siguientes:

$$P^{fv}[kW] = N_{fv} \frac{P_{unidad}^{fv}}{1000} \left[ \frac{kW}{unidad} \right] \quad (65)$$

$$E^{Baterias}[kWh] = N_{bateria} \frac{E_{unidad}^{Bateria}}{1000} \left[ \frac{kWh}{unidad} \right] \quad (66)$$

$$P^{inversor} [kW] = N_{inversor} \frac{p_{unidad}^{inversor}}{1000} \left[ \frac{kW}{unidad} \right] \quad (67)$$

$$P^{controlador} [kW] = N_{controlador} \frac{p_{unidad}^{controlador}}{1000} \left[ \frac{kW}{unidad} \right] \quad (68)$$

donde  $p_{unidad}^k$  es la potencia unitaria de cada uno de los componentes y  $N_k$  es el número unitario de sus componentes.

### 3.6.5.1.2 Formulación de la optimización

El CPN como función objetivo tiene variables de la optimización conformado por números unitarios de cada componente (fotovoltaico, inversor, batería e inversor) y variables constantes de costos, entre ellos los costos capitales, costos O&M y entre otros.

La función objetivo con estas variables es mostrada en la siguiente ecuación:

$$\begin{aligned} CPN = & N_{fv} p_{unidad}^{fv} \left( C^{fv} \left[ \frac{\$}{kWp} \right] + \sum_{t=1}^{t=20} \frac{C_{O\&M}^{fv} \left[ \frac{\$}{kWp - año} \right]}{(1+i)^t} \right) \\ & + N_{bateria} E_{unidad}^{bateria} \left( C^{bateria} \left[ \frac{\$}{kWh} \right] + \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{bateria} \left[ \frac{\$}{kWh - año} \right]}{(1+i)^t} + C_{Reemplazo}^{bateria} \left[ \frac{\$}{kWh} \right] \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{bateria} \left[ \frac{\$}{kWh - año} \right]}{(1+i)^t} - C_{Reemplazo}^{bateria} \left[ \frac{\$}{kWh - año} \right] \left( \frac{R_{Rem}^{Baterias}}{R_{Com}^{Baterias}} \right) \right) \\ & + N_{inversor} D_{unidad}^{inversor} \left( C^{inversor} \left[ \frac{\$}{kW} \right] + \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{inversor} \left[ \frac{\$}{kW - año} \right]}{(1+i)^t} + C_{Reemplazo}^{inversor} \left[ \frac{\$}{kW} \right] \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{inversor} \left[ \frac{\$}{kW - año} \right]}{(1+i)^t} - C_{Reemplazo}^{inversor} \left[ \frac{\$}{kW - año} \right] \left( \frac{R_{Rem}^{Inversor}}{R_{Com}^{Inversor}} \right) \right) \\ & + N_{control} P_{unidad}^{control} \left( C^{control} \left[ \frac{\$}{kW} \right] + \sum_{t=1}^{t=20} \frac{C_{O\&M}^{control} \left[ \frac{\$}{kW - año} \right]}{(1+i)^t} + C_{Reemplazo}^{control} \left[ \frac{\$}{kW} \right] \sum_{t=1}^{t=TP} \frac{C_{O\&M}^{control} \left[ \frac{\$}{kW - año} \right]}{(1+i)^t} - C_{Reemplazo}^{control} \left[ \frac{\$}{kW - año} \right] \left( \frac{R_{Rem}^{Control}}{R_{Com}^{Control}} \right) \right) \end{aligned} \quad (69)$$

### 3.6.5.1.3 Restricciones que sujetan la función objetivo CPN

La función objetivo presenta restricciones para encontrar valores óptimos de las variables. Estas restricciones están clasificadas en tres tipos, entre ellas son las restricciones de variables límites, las restricciones de igualdad y desigualdad:

- 1) Las restricciones de variables límites se basan en limitar la cantidad de número mínimo y máximo de cada uno de los componentes hasta encontrar la cantidad óptima. Estas restricciones son mostradas a continuación:

$$N_{PVmin} \leq N_{PV} \leq N_{PVmax} \quad (70)$$

$$N_{Batmin} \leq N_{Bat} \leq N_{Batmax} \quad (71)$$

$$N_{Contrmin} \leq N_{Contr} \leq N_{Contrmax} \quad (72)$$

$$N_{Invmin} \leq N_{Inv} \leq N_{Invmax} \quad (73)$$

- 2) Así mismo, las restricciones de igualdad se basan en las relaciones que tiene cada una de las variables en estudio. Entre ellas tenemos que la cantidad de controladores es igual a la cantidad de paneles fotovoltaicos y el número de inversores es igual a un factor de corrección de 1.1 de la potencia del fotovoltaico dividida por la potencia del inversor:

$$N_{Contr} = N_{PV} P_{PV} \quad (74)$$

$$N_{Inv} = N_{PV} \frac{1.1 P_{PV}}{P_{Inv}} \quad (75)$$

- 3) Por otro lado, las restricciones de desigualdad se componen por la desigualdad de las potencias, el límite del estado de carga o SOC y el proceso de carga y descarga:

- La restricción de la desigualdad de la potencia sucede cuando la potencia total del sistema fotovoltaico es mayor que la potencia de consumo del lugar, o viceversa, cuando la potencia de consumo es mayor que la potencia fotovoltaica. Esto es expresado de la siguiente forma:

$$P_{P(t)} u_{inv} \geq P_{l(t)} \quad (76)$$

$$N_{sol} P_{sol(t)} u_{inv} \geq P_{l(t)} \quad (77)$$

$$P_{P(t)}u_{inv} < P_{l(t)} \quad (78)$$

$$N_{sol}P_{sol(t)}u_{inv} < P_{l(t)} \quad (79)$$

- El límite del estado de carga o SOC se basa en limitar la capacidad de energía almacenada en el banco de baterías en un mínimo y máximo. Esta restricción de desigualdad es mostrada en la siguiente ecuación:

$$E_{batmin} \leq SOC \leq E_{batmax} \quad (80)$$

$$N_{bat}V_{bat}C_{bat_{nom}}(1 - DOD) \leq \frac{E_{batmin}(100\%)}{E_{batmax}} \leq N_{bat}V_{bat}C_{bat_{nom}} \quad (81)$$

- La restricción del proceso de carga se basa en considerar la energía mínima para almacenar en el banco de batería:

$$E_{chargebat(t)} \leq E_{batmax} - E_{bat(t-1)} \quad (82)$$

$$\left(N_{sol}P_{sol} - \frac{P_{l(t)}}{u_{inv}}\right)n_{bat} \leq N_{bat}V_{bat}C_{bat_{nom}} - E_{bat(t-1)} \quad (83)$$

- Así mismo, la restricción del proceso de descarga considera la energía mínima para descargar del banco de batería:

$$E_{bat(t-1)} - E_{batmin} \geq E_{DischargeBat(t)} \quad (84)$$

$$E_{bat(t-1)} - N_{bat}V_{bat}C_{bat_{nom}}(1 - DOD) \geq \frac{P_{l(t)} - N_{sol}P_{sol}u_{inv}}{u_{conv}} \quad (85)$$

### 3.6.5.1.2 Modelamiento de las componentes para el diseño

El modelamiento de cada componente (fotovoltaico, batería, controlador e inversor) del sistema de generación híbrido es descritas a continuación.

### 1) Modelamiento del sistema fotovoltaico.

El sistema fotovoltaico tiene una potencia de salida  $P_{PV(t)}$  para un determinado tiempo  $t$ . Este es determinado por la siguiente ecuación matemática:

$$P_{PV(t)}(kW) = N_{PV} V_{OC(t)} I_{SC(t)} \eta_{PV} \eta_{inv} FF \quad (86)$$

$$P_{PV(t)}^{modulo} = V_{OC(t)} I_{SC(t)} \quad (87)$$

donde  $N_{PV}$  es el número de módulos fotovoltaico en un arreglo.  $V_{OC(t)}$  y  $I_{SC(t)}$  son el voltaje de circuito abierto y la corriente de corto circuito de un módulo fotovoltaico en un tiempo  $t$ , que a su vez su multiplicación representa la potencia de salida de un módulo fotovoltaico  $P_{PV(t)}^{modulo}$  para un valor de tiempo  $t$ .  $\eta_{PV}$ ,  $\eta_{inv}$  y  $FF$  son la eficiencia del fotovoltaico e inversor y el factor de llenado en %, respectivamente.

En un módulo fotovoltaico, la corriente de corto circuito  $I_{SC(t)}$  y el voltaje de circuito abierto  $V_{OC(t)}$  tienen dependencia de la temperatura de la celda y la radiación solar. Estas dependencias son expresadas en la siguiente expresión:

$$V_{OC(t)} = V_{OC(t),stc} + K_v (T_{C(t)} - T_{ref(t)}) \quad (88)$$

$$I_{SC(t)} = \left( I_{SC(t),stc} + K_I (T_{C(t)} - T_{ref(t)}) \right) \frac{G(t)}{1000} \quad (89)$$

Donde  $V_{OC(t),stc}$ ,  $I_{SC(t),stc}$ , y  $T_{ref(t)}$  son el voltaje de circuito abierto, la corriente corto circuito y la temperatura ambiente a condiciones estándar, respectivamente.  $K_v$  y  $K_I$  son los coeficientes de temperatura en unidades  $\%/^{\circ}C$ . Estos parámetros mencionados son obtenidos de las especificaciones técnicas del fabricante del módulo.  $G(t)$  es la radiación solar en el lugar de estudio, los cuales son obtenidos a través de medidas tomadas en unidades temporales, y  $T_{C(t)}$  es la temperatura de la celda fotovoltaica.

Por otro lado, la temperatura de la celda fotovoltaica  $T_{C(t)}$  es determinada por la ecuación:

$$T_{c(t)} = T_{a(t)} + \left( \frac{NCOT - 20^{\circ}}{800} \right) G_{(t)} \quad (90)$$

donde  $G_{(t)}$  es la radiación solar,  $T_{a(t)}$  es la temperatura ambiente y TONC es la Temperatura de Operación Nominal de la Célula fotovoltaica.

La tasa de potencia del arreglo fotovoltaico, su energía de producción diaria y anual son descrito por las siguientes expresiones matemáticas:

$$P_{PVrate}(kW) = N_{PV\acute{o}ptimo} P^{nom}_{PV}(kW) \quad (91)$$

$$E_{Produccion\ Total} \left( \frac{kWh}{d} \right) = N_{PV\acute{o}ptimo} P^{nom}_{PV}(kW) HSP \left( \frac{h}{d} \right) \quad (92)$$

$$E_{Produccion\ Total} \left( \frac{kWh}{a\tilde{n}o} \right) = P^{nom}_{PV}(kW) N_{PV} HSP \left( \frac{h}{d} \right) \left( \frac{24h}{d} \right) \left( \frac{365d}{a\tilde{n}o} \right) \quad (93)$$

$$HSP \left( \frac{h}{d} \right) = \frac{\bar{G} \left( \frac{W}{m^2} \right)}{G_{ref} \left( \frac{W}{m^2} \right)} \left( \frac{24h}{d} \right) \quad (94)$$

donde  $N_{PV\acute{o}ptimo}$  es el número de paneles fotovoltaicos que comprende el arreglo fotovoltaico, el cual es obtenido de la optimización del sistema.  $HSP$  es la hora solar pico y  $P^{nom}_{PV}$  es la potencia nominal de un módulo fotovoltaico.  $\bar{G}$  es el promedio de la radiación solar a lo largo del año y  $G_{ref}$  es la radiación solar constante referencial con un valor de  $1000\ W/m^2$ .

## 2) Modelamiento del banco de baterías

La energía total almacenada y la energía usable o de descarga en el banco de baterías son determinadas por las ecuaciones:

$$E_{Banco\ Baterias}(kWh) = \left( \frac{N_{Bat\acute{o}ptimo}}{N_{Bat\acute{serie}}} \right) C_{batNom}(Ah) V_{sist}(V) \quad (95)$$

$$E_{Usable\ Baterias}(kWh) = \left( \frac{N_{Bat\ \acute{o}ptimo}}{N_{Bat\ serie}} \right) C_{bat\ Nom}(Ah) V_{sist}(V) DOD \quad (96)$$

donde  $N_{Bat\ \acute{o}ptimo}$  es el nmero optimo de bateras que conforman el banco de bateras, el cual es dado por la optimizacin.  $N_{Bat\ serie}$ ,  $C_{bat\ Nom}$ ,  $DOD$  y  $V_{sist}$  son el nmero de bateras en serie, la capacidad nominal de una batera, la profundidad de descarga y el voltaje del sistema de generacin hbrido, respectivamente.

Por otro lado, el nmero de bateras en serie  $N_{Bat\ serie}$  y el nmero de horas de autonoma  $HA$  del banco de bateras son determinados por las siguientes ecuaciones:

$$HA(h) = \frac{N_{Bat\ \acute{o}ptimo} C_{bat\ Nom}(kWh) DOD}{E_{Consumo} \left( \frac{kWh}{d} \right)} \left( \frac{24h}{d} \right) \quad (97)$$

$$E_{Consumo} \left( \frac{kWh}{d} \right) = \bar{P}_{Consumo}(kW) \left( \frac{24h}{d} \right) \quad (98)$$

$$N_{Bat\ serie} = \frac{V_{sist}}{V_{bat}} \quad (99)$$

donde  $E_{Consumo}$  es el consumo de energa del lugar en estudio.  $\bar{P}_{Consumo}$  es la potencia de consumo promedio durante el ao,  $DOD$  es la profundidad de descarga de la batera y  $V_{bat}$  es el voltaje de la batera.

As mismo, el tiempo de vida de duracin de las bateras es expresada por la siguiente expresin matemtica:

$$Vida_{Baterias} = \frac{E_{baterias\ Rendimiento}}{E_{Descarga\ Anual}} \quad (100)$$

donde  $E_{baterias\ Rendimiento}$  es el rendimiento energtico, el cual determina la cantidad total de energa que se espera que las bateras almacenen y suministren durante toda su vida til.  $E_{Descarga\ Anual}$  es la energa de descarga de las bateras durante un ao.

El rendimiento energético del banco de baterías  $E_{bateríasRendimiento}$  y el rendimiento de la batería  $E_{batRendimiento}$  son determinados por:

$$E_{bateríasRendimiento} = N_{Baterías\acute{o}ptimo} \left( C_{batNom} \eta^{eficiencia} \left( \frac{carga}{descarga} \right) DOD N_{Ciclos} \right) \quad (101)$$

$$E_{batRendimiento} = C_{batNom} \eta^{eficiencia} \left( \frac{carga}{descarga} \right) DOD N_{Ciclos} \quad (102)$$

donde  $N_{Baterías\acute{o}ptimo}$  es el número de baterías del sistema,  $C_{batNom}$  es la capacidad de una batería,  $\eta^{eficiencia}$  es la eficiencia de descarga y carga de la batería,  $N_{Ciclos}$  es el número de ciclos de descarga y carga de la batería y  $DOD$  es la profundidad de descarga. Estos 4 parámetros son obtenidos de las fichas técnicas del fabricante de baterías.

### 3) Indicadores de confiabilidad del sistema

La energía déficit determina la energía que falta para cubrir la potencia de consumo de energía. Esta energía está determinada por la siguiente expresión matemática:

$$E_{Deficit\_sistema(t)} = \frac{P_{L(t)} - (Pp(t) + E_{dch(t)})\eta_{inv}}{\eta_{conv}} \quad (103)$$

Por otro lado, la probabilidad de pérdida de energía (en inglés *Loss of Power probability-LPP*) determina la confiabilidad del sistema y es determinado por la siguiente expresión:

$$LPP = \left( \frac{\sum_1^{8760} E_{Deficit} (kWh)}{\sum_1^{8760} E_{Consumo} (kWh)} \right) 100\% \quad (104)$$

Un valor LPP nos indica menor confiabilidad del sistema, por lo que este es un indicador que nos permite determinar la capacidad del sistema de cubrir el consumo de energía, siendo esto debido a muchos factores tanto técnico, climáticos y costos del sistema.

#### 4) Parámetros tecno-económicos

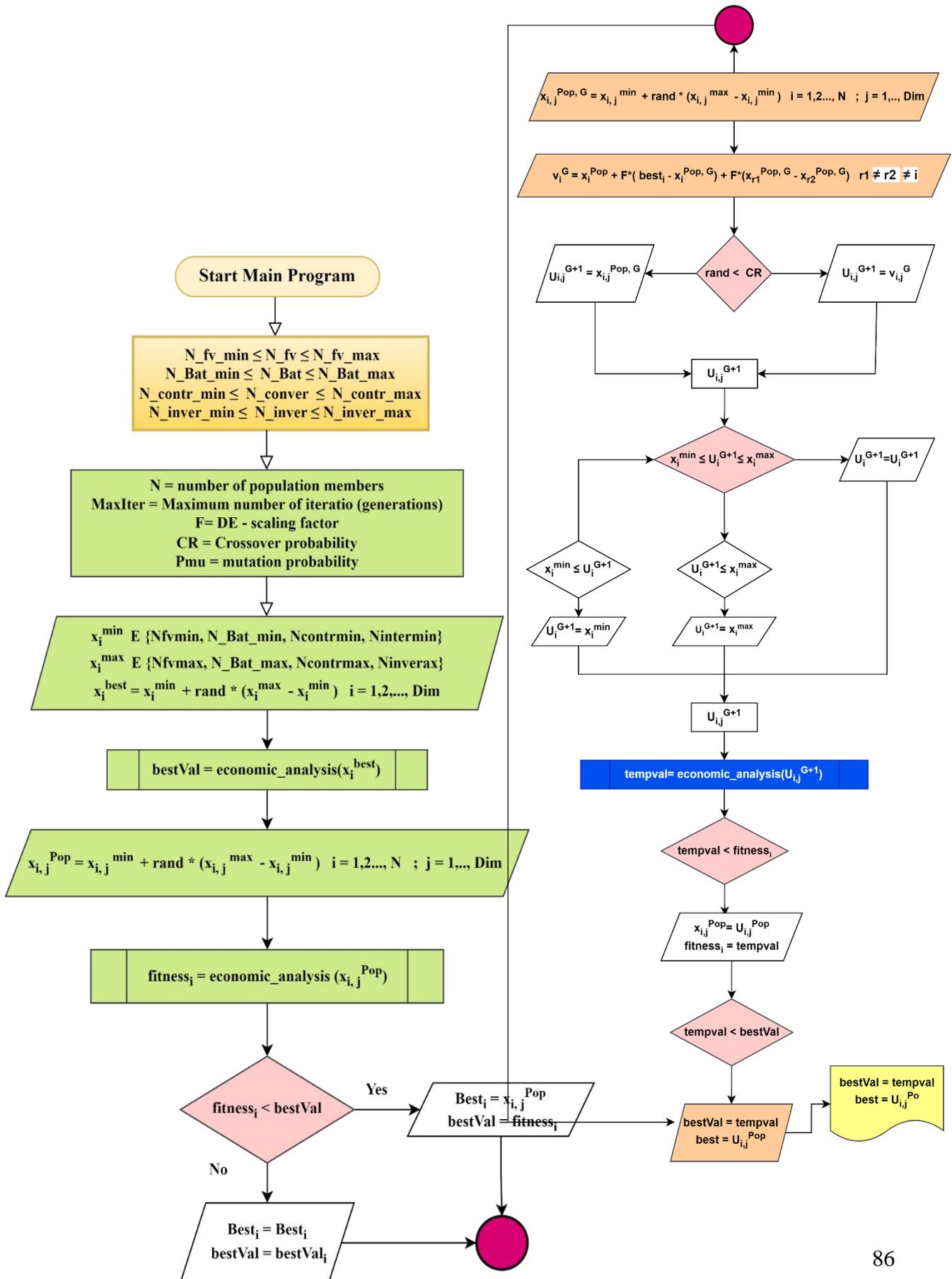
De los parámetros económicos mencionados anteriormente, como el LCC y CPN, el nivel del Costo de Energía (COE) es un parámetro importante a considerar en un proyecto energético. Este parámetro permite identificar el costo por kWh del sistema, haciéndolo comparable contra otros tipos de sistemas. El COE indica que tan económico es su producción de energía y es expresado por la siguiente expresión:

$$COE = \frac{LCC \left( \frac{\$}{\text{año}} \right)}{E_{\text{consumo}} \left( \frac{kWh}{d} \right) \left( \frac{d}{365 \text{ año}} \right)} \quad (105)$$

##### 3.6.5.1 Diagrama de flujo del diseño de generación fotovoltaico híbrido

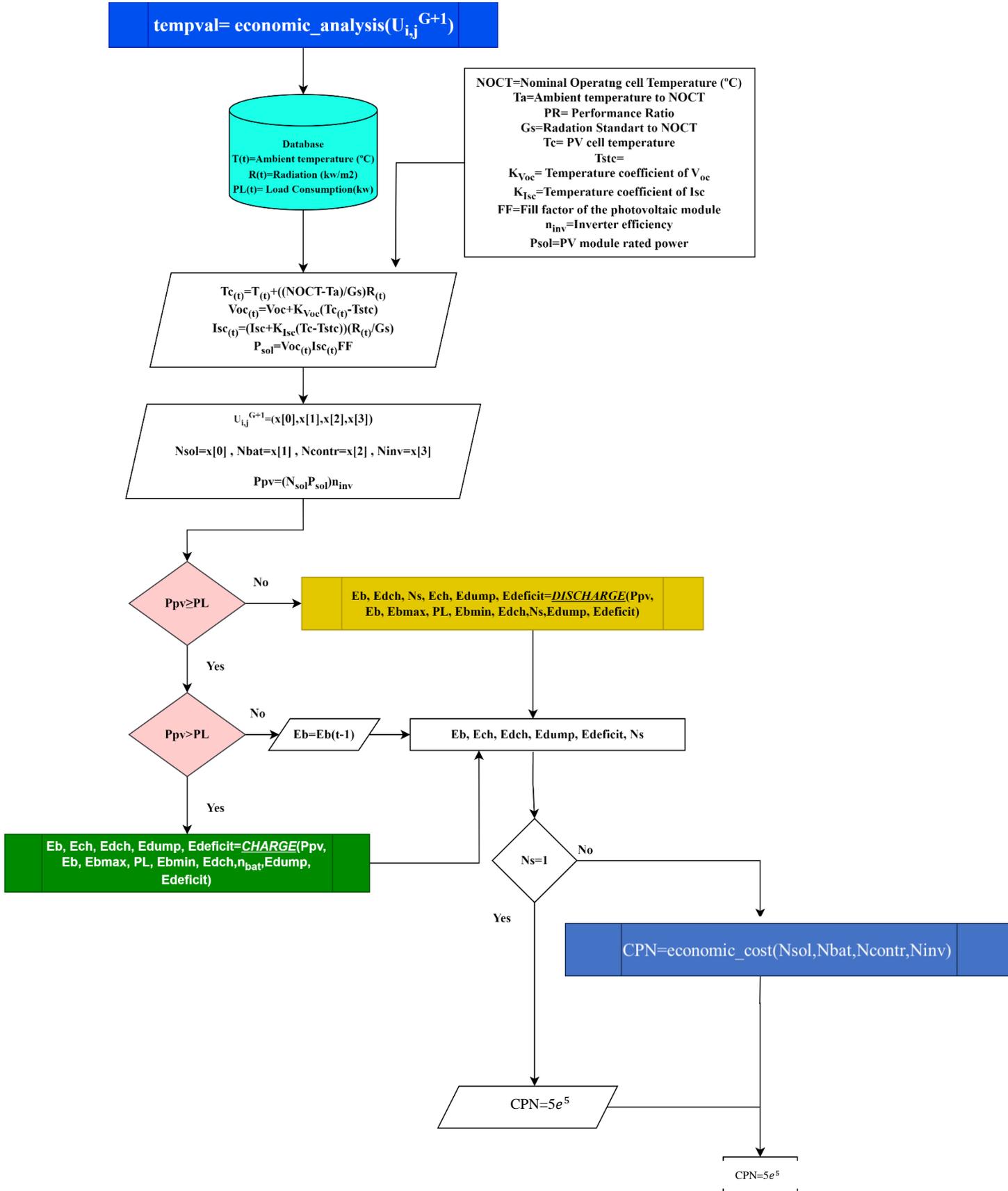
El diagrama de flujo del algoritmo de optimización del tipo Evolución Diferencial, con sus estructuras de inicialización, mutación, cruzamiento y selección, que se ha propuesto en esta investigación es mostrada en la Figura 12. Así mismo, en la Figura 13 se muestra el algoritmo de la operatividad del sistema Fotovoltaico y batería, a través de la lectura de la información del consumo eléctrico, la temperatura ambiente y radiación solar. Finalmente, en la Figura 14 se muestra el algoritmo de la operatividad de la batería en sus procesos de carga y descarga. Los algoritmos desarrollados para cada flujo son mostrado en el Anexo 6.

**Figura 12**  
 Algoritmo de optimización Evolución Diferencial



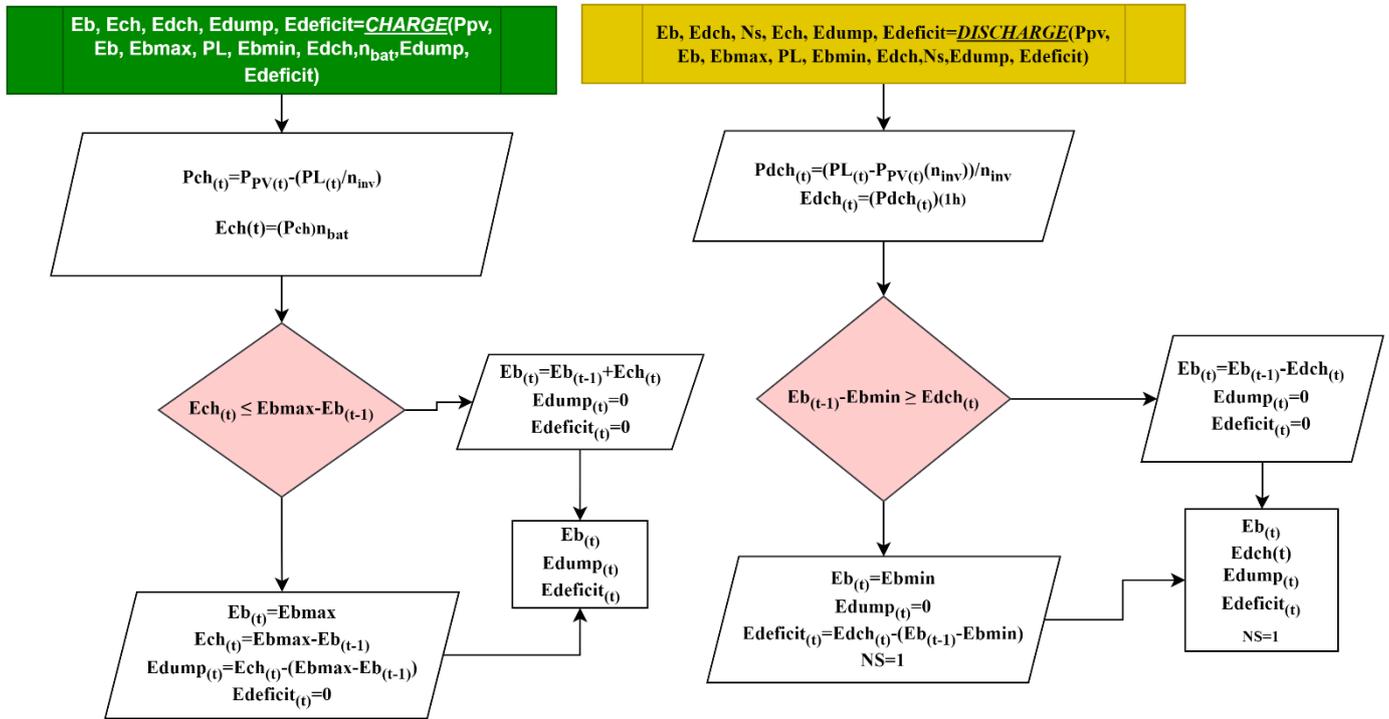
**Figura 13**

Diagrama de flujo del ingreso de la base de datos y la operatividad del sistema considerando el costo presente neto como función objetivo



**Figura 14**

Diagrama de flujo de los programas que gestionan la descarga y carga del banco de baterías



### 3.6.5.3 Modelo de inteligencia artificial para la optimización del sistema

#### 3.1.5.5.1 Algoritmo evolución diferencial

El modelo de optimización implementada para la operatividad de nuestro sistema propuesto (fotovoltaico, batería, controlador e inversor) se basa en un modelo de inteligencia artificial del tipo Evolución Diferencial (en inglés *Differential Evolution-DE*). El DE es un algoritmo propuesto por (Storn y Price, 1995) y tiene la capacidad de resolver problemas de optimización muy complejos y tiene una similitud a los algoritmos evolutivos inteligentes (Zhong y Cheng, 2020).

Este algoritmo de optimización tiene una estructura y es descrito por el mismo autor y otros autores que lo utilizan para diversas aplicaciones (Chakraborty *et al.*, 2023):

- 1) La formulación general de la función objetivo dependiente de sus variables, restringidas entre sus límites, es descrito en la siguiente forma (Zhong y Cheng, 2020):

$$\min f = f(x_1, x_1, \dots, x_D) \quad (106)$$

$$\text{s. t: } x_j^L \leq x_j \leq x_j^U, j = 1, 2, \dots, D. \quad (107)$$

donde  $f$  denota la función a optimizar.  $D$  representa la dimensión del problema y  $x_j^L$  y  $x_j^U$  son las restricciones de la variable  $x_j$ , los cuales son valores mínimo y máximo, respectivamente.

2) El algoritmo DE tiene una estructura de procesamiento basado en las siguientes características:

- **Inicialización.** Se genera un vector de dimensión  $D$  que comprende a los individuos, los cuales son calculados por los límites de las variables  $x_j^L$  y  $x_j^U$  (Bilal *et al.*, 2020):

$$x_{i,j}^G = x_{i,j}^{L,G} + \text{rand}(0,1) * (x_{i,j}^{U,G} - x_{i,j}^{L,G}) \quad (108)$$

$$X_j^G = (x_{i,j}^G, v, x_{i,j}^G, \dots, x_{i,j}^G), i = 1, 2, \dots, NP \quad (109)$$

donde el parámetro  $\text{rand}$  establece valores entre 0 y 1. Los límites superiores e inferiores de  $x$  son representados por  $L$  y  $U$ .  $i$ -ésimo representa las componentes de la dimensión, tomando valores de  $i=1, 2, \dots, D$ , y  $j$ -ésimo individuos tiene valores de  $j=1, 2, \dots, NP$  que representa el número de población.  $S^G$  es la población para cada generación  $G$ .

- **Mutación.** Un vector mutación  $V_i^G$  es generado a partir de dos individuos de la población y su forma base, denominada “DE/rand/1”, es expresada de la forma (Ahmad *et al.*, 2022):

$$v_{i,j}^G = x_{r1,j}^G + F(x_{r2,j}^G - x_{r3,j}^G) \quad (110)$$

$$V_j^G = \{v_{1,j}^G, v_{2,j}^G, \dots, x_{D,j}^G\} \quad (111)$$

donde  $x_{rn}^G = x_{i,j}^G$  es el individuo  $n$  de la población y los individuos considerados deben ser distintos,  $r1 \neq r2 \neq r3$ .  $F$  es un factor de escala que toma valores de 1 a 0, que controla la ampliación del vector.

- **Crossover.** El vector de prueba  $U_i^G$  es generado a través del cruce del entre el vector mutado  $V_i^G$  y el vector objetivo o individuo  $x_{i,j}^G$ , el cual es representado por la siguiente expresión (Chakraborty *et al.*, 2023):

$$u_{i,j}^G = \begin{cases} v_{i,j}^G, & \text{rand} < CR \text{ or } j = j_{rand} \\ x_{i,j}^G, & \text{otherwise} \end{cases} \quad (112)$$

$$U_j^G = \{u_{1,j}^G, u_{2,j}^G, \dots, u_{D,j}^G\} \quad (113)$$

donde  $Cr$  es la probabilidad de cruce que tiene valores entre 0 y 1,  $\text{rand}$  (0,1) tiene un valor aleatorio y  $j_{rand} (\in [1, D])$  es un número entero aleatorio uniformemente distribuido. El vector mutado depende de la desigualdad entre los parámetros mencionados.

- **Selección.** La comparación entre el vector de prueba  $U_{i,j}^G$  y el vector objetivo  $X_j^G$  es a través de su evaluación en la función objetivo  $f$ . El individuo que tenga mejor valor de aptitud se seleccionará como el individuo descendiente  $X_{i,j}^{G+1}$  (Bilal *et al.*, 2020):

$$X_{i,j}^{G+1} = \begin{cases} U_j^G, & \text{si } f(U_{i,j}^{G+1}) < f(X_{i,j}^G) \\ X_j^G, & \text{otherwise} \end{cases} \quad (114)$$

donde  $X_{i,j}^{G+1}$  es la nueva población de generación  $G+1$  que está conformado de individuos que tengan mayor valor de aptitud al ser evaluado en la función objetivo  $f$ .

El algoritmo DE tiene diferentes variaciones de la creación de su vector mutado. En la literatura científica se muestran las siguientes variaciones (Qin *et al.*, 2009):

- 1) “DE/rand/1”:

$$v_{i,j}^G = x_{r_1^i,j}^G + F(x_{r_1^i,j}^G - x_{r_3^i,j}^G) \quad (115)$$

2) “DE/best/1”:

$$v_{i,G}^G = x_{best,G}^G + F(x_{r_1^i,j}^G - x_{r_3^i,j}^G) \quad (116)$$

3) “DE/rand-to-best/1”:

$$v_{i,j}^G = x_{i,j}^G + F(x_{best,j}^G - x_{i,j}^G) + F(x_{r_1^i,j}^G - X_{r_2^i,j}^G) \quad (117)$$

4) “DE/best/2”:

$$v_{i,j}^G = x_{best,j}^G + F(x_{r_1^i,j}^G - X_{r_2^i,j}^G) + F(X_{r_3^i,j}^G - X_{r_4^i,j}^G) \quad (118)$$

5) “DE/rand/2”:

$$v_{i,j}^G = x_{r_1^i,j}^G + F(x_{r_2^i,j}^G - x_{r_3^i,j}^G) + F(x_{r_4^i,j}^G - X_{r_5^i,j}^G) \quad (119)$$

Los índices representan los individuos ubicados en diferentes posiciones ( $r_1^i, r_2^i, r_3^i, r_4^i, r_5^i$ ) y son distintos cuando se quiere calcular el vector mutado.  $F$  es el factor de escala, el cual controla la diferencia y amplitud del vector. Además, el individuo  $X_{best,j}^G$  representa el mejor vector individual con el mejor valor de aptitud en la población y en la generación  $G$ .

## CAPÍTULO IV - RESULTADOS

### 4.1 Análisis de la base de datos

La potencia de consumo y la radiación solar son estudiados para un centro poblado que se ubica en la parte central del distrito de San Juan de Tarucani, provincia de Arequipa. Este se ubica al sur del Perú a 4 210 metros de altura sobre el nivel del mar con coordenadas de latitud  $16^{\circ} . 11 'S$  y longitud  $71^{\circ} 3.6' E$ .

El distrito de San Juan de Tarucani es una comunidad con una población total de 2.179 habitantes y 945 viviendas, según el último censo del año 2017 y el reporte técnico del Instituto Geológico, Minero y Metalúrgico (INGEMMET, 2019). La información de radiación solar es obtenida de la predicción de recursos energéticos globales de la NASA (NASA, 2020), utilizando la latitud y longitud de la ubicación del estudio. La Figura 15 (a) muestra la radiación solar durante el año. La energía de la radiación solar promedio mensual está entre  $5.45 \text{ kWh/m}^2/\text{día}$  y  $7.894 \text{ kWh/m}^2/\text{día}$  y un promedio anual de  $6.71 \text{ kWh/m}^2/\text{día}$ .

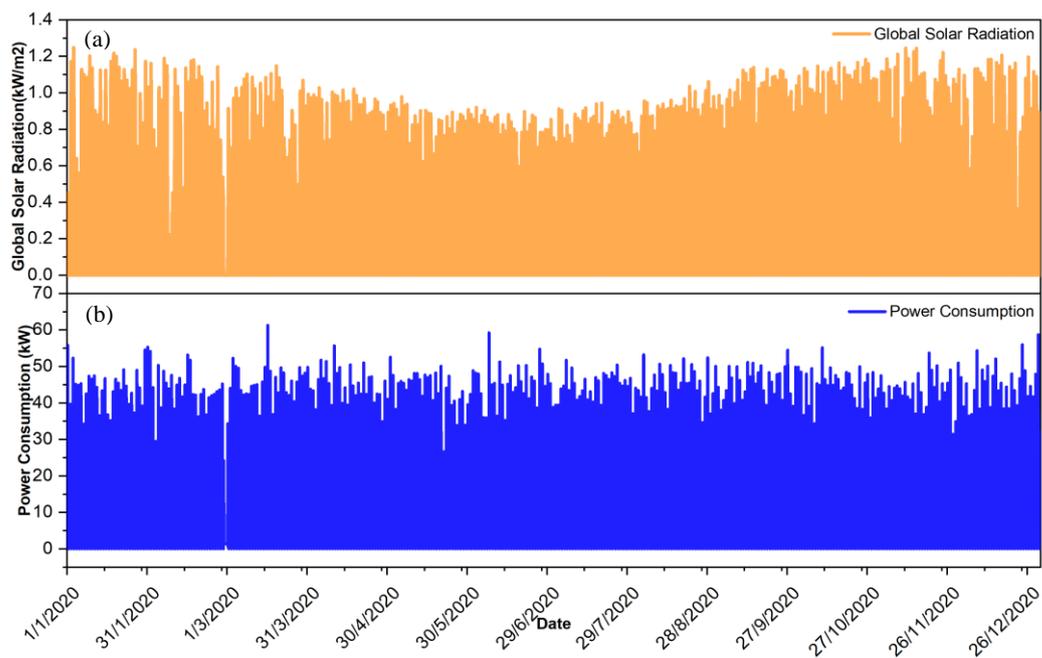
En este estudio se ha considerado solo 200 viviendas, cercanas entre ellas, una escuela, una comisaría, un puesto de salud, una iglesia y las iluminaciones de la calle. El perfil de consumo eléctrico del lugar de estudio se estimó a lo largo del día para las viviendas e instalaciones, en el cual se considera el uso de radio, televisión, focos, carga de teléfono móvil y otros electrodomésticos. El consumo eléctrico de las otras instalaciones es mostrado en la Tabla 1, así como en el Anexo 1.

La información del perfil de carga o consumo eléctrico diaria del lugar se ingresa al programa HOMER para generar el perfil de consumo eléctrico en forma de series de tiempo para todo el año. En la generación se consideró una variabilidad aleatoria de día a día y pasos de tiempo de 10% y 5%, respectivamente, para hacerla más realista. En la Figura 15 (b) se observa la serie de tiempo del consumo del lugar generada a lo largo de del año.

**Tabla 1**  
*Detalles del lugar de estudio y su consumo de energía*

Lugar de estudio	Detalles	Lugares de consumo de energía	Consumo de energía diario (Wh/d)
Comunidad Rural	Tarucani	200 viviendas	137 000
Distrito	San Juan de Tarucani	Escuela	4 490
Región y provincia	Arequipa	Estación Policial	4 088
País	Perú	Posta de salud	1 980
Longitud y Latitud	16°11' S, 71° 3.6' E	Iglesia	240
Energía de consumo diaria en el lugar (kWh/d)	165.798	Alumbrado público	18 000

**Figura 15**  
*Serie de tiempo de la radiación solar global y la potencia de consumo del lugar de estudio desconectada de la red eléctrica*



## 4.2 Modelos predictivos de las series de tiempo

### 4.2.1 Modelo predictivo SARIMA

#### 4.2.1.1 Pre-Procesamiento de la data, configuración e identificación de parámetros

Las series de tiempo de la potencia de consumo y radiación solar son datos de un año con registros horarios de 8760. Estos datos fueron divididos en data de entrenamiento y prueba con registros horarios de 7884 y 876, respectivamente, siendo la data de prueba el 10% de toda la data.

El modelo SARIMA requiere que la serie de tiempo sea estacionaria por lo que se realizó la prueba de Dickey-Fuller (DF):

- Los resultados de la prueba de DF para la radiación solar se obtiene un test de -11.8 y valores críticos de 1%, 5% y 10% con valores de -3.43, -2.86 y -2.57, respectivamente. Similarmente, para la potencia de consumo se encontró un test estadístico de -17 con valores de -3.43, -2.86 y -2,57, respecto a sus valores críticos.
- Para ambas series de tiempo se tienen valores de test estadísticos altos con respecto a sus valores críticos. Esto quiere decir que las series de tiempos no presentan raíz unitaria, por lo tanto, estas series son estacionarias y no requieren diferenciación. Esta información nos permite identificar el orden “d” del término no estacionario del modelo SARIMA. Por lo tanto, las series de tiempo son estacionarias ( $d=0$ ,  $D=0$ ).

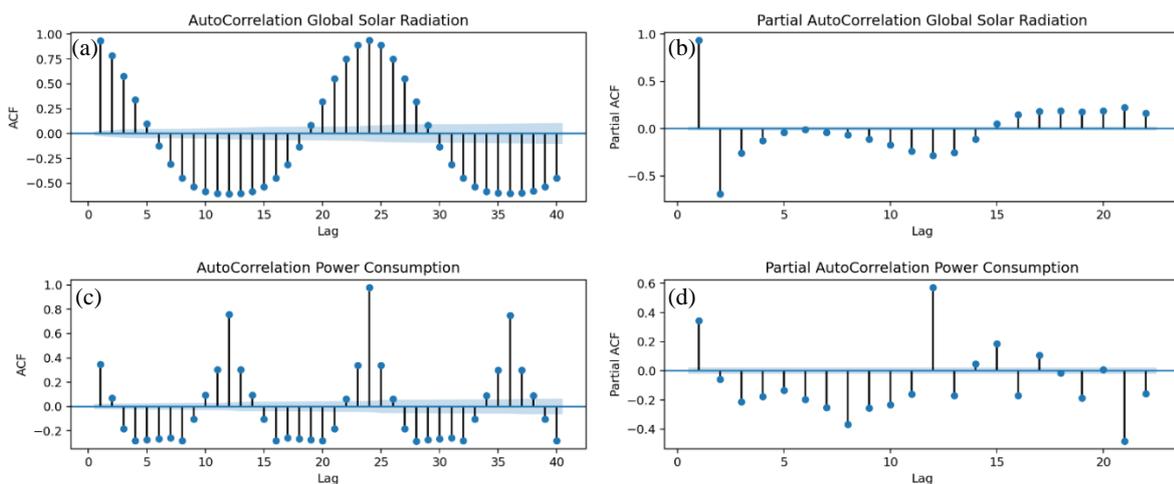
Para determinar las otras órdenes no estacionales y estacionales se requiere realizar las pruebas de Auto correlación y Parcial Auto correlación. En la Figura 16, se muestran los resultados de la prueba ACF y PACF para las series de tiempo. Estas pruebas han sido realizadas hasta un valor lag de 50 para el ACF y 25 para PACF y un límite de confianza del 95% (líneas celestes). En el ACF y PACF se observan picos fuera del límite de confianza y en forma de ondas y movimientos oscilatorios a lo largo de los lags. Esto indica que las series no tienen tendencia y se confirma la estacionalidad en ellas.

En la Figura 16 (a) y (b) muestra las pruebas de ACF y PACF para la radiación solar horaria. En el ACF se observan 5 picos positivos fuera del límite de confianza en el cual tiene un pico significativo en 5, esto indica el orden del término no estacional MA(q). También, hay un patrón cada 12 lag o su múltiplo, el cual determina el término seasonal MA(Q) con un orden de 1 y a su vez un valor de  $s=12$ . Por otro lado, en el PACF se observan un positivo pico significativo en el lag 1, el cual indica la posibilidad del término no-estacional AR(p) con un orden de 1 y picos significativos en lags múltiplo de 12 por lo que el término estacional AR(P) tiene un orden de 1. Este análisis sugiere un modelo SARIMA de  $(1,0,5) (1,0,1)_{12}$  para la serie de tiempo de la radiación solar.

Similarmente, en la Figura 16 (c) y (d) se muestra el ACF y PACF para la potencia de consumo. Por medio del ACF se pueden determinar el término no estacional y estacional para el MA con el orden de 2 y 1. Además, hay un patrón cada 12 lag lo que determina el valor de la orden s. En el PACF se observa el orden de los términos no estacional y estacional para AR con valor de 1. Este análisis sugiere un modelo SARIMA de  $(1,0,2) (1,0,1)_{12}$  para la serie de tiempo de la potencia de consumo. Los modelos obtenidos por las gráficas de ACF y PACF son estimaciones iniciales, por lo que se probarán varios valores de los parámetros p, d, q, P, D y Q para encontrar un mejor modelo con bajo error predictivo mediante las métricas de error.

**Figura 16**

*Función de Autocorrelación y Parcial Autocorrelación para la radiación solar y potencia de consumo.*



Nota: Las barras con puntos representan los picos de significancia y la franja celeste es el intervalo de confianza.

#### 4.2.1.2 Evaluación de los parámetros SARIMA

Los valores predichos del modelo SARIMA y los valores observables de la data de entrenamiento son evaluados bajo las métricas de error (MSE, RSME, MAE y  $R^2$ ). Estos resultados son mostrados en la Tabla 2. Los mejores valores de estas métricas determinan el modelo adecuado para describir y predecir ambas series de tiempos.

Para la radiación solar, se puede observar que los modelos  $(1,0,0) (3,1,3)_{12}$ ,  $(1,0,0) (4,0,6)_{12}$  y  $(1,0,0) (6,0,4)_{12}$  tienen mayores valores entre todos los modelos evaluados. Sin embargo, el modelo  $(1,0,0) (6,0,5)_{12}$  tiene el mayor valor de  $R^2$  (0.8875) y es el modelo más adecuado, ya que tiene mayor precisión en las predicciones de la serie de tiempo. Por otro lado, en la potencia de consumo se observan modelos  $(1,0,0) (1,1,3)_{12}$ ,  $(1,0,0) (2,1,3)_{12}$  y  $(1,0,0) (5,0,4)_{12}$  con valores altos de  $R^2$ . Este último modelo tiene el mayor valor de  $R^2$  y menores valores en las otras métricas, por lo que el modelo puede describir la serie de tiempo de forma más precisa con un bajo error de predictivo.

**Tabla 2**  
*Diferentes combinaciones del modelo SARIMA con sus respectivos errores métricos*

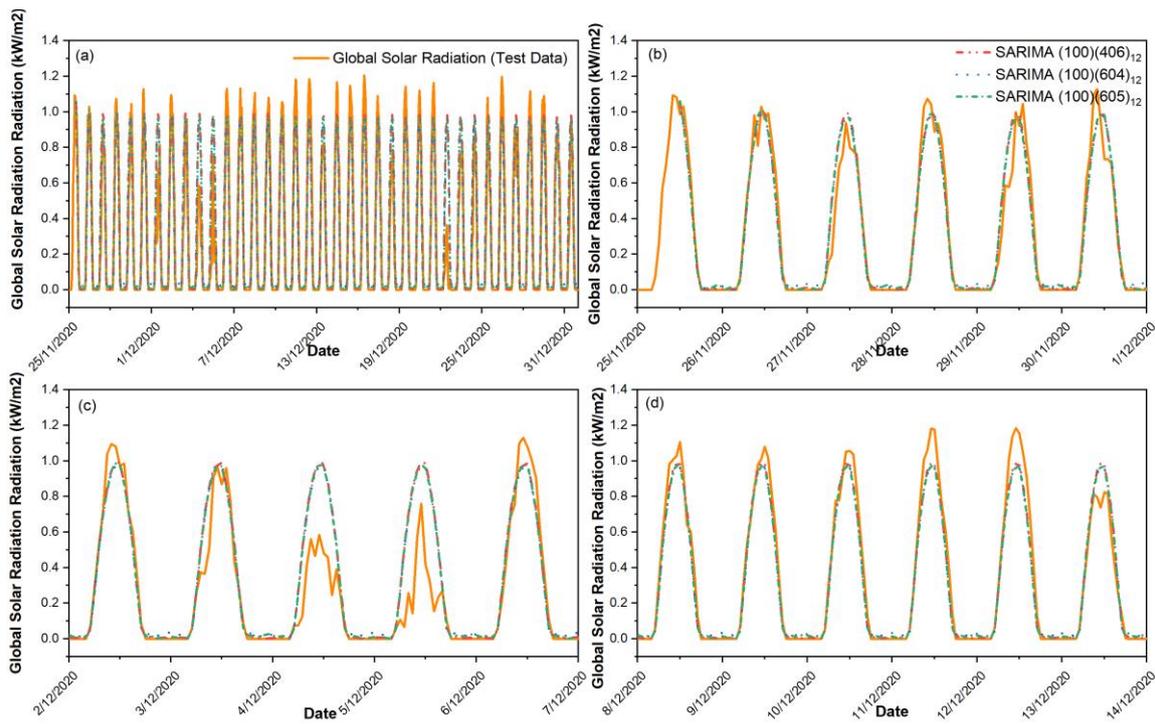
Radiación solar global (Data de prueba)					Potencia de consumo (Data de prueba)				
$(p,q,d)(P,Q,D)_s$	MSE	RSME	MAE	$R^2$	$(p,q,d)(P,Q,D)_s$	MSE	RSME	MAE	$R^2$
$(1,0,5)(1,0,1)_{12}$	0.47089	0.68622	0.59160	<b>-2.19578</b>	$(1,0,2)(1,0,1)_{12}$	15.58711	3.94805	2.06203	0.85644
$(1,0,0)(2,0,2)_{12}$	0.02219	0.12924	0.07229	0.85665	$(1,0,0)(1,1,2)_{12}$	2.58395	1.60747	0.68344	0.97620
$(1,0,0)(2,0,6)_{12}$	0.02111	0.14528	0.08504	0.85675	$(1,0,0)(1,1,3)_{12}$	2.58087	1.60651	0.68248	<b>0.97623</b>
$(1,0,0)(2,1,1)_{12}$	0.01970	0.14036	0.09436	0.86630	$(1,0,0)(2,1,1)_{12}$	2.66298	1.63186	0.71375	0.97547
$(1,0,0)(2,1,3)_{12}$	0.01689	0.12994	0.07696	0.88541	$(1,0,0)(2,1,3)_{12}$	2.58123	1.60662	0.68268	<b>0.97621</b>
$(1,0,0)(2,1,4)_{12}$	0.01689	0.12996	0.07720	0.88537	$(1,0,0)(2,1,4)_{12}$	2.59511	1.61093	0.68578	0.97610
$(1,0,0)(3,1,3)_{12}$	0.01683	0.12972	0.07455	<b>0.88580</b>	$(1,0,0)(4,1,2)_{12}$	2.58142	1.60668	0.68274	0.97623
$(1,0,0)(4,0,6)_{12}$	0.01683	0.12971	0.06878	<b>0.88582</b>	$(1,0,0)(4,1,5)_{12}$	2.58208	1.60689	0.68799	0.97622
$(1,0,0)(6,0,4)_{12}$	0.01670	0.12922	0.07621	<b>0.88668</b>	$(1,0,0)(4,1,6)_{12}$	2.58690	1.60838	0.68592	0.97618
$(1,0,0)(6,0,5)_{12}$	<b>0.01657</b>	<b>0.12874</b>	<b>0.07122</b>	<b>0.88752</b>	$(1,0,0)(4,2,4)_{12}$	2.58372	1.60740	0.68711	0.97620
$(1,0,0)(6,0,7)_{12}$	0.01682	0.12971	0.07969	0.88572	$(1,0,0)(5,0,4)_{12}$	<b>2.53350</b>	<b>1.59170</b>	<b>0.67842</b>	<b>0.97667</b>

En la Figura 17 y 18 se visualizan los resultados de las predicciones de los modelos con mayor rendimiento y bajo error predictivo para ambas series de tiempo. En la Figura 17 (a) y 18 (a), se observan las predicciones de los modelos SARIMA (líneas punteadas coloreadas) sobre la data de prueba (línea continua) de las series de tiempo, comprendidas

entre el 25 de noviembre hasta el 31 de diciembre. En la Figura 17 (b) se observa un buen ajuste de los modelos entre las predicciones y la data de prueba de la radiación solar desde el 25 de noviembre hasta el 1 de diciembre. En la Figura 17 (c) y (d), comprendidos entre el 2 de diciembre hasta 14 de diciembre, se observan buen ajuste en algunos picos, sin embargo, cuando hay presencia de nubes no llegan a ser predichos por el modelo.

**Figura 17**

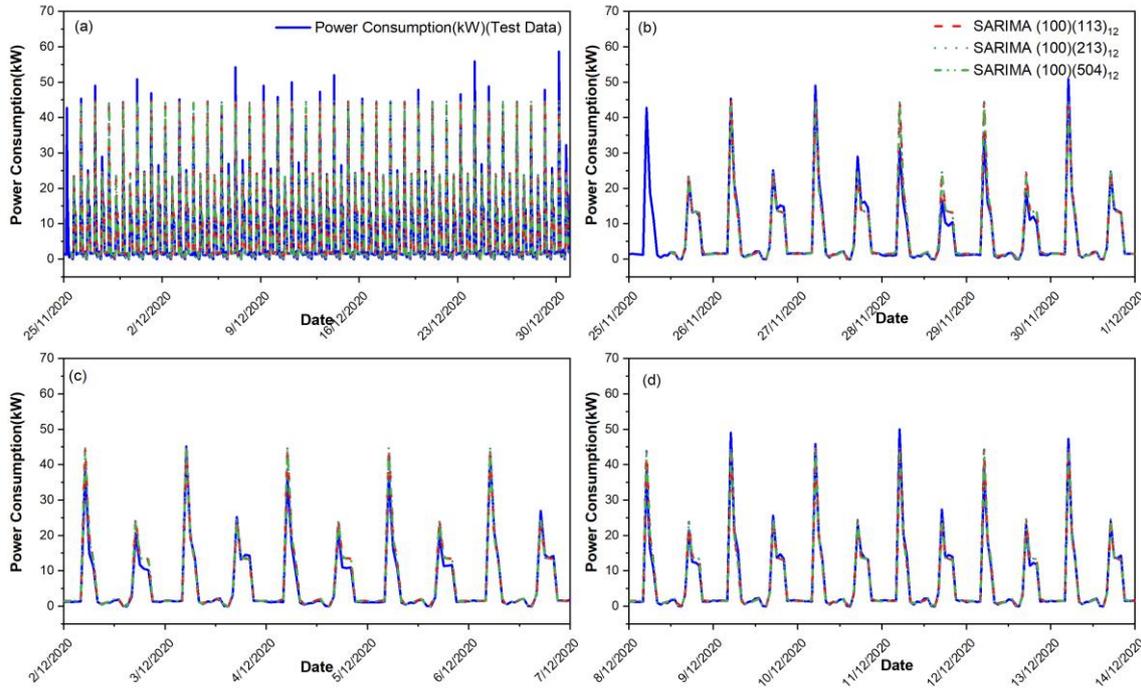
*Predicción del modelo SARIMA sobre la data de prueba de la radiación solar*



Por otro lado, en la potencia de consumo se muestra un buen ajuste para los modelos SARIMA. Esto es debido a la poca variabilidad que tiene esta serie de tiempo que permite una mejor precisión en las predicciones del modelo. En la Figura 18 (a) se observa las predicciones de los modelos sobre la data de prueba y en las Figura 18 (b), (c) y (d) muestran una vista ampliada de las predicciones.

**Figura 18**

Visualización de la data de prueba y las predicciones del modelo SARIMA con sus diferentes configuraciones en la potencia de consumo



## 4.2.2 Modelo predictivo ANN

### 4.2.2.1 Pre-procesamiento

La partición de la data en data de entrenamiento y prueba también fue utilizado para los modelos supervisados del tipo Redes Neuronales Perceptron Multicapa (ANN-MLP). Para la construcción del modelo ANN fue necesario escalar o normalizar la data de entrenamiento, entre valores de 0 y 1, para que sean interpretados por la función de activación de los modelos ANN.

A su vez la data de entrenamiento de las series de tiempo univariable (radiación solar, potencia de consumo) fueron transformadas mediante la técnica de desplazamiento de ventanas para ser interpretadas por el modelo ANN, donde los valores observables son desplazados en pasos de tiempo anteriores, llamándose lag de observaciones, y formándose así varias series de tiempo conformado por valores observables desplazadas que ingresan a las entradas (características) y salida (objetivo) del modelo. Estas lags

observables para la data de prueba fue un valor de 876 que comprende toda nuestra data de prueba para ser predicha por el modelo. Una vez la data es pre procesada esta lista para ser evaluada por el modelo.

#### 4.2.2.2 Evaluación de hiperparámetros y desempeño del modelo ANN

Para obtener un modelo ANN adecuado es importante estudiar sus hiperparámetros y el desempeño de sus predicciones. Para ello se evalúan diferentes hiperparámetros que generan diversas configuraciones del modelo, entre ellos tenemos algoritmo de aprendizaje, función de activación, abandono de neuronas y número de neuronas en la capa oculta. Hay otros hiperparámetros, como ventanas de entrada o neuronas de entrada, tasa de aprendizaje, número de neuronas de salida y otros parámetros que se mantienen fijos en este estudio.

En la Tabla 3 se muestran los diferentes valores de estos hiperparámetros utilizados en la configuración ANN, los cuales generan 216 configuraciones a evaluar. Estas configuraciones se representan como arquitecturas del tipo (número de neuronas, optimizador, función de activación, con y sin caída de neuronas, entre otras). Las arquitecturas se entrenan con datos de entrenamiento a una tasa de aprendizaje de 0,001 y con 80 épocas de entrenamiento para cada optimizador.

**Tabla 3**

*Hiperparámetros y sus valores para diferentes arquitecturas del modelo de red neuronal*

Hiperparámetros	Valores posibles	Cantidad
Número de etapas de tiempo o neuronas de entradas	876	1
Algoritmo de aprendizaje u optimizador	SGD, Adam, RMSprop	3
Tasa de aprendizaje	0.001	1
Funciones de activación	Sigmoid, Relu and Tanh	3
Número de capa oculta	1	1
Número de neuronas en la capa oculta	2, 3, 4, 5,6, 10, 20, 40, 60, 80, 100, 150	12
Porcentaje de anulación de neuronas	0% (dropout=0) and 20% (dropout=0.2)	2
Número de etapas de tiempo de salida o neurona de salida	1	1
Número total de hiperparámetros		216

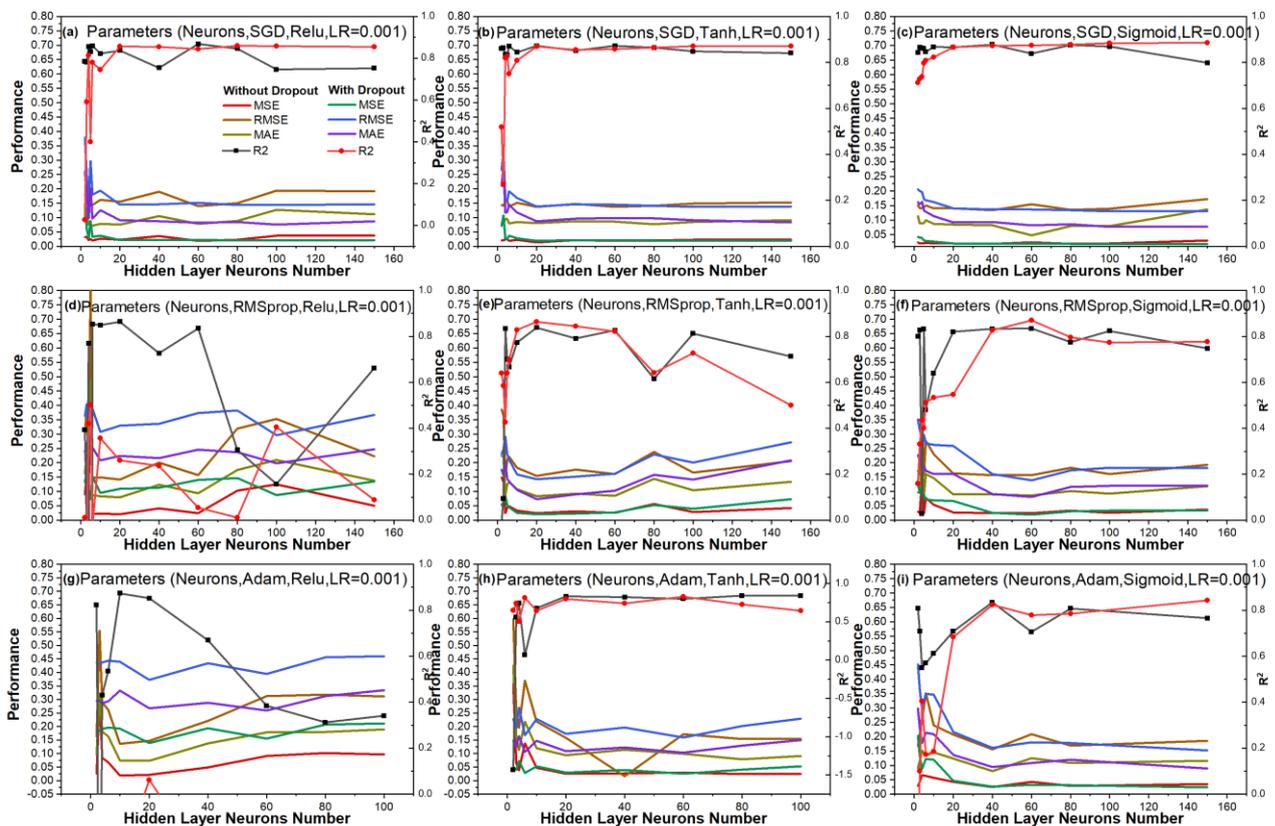
La Figura 19 muestra los desempeños de las arquitecturas del modelo ANN en función del número de neuronas y la tasa de aprendizaje ( $Lr = 0.001$ ). Estos desempeños se obtienen mediante las métricas de error del tipo MSE, RMSE, MAE y  $R^2$ .

Las arquitecturas con optimizador SGD y diferentes funciones de activación (Relu, Tanh y Sigmoid) tienen mejor desempeño que con los otros optimizadores, como se muestra en la Figura 19 (a) y (c). Por otro lado, en la Figura 19 (d) y (f), las arquitecturas con RMSprop optimizado y diferentes funciones de activación tienen un rendimiento bajo cuando se aumentan las neuronas para casos con y sin caída de neuronas. También se observa que el optimizador RMSprop con Tanh y Sigmoid tienen menor rendimiento que el optimizador SGD y las funciones de activación estudiadas.

En la Figura 19 (g), las arquitecturas con Adam y Relu tienen rendimientos similares al optimizador RMSprop y la función Relu. Sin embargo, con las funciones de activación de Tanh y Sigmoid, estos rendimientos tienden a aumentar cuando las neuronas aumentan, como se muestra en la Figura 19 (h) y (i).

**Figura 19**

*Métricas de error para diferentes arquitecturas con SGD, RMSprop, Adam Optimizer y funciones de activación (Relu, Tanh y Sigmoid) para diferentes casos con y sin neurona dropout.*



#### 4.2.2.3 Predicción y rendimiento del modelo ANN sobre las bases de datos

Los datos de la serie tiempo de la radiación solar son de un año de registro en unidades de kW/m<sup>2</sup> por hora y tiene 8760 valores observables que se dividen en datos de entrenamiento y prueba en una relación del 90% y 10%, respectivamente. Así mismo, los valores mínimo y máximo de la data de entrenamiento de la radiación solar son de 0 y 1.2 kW/m<sup>2</sup>, los cuales son escalado o normalizados a 0 y 1 según la función de activación del modelo.

Para la preparación de datos, los valores observables de la serie temporal se dividieron en 7884 muestras con 876 pasos de tiempo como variable de entrada y un paso de tiempo como variable de salida. Estos datos de entrada y salida ingresan al modelo para entrenarlo y hacer predicciones.

Las predicciones generadas por las arquitecturas del modelo ANN se evalúan con los datos de prueba utilizando métricas de error (MSE, RMSE, MAE, R<sup>2</sup>) para indicar qué tan bueno es su rendimiento. Posteriormente se guardan las arquitecturas ANN para posteriormente entrenarlas con toda la serie temporal y predecir valores futuros fuera de la serie.

La configuración de los modelos ANN dependen de sus parámetros, los cuales son el número de entradas, tasa de aprendizaje, número de neuronas en la capa oculta, optimizadores y varios parámetros para el procesamiento de la data en el modelo. Las configuraciones que se realizaron para nuestros modelos ANN fueron 876 datos de entrada o nodos, una tasa de aprendizaje de 0.001, neuronas entre 2 hasta 500 neuronas, optimizador del tipo gradiente descendiente estocástico con un epoch de aprendizaje de 80 y funciones de activación del tipo sigmoid para la radiación solar y sigmoid, relu y tanh para la potencia de consumo.

Después de varias pruebas con diferentes configuraciones de los modelos, las mejores configuraciones ANN con mayor precisión predictiva para ambas series de tiempo fueron estudiadas mediante sus métricas de error sobre la data de prueba. En la Tabla 4 se muestran los resultados de diferentes configuraciones del modelo. Las configuraciones del modelo ANN para la radiación solar tienen bajos error predictivo de

MSE, RMSE, MAE y alto  $R^2$  en las configuraciones con función de activación sigmoid, optimizador SGD y neuronas de 200, 500 y 600.

Por el lado de la potencia de consumo, se observa que las configuraciones presentan mejores predicciones con bajos errores predictivos y altos  $R^2$  cuando su configuración está comprendida por el optimizador SGD y distintas funciones de activación (sigmoid, relu y tanh) con neuronas de 80, 100 y 150.

**Tabla 4**  
*Diferentes configuraciones del modelo ANN para la radiación solar y la potencia de consumo con bajos errores predictivos*

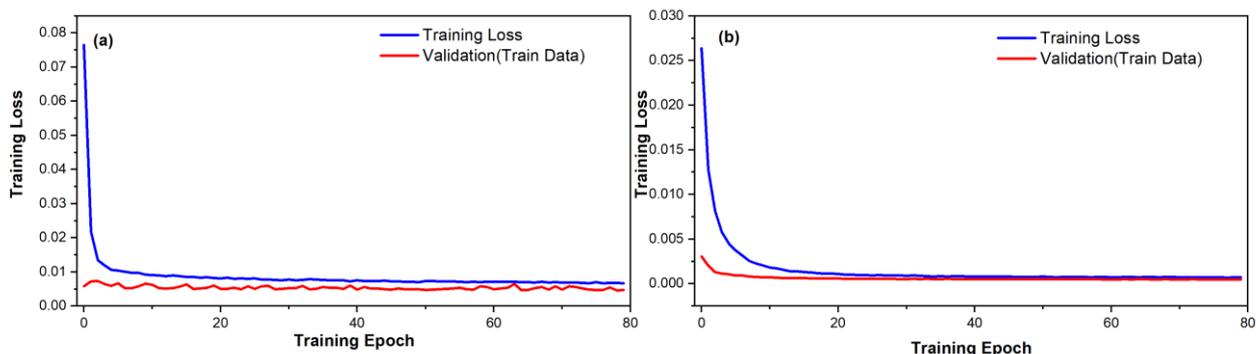
Radiación solar ( Data de prueba)					Potencia de consumo (Data de prueba)				
Configuración ANN	MSE	RSME	MAE	$R^2$	Configuración ANN	MSE	RSME	MAE	$R^2$
(876,600,SGD,Sigmoid)	0.01737	0.13256	0.07993	0.88075	(876,150,SGD,Sigmoid)	6.60237	2.56951	1.40275	0.93919
(876,550,SGD,Sigmoid)	0.02977	0.17255	0.14071	0.79794	(876,150,SGD,Relu)	5.04491	2.24609	1.14011	0.95354
(876,500,SGD,Sigmoid)	0.01696	0.13022	0.07599	<b>0.88493</b>	(876,150,SGD,Tanh)	<b>2.77300</b>	<b>1.66523</b>	<b>0.82319</b>	<b>0.97446</b>
(876,450,SGD,Sigmoid)	0.01850	0.13602	0.09698	0.87444	(876,100,SGD,Sigmoid)	10.48545	3.23812	2.45476	0.90343
(876,400,SGD,Sigmoid)	0.01673	0.12936	0.07988	<b>0.88644</b>	(876,100,SGD,Relu)	4.81735	2.19985	1.08520	0.95563
(876,350,SGD,Sigmoid)	0.01710	0.13076	0.07389	0.88396	(876,100,SGD,Tanh)	2.83562	1.68393	0.86494	<b>0.97388</b>
(876,300,SGD,Sigmoid)	0.01990	0.14108	0.10044	0.86492	(876,80,SGD,Sigmoid)	5.04491	2.24609	1.14011	0.95354
(876,250,SGD,Sigmoid)	0.02399	0.15488	0.12047	0.83721	(876,80,SGD,Tanh)	5.65534	2.37810	1.19221	0.94792
(876,200,SGD,Sigmoid)	<b>0.01684</b>	<b>0.12976</b>	<b>0.07728</b>	<b>0.88574</b>	(876,80,SGD,Relu)	4.79921	2.19071	1.09551	<b>0.95580</b>
(876,150,SGD,Sigmoid)	0.01697	0.13026	0.07744	0.88484	(876,70,SGD,Sigmoid)	16.46820	4.05810	2.27594	0.84833

En la Figura 20 se muestran las curvas de entrenamiento de las distintas configuraciones del modelo ANN con mayor rendimiento en sus predicciones para la radiación solar y potencia de consumo. En la Figura 20 (a) se observa el aprendizaje de la configuración (876, 400, SGD, Sigmoid) para la radiación solar con una adecuada curva de pérdida de entrenamiento, alcanzando valores mínimos a medida que aumenta los epoch o iteraciones. Similarmente, se observa en la validación de la data de entrenamiento, pero con ligeros saltos indicando pequeños errores de las predicciones sobre la data de entrenamiento.

En la Figura 20 (b) se observa una mejor curva de aprendizaje en la pérdida de entrenamiento y validación para la configuración (876,150, SGD, Tanh) en la potencia de consumo, lo que indica un buen ajuste de sus parámetros, un buen entrenamiento y una mejor predicción.

**Figura 20**

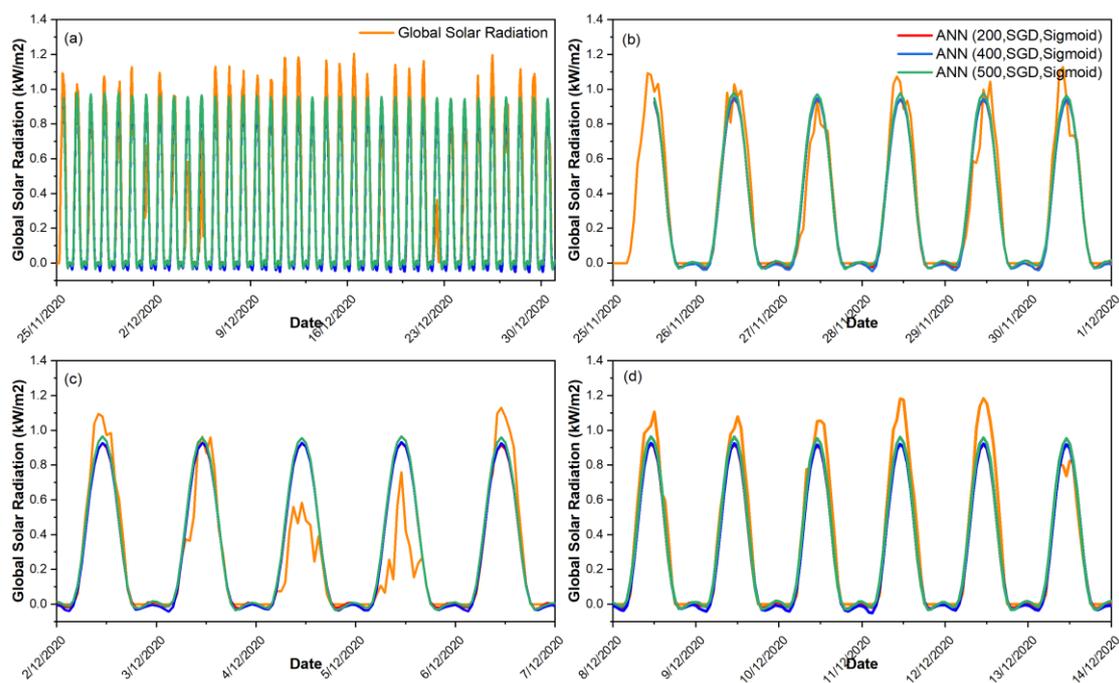
Curva de entrenamiento del modelo ANN: (a) Entrenamiento de la data para la radiación solar. (b) Entrenamiento de la data para potencia de consumo



En la Figura 21 y 22, las predicciones de los modelos de la radiación solar y la potencia de consumo son comparados en sus datas de pruebas. Para ambos casos, las predicciones se ajustan bastante bien y hay excepciones en algunos picos con altos y bajos valores predictivos. Se observa también que las predicciones de las mejores configuraciones del modelo ANN para la radiación solar no varían mucho y similarmente se observan en la potencia de consumo. En la Figura 21 (c) se muestra que las predicciones no pueden predecir esas variaciones fluctuantes de la radiación solar al igual como ocurre en el modelo SARIMA.

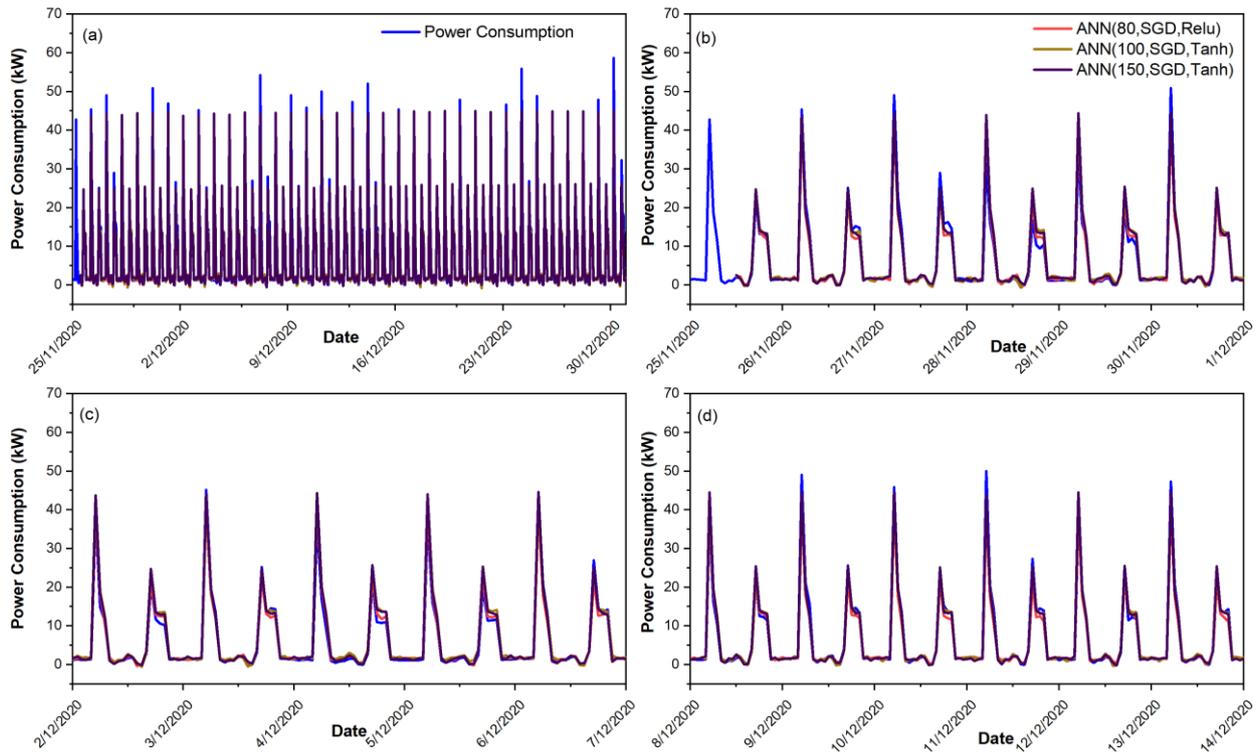
**Figura 21**

Visualización de las predicciones de las diferentes configuraciones de los modelos ANN sobre la data de prueba de la radiación solar



**Figura 22**

Visualización de las predicciones de los modelos ANN sobre la data de prueba de la potencia de consumo



## 4.2.3 Modelo LSTM

### 4.2.3.1 Preprocesamiento del modelo

El preprocesamiento del modelo LSTM es similar al modelo ANN en base a la partición de la data y la normalización que esta entre valores de 0 y 1. Así mismo, la data es convertida en valores observables desplazados con la misma técnica de Desplazamiento de Ventanas con lags, llamadas etapas de tiempo o pasos, de 50 para que ingresen al modelo como datos de entradas y salidas.

### 4.2.3.2 Evaluación de los hiperparámetros y desempeño del modelo

El modelo LSTM considera los hiperparámetros: Las celdas de memorias en la capa de entrada del modelo, el algoritmo de aprendizaje, las funciones de activación, anulación de celdas de memoria y otros parámetros. En la Tabla 5 se muestra las distintas combinaciones que se han realizado en la evaluación de los hiperparámetros, en el cual

se ha utilizado una librería llamada GridSearch en Python para evaluarlos y obtener sus resultados de forma automatizada.

**Tabla 5**

*Hiperparámetros y sus valores para cada configuración del modelo LSTM*

<b>Hiperparámetros</b>	<b>Valores posibles</b>	<b>Cantidad</b>
Número de etapas de tiempo o de la celda de memoria	50	1
Algoritmo de aprendizaje u optimizador	SGD, Adam, RMSprop	3
Tasa de aprendizaje	0.001	1
Funciones de activación	Sigmoid, Relu and Tanh	3
Número de capas ocultas	4	4
Número de celda de memoria en la capa oculta	50	12
Porcentaje de anulación de neuronas	0% (dropout=0) and 20% (dropout=0.2)	2
Número de etapas de tiempo de salida o celda de memoria	1	1
Número total de hiperparámetros		864

#### **4.2.3.3 Predicción y rendimiento del modelo**

En modelo LSTM se ha utilizado para predecir la radiación solar, temperatura ambiente y potencia de consumo para dos localidades, entre ellas la Comunidad rural de Tarucani y la Biblioteca central de la Universidad Nacional de Ingeniería.

Los resultados obtenidos al evaluar el modelo LSTM para ambos casos de estudios son descritos a continuación:

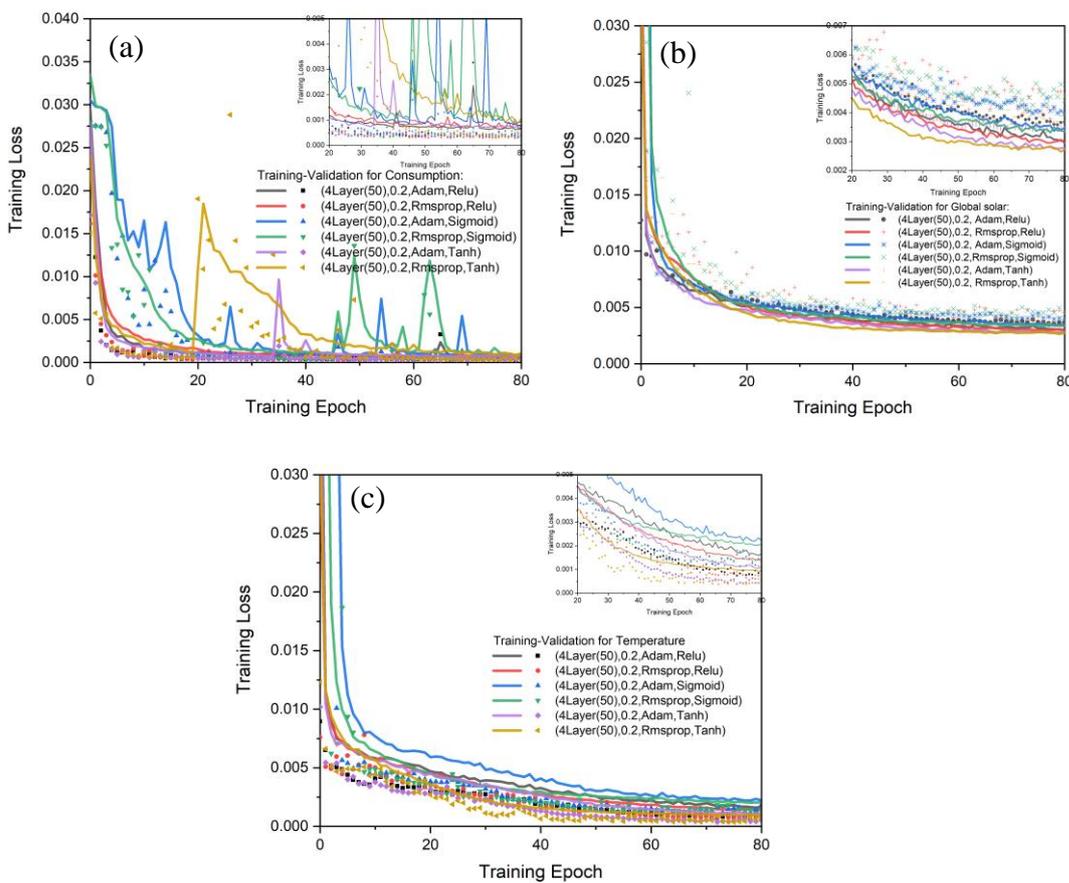
- 1) El primer caso de estudio, comunidad de Tarucani:

La Figura 23 muestra la curva de aprendizaje del modelo LSTM con diferentes configuraciones para las series temporales de la radiación solar, temperatura ambiente y potencia de consumo. En la Figura 23 (b) y (c), la curva de aprendizaje y validación de la radiación solar y potencia de consumo tienden a converger para todas las configuraciones del modelo, lo que indica que el rendimiento del modelo está evolucionando a medida que aumenta las iteraciones o épocas. Así mismo, la curva de entrenamiento refleja el

ajuste de las configuraciones del modelo con la data de entrenamiento. Por otro lado, la curva de validación indica el rendimiento de las configuraciones alimentadas con nuevos datos. En la Figura 23 (a) se muestra la curva de entrenamiento de la potencia de consumo del lugar en estudio, el cual presenta grande variabilidad con respecto a distintas configuraciones del modelo LSTM, sin embargo, en algunas configuraciones presenta un buen aprendizaje convergiendo a medida que se va entrenando.

**Figura 23**

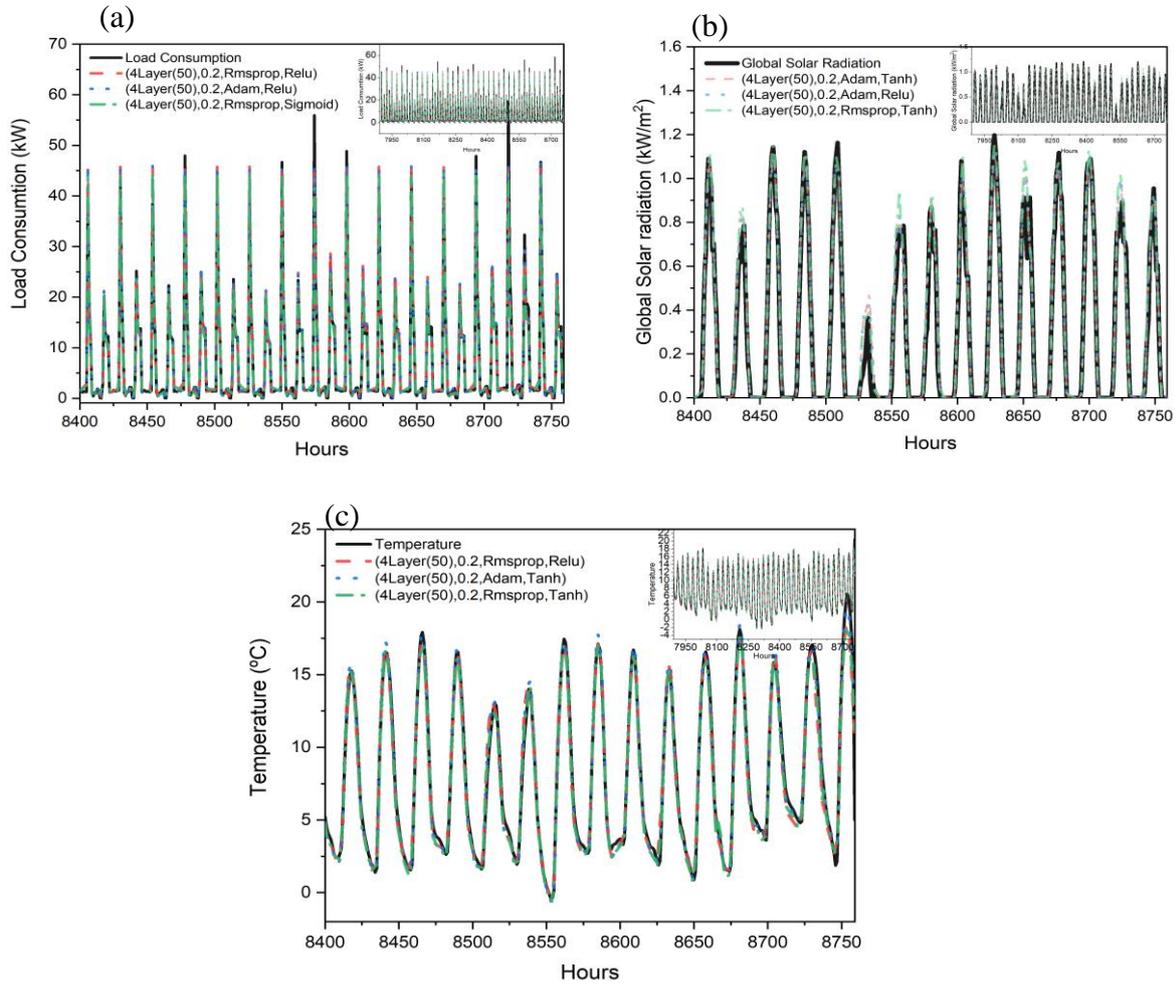
*Curva de entrenamiento de las diferentes configuraciones del modelo LSTM para la radiación sola, potencia de consumo y temperatura ambiente*



En la Figura 24 se muestran la predicción de las distintas configuraciones del modelo LSTM para las series de tiempo. En esta figura se muestra la potencia de consumo, radiación solar y temperatura ambiente observable con líneas de color negro, mientras que las predicciones de sus configuraciones del modelo tienen diferentes colores. Estas configuraciones mostradas en las gráficas tienen un mayor rendimiento determinado por sus métricas de error, como se muestra en la Tabla 6.

**Figura 24**

*Predicciones de las configuraciones del modelo LSTM con mayores rendimientos sobre la radiación, potencia de consumo y temperatura ambiente*



**Tabla 6**

*Resultado de la evaluación de las distintas configuraciones del modelo LSTM sobre las series de tiempo estudiadas*

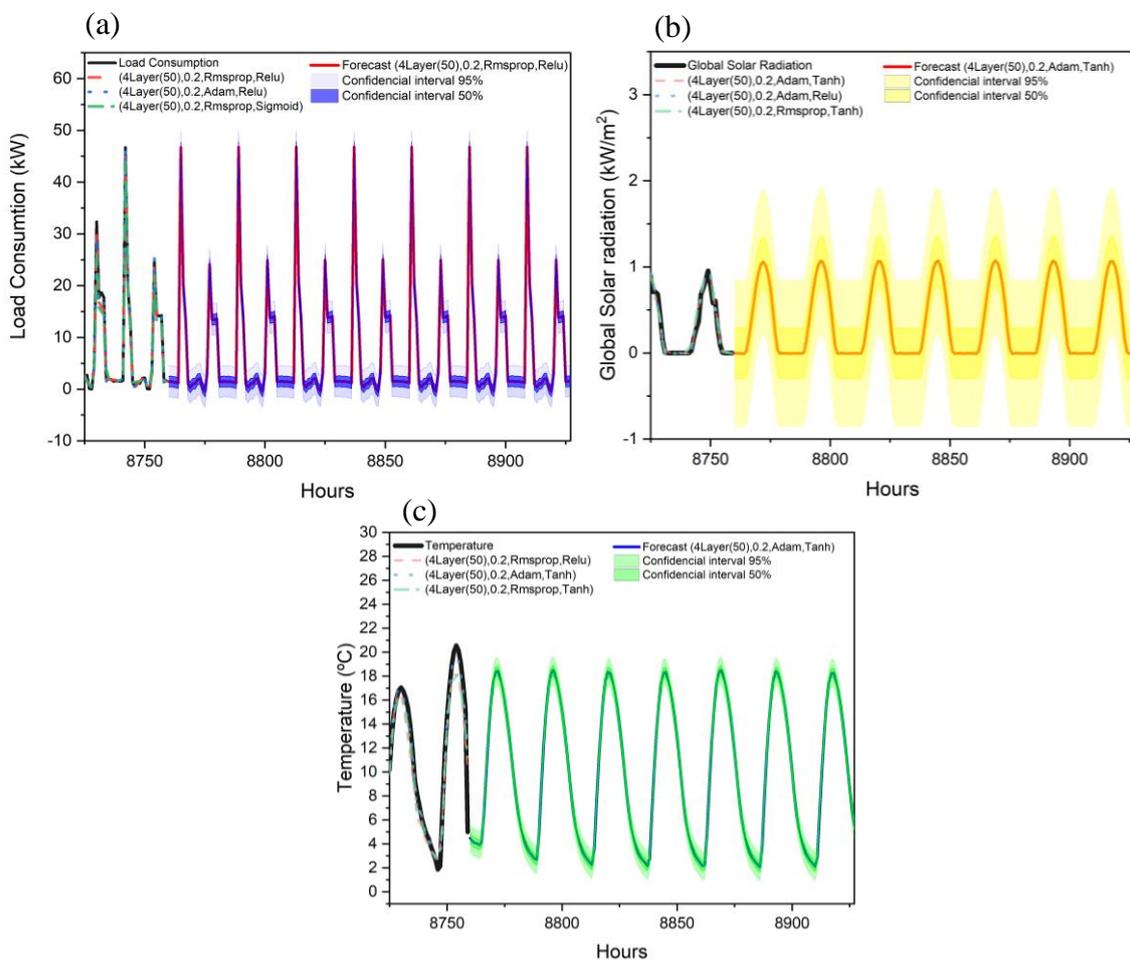
Radiación solar (Data de prueba)				Potencia de consumo (Data de prueba)			
Configuración LSTM	MSE	RSME	R <sup>2</sup>	Configuración LSTM	MSE	RSME	R <sup>2</sup>
(50,50,50,50,Tanh,Adam)	0.00539	0.07338	<b>0.96335</b>	(50,50,50,50,Relu,Rmsprop)	1.95645	1.39873	<b>0.98208</b>
(50,50,50,50,Relu,Adam)	0.00564	0.07507	<b>0.96165</b>	(50,50,50,50,Relu,Adam)	2.03406	1.42620	<b>0.98137</b>
(50,50,50,50,Tanh,Rmsprop)	0.00578	0.07601	<b>0.96068</b>	(50,50,50,50,Sigmoid,Rmsprop)	2.15872	1.46926	<b>0.98023</b>
(50,50,50,50,Relu,Rmsprop)	0.00593	0.07699	<b>0.95967</b>	(50,50,50,50,Tanh,Adam)	<b>2.20254</b>	<b>1.48410</b>	<b>0.97983</b>
(50,50,50,50,Sigmoid,Adam)	0.00686	0.08280	0.95335	(50,50,50,50,Tanh,Rmsprop)	2.45531	1.56694	0.97752
(50,50,50,50,Sigmoid,Rmsprop)	0.00712	0.08439	0.95154	(50,50,50,50,Sigmoid,Adam)	2.49324	1.57900	0.97717

Temperatura ambiente (Data de prueba)			
Configuración LSTM	MSE	RSME	R <sup>2</sup>
(50,50,50,50,Tanh,Adam)	0.27343	0.52291	0.99008
(50,50,50,50,Tanh,Rmsprop)	0.40389	0.63552	0.98534
(50,50,50,50,Relu,Rmsprop)	0.54586	0.73882	0.98019
(50,50,50,50,Relu,Adam)	0.71947	0.84821	0.97389
(50,50,50,50,Sigmoid,Adam)	0.94389	0.97154	0.96575
(50,50,50,50,Sigmoid,Rmsprop)	1.13518	1.06545	0.95881

En la Figura 25 se muestran las predicciones a futuro de las configuraciones que presentaron mayor rendimiento en el entrenamiento y predicción. Estas predicciones del modelo LSTM se realizaron para la semana siguiente con sus intervalos de confianza al 95% y 50%. En la Figura 25 (a) y (c) presentaron intervalos de confianza ajustados, cercanos a los valores predichos; mientras que la Figura 25 (b) se muestra que el intervalo de 95% es más amplio. Esto puede deberse a la variabilidad que presenta la radiación solar a lo largo del año que conlleva a un error grande en sus predicciones a futuro.

**Figura 25**

*Predicción de valores futuros de la configuración con mayor rendimiento en las series de tiempo*



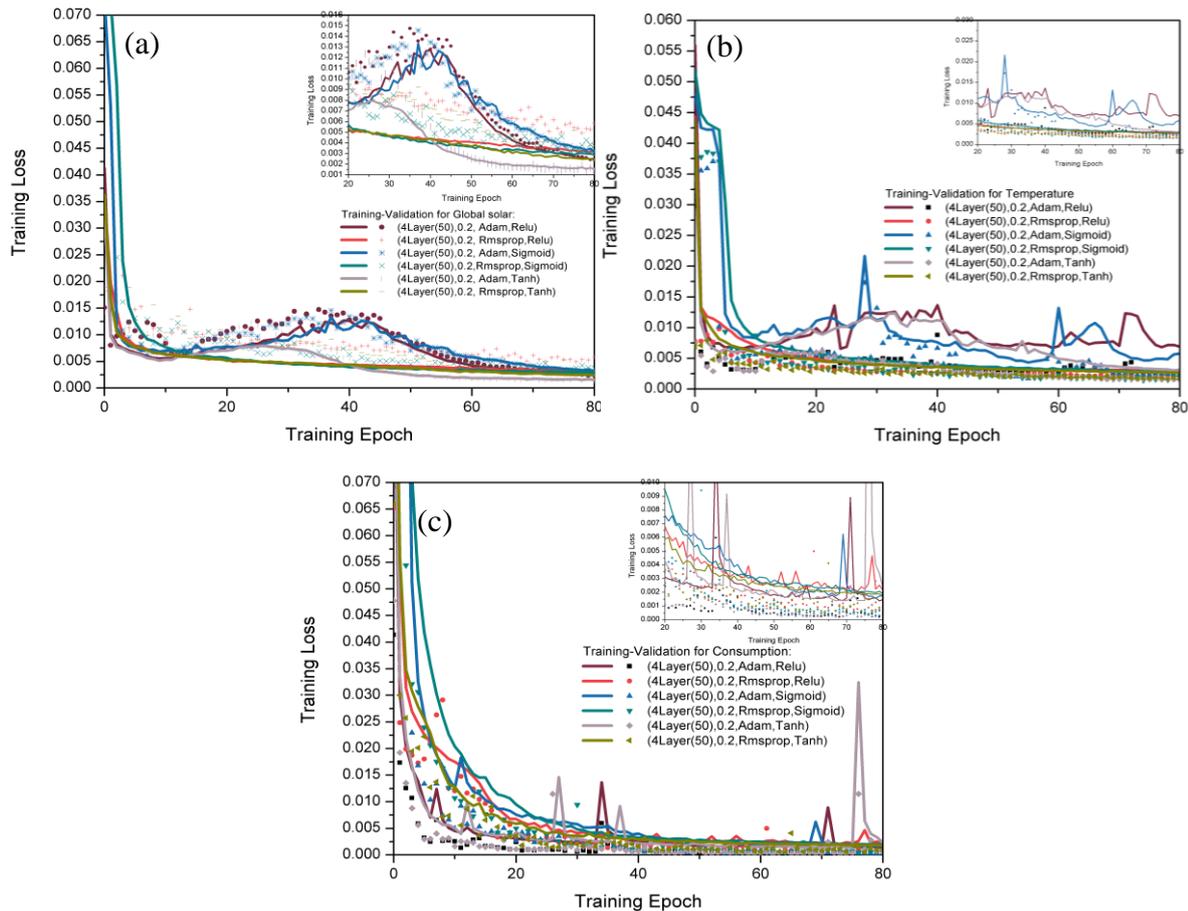
2) El segundo caso de estudio, Biblioteca Central de la Universidad:

En la Figura 26 se muestran las curvas de aprendizaje de las distintas configuraciones del modelo para las tres series de tiempo. En ellas se observan una convergencia para las distintas configuraciones, tanto para las predicciones de entrenamiento y validación, sin embargo, también se observan que a medida que aumenta

las épocas o iteraciones del aprendizaje hay algunas configuraciones que difieren entre la predicción de validación y entrenamiento teniendo valores altos, como picos, lo que indican una reducción del rendimiento del modelo en su aprendizaje.

**Figura 26**

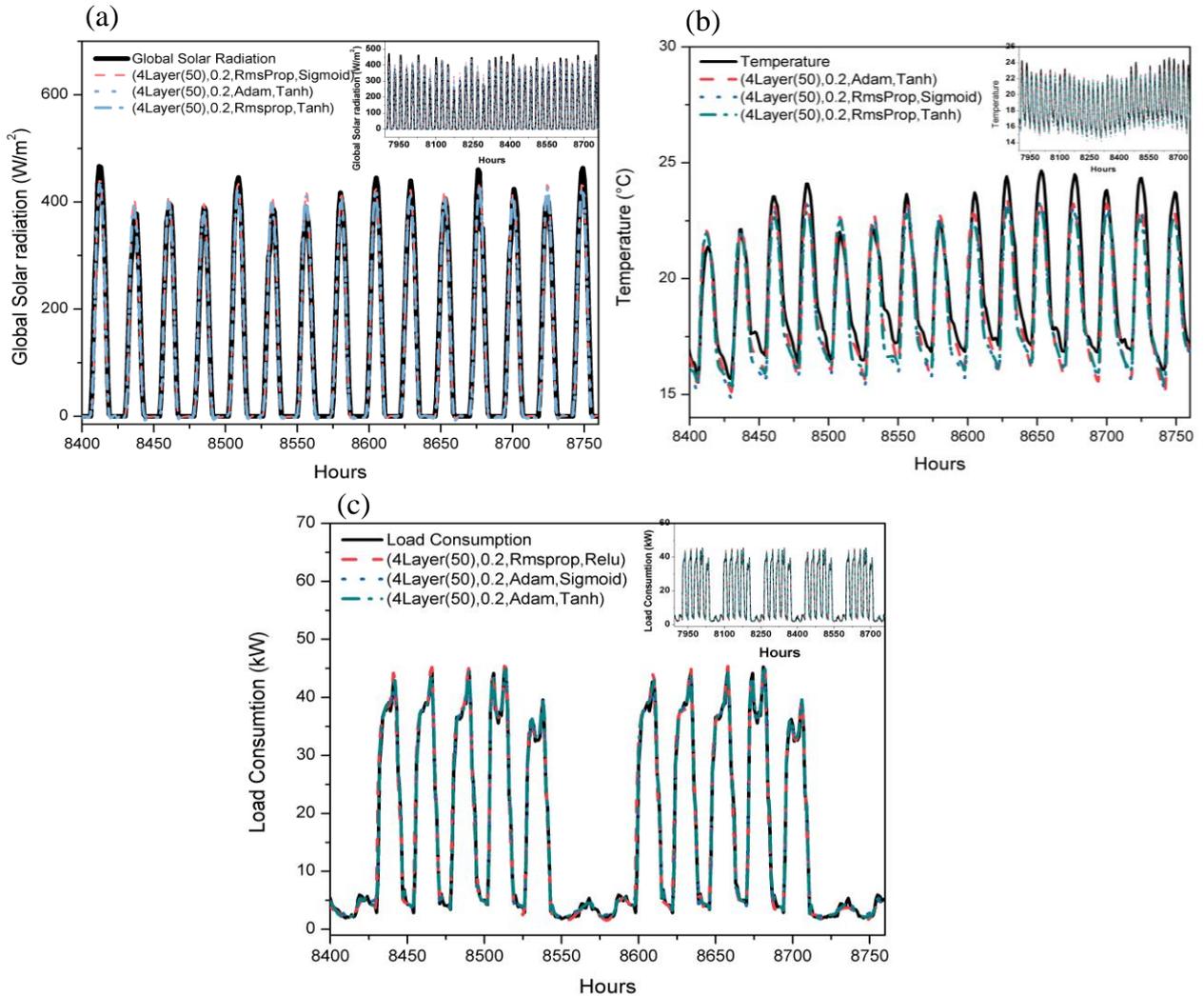
*Curva de entrenamiento de todas las configuraciones del LSTM para las series de tiempo de la Biblioteca Universitaria*



En la Figura 27 se muestran las predicciones de las configuraciones del modelo LSTM sobre la data de prueba. La radiación solar, temperatura ambiente y potencia de consumo están representados por la línea negra y las predicciones de las diferentes configuraciones del modelo están de colores, siendo estas configuraciones las que tuvieron mayor rendimiento según las métricas de error. En la Figura 27 (a) y (c) se observan que los valores de las predicciones se ajustan a los valores observables de la serie de tiempo, mientras que en la Figura 27 (b), la serie de tiempo de la temperatura ambiente tienen predicciones por debajo de los valores esperados. Así mismo, en la Tabla 7 se muestran los rendimientos de las distintas configuraciones para las tres series temporales.

**Figura 27**

*Predicción de las 3 mejores configuraciones del LSTM para las series de tiempo de la biblioteca universitaria*



**Tabla 7**

*Resultado de la evaluación de todas las configuraciones del LSTM para las series de tiempo de la biblioteca universitaria*

Radiación solar (Data de prueba)				Potencia de consumo (Data de prueba)			
Configuración LSTM	MSE	RSME	R <sup>2</sup>	Configuración LSTM	MSE	RSME	R <sup>2</sup>
(50,50,50,50,Tanh,Adam)	361.41297	19.01086	0.98520	(50,50,50,50,Relu,Rmsprop)	2.93754	1.71392	0.98814
(50,50,50,50,Relu,Adam)	859.89965	29.32404	0.96479	(50,50,50,50,Relu,Adam)	1.28004	1.13138	0.99483
(50,50,50,50,Tanh,Rmsprop)	573.06020	23.93867	<b>0.97653</b>	(50,50,50,50,Sigmoid,Rmsprop)	1.22889	1.10855	0.99503
(50,50,50,50,Relu,Rmsprop)	716.50059	26.76752	0.97066	(50,50,50,50,Tanh,Adam)	<b>0.37583</b>	<b>0.61305</b>	<b>0.99848</b>
(50,50,50,50,Sigmoid,Adam)	1335.5714	36.54547	0.94532	(50,50,50,50,Tanh,Rmsprop)	0.89383	0.94542	0.99639
(50,50,50,50,Sigmoid,Rmsprop)	619.09950	24.88171	<b>0.97465</b>	(50,50,50,50,Sigmoid,Adam)	0.48537	0.69668	0.99804

Temperatura ambiente (Data de prueba)			
Configuración LSTM	MSE	RSME	R <sup>2</sup>
(50,50,50,50,Tanh,Adam)	<b>0.55443</b>	<b>0.74460</b>	<b>0.91448</b>
(50,50,50,50,Tanh,Rmsprop)	0.75780	0.87051	0.88312
(50,50,50,50,Relu,Rmsprop)	1.15335	1.07394	0.82211
(50,50,50,50,Relu,Adam)	1.44841	1.20349	0.77661
(50,50,50,50,Sigmoid,Adam)	1.40943	1.18719	0.78262
(50,50,50,50,Sigmoid,Rmsprop)	0.78835	0.88789	0.87841

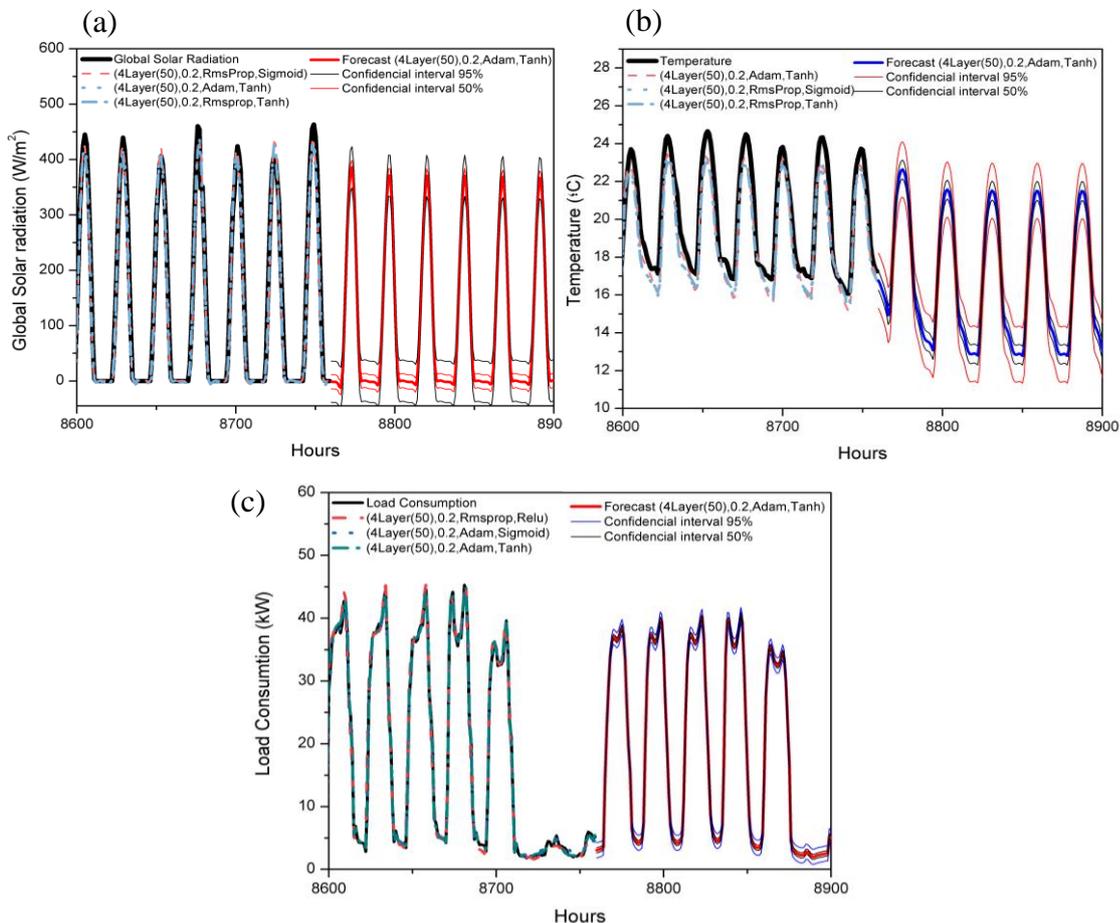
Las predicciones a futuro de las distintas configuraciones del modelo LSTM con mejor rendimiento por las métricas de error son mostradas en la Figura 28. La configuración (4Layer(50), 0.2, Adam, Tanh) tuvo un valor de  $R^2$  de 0.98520 y su predicción a futuro de la radiación solar es mostrado en la Figura 28 (a).

Así mismo, la predicción de la temperatura ambiente, con la configuración (4Layer(50), 0.2, Adam, Tanh), tuvo un valor de  $R^2$  de 0.91448 donde su predicción a futuro e intervalo de confianza son mostrados en la Figura 28 (b).

Por otro lado, un valor de  $R^2$  de 0.99848 es obtenido por la configuración (4Layer(50), 0.2, Adam, Tanh) para la potencia de consumo, en el cual esta configuración es usada para predecir a futuro la potencia de consumo con su intervalo de confianza como se observa en la Figura 28 (c). Las predicciones del modelo LSTM presentan el intervalo de confianza cercanos a los valores predicho, lo que indica un error muy bajo.

**Figura 28**

*Predicción a futuro de la configuración del LSTM con mayor rendimiento para las series de tiempo de la Biblioteca Universitaria*



#### 4.2.4. Análisis comparativo de los rendimientos de los modelos SARIMA, ANN y LSTM

Las métricas de error (MSE, RSME y  $R^2$ ) son indicadores que miden el rendimiento en base a la precisión y confiabilidad de cada modelo y sus configuraciones. En este estudio se evaluaron 3 modelos predictivos, tales como el modelo SARIMA, ANN y LSTM para la predicción de las series de tiempo de la potencia de consumo, radiación solar y temperatura ambiente de un lugar de estudio.

Los resultados obtenidos al evaluar las diferentes configuraciones de los tres modelos estudiados para la comunidad de Tarucani son mostrados en la siguiente Tabla 8. Es notable observar que el modelo LSTM presenta mayor  $R^2$ , obteniendo un valor por encima de 0.96 y un menor valor de MSE y RSME para las tres series temporales. Esto indica que el modelo LSTM es más preciso y confiable que los modelos SARIMA y ANN.

**Tabla 8**

*Comparación de las configuraciones con mayor rendimiento del modelo LSTM para las distintas series de tiempo en la comunidad rural*

<b>Radiación solar (Data de prueba)</b>	<b>MSE</b>	<b>RSME</b>	<b><math>R^2</math></b>
SARIMA (1,0,0)(6,0,4) <sub>12</sub>	0.01670	0.12922	<b>0.88668</b>
ANN (876,200,SGD,Sigmoid)	0.01684	0.12976	<b>0.88574</b>
LSTM (866,50,50,50,Tanh,Adam)	0.00539	0.07338	<b>0.96335</b>
<b>Consumo eléctrico (Data de prueba)</b>	<b>MSE</b>	<b>RSME</b>	<b><math>R^2</math></b>
SARIMA (1,0,0)(5,0,4) <sub>12</sub>	2.53350	1.59170	<b>0.97667</b>
ANN (876,150,SGD,Tanh)	2.77300	1.66523	<b>0.97446</b>
LSTM (866,50,50,50,Relu,Rmsprop)	1.95645	1.39873	<b>0.98208</b>
<b>Temperatura ambiente (Data de prueba)</b>	<b>MSE</b>	<b>RSME</b>	<b><math>R^2</math></b>
LSTM (866,50,50,50,Tanh,Adam)	0.27343	0.52291	0.99008

Por otro lado, se evaluaron distintas configuraciones del modelo LSTM para las series temporales (radiación solar, temperatura ambiente y potencia de consumo) de la Biblioteca Universitaria, en donde se obtuvieron los siguientes resultados de sus métricas de error que son mostrados en la Tabla 9. Los valores de  $R^2$  estuvieron altos para la serie temporal de la radiación solar y potencia de consumo, mientras que para la temperatura ambiente fue menor.

**Tabla 9**

*Comparación de las configuraciones con mayor rendimiento del modelo LSTM para las distintas series de tiempo en la Biblioteca Universitaria*

<b>Radiación solar (Data de prueba)</b>	<b>MSE</b>	<b>RSME</b>	<b>R<sup>2</sup></b>
LSTM (50,50,50,50,Tanh,Adam)	361.41297	19.01086	0.98520
<b>Potencia de consumo (Data de prueba)</b>	<b>MSE</b>	<b>RSME</b>	<b>R<sup>2</sup></b>
LSTM (50,50,50,50,Tanh,Adam)	0.37583	0.61305	0.99848
<b>Temperatura ambiente (Data de prueba)</b>	<b>MSE</b>	<b>RSME</b>	<b>R<sup>2</sup></b>
LSTM (50,50,50,50,Tanh,Adam)	0.78835	0.88789	0.87841

### **4.3 Modelamiento y simulación del sistema de generación fotovoltaica mediante el programa HOMER**

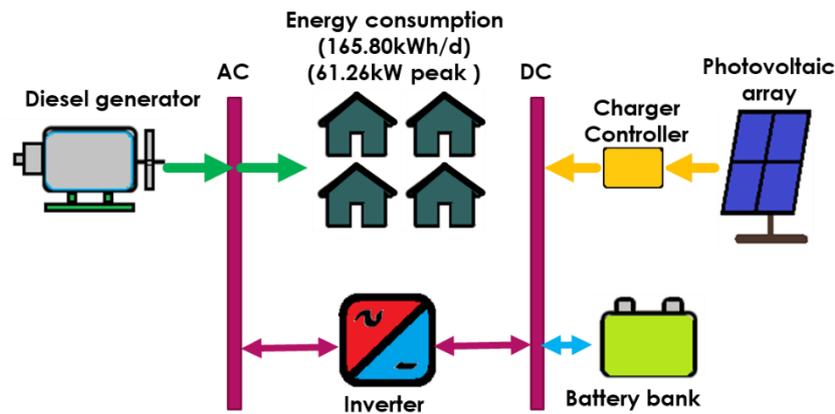
#### **4.3.1 Diseño el sistema de generación**

El sistema de energía híbrido en esta investigación consta de cuatro componentes tales como: El sistema fotovoltaico, convertidor, baterías y generador diésel para satisfacer el consumo eléctrico de la comunidad de Tarucani. En la Figura 29 se ilustra un diagrama esquemático del Sistema Fotovoltaico Híbrido (SFH) propuesto.

En esta estructura del sistema, las cargas están conectadas al bus de Corriente Alterna AC, mientras que un sistema de energía fotovoltaica y una unidad de almacenamiento de energía están conectados al bus de Corriente Directa DC. Un convertidor de potencia bidireccional une al bus de CA a DC para la transferencia de energía. Además, las baterías almacenan el exceso de electricidad cuando el sistema fotovoltaico no puede servir la carga requerida, mientras que el generador diésel sirve de respaldo cuando el sistema fotovoltaico y las baterías no cumplen con la carga.

**Figura 29**

*Diseño del Sistema fotovoltaico híbrido para la comunidad desconectada de la red en estudio*



#### 4.3.2 Datos climáticos y consumo de energía del lugar en estudio

El centro poblado se encuentra ubicado en la parte central del distrito de San Juan de Tarucani, provincia de Arequipa y se ubica al sur del Perú en las coordenadas de latitud  $16^{\circ}11'S$  y longitud  $71^{\circ} 3.6' E$ . La comunidad tiene una población total de 2.179 y 945 viviendas, según el último censo de 2017 (Municipalidad Provincial de Arequipa, 2021), y la ganadería es su principal actividad económica, como la crianza de camélidos, llamas y alpacas; y otras actividades relacionadas con la agricultura.

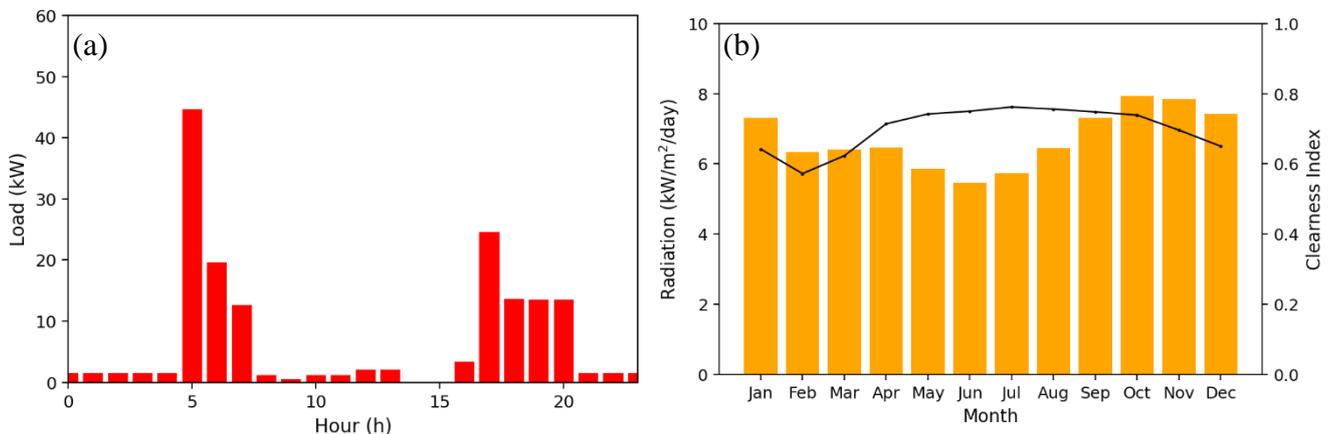
Esta comunidad rural es considerada en nuestra investigación para diseñar un sistema fotovoltaico híbrido aislado (Fotovoltaico, Controlador de carga, Inversor de batería y diésel). En este sistema se evalúa el potencial de la radiación solar del lugar, así como su análisis técnico-económico de sus componentes, seguridad para cubrir el consumo energético e impacto ambiental.

En esta investigación se ha considerado el consumo energético de la comunidad rural de San Juan de Tarucani para 200 viviendas, escuela, comisaría, puesto de salud, iglesia y alumbrado público. La Figura 30 (a) muestra la representación del perfil de carga por hora. El consumo energético medio presenta un valor de 165.80 kWh/día y la carga pico durante el año de 61.26 kW, mediante un factor de carga de 0.11. Estos datos son mostrados en la Tabla 10.

La Figura 30 (b) muestra los datos de radiación solar diaria, el mes y el índice de claridad. El gráfico de barras muestra la radiación solar diaria promedio mensual ( $\text{kWh}/\text{m}^2/\text{d}$ ) y la curva punteada representa el índice de claridad para la comunidad de Tarucani. El rango de radiación solar está entre  $5.45 \text{ kWh}/\text{m}^2/\text{día}$  a  $7.94 \text{ kWh}/\text{m}^2/\text{día}$  en los meses del año y una radiación solar media anual de  $6.71 \text{ kWh}/\text{m}^2/\text{día}$ .

En la Tabla 11 se muestran los valores mensuales de radiación solar y temperatura ambiente. Esta tabla muestra que los valores más bajos de radiación solar se encuentran entre los meses de mayo, junio y julio con valores de  $5.85$ ,  $5.45$  y  $5.73 \text{ Wh}/\text{m}^2/\text{día}$ , respectivamente. El índice de claridad medio es de  $0.70$  y la temperatura ambiente media es de  $6.11 \text{ }^\circ\text{C}$ .

**Figura 30**  
*Perfil de potencia de consumo horaria y radiación solar mensual de la comunidad San Juan de Tarucani*



**Tabla 10**  
*Energía y potencia de consumo promedio del lugar de estudio*

Métrica	Detalles
Energía promedio	165.80 kWh/d
Potencia promedio	6.91 kW
Pico máximo de potencia	61.26 kW
Factor de carga	0.11

**Tabla 11***Radiación solar (kWh/m<sup>2</sup>/d) y temperatura ambiente (°C) en la comunidad*

Mes	Radiación solar diaria horizontal	Índice de claridad	Temperatura ambiente
Enero	7.32	0.642	6.772
Febrero	6.34	0.573	7.221
Marzo	6.41	0.625	6.771
Abril	6.47	0.717	6.020
Mayo	5.85	0.745	5.105
Junio	5.45	0.75	3.722
Julio	5.73	0.76	4.518
Agosto	6.44	0.752	4.675
Setiembre	7.31	0.746	7.085
Octubre	7.94	0.737	6.913
Noviembre	7.84	0.695	7.364
Diciembre	7.43	0.649	7.298
Promedio anual	6.71	0.70	6.11

### 4.3.3 Parámetros técnicos de los componentes del sistema basados en sus tecnologías y costos

Los parámetros técnicos y económicos de los componentes del Sistema Fotovoltaico Híbrido (SFH) consideran parámetros de entrada para la simulación y optimización del sistema. En la Tabla 12 se describen los parámetros técnico-económicos y los componentes del SFH (solar fotovoltaica, batería, regulador de carga, inversor y generador diésel).

Los parámetros técnico-económicos describen:

- 1) El tiempo de vida operacional y parámetros técnicos de cada componente son obtenidos desde la ficha técnica de cada componente.
- 2) El tamaño considerado del dimensionamiento de cada componente para encontrar la óptima capacidad es definido a través de un rango de potencial total mínimo y máximo de cada componente del sistema ( $P_k = N_k P_{nom,k}$ ). La cantidad de cada componente  $N_k$  está sujeta a la restricción de mínimo y máximo ( $N_{PV,min} \leq N_{PV} \leq N_{PV,max}$ ).
- 3) La información de los costos de cada componente es: 450W de módulo fotovoltaico (\$251.49), 150 Ah de batería (\$351.66) con una capacidad de energía almacenada de 7 200 Wh (7.2 kWh) para 48 V del sistema, 4 kW de inversor (\$1 530.51), 3.44 kW de controlador de carga (\$685.89) y 50 kW de generador diésel (\$13 000).
- 4) El costo capital por kilowatt (\$/kW) para cada componente está definido por la división entre el costo unitario y la potencia del componente. El costo de operación unitario por

año (\$/año) para cada componente  $C_{O\&M,k}$  es estimado por 1.5% del costo capital  $C_{capital,k}$ . También es denotado por costo de operación unitario por kilowatts y año (\$/kw-año) que resulta de la división del  $C_{O\&M,k}$  por la capacidad unitaria del componente en kilowatts. Por otro lado, el costo de reemplazo  $C_{reemp,k}$  se ha considerado igual al costo capital  $C_{capital,k}$  de cada componente.

**Tabla 12**  
*Parámetros técnico-económicos y los componentes del SFH*

<b>Parámetros de entrada técnico y económico</b>	<b>Fotovoltaico</b>	<b>Controlador de carga</b>	<b>Batería</b>	<b>Inversor</b>	<b>Generador diésel</b>
<b>Tipo de tecnología y fabricante</b>	Monocrystalline-PERC (JA SOLAR)	MPPT Blue Solar (VICTRON ENERGY)	Deep Cycle Gel (RITAR)	Phoenix (VICTRON ENERGY)	GP-50-UCI 224D (STAMFORD)
<b>Tiempo de vida operacional</b>	25 años	10 años	15 años	10 años	20 000 hrs
<b>Espacio de búsqueda</b>	0-70kW	0-90kW	0-400bty	0-300kW	0-50Kw
<b>Costo capital</b>	\$558.87/kW	\$199.39/kW	\$351/bty.	\$382.625/kW	\$260/kW
<b>Costo de reemplazo</b>	\$0/kW	\$199.39/kW	\$351/bty.	\$382.625/kW	\$0/kW
<b>Costo O&amp;M</b>	\$3.77/año. (\$8.377/kW-año)	\$10.28/año.	\$5.27/año.	\$22.96/año.	\$0.79/año.
<b>Parámetros técnicos adicionales</b>	Pot. Max nominal: 450W	Pot. Max nominal: 3.44Kw	Voltaje nominal: 12V	Entrada (DC)	Porcentaje de carga Min.: 25%
	Voltaje de Operación óptima (Vmp): 41.52V	Voltaje de la batería: 48V	Capacidad nominal: 150Ah	Potencia DC usable: 4kW	Recurso del combustible: Diesel
	Corriente de Operación Óptima (Isc): 11.36A	Corriente de carga nominal: 60A	Corriente de carga Max.: 30A	Voltaje DC Max.: 38-66V	Precio del combustible: \$1.347/L
	Coefficiente de temperatura sobre la potencia: -0.350%/°C	Eficiencia Max.: 98%	Estado inicial de carga: 100%	Salida (AC)	Consumo promedio del combustible: 13L/hrs
	Temperatura Operacional Nominal de la celda (TONC): 45°C	Voltaje de circuito Abierto Max.: 150V	Estado de carga Min.: 20%	Potencia nominal AC: 4kW	Potencia nominal max.: 50kW
	Eficiencia a condiciones estándar: 20.2%		Eficiencia energética: 1044.29kWh	Voltaje AC: 220V	Coefficiente de intercepción: 0.0220
			Tamaño de String: 4 (48V)	Eficiencia (95%)	L/hrs/kWrated Pendiente: 0.2377 L/hrs/kWoutput

#### 4.3.4 Optimización del sistema de generación

Las simulaciones y optimizaciones llevadas a cabo por el programa HOMER para determinar la combinación óptima del sistema fotovoltaico híbrido se limitan a los costos y parámetros técnicos que se han dado como datos de entrada. La secuencia de la simulación y optimización se basan en lo siguiente:

- 1) Para optimizar el tamaño o capacidades del sistema mediante la simulación, se considera un espacio de búsqueda para las capacidades de cada componente, que consiste en variar el número de baterías, la capacidad del generador fotovoltaico y diésel, entre otros. El espacio de búsqueda o tamaño considerado para cada componente en la simulación se muestra en la Tabla 13. Estos ajustes se dieron para encontrar los mejores resultados de optimización.
- 2) Después de simular y optimizar el sistema a través de los valores de CPN, COE y RF para la ubicación seleccionada, el optimizador encontró valores mínimos del CPN y COE. Esto indica que la configuración del sistema fotovoltaico híbrido es económicamente viable para la comunidad.

En la Tabla 13 se muestra el resumen de los resultados de la optimización y los resultados de las combinaciones de capacidades para cada componente. En el resumen de la optimización se puede observar que el algoritmo de optimización del programa encontró dos sistemas posibles: El sistema A que comprende al sistema Fotovoltaico/Controlador de carga/Inversor/Diésel/Baterías y un sistema B que está formado por Fotovoltaico/Controlador de carga/Inversor/Baterías que cumplen con el valor mínimo del CPN, las restricciones dadas en la simulación y la fiabilidad para cubrir el consumo de energía.

Los resultados obtenidos en la simulación y optimización del sistema de generación híbrido son descritos a continuación:

- 1) Los resultados de la optimización para el sistema A son: El sistema óptimo consiste de una capacidad de 40 kW para el arreglo fotovoltaico, 20 kW del controlador de carga, 50 kW de inversor, 50 kW de generador diésel y 124 de baterías. Los valores

de CPN, COE y RF fueron de \$ 304 308, \$ 0.471/kWh y 90.60% de energía renovable.

- 2) Por otro lado, el sistema B presenta una capacidad de 50 kW para el arreglo fotovoltaico, 24 kW del controlador de carga, 56.3 kW del inversor y 196 baterías. Además, los valores de CPN, COE y RF del sistema B fueron de \$314 316, \$0.487/kWh y 100.00%, respectivamente, siendo este sistema con mayor valor de RF en la optimización.

**Tabla 13**

*Resultado de la optimización del tipo categorizada y general acorde al Costo de Energía (COE) y el Costo Presente Neto (CPN) o NPC*

Resultado de la optimización																
Resultado detallado de la simulación														✓	Categorizado	General
Arquitecturas										Cost				System		
P V	DG	BAT	INV	PV PERC (kW)	MPPT (kW)	GEN50 (kW)	BAT. GEL	INV. (KW)	DISPATCH	COE (US\$ )	NPC (US\$)	Operating Cost (US\$/yr)	Initial Capital (US\$)	Ren. Frac (%)	Total Fuel (L/yr)	
✓	✓	✓	✓	40.00	22.00	50.00	124	50.00	LF	0.471	304,308	18,907	102,478	90.60	1,496	
✓		✓	✓	50.00	24.00		196	56.30	LF	0.487	314,316	17,903	123,210	100.00	0	
Resultado detallado de la simulación														Categorizado	✓	General
Arquitecturas										Cost				System		
P V	DG	BAT	INV	PV PERC (kW)	MPPT (kW)	GEN50 (kW)	BAT .GEL	INV.	DISPATCH	COE (US\$)	NPC (US\$)	Operating Cost (US\$/yr)	Initial Capital (US\$)	Ren. Frac (%)	Total Fuel (L/yr)	
✓	✓	✓	✓	40	22	50	124	50	LF	0.471	304,308	18,907	102,478	90.64	1,495.94	
✓	✓	✓	✓	40	24	50	124	50	LF	0.472	304,968	18,931	102,877	90.69	1,486.45	
✓	✓	✓	✓	50	18	50	116	50	LF	0.473	305,607	18,843	104,456	90.29	1,572.89	
✓	✓	✓	✓	40	20	50	132	50	LF	0.473	305,636	18,805	104,892	90.74	1,481.27	
✓	✓	✓	✓	40	26	50	116	50	LF	0.473	305,706	19,226	100,462	90.03	1,614.66	
✓	✓	✓	✓	40	26	50	124	50	LF	0.473	305,729	18,965	103,275	90.71	1,484.27	
✓	✓	✓	✓	50	20	50	116	50	LF	0.473	305,819	18,826	104,855	90.74	1,493.40	
✓	✓	✓	✓	40	22	50	124	51.5625	LF	0.473	305,849	18,995	103,076	93.56	1,040.33	
✓	✓	✓	✓	40	20	50	124	51.5625	LF	0.474	305,959	19,043	102,677	93.21	1,107.58	
✓	✓	✓	✓	40	18	50	132	50	LF	0.474	305,981	18,875	104,494	90.03	1,613.61	
✓	✓	✓	✓	40	22	50	132	50	LF	0.474	306,049	18,806	105,291	90.88	1,454.82	
✓	✓	✓	✓	50	18	50	124	50	LF	0.474	306,079	18,624	107,269	90.92	1,456.20	

### 4.3.5 Capacidades y potencias óptimas de los componentes del sistema

La Tabla 14 detalla los resultados de la optimización para los sistemas A y B en donde se observa las capacidades óptimas de sus componentes, sus potencias mínimas y máximas y media de la potencia de salida a lo largo del año, las cuales fueron obtenidas del modelamiento de las componentes. Asimismo, se observa el tiempo de vida operacional obtenida de su ficha técnica y la capacidad de las baterías, su hora de autonomía y el tiempo de vida esperada.

Los resultados de la simulación y optimización obtenidos por el programa HOMER son las siguientes:

- 1) En el sistema A, el arreglo fotovoltaico a lo largo del año tiene una potencia de salida media de 8 kW, una energía media de 192 kWh/d en el día y un factor de capacidad de 20%, obtenida de dividir entre la potencia de salida media y la capacidad nominal (rated).
- 2) El tiempo de vida operacional es de 25 años, el cual fue asignado en la simulación por la ficha técnica, y el controlador de carga e inversor tienen una potencia máxima de 22 kW y 47.6 kW, respectivamente, y un tiempo de vida operacional de 10 años cada una.
- 3) El generador diésel tiene una potencia nominal (rated) de 50 kW, una potencia máxima de salida de 47.6 kW y una vida operacional de 20,000 hr que equivale a 2.28 años. El banco de baterías tiene una capacidad de 223 kWh que equivale a 124 baterías, una capacidad usable o descargable de 179 kWh debido a la profundidad de descarga (DOD=80%), una autonomía de 25.8 h equivalente a 1.075 días y un tiempo de vida esperado de 2.77 años.
- 4) Por otro lado, el sistema B presenta las siguientes características: El arreglo fotovoltaico tiene una potencia de salida media de 9.15kW, una energía media de 219.6 kWh/d en el día, un factor de capacidad de 18.3% y una capacidad nominal (rated) y un tiempo de vida operacional de 25 años.
- 5) El controlador de carga e inversor tienen una potencia máxima de 24 kW y 56.3 kW, respectivamente, y un tiempo de vida operacional de 10 años cada una. El banco de baterías tiene una capacidad de 353 kWh que equivale a 196 baterías, una

capacidad usable de 282 kWh, una autonomía de 40.9 h equivalente a 1.7 días y un tiempo de vida esperado de 3.96 años.

**Tabla 14**

*Resultados de la optimización del sistema mediante el programa HOMER*

<b>Sistema A: PV/Diésel/Baterías/Inversor/Controlador de carga</b>				
<b>Componentes</b>	<b>Capacidad nominal (kW)</b>	<b>Potencia mínima y máxima (kW)</b>	<b>Media de la potencia de salida (kW)</b>	<b>Tiempo de vida operación (año)</b>
<b>PV</b>	40	0 – 40	8	25
<b>Controlador de carga</b>	22	0 – 22	22	10
<b>Diesel</b>	50	12.5 – 50	41.6	2.28
<b>Inversor</b>	50	0 – 47.6	6.26	10
	Capacidad Nominal (kWh)	Usable nominal capacity (kWh)	Autonomía (hr)	Tiempo de vida esperada (año)
<b>Baterías</b>	223	179	25.8	2.77
<b>Sistema B: PV/Baterías/Inversor/Controlador de carga</b>				
<b>Componentes</b>	<b>Capacidad nominal (kW)</b>	<b>Potencia mínima y máxima (kW)</b>	<b>Media de la potencia de salida (kW)</b>	<b>Tiempo de vida operación (año)</b>
<b>PV</b>	50	0 – 50	9.15	25
<b>Controlador de carga</b>	24	0 – 24	24	10
<b>Inversor</b>	56.3	0 – 56.3	6.91	10
	Capacidad nominal (kWh)	Capacidad nominal usable (kWh)	Autonomía (hr)	Tiempo de vida esperada (año)
<b>Baterías</b>	353	282	40.9	3.96

En la Figura 31 y 32 se muestran las potencias de salida de cada componente y la confiabilidad del SFH obtenidas del modelamiento y simulación. En estas gráficas podemos observar lo siguiente:

- 1) En la Figura 31 se observa la operación dinámica del sistema A. En estos resultados se pueden identificar:
  - La salida de potencia del controlador de carga toma un valor máximo de 22 kW, permitiendo así controlar la potencia de salida del fotovoltaico y proteger al banco de baterías mediante la carga.
  - La potencia de descarga de las baterías es variante debido a que la potencia del arreglo fotovoltaico no es suficiente para cubrir la potencia de consumo.

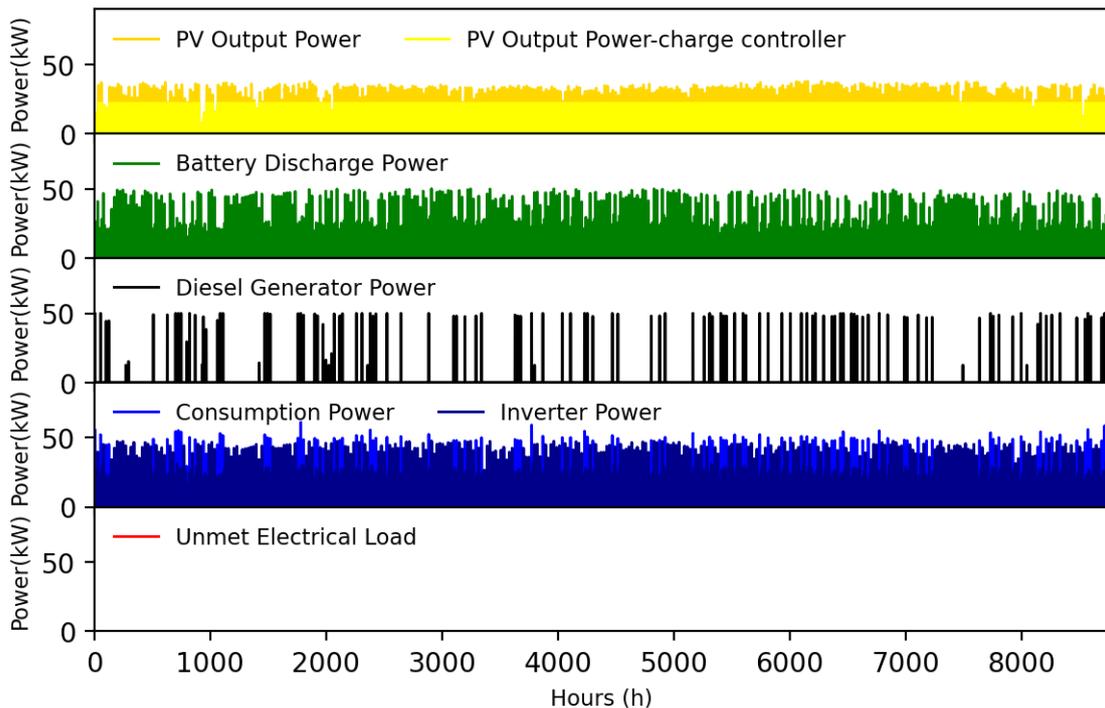
- La potencia de consumo no es cubierta por la potencia de salida del inversor por lo que se necesita el apoyo del generador diésel con una potencia de salida máxima de 50 kW.
- La carga insatisfecha tiene valores de ceros a lo largo del año lo que representa la confiabilidad del sistema.

2) En la Figura 32 se muestra la operación dinámica del sistema B, en el cual se puede identificar lo siguiente:

- Se observa una mayor potencia de salida del controlador con un valor de 24 kW debido al aumento en capacidad del arreglo fotovoltaico. Además, se muestra que este sistema tiene máximos valores de potencia del banco de baterías a lo largo del año.
- Se observa 4 picos en la carga insatisfecha, el cual indica que el sistema presenta déficit para ciertos días.

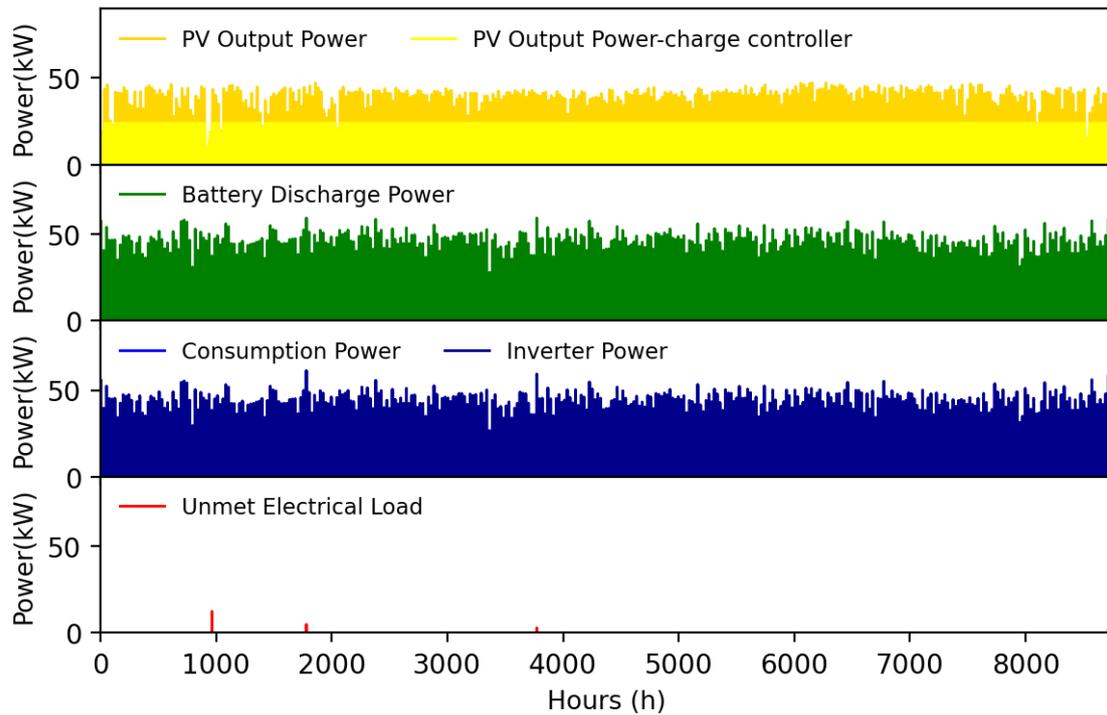
**Figura 31**

*Potencia de salida de los componentes del Sistema A durante el año*



**Figura 32**

Potencia de salida de los componentes del Sistema B a lo largo del año



#### 4.3.6 Análisis tecno-económico y producción eléctrica del sistema

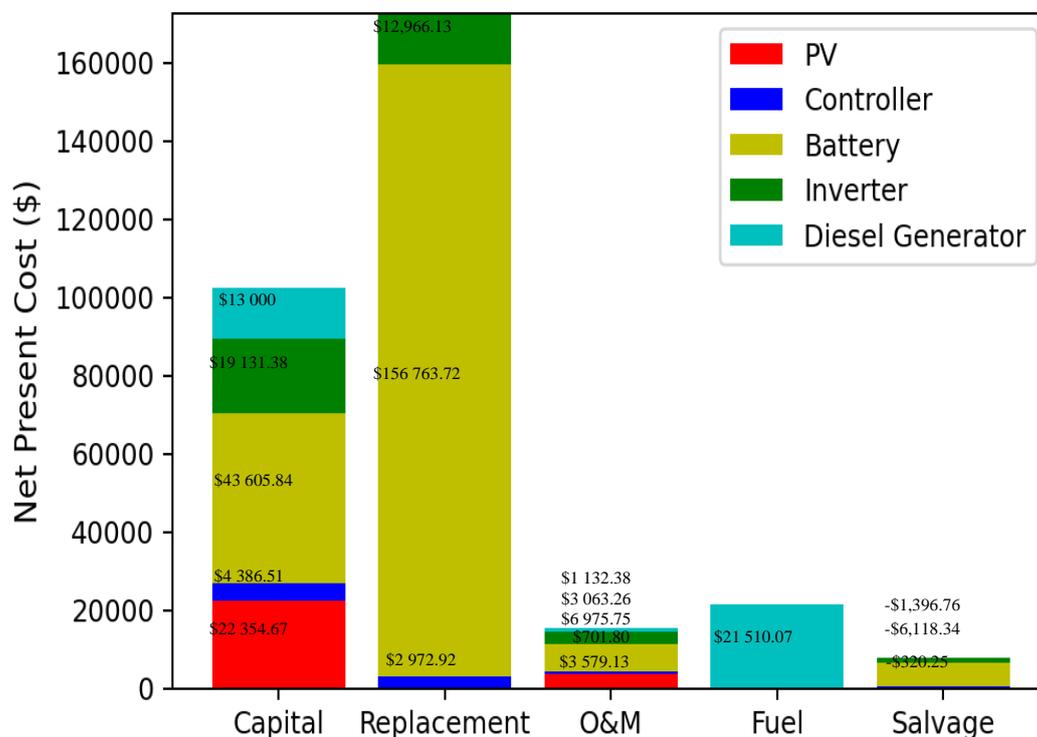
El análisis tecno-económico fue resultado de la simulación bajo los parámetros de tasa de descuento nominal ( $i$ ) con valor de 8% y un proyecto de vida ( $T_p$ ) de 25 años. Asimismo, el tiempo de vida operacional y tiempo de vida esperado fueron considerados para calcular el costo total capital ( $C^{total}_{capital,k}$ ), el costo total de O&M ( $C^{total}_{O\&M,k}$ ) y el costo total de reemplazo ( $C^{total}_{remp,k}$ ) para cada componente.

Los resultados del análisis tecno-económico para los sistemas A y B son mostrados en la Tabla 15. En resumen, el valor de CPN del sistema B está comprendido por la suma de todos los costos acumulados de cada componente, teniendo un valor de \$314 315.79.

Los costos capitales, costo de reemplazo, costo de O&M, costo de combustible y valor residual o salvage de cada una de las componentes del sistema A son mostrados en la Figura 33. En esta gráfica se observa que los costos de reemplazo de la batería para este sistema son altos con respecto a otras componentes.

**Figura 33**

Flujo de caja del Sistema A (Fotovoltaico, controlador, baterías, inversor y generador diésel) obtenido del programa HOMER



**Tabla 15**

Resumen de los costos de cada uno de los componentes para los sistemas A y B

Sistema A: PV/Diésel/Baterías/Inversor/Controlador de carga						
Componentes	Costo Capital (\$)	Costo acumulado de reemplazo (\$)	Costo acumulado total del O&M (\$)	Costo del combustible (\$)	Valor residual (\$)	Total (\$)
<b>Baterías</b>	43 605.84	156 763.72	<b>6 975.75</b>	0	-6 118.34	201 226.97
<b>Diesel</b>	<b>13 000.00</b>	0.00	1 132.38	<b>21 510.07</b>	0.00	35 642.45
<b>Inversor</b>	19 131.38	12 966.13	3 063.26	0.00	-1 396.76	33 764.00
<b>PV</b>	22 354.67	0.00	3 579.13	0.00	0.00	25 933.80
<b>Controlador</b>	<b>4 386.51</b>	<b>2 972.92</b>	<b>701.80</b>	<b>0.00</b>	<b>-320.25</b>	7740.97
<b>Sistema</b>	102 478.39	172 702.77	15 452.33	21 510.07	-7 835.36	<b>CPN</b>
						<b>304 308.20</b>

Escenario B: PV/Baterías/Inversor/Controlador de carga						
Componentes	Costo Capital (\$)	Costo acumulado de reemplazo (\$)	Costo acumulado total del O&M (\$)	Costo del combustible (\$)	Valor residual (\$)	Total (\$)
<b>Baterías</b>	68 925.36	162 370.97	11 026.19	0	-6 911.92	235 410.59
<b>Inversor</b>	21 556.08	14 609.45	3 451.50	0	-1 573.79	38 043.24
<b>PV</b>	27 943.33	0	4 473.92	0	0	32 417.25
<b>Controlador</b>	4 785.28	3 243.18	765.60	0	-349.37	8 444.70
<b>Sistema</b>	123 210.05	180 223.61	19 717.21	0	-8 835.08	<b>CPN</b>
						<b>314 315.79</b>

La Tabla 16 muestra la producción eléctrica anual de la configuración óptima del SFH, con el sistema A y B. Los resultados del análisis eléctrico para los sistemas A y B tienen las siguientes características:

- El sistema A indica que el generador fotovoltaico y el generador diésel tienen 70.053 kWh/año y 5.663 kWh/año, que representan el 92,5% y el 7,48%, respectivamente. Además, el sistema da como resultado una producción eléctrica total de 75.716 kWh/año (100%), con un total de 24.222 kWh/año (32%) de exceso de electricidad y cero cargas eléctricas insatisfechas y cero escaseces de capacidad. Esto indica que la energía generada por el sistema A cubre completamente el consumo de energía.
- Por otro lado, el sistema B presenta mayor fracción renovable debido al fotovoltaico, obteniendo una producción de electricidad de 80.140 kWh/año y una electricidad insatisfecha de 25.4 kWh/año lo que indica que el sistema tiene déficit en ciertos días del año.
- Además, se observa una reducción en la cantidad y tiempo de vida de baterías cuando está presente el generador diésel.

**Tabla 16**

*Producción de energía eléctrica y consumo anual, exceso de electricidad y fracción renovable*

<b>Escenario A: FV/Diésel/Baterías/Inversor/Controlador de carga</b>					
<b>Componente</b>	<b>Producción (kWh/yr)</b>	<b>Fracción (%)</b>	<b>Carga eléctrica</b>	<b>Consumo (kWh/yr)</b>	<b>Fracción (%)</b>
Fotovoltaico	70,053	92.5	AC Primary load	60,516	100
Generador diésel	5,663 (1,496 L)	7.48	DC Primary load	0	0
Total	75,716	100	Total	60,516	100
<b>Cantidad</b>	<b>kWh/yr</b>	<b>Fracción (%)</b>	<b>Cantidad</b>	<b>Fracción (%)</b>	
Exceso eléctrico	24,222	32	Fracción renovable	90.6	
Electricidad no satisfecha	0	0	Max. Penetración renovable	45,832	
Escasez de capacidad	0	0			
<b>N° Baterías</b>	<b>Autonomía (h)</b>	<b>Energía de descarga (kWh/ty)</b>	<b>Energía de carga (kWh/yr)</b>	<b>Vida esperada (yr)</b>	
124	25.8	46,349	47,210	2.77	
<b>Escenario B: FV /Baterías/Inversor/Controlador de carga</b>					
<b>Componente</b>	<b>Producción (kWh/yr)</b>	<b>Fracción (%)</b>	<b>Carga eléctrica</b>	<b>Consumo (kWh/yr)</b>	<b>Fracción (%)</b>
Fotovoltaico	80,130	100	AC Primary load	60,491	100
Total	80,130	100	DC Primary load	0	0
			Total	60,491	100
<b>Cantidad</b>	<b>kWh/yr</b>	<b>Fracción (%)</b>	<b>Cantidad</b>	<b>Fracción (%)</b>	
Exceso eléctrico	38,886	48.5	Fracción renovable	100	
Electricidad no satisfecha	25.4	0.0420	Max. Penetración renovable	49,999	
Escasez de capacidad	72.1	0.119			
<b>N° Baterías</b>	<b>Autonomía (h)</b>	<b>Energía de descarga (kWh/ty)</b>	<b>Energía de carga (kWh/yr)</b>	<b>Vida esperada (yr)</b>	
196	40.9	52,139	51,168	3.96	

#### 4.3.5 Operación e impacto medio ambiental de los sistemas

##### 4.3.5.1 Operación dinámica de los componentes del sistema

La Figura 34 muestra la óptima operación de SFH para cubrir el consumo de energía del sitio. La potencia óptima suministrada del Sistema A (Fotovoltaico/Diésel/Baterías/Inversor/Controlador de carga) y sistema B (Fotovoltaico/Baterías/Inversor/Controlador de carga) a la demanda energética depende de las capacidades de cada componente.

La operatividad de ambos sistemas es mostrada en la Figura 34 (a) y Figura 34 (b) y descrita de la siguiente manera:

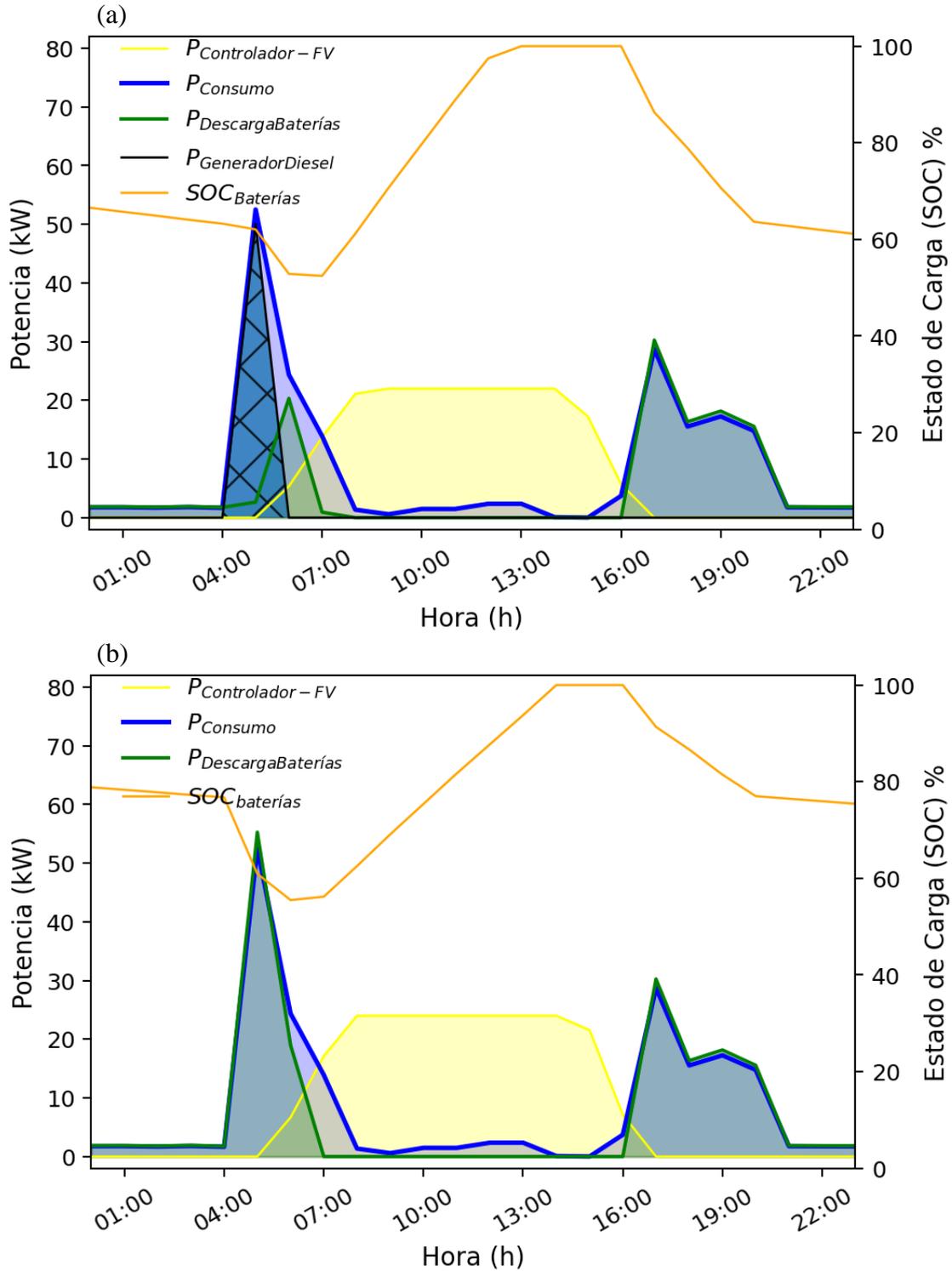
1) Operación de los componentes del sistema A:

- En el rango de 4:00 a 9:00 horas, el consumo de energía es cubierto en gran parte por el generador diésel y la otra parte por la descarga de energía de las baterías junto con la potencia de salida fotovoltaica.
- Horas más tarde, en el rango de 9:00 a 16:00 horas, se produce un exceso de energía procedente de la generación fotovoltaica por lo que esta energía se almacena en las baterías, generando un aumento de hasta el 100% del estado de carga (SOC).
- Entre las 17:00 y las 22:00 horas, el consumo de energía vuelve a aumentar y se cubre en gran medida con la potencia de descarga de la batería, alcanzando un valor del 62% del SOC.

2) Operación de los componentes del sistema B:

- Entre las horas de 9:00 a 22 horas, la operatividad del sistema B tiene similitud al sistema A, excepto del estado de descarga y carga de las baterías que tiende a cargarse lentamente.
- En el rango de 4:00 a 9:00 horas, se observa que gran parte del consumo eléctrico es cubierto por la descarga de la batería y la otra parte por el sistema fotovoltaico.

**Figura 34**  
 Flujo de potencia del SHF para cubrir la potencia de consumo



### 4.3.5.2 Impacto medio ambiental del sistema

La Tabla 17 muestra la comparación de los dos sistemas con respecto a sus emisiones. El sistema A presenta mayores emisiones de gases de efecto invernadero con un total de 3 974.217 kg/año.

Por otro lado, el sistema B (fotovoltaico/batería) presenta cero emisiones de gases contaminantes debido a que el sistema funciona con energía 100% renovable. En cuanto a la fracción renovable para los dos sistemas, el sistema A presenta una fracción del 90.6% menor que el segundo que presenta el 100%. Por lo tanto, estos sistemas son considerados los más ecológicos.

**Tabla 17**  
*Emisión de gases contaminantes para ambos sistemas A y B*

Contaminantes (kg/yr)	Emisión de gases contaminantes	
	Escenario A: PV/Diésel/Baterías/Inversor /Controlador de carga	Escenario B: PV/Baterías/Inversor/ Controlador de carga
Emisión de cantidad de dióxido de carbono (CO <sub>2</sub> )	3 916	0
Emisión de cantidad de monóxido de carbono (CO)	24.4	0
Cantidad de emisiones de hidrocarburos no quemados	1.08	0
Cantidad de emisión de partículas	0.147	0
Cantidad de emisiones de dióxido de azufre (SO <sub>2</sub> )	9.59	0
Cantidad de emisiones de óxidos de nitrógeno (NO <sub>2</sub> )	23.0	0
Total	3 974.217	0

## 4.4 Optimización del sistema híbrido fotovoltaico

### 4.4.1. Área de estudio y diseño del sistema

La simulación del sistema de generación híbrido (Fotovoltaico, controlador, inversor y baterías) para la comunidad de Tarucani ha sido desarrollado en el lenguaje de programación Python. En el cual, se ha considerado las series temporales de la temperatura ambiente, radiación solar y potencia de consumo del lugar para cubrir su consumo de energía. Así mismo, se ha diseñado el sistema de generación híbrido basado en las componentes y sus características técnicas y económicas.

#### 4.4.3 Diseño del algoritmo de optimización y gestión energética del sistema fotovoltaico híbrido

La optimización tuvo como objetivo encontrar valores óptimos de la cantidad de componentes del sistema, minimizando el Costo Presente Neto (*en inglés Net Present Cost-NPC*) de la ecuación (69) y el costo eléctrico nivelado (*en inglés Levelized Cost of Energy-COE*) de la ecuación (105). Estas funciones objetivas tuvieron como restricciones a las variables límites, denotadas por la cantidades mínima y máxima de  $N_{PV}$ ,  $N_{Bat}$ ,  $N_{Contr}$ ,  $N_{Inv}$ , y las restricciones de igualdad y desigualdad descritas en la ecuación (70) hasta la ecuación (85).

En la Tabla 18 se presentan los costos y parámetros técnicos de las componentes del sistema, parámetros económicos y configuración del optimizador: El tiempo de vida del proyecto fue de 20 años y el tiempo de vida del controlador e inversor fue de 10 años. Así mismo, la tasa de interés es del 8% y las operaciones y mantenimiento (O&M) son anuales.

Los parámetros y configuraciones del optimizador, Evolución Diferencial, presentan un número de población de 20 y un número de generaciones de 50 con un factor de escala de 0.85. Así mismo, se muestra los límites mínimo y máximo de cada una de las componentes que son usado como restricción del optimizador.

**Tabla 18***Parámetros tecno-económicos y parámetros del optimizador Evolución Diferencial*

<b>Costos de componentes</b>	<b>Fotovoltaico</b>	<b>Inversor</b>	<b>Controlador</b>	<b>Costos de componente</b>	<b>batería</b>
Potencia (W/unidad)	450	4000	2880	Capacidad (Wh/unidad)	1800
Costo capital (\$/kW)	558.87	382.63	29.57	Costo Capital (\$/kWh)	195.37
Costo O&M (\$/kWp-año)	8.38305	5.73945	0.44355	Costo O&M (\$/kWh-año)	2.93055
Costo Reemplazo (\$/kW)	558.87	382.63	29.57	Costo Reemplazo (\$/kwh)	195.37
Tiempo de vida (año)	25	10	10	Vida esperada	*
<b>Parámetros económicos</b>	<b>Tiempo de vida del proyecto (año)</b>	<b>Tiempo de O&amp;M (año)</b>		<b>Tasa de interés (%)</b>	
	20	1		8	
<b>parámetros técnicos de los componentes</b>	<b>parámetros fotovoltaicos</b>			<b>parámetros de la batería</b>	
NOCT (°C)	43.6	Ki (A/°C)	0.0005	Eficiencia (%)	0.9
Radiación Standard (W/m <sup>2</sup> )	1000	Kv (V/°C)	-0.0028	DOD(%)	0.8
Coefficiente gamma (%/°C)	-0.41	Isc(A)	10.39	Capacidad(Ah)	150
Potencia Fotovoltaico (W/unidad)	450	Tstc (°C)	25	n_rountrip(%)	0.98
GNOCT	800	Voc (V)	50	Ncycle	740
FF (%)	0.865	Vmpp (V)	41.6	Voltaje (v)	12
TaNoct (°c)	20	Imp (V)	10.812		
<b>Parámetro del inversor y controlador</b>			<b>Sistema</b>		
Eficiencia	0.9	Voltaje	48	Factor de seguridad	1.1
<b>Optimizador DE</b>					
Número de población	20	Número máximo de interacciones (Generaciones)	50	Factor de escala DE	0.85
Probabilidad de cruzamiento	0.7	Probabilidad de mutación	0.5		
<b>Espacio de búsqueda</b>					
<b>Número de fotovoltaico</b>	<b>Número de baterías</b>	<b>Número de controladores</b>	<b>Número de inversores</b>	<b>Número de variables o dimensiones</b>	<b>Penalización</b>
[50,300]	[50,300]	[7,47]	[6,38]	4	5.00E+05

Nota: La vida esperada de la batería \* se calcula al definir una cierta cantidad de baterías de prueba en la simulación.

Los resultados de la optimización y la simulación de la operatividad del sistema (Fotovoltaico, baterías, controlador e inversor) para el caso de estudio son mostrados y descritos en las siguientes figuras:

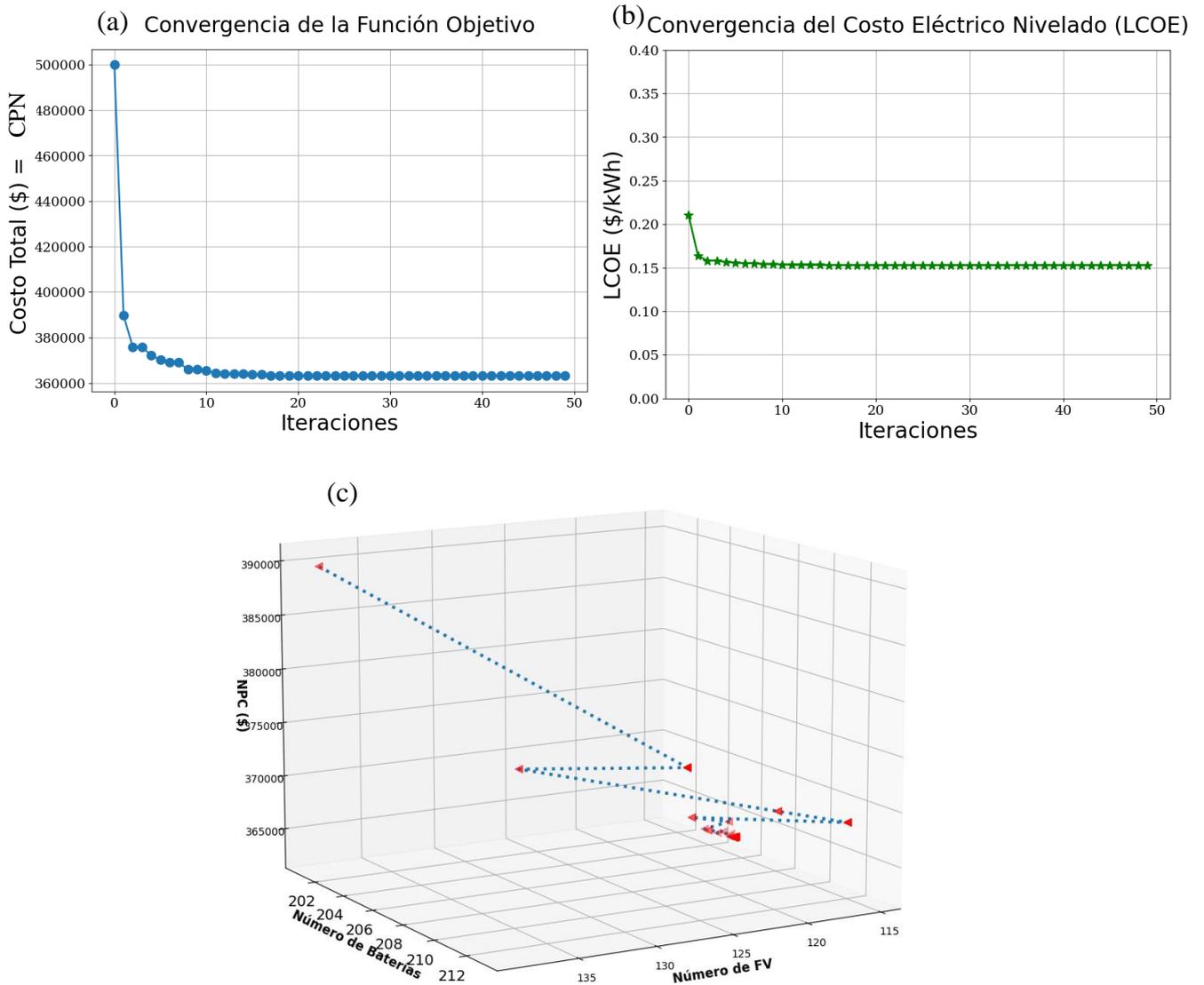
- Caso de estudio para la comunidad de Tarucani:

En la Figura 35 se muestran los valores óptimos de CPN, COE y la cantidad óptima de módulos fotovoltaicos, baterías, inversores y controladores. En la Figura 35 (a), el valor de CPN va convergiendo hasta encontrar un valor óptimo a medida que aumentan las iteraciones o generaciones (Máxima Generación = 50). El valor óptimo del CPN encontrado fue de \$ 363 191.406364, bajo los límites mínimo y máximo de la cantidad de componentes ( $50 \leq N_{PV} \leq 300$ ,  $50 \leq N_{Bat} \leq 300$ ,  $7 \leq N_{Contr} \leq 47$ ,  $6 \leq N_{Inv} \leq 38$ ).

Por otro lado, el valor óptimo de COE tuvo un valor de \$ 0.15273467/kWh en la última iteración de la optimización, tal como se muestra en la Figura 35 (b). Así mismo, en la Figura 35 (c) se muestra como el valor CPN va disminuyendo acorde a cada cantidad de módulos de fotovoltaico y baterías, encontrando valores óptimos de 115 y 206, respectivamente. Además, en la optimización se determinaron que los números de inversores y controladores óptimos tuvieron valores de 18 y 14.

**Figura 35**

Convergencia del CPN, COE y el número de baterías y fotovoltaicos

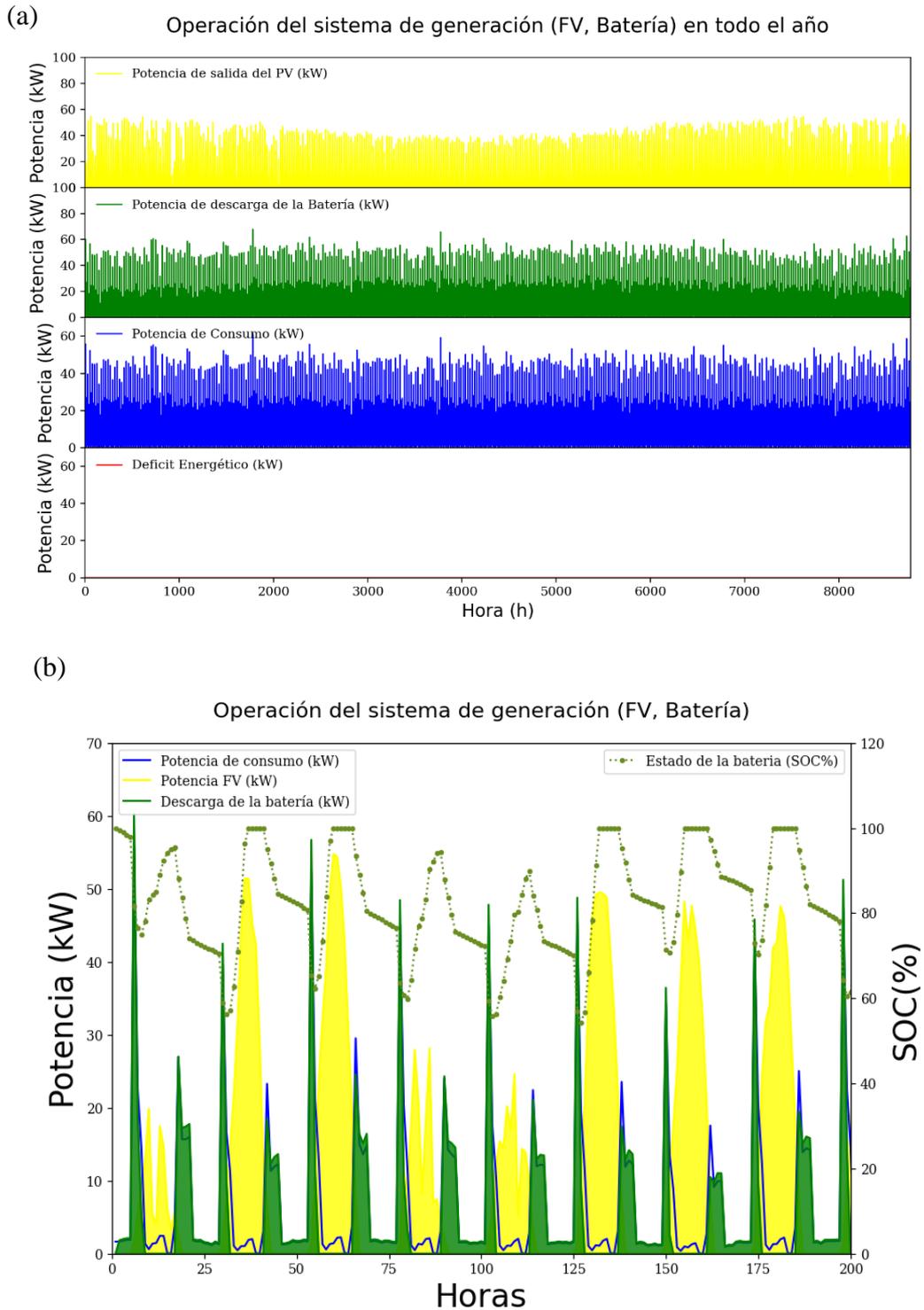


En la Figura 36 (a) presenta la operación del sistema fotovoltaico y batería a lo largo del año. La potencia de salida del fotovoltaico, la descarga de las baterías y el consumo energético en función a sus potencias son visualizados en esta figura. Además, se puede notar que el sistema no presenta déficit incluso cuando la radiación es mínima.

Por otro lado, en la Figura 36 (b) se muestra los primeros ocho días de operación del sistema. La radiación solar en el primer, cuarto y quinto día tienen valores bajos debido a nubes, ocasionando una baja generación fotovoltaica. Por lo que las baterías descargan su energía para cubrir gran parte del consumo eléctrico del lugar, generando

una gran disminución del estado de carga (SOC). Además, se muestra que el SOC alcanza valores altos cuando la radiación solar llega a sus máximos.

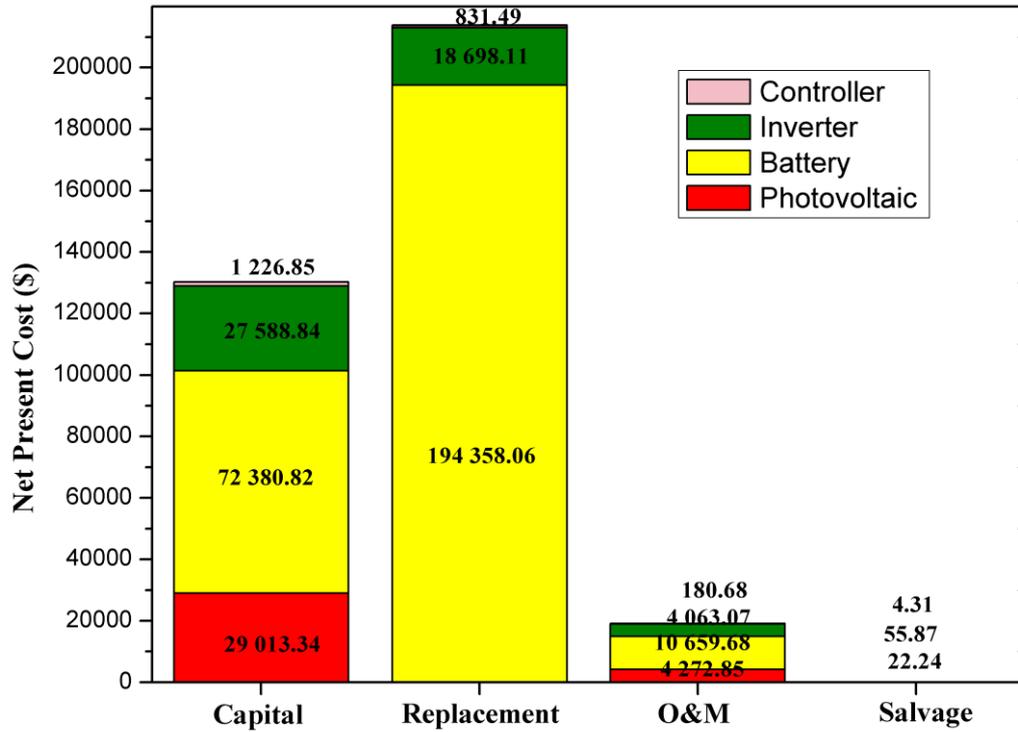
**Figura 36**  
*Operación horaria del sistema de generación*



En la Figura 37 y Tabla 19 se muestran los costos acumulados del sistema (Fotovoltaico y Batería) durante su vida útil. El costo capital, reemplazo, O&M y residual o salvage tienen valores de \$ 130 209.877, \$ 213 887.667, \$ 19 176.296, -\$ 82.434, respectivamente. La suma total acumulado de los costos incurridos por el sistema tiene un valor de \$ 363 191.406, el cual representa el Costo Presente Neto (CPN)

**Figura 37**

*Costo Capital, reemplazo, O6M y salvage obtenidos de la simulación y operación del sistema*



**Tabla 19**

*Costos de cada uno de los componentes para el cálculo del CPN*

Componentes	Costo Capital (\$)	Costo acumulado de reemplazo (\$)	Costo acumulado total del O&M (\$)	Valor residual (\$)	Total (\$)
<b>Fotovoltaico</b>	29 013.345	0.000	4 272.859	0.000	33 286.205
<b>Baterías</b>	72 380.829	194 358.065	10 659.685	-22.245	277 376.334
<b>Inversor</b>	27 588.845	18 698.111	4 063.070	-55.871	50 294.156
<b>Controlador</b>	1 226.857	831.492	180.682	-4.318	2 234.713
<b>Sistema</b>	130 209.877	213 887.667	19 176.296	-82.434	<b>CPN</b> <b>363 191.406</b>

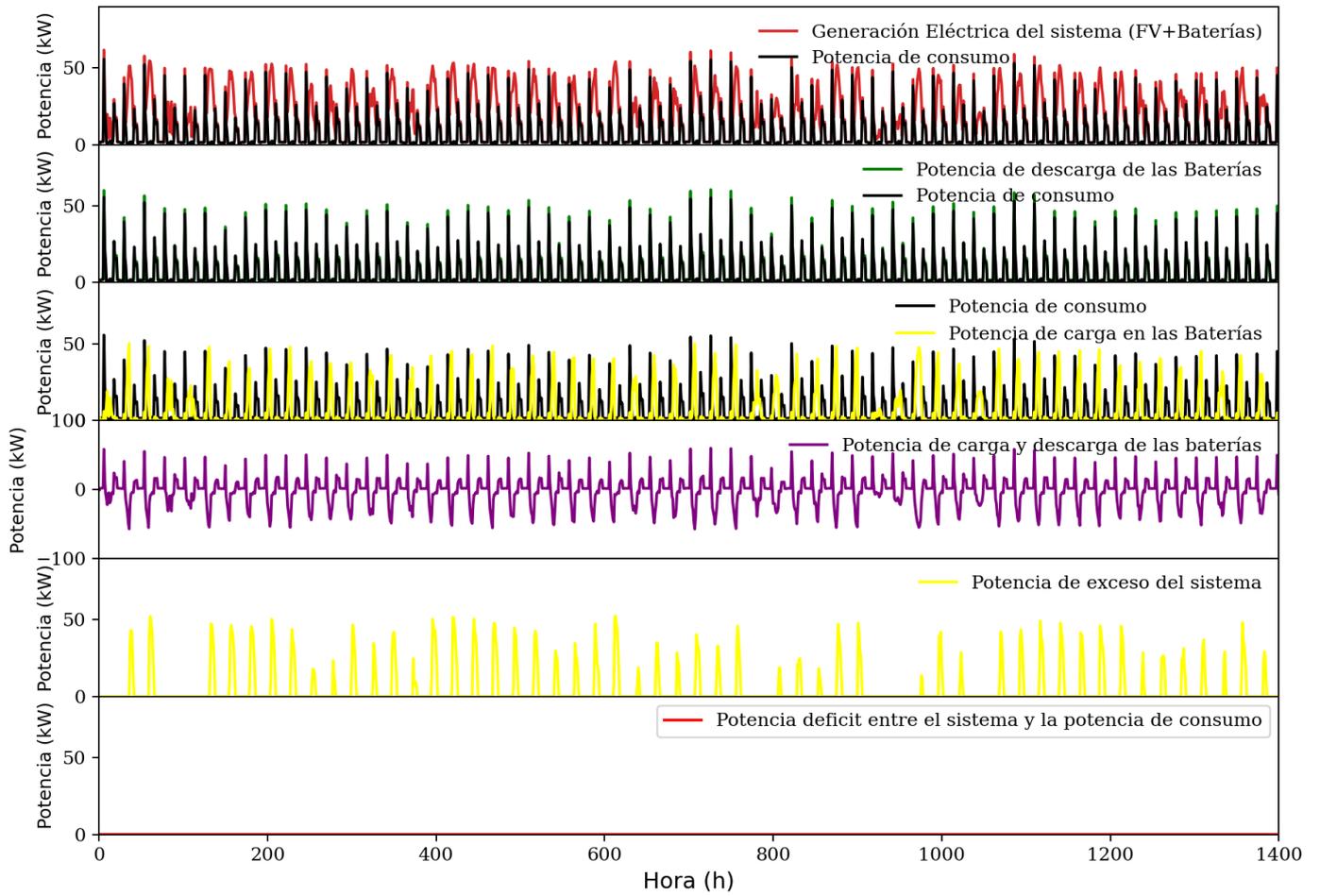
En la Figura 38 presenta la potencia de generación eléctrica del sistema (Fotovoltaico y Baterías), la potencia de consumo, la potencia de descarga y carga de las baterías, la potencia de exceso y la energía déficit del sistema. En esta figura se observa las siguientes características:

- 1) La potencia de generación eléctrica del sistema está cubriendo la potencia de consumo a lo largo del año.
- 2) La potencia de descarga y carga de la batería es continua en cada hora del año. Existe un aumento de la carga de la batería cuando hay exceso de la generación fotovoltaica y una disminución cuando la generación fotovoltaica no cubre la potencia de consumo.
- 3) La confiabilidad del sistema se muestra en la potencia déficit. En ella se observa que el sistema está cubriendo por completo la potencia de consumo a lo largo del año.

Con los valores óptimos y la operatividad del sistema se pudo determinar los valores de energía a lo largo del año, teniendo los siguientes resultados: La energía del consumo tiene un valor de 60 516.27 kWh/año y el sistema fotovoltaico tiene una generación de 107 521.28 kWh/año. El sistema tiene una energía de exceso de 41 705.11 kWh/año, una energía de descarga y carga de las baterías de 59.0344 kWh/año y 53,734.502 kWh/año, respectivamente. Así mismo, la vida esperada de la batería presenta un tiempo esperado de vida de 3 años, una autonomía de 32.17 horas y una capacidad nominal de 370.4807 kWh.

**Figura 38**

*Potencias de descarga y carga de las baterías, exceso y déficit del sistema*



## DISCUSIONES

1) En esta investigación ha sido posible implementar, evaluar y comparar modelos predictivos con la finalidad de predecir el comportamiento fluctuante de la temperatura ambiente, radiación solar y consumo eléctrico de un lugar en estudio. A través de este estudio, indicamos lo siguiente:

- Al analizar cada uno de los modelos propuestos (SARIMA, ANN y LSTM) se ha identificado que el modelo SARIMA tiene más procesos para la identificación de sus parámetros que los modelos ANN y LSTM. Estos procesos consideran un análisis de Dickey-Fuller y los análisis de prueba de Auto correlación (ACF) y parcial Autocorrelación (PACF) para la identificación de sus parámetros y estacionalidad de la información. Mientras que los modelos ANN y LSTM tienen un preprocesamiento y procesamiento similar y varios hiperparámetros. Es notable que la estructura interna y en sus predicciones de estos modelos presentan grandes diferencias.
- La evaluación de los distintos parámetros de los modelos SARIMA, ANN y LSTM fueron de gran importancia para determinar sus comportamientos en las predicciones para cada serie de tiempo. En la Tabla 8 y Tabla 9 se realizó un análisis comparativo de los mejores rendimientos de los modelos mencionados para cada serie de tiempo en estudio. En estos resultados se ha observado que de todas las configuraciones evaluadas del modelo SARIMA y ANN han llegado alcanzar valores de  $R^2$  aproximadamente 0.88 para la serie de tiempo de la radiación solar, mientras que de todas las configuraciones del modelo LSTM alcanzaron valores superiores a ellas, tomando un valor de  $R^2$  máximo de 0.96.
- Por otro lado, los modelos SARIMA y ANN, alcanzaron valores de  $R^2$  cercanos a 0.97 para el consumo eléctrico. Sin embargo, el modelo LSTM para estas informaciones, obtuvieron valores cercanos de  $R^2$  a 0.982. La serie de tiempo de la temperatura ambiente también fue evaluada únicamente por el LSTM, en el cual obtuvo un valor alto de  $R^2=0.99$ .

- Los modelos SARIMA y ANN presentaron predicciones cercanas sobre los valores observables de la data de prueba, sin embargo, para la radiación solar no pudieron predecir de forma exacta los días nublados o los días donde hay mayor radiación, como se muestran en la Figura 17 y 21. El modelo LSTM pudo predecir de forma exacta esos días donde no hay radiación con una buena precisión como se muestra en la Figura 24 (b). Así mismo, se ha realizado un nuevo estudio para otra localidad donde se utilizó el modelo LSTM para predecir las series temporales y evaluar sus precisiones. Estas evaluaciones resultaron con rendimientos precisos en todas las configuraciones del LSTM, como se muestran en las Figuras 27 (a), (b) y (c).
- La predicción a futuro del modelo LSTM para las series de tiempo de la radiación solar, potencia de consumo y temperatura ambiente han sido realizados para los dos casos de estudio, a través del entrenamiento de toda la data de un año. En estas predicciones se observan sus intervalos de predicción o de confianza cercanos a los valores predicho para la potencia de consumo y temperatura ambiente, sin embargo, en la radiación solar en la comunidad de Tarucani hay altos valores de sus intervalos de confianza al 50%. Por otro lado, en el segundo caso de estudio de la Biblioteca Universitaria se han mostrado sus intervalos de predicción muy cercanas a los valores predichos a futuro, indicando un bajo error predictivo para las tres series temporales, como se muestran en la Figura 28.

En base a los resultados obtenidos es notable indicar que el modelo LSTM tiene un gran potencial en el procesamiento de la información con predicción más precisas y confiables. En la Tabla 20 se muestran los resultados obtenidos del modelo LSTM para las diferentes variables de series temporales, las cuales son comparadas con resultados de otros autores.

**Tabla 20**

Resultados obtenidos de las variables temporales en estudio y comparadas con otros resultados

		Comunidad Rural (San Juan de Tarucani)				Institucional (Biblioteca Universitaria)					
Autores	Modelo	Variables	MSE	RMSE	R <sup>2</sup>	Autores	Modelo	Variables	MSE	RMSE	R <sup>2</sup>
Este trabajo	LSTM (50,50,50,50,Tanh,Adam)	Radiación solar (W/m <sup>2</sup> )	0.00539	0.07338	0.96335	Este trabajo	LSTM (50,50,50,50,Tanh,Adam)	Radiación solar (W/m <sup>2</sup> )	361.41297	19.01086	0.98520
	LSTM (50,50,50,50,Tanh,Adam)	Consumo eléctrico	1.95645	1.39873	0.98208		LSTM (866,50,50,50,Relu,Rmsprop)	Consumo eléctrico	0.37583	0.61305	0.99848
	LSTM (50,50,50,50,Tanh,Adam)	Temperatura ambiente	0.27343	0.52291	0.99008		LSTM (866,50,50,50,Tanh,Adam)	Temperatura ambiente	0.78835	0.88789	0.87841
Medina y Cárdenas (2022)	RNN (20, 16,32,64,-,-)	Radiación solar	28911.22	170.033	N.a	Gasparin <i>et al.</i> (2022)	LSTM- Multiple entrada y salida	Consumo eléctrico (kWh)	N.a	18.1	0.895
	RNN (48,16,32,64,-,-)	Temperatura ambiente	N.a	2.057	N.a		LSTM-recurrente	Consumo eléctrico (kWh)	N.a	13.8	0.936
Paoli <i>et al.</i> (2010)	ANN (8,3,1,Levenberg-Marquardt,linear)	Radiación solar (MJ/m <sup>2</sup> )	N.a	3.73	0.79	Farsi <i>et al.</i> (2021)	Red Paralela LSTM-CNN	Consumo eléctrico (GWh)	N.a	0.061	0.9118
	ANN (1-D Filter 1/NN1)	Radiación solar (W/m <sup>2</sup> )	1583.24	39.79	0.98	Rao <i>et al.</i> (2022)	ANN	Consumo eléctrico (W)	1.8151 × 10 <sup>5</sup>	426.04	0.79
Alsharif y Younes (2019)	ANFIS	Radiación solar (Wh/m <sup>2</sup> )	N.a	55.4	0.80		ANN (,Levenberg-Marquardt)	Consumo eléctrico (W)	68722	262.148	0.96858
Azimi <i>et al.</i> (2016)	Transformación basada en K-means	Radiación solar (W/m <sup>2</sup> )	N.a	54.69	N.a	Benamrou <i>et al.</i> (2020)	Extracción y Selección de características-Cross Validation-Xgbosst-LSTM (4,10,40,40,Adam, Relu)	radiación solar (MJ/m <sup>2</sup> )	0.0841	0.29	0.916

2) El diseño y la simulación de la operación horaria de un sistema fotovoltaico híbrido, conformado por baterías y generador diésel, para cubrir la demanda energética de la comunidad de Tarucani ha sido exitoso mediante el programa HOMER. El programa mediante su algoritmo de optimización encontró valores óptimos del dimensionamiento de la generación del sistema, determinando dos sistemas posibles que pueden cubrir la demanda energética del lugar:

- El sistema A, conformado por Fotovoltaico/controlador de carga/Inversor/Diésel/Baterías, tienen valores óptimos de la capacidad de 40 kW para el sistema fotovoltaico, 2.2 kW del controlador de carga y 124 baterías. Estas capacidades son suficiente para cubrir el consumo del lugar. Además, se obtuvo un Costo de Presente Neto (CPN) de \$304 308.00, que incurre el sistema a lo largo de su vida útil, y un costo de la energía COE de \$ 0.471/kWh y un porcentaje renovable de 90 %.
- Por otro lado, el segundo sistema B, que está conformado por Fotovoltaico/controlador de carga/Inversor/Baterías, tuvo valores óptimos de 50 kW para el arreglo fotovoltaico, 24 kW del controlador de carga, 56.3 kW del inversor y 196 baterías; y costos del CPN, COE y RF de \$314 316, \$0.487/kWh y 100%, respectivamente. Este sistema resulto más costoso que el sistema A. Esto puede deberse al uso de más baterías que hacen su incremento en la inversión e incluso en el costo de reemplazo durante la vida útil del sistema. Sin embargo, su fracción renovable de este sistema es del 100% que lo hacen más amigable con el medio ambiente.
- La operatividad diaria entre los sistemas A y B, es notable que el sistema B descarga mayor energía de las baterías para cubrir la demanda energética, presentando una autonomía de 40.9 h con un tiempo de vida esperado de 3.96 años. Mientras que el sistema A, la autonomía de las baterías es de 25.8h con un tiempo de vida esperado de 2.77 años. Por lo que las baterías tienen más tiempo de vida esperada en el sistema B, generando una reducción de los costos de reemplazo en las baterías. Además, el sistema A se apoya del generador diésel que produce 5 663 kWh/año, con un consumo total de 1 496 L de combustible, para cubrir la demanda energética.

- Es notable observar que el sistema B presenta 3 días donde el sistema no cubre por completo el consumo eléctrico, con una electricidad insatisfecha de 25.4 kWh/año, como se muestra en la Figura 32. Mientras que el sistema A cumple con la confiabilidad del sistema presentado un exceso eléctrico de 24 222 kWh/año. Con respecto a la emisión de gases contaminantes al ambiente, el sistema A tiene una emisión de 3 974.217 kg/año y el sistema B tiene cero emisiones.

Podemos indicar que la simulación y operatividad de un sistema fotovoltaico híbrido, mediante el programa HOMER, puede dar información relevante de la confiabilidad con respecto al consumo eléctrico del lugar, los costos totales que incurre el proyecto a lo largo de la vida útil y las emisiones de gases contaminantes. En este estudio podemos indicar que el sistema A (Fotovoltaico/controlador de carga/Inversor/Diésel/Baterías) presenta un menor costo presente neto (CPN), COE y una mejor confiabilidad del sistema con emisiones de gases contaminantes muy bajas. Así mismo, el apoyo del generador diésel al sistema permite aumentar el tiempo de vida esperado de la batería y por lo tanto reducir los costos de reemplazo generando una reducción del CPN.

- 3) La implementación de un algoritmo de optimización inteligente, llamado Evolución Diferencial, a la operatividad dinámica del sistema de generación híbrido (Fotovoltaico y Batería) para cubrir el consumo de energía de la comunidad de Tarucani ha sido lograda en el lenguaje de programación Python.

El costo presente neto CPN es la función objetivo en esta optimización, el cual está comprendido por los costos que incurren todo el sistema a lo largo de su vida útil. Estos costos fueron los costos de capital, costo de O&M, costos de reemplazo y el valor residual, denominada salvage, tanto anual y a lo largo de toda la vida útil del proyecto. Esta función objetivo está comprendida por variables de decisión ( $N_{PV}$ ,  $N_{Bat}$ ,  $N_{Contr}$ ,  $N_{Inv}$ ) están sujetas a restricciones de variables límite, de igualdad y desigualdad.

Los resultados encontrados en la simulación y optimización son las siguientes:

- La optimización del costo presente neto del sistema fotovoltaico y batería presentó valores óptimos de la cantidad de cada uno de los componentes: 115 módulos fotovoltaico, 206 baterías, 18 inversores y 14 controladores son las cantidades óptimas encontradas por el optimizador, obteniendo un mínimo valor de CPN de \$ 363 191.406364 y COE de \$ 0.15273467/kWh en la máxima iteración o generación del algoritmo (Max. Generación=50).
- La configuración del optimizador Diferencial Evolutivo tuvo las siguientes características: La población de NP=20, un factor de escala F de 0.87 y una probabilidad de crossover CR, o cruzamiento, de 0.7 con variables límites o espacio de búsqueda de ( $50 \leq N_{PV} \leq 300$ ,  $50 \leq N_{Bat} \leq 300$ ,  $7 \leq N_{Contr} \leq 47$ ,  $6 \leq N_{Inv} \leq 38$ ).
- Los resultados mostrados en las Figuras 35 (a) y (b) tiene la curva características de convergencia de los parámetros económicos CPN y COE. Así mismo, en la Figura 35 (c) se muestra como el algoritmo busca los valores óptimos de  $N_{PV}$  y  $N_{Bat}$  para valores mínimos de CPN.
- En la operatividad horaria del sistema fotovoltaico híbrido (Fotovoltaico y Batería) ha logrado la confiabilidad a lo largo del año mediante el algoritmo Evolución Diferencial, sin presentar déficit energético, como se muestra en la Figura 36 (a). Es notable en esta simulación que la generación fotovoltaica de exceso es almacenada en la batería, y cuando hay déficit en la generación, la energía almacenada es descargada cubriendo por completo el consumo eléctrico. Estos procesos de carga y descarga son observados en la curva del estado de carga (SOC) de las baterías.
- La energía anual que produce el sistema es obtenida a través de los valores óptimos de cada uno de los componentes: La generación fotovoltaica produce una energía alrededor de 107 521.28 kWh/año, las baterías tienen una energía de carga y descarga de 59.0344 kWh/año y 53 734.502

kWh/año, respectivamente. Esto indica un exceso de la energía del sistema de 41 705.11kWh/año y un déficit y probabilidad de pérdida de carga (LPP) de cero, lo que indica que el sistema cubre toda la energía de consumo del lugar (60 516.27kWh/año).

- Las baterías en este sistema presentan una capacidad nominal de 370.4807 kWh y una duración o tiempo de vida esperada de aproximadamente 3 años y una autonomía de 32.17 horas. Este tiempo de duración conlleva a que los costos de reemplazo de las baterías aumenten y por ende los costos totales aumenten a lo largo de proyecto.

En este estudio hemos podido implementar un algoritmo de optimización llamado Evolución Diferencial en la operatividad del sistema, el cual ha determinado las cantidades óptimas de los componentes. El sistema fotovoltaico apoyado con baterías ha obtenido un mínimo costo de inversión y una gran confiabilidad. Hemos identificado que el tiempo de vida esperado de la batería es de suma importancia, debido a que una reducción de este tiempo genera mayores costos de reemplazo produciendo un aumento de los costos de inversión del proyecto.

En la Tabla 21 se muestran los resultados obtenidos de la simulación y optimización del programa HOMER y la implementación propuesta. Estos resultados son comparados con resultado de otros autores.

**Tabla 21***Diferentes sistemas fotovoltaicos híbridos obtenido en esta investigación y comparación con resultados de otros autores*

<b>Autores</b>	<b>Sistema</b>	<b>Consumo de energía (Wh/año)</b>	<b>Programa simulación (optimizador)</b>	<b>CPN (\$)</b>	<b>COE (\$/kWh)</b>	<b>energía de exceso (kWh/año)</b>	<b>energía déficit (kWh/año)</b>	<b>Vida de batería (años)</b>	<b>Autonomía de la batería (horas)</b>
Este trabajo	Fotovoltaico, baterías	60,516.27	Algoritmo (Evolución Diferencial)	363,191.406364	0.15273467	41,705.11	0	3	32.17
Este trabajo	Fotovoltaico, baterías	60,516.27	HOMER	314,316.79	0.487	38,886	25.4	3.96	40.9
Este trabajo	Fotovoltaico, baterías, diésel	60,516.27	HOMER	304,308.20	0.471	24,222	0	2.77	25.8
Shezan (2019)	Fotovoltaico, baterías, diésel	5,126.1	PVSYST	157,907.62	0.885	397.7	125.3	N.a	N.a
Anand <i>et al.</i> (2022)	Fotovoltaico, baterías, biomasa	809,002.4	Algoritmo (Búsqueda exploratoria aleatoria del halcón de Harris híbrido)	781,000	0.115	N.a	0	N.a	N.a
Mahmoud <i>et al.</i> (2022)	Fotovoltaico, aerogenerador, baterías, diésel	N.a	Algoritmo (Lobo gris)	6,147,058.82	0.2098	8,925	0	N.a	N.a
Hassan (2021)	Fotovoltaico, baterías	2.485	Algoritmo (Iterativo)	6,244	0.196	1,468	N.a	N.a	22.5
Bedadi y GebreMichael (2021)	Fotovoltaico, baterías, micro hidro	66.06	HOMER	78,763	0.075	N.a	N.a	N.a	3.79

## CONCLUSIONES

En esta investigación se ha utilizado técnicas de inteligencia artificial para aplicaciones en energías renovables. El desarrollo e implementación de modelos predictivos y optimizadores inteligentes ha sido el enfoque de esta investigación y tienen como objetivo obtener la óptima operatividad del sistema de generación fotovoltaico híbrido y resolver problemas que el sistema presenta.

El diseño e implementación de un sistema fotovoltaico híbrido requiere de un estudio de la radiación solar, temperatura ambiente y el consumo eléctrico del lugar a instalar. Por lo que la evaluación de modelos predictivos, como SARIMA, ANN y LSTM, para predecir la información del lugar en estudio de forma precisa a través de las métricas de error (MSE, RMSE, MAE,  $R^2$ ) son de gran importancia ya que nos permite conocer de manera precisa sus comportamientos cambiantes. Los resultados en estos estudios muestran un notable potencial predictivo del modelo LSTM, con respecto al SARIMA y ANN, demostrando valores por encima de 0.97 con respecto a su  $R^2$  para todas las series de tiempo estudiadas. El modelo LSTM tuvo resultados de forma precisa en los días donde la radiación, temperatura ambiente y consumo de energía son altas y bajas.

Antes de la implementación del sistema de generación, el diseño y la simulación de la operación de un sistema fotovoltaico híbrido es requerido para estudiar su confiabilidad del sistema con respecto a la demanda energética del lugar en estudio. La simulación a través del programa HOMER determinó la capacidad óptima de las componentes del sistema bajo un Costo Presente Neto mínimo durante todo el proyecto.

Los resultados de la optimización realizados en el programa HOMER, encontraron dos posibles sistemas conformado por Fotovoltaico/controlador de carga/Inversor/Diésel/Baterías y Fotovoltaico/controlador de carga/Inversor/Baterías. Entre estos dos sistemas, los resultados indican que el sistema apoyado con baterías y diésel presentan un menor CPN y una mayor confiabilidad que el que es apoyado con solo baterías, teniendo el valor de cero en la energía déficit del sistema y bajos valores de gases contaminantes. También, se identificó que el apoyo del generador diésel al sistema, fotovoltaico y batería, aumenta el tiempo de vida esperada de las baterías y por ende una reducción del costo de reemplazo, generando una disminución en el CPN.

La implementación de un sistema inteligente que permita analizar y optimizar distintas variables del sistema de generación híbrido con la finalidad de hacerla confiable y a costos de inversión mínimas durante la vida del proyecto es de gran importancia. Los resultados de nuestra implementación del algoritmo de Evolución Diferencial a la operatividad del sistema, fotovoltaico y batería, permitió obtener los valores óptimos de las cantidades de componentes del sistema obteniendo un valor mínimo de CPN y la confiabilidad del sistema, fotovoltaico y baterías, con un valor de probabilidad de pérdida de cero.

El algoritmo implementado de carga y descarga de la batería, permitió identificar la autonomía y la vida esperada de las baterías, identificándolas como factores importantes que hace que el CPN del sistema aumente o disminuye en función al uso de las baterías.

## RECOMENDACIONES Y TRABAJOS FUTUROS

Esta investigación ha limitado varios aspectos en la predicción de series temporales, debido a que la información estudiada son series de tiempos univariados de un año que han ingresado a los modelos predictivos inteligentes. Por lo que, sería de gran interés y como trabajo futuro poder utilizar las series de tiempo multivariadas, de otros factores climáticos, que ingresen en forma conjunta al modelo y se entrenen. Esto podría generar un modelo inteligente más completo que analice más variables en conjunto, como la radiación solar, temperatura ambiente, el consumo eléctrico y adicionar otras variables del lugar. Esto permitiría una mayor precisión de cada una de las predicciones y, por lo tanto, una previsión exacta de sus comportamientos a futuro de estas variables que afectan a la producción fotovoltaica.

En la simulación se ha diseñado un sistema de generación híbrido (Fotovoltaico, Batería y Diésel), sin embargo, se podría realizar como trabajo futuro la combinación de otros tipos de tecnologías, como supercapacitores, almacenamiento de hidrógeno, sistema de conexión a la red y otras configuraciones, a fin de obtener un mejor diseño del sistema con un trabajo cooperativo entre tecnologías y aprovechar el recurso natural del lugar.

En la operatividad del sistema de generación fotovoltaico híbrido apoyada con baterías se ha usado el algoritmo Evolución Diferencial para encontrar los valores óptimos de las cantidades de los componentes del sistema que determinan un mínimo CPN. Por lo que se puede proponer como trabajo futuro la implementación de otro algoritmo de inteligencia artificial que no considere una sola función objetivo, sino que considere diferentes funciones a minimizar, como la probabilidad de pérdida de carga (LPP), el tiempo de vida esperada y el estado de carga (SOC) de la batería, el costo de la energía (COE), entre otras variables tecno-económicas. Así mismo, se propone desarrollar un algoritmo de operación basado en la optimización, donde se considere más sistemas de generación como fotovoltaico, baterías, diésel, supercapacitores, almacenamiento de hidrógeno, celda combustible, concentradores solares, vehículos eléctricos, entre otros.

## CONTRIBUCIONES CIENTÍFICAS Y PARTICIPACIÓN EN LA COMUNIDAD ACADÉMICA

En el desarrollo de esta investigación doctoral se ha participado en eventos científicos y realizado contribuciones científicas:

En la Jornadas Peruanas de Energía Solar-JOPES 2021, los resultados preliminares de la investigación titulada “*Predicción de la serie de tiempo de la radiación solar utilizando perceptrón multicapa*” se expusieron de manera oral. Así mismo, la investigación titulada “*Optimización de un sistema Fotovoltaico, Batería y Diésel para una comunidad rural*” fue presentada en la Jornadas Peruana de Energía Solar-JOPES 2022.

La investigación realizada en esta tesis doctoral ha presentado contribuciones científicas en el campo de las predicciones de variables climáticas y la optimización de sistemas fotovoltaico híbrido. El primer artículo titulado “*Global solar radiation time serie forecasting using different architectures of the multilayer perceptrón model*” fue aceptado en la revista IOP Publishing, a través de Peruvian Workshop on Solar Energy (JOPES 2021). Este artículo se muestra en el Anexo 7. Por otro lado, el segundo artículo titulado “*Optimization of hybrid Photovoltaic/Battery/Diesel system for electrification in a rural community*” está siendo sometida a una revista científica internacional.

## REFERENCIA BIBLIOGRÁFICA

- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Umar, A. M., Linus, O. U., Arshad, H., Kazaure, A. A., Gana, U., y Kiru, M. U. (2019). Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition. *IEEE Access*, 7(February 2017), 158820-158846. <https://doi.org/10.1109/ACCESS.2019.2945545>
- Adebiyi, A. A., Adewumi, A. O., y Ayo, C. K. (2014). Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. *Journal of Applied Mathematics*, 2014, 614342. <https://doi.org/10.1155/2014/614342>
- Agencia Andina. (2023). *Inteligencia artificial avanza a grandes pasos y estos son algunos proyectos peruanos*. <https://andina.pe/agencia/noticia-inteligencia-artificial-avanza-a-grandes-pasos-y-estos-son-algunos-proyectos-peruanos-923063.aspx%22>
- Ahmad, M. F., Isa, N. A. M., Lim, W. H., y Ang, K. M. (2022). Differential evolution: A recent review based on state-of-the-art works. *Alexandria Engineering Journal*, 61(5), 3831-3872. <https://doi.org/10.1016/j.aej.2021.09.013>
- Ahmed, N. M., Farghally, H. M., y Fahmy, F. H. (2017). Optimal sizing and economical analysis of PV-wind hybrid power system for water irrigation using genetic algorithm. *International Journal of Electrical and Computer Engineering*, 7(4), 1797-1814. <https://doi.org/10.11591/ijece.v7i4.pp1797-1814>
- Alsharif, M. H., y Younes, M. K. (2019). Evaluation and forecasting of solar radiation using time series adaptive neuro-fuzzy inference system: Seoul city as a case study. *IET Renewable Power Generation*, 13(10), 1711-1723. <https://doi.org/10.1049/iet-rpg.2018.5709>
- Anand, P., Kamboj, V. K., Alaraj, M., Rizwan, M., y Mwakitalima, I. J. (2022). Optimal Sizing of Hybrid Energy System Using Random Exploratory Search-Centred Harris Hawks Optimizer with Improved Exploitation Capability. *Mathematical Problems in Engineering*, 2022. <https://doi.org/10.1155/2022/5348017>
- Apeh, O. O., Meyer, E. L., y Overen, O. K. (2021). Modeling and experimental analysis of battery charge controllers for comparing three off-grid photovoltaic power plants. *Heliyon*, 7(11), 13. <https://doi.org/10.1016/j.heliyon.2021.e08331>
- Arif, S., Taweekun, J., Ali, H. M., Ahmed, A., y Bhutto, A. A. (2023). Building Resilient communities: Techno-economic assessment of standalone off-grid PV powered net zero energy (NZE) villages. *Heliyon*, 9(11), e21426.

- <https://doi.org/10.1016/j.heliyon.2023.e21426>
- Ashari, M., y Nayar, C. V. (1999). An optimum dispatch strategy using set points for a photovoltaic (PV)-diesel-battery hybrid power system. *Solar Energy*, 66(1), 1-9. [https://doi.org/10.1016/S0038-092X\(99\)00016-X](https://doi.org/10.1016/S0038-092X(99)00016-X)
- Aslam, M., Lee, S. J., Khang, S. H., y Hong, S. (2021). Two-Stage Attention over LSTM with Bayesian Optimization for Day-Ahead Solar Power Forecasting. *IEEE Access*, 9, 107387-107398. <https://doi.org/10.1109/ACCESS.2021.3100105>
- Asociación Peruana de Energía Renovable. (2023). *Reporte Estadístico Sector Eléctrico*. <https://www.spr.org.pe/reportes-estadisticos/reportes-estadisticos-2023/>
- Azimi, R., Ghayekhloo, M., y Ghofrani, M. (2016). A hybrid method based on a new clustering technique and multilayer perceptron neural networks for hourly solar radiation forecasting. *Energy Conversion and Management*, 118, 331-344. <https://doi.org/10.1016/j.enconman.2016.04.009>
- Bedadi, L. A., y GebreMichael, M. G. (2021). Design and optimization of off-grid hybrid renewable power plant with storage system for rural area in Rwanda. *IET Renewable Power Generation*, 15(15), 3715-3731. <https://doi.org/10.1049/rpg2.12267>
- Benamrou, B., Ouardouz, M., Allaouzi, I., y Ahmed, M. Ben. (2020). A proposed model to forecast hourly global solar irradiation based on satellite derived data, deep learning and machine learning approaches. *Journal of Ecological Engineering*, 21(4), 26-38. <https://doi.org/10.12911/22998993/119795>
- Bilal, Pant, M., Zaheer, H., Garcia-Hernandez, L., y Abraham, A. (2020). Differential Evolution: A review of more than two decades of research. *Engineering Applications of Artificial Intelligence*, 90(January), 103479. <https://doi.org/10.1016/j.engappai.2020.103479>
- BloombergNEF. (2023). *Cost of Clean Energy Technologies Drop as Expensive Debt Offset by Cooling Commodity Prices | BloombergNEF*. <https://about.bnef.com/blog/cost-of-clean-energy-technologies-drop-as-expensive-debt-offset-by-cooling-commodity-prices/>
- Bortolini, M., Gamberi, M., y Graziani, A. (2014). Technical and economic design of photovoltaic and battery energy storage system. *Energy Conversion and Management*, 86, 81-92. <https://doi.org/10.1016/j.enconman.2014.04.089>
- Bortolini, M., Gamberi, M., Graziani, A., y Pilati, F. (2015). Economic and environmental bi-objective design of an off-grid photovoltaic-battery-diesel generator hybrid energy system. *Energy Conversion and Management*, 106, 1024-1038.

<https://doi.org/10.1016/j.enconman.2015.10.051>

- Chakraborty, S., Saha, A. K., Ezugwu, A. E., Agushaka, J. O., Zitar, R. A., y Abualigah, L. (2023). Differential Evolution and Its Applications in Image Processing Problems: A Comprehensive Review. En *Archives of Computational Methods in Engineering* (Vol. 30, Número 2). Springer Netherlands. <https://doi.org/10.1007/s11831-022-09825-5>
- Charfi, S., Brahmi, N., Atieh, A., y Chaabene, M. (2018). Energy scheduling strategy for a photovoltaic/batteries bank/diesel generator power supply system for an off-grid house. *Journal of Renewable and Sustainable Energy*, 10(1). <https://doi.org/10.1063/1.5020831>
- Chatfield, C., y Yar, M. (1991). Prediction intervals for multiplicative Holt-Winters. *International Journal of Forecasting*, 7(1), 31-37. [https://doi.org/10.1016/0169-2070\(91\)90030-Y](https://doi.org/10.1016/0169-2070(91)90030-Y)
- Chi, D. (2022). Research on electricity consumption forecasting model based on wavelet transform and multi-layer LSTM model. *Energy Reports*, 8, 220-228. <https://doi.org/10.1016/j.egy.2022.01.169>
- COES. (2023). *Comité de Operación Económica del Sistema (COES)*. <https://www.coes.org.pe/Portal/>
- Colla, M., Ioannou, A., y Falcone, G. (2020). Critical review of competitiveness indicators for energy projects. *Renewable and Sustainable Energy Reviews*, 125(February), 109794. <https://doi.org/10.1016/j.rser.2020.109794>
- Comello, S. D., Glenk, G., y Reichelstein, S. (2017). Levelized Cost of Electricity Calculator: A User Guide. *Sustainable Energy Initiative*, March, 1-14. [http://stanford.edu/dept/gsb\\_circle/sustainable-energy/lcoe](http://stanford.edu/dept/gsb_circle/sustainable-energy/lcoe)
- Cruz, E. G. (2010). Diseño y simulación del comportamiento de un sistema híbrido de potencia. En *Universidad de Piura*. <https://pirhua.udep.edu.pe/handle/11042/1448>
- Cui, H., Wei, P., Mu, Y., y Peng, X. (2016). *SARIMA-Orthogonal Polynomial Curve Fitting Model for Medium-Term Load Forecasting*. 2016.
- Dalton, G. J., Lockington, D. A., y Baldock, T. E. (2009). Feasibility analysis of renewable energy supply options for a grid-connected large hotel. *Renewable Energy*, 34(4), 955-964. <https://doi.org/10.1016/j.renene.2008.08.012>
- Das, V., Karuppanan, P., Singh, A. K., y Thakur, P. (2021). Optimal Sizing and Control of Solar PV-PEMFC Hybrid Power. *International Journal of Mathematical, Engineering and Management Sciences*, 6(4), 1137-1156.

<https://doi.org/10.33889/IJMEMS.2021.6.4.068>

- Diaf, S., Diaf, D., Belhamel, M., Haddadi, M., y Louche, A. (2007). A methodology for optimal sizing of autonomous hybrid PV/wind system. *Energy Policy*, 35(11), 5708-5718. <https://doi.org/10.1016/j.enpol.2007.06.020>
- Do, Q. H., Lo, S., Chen, J., Le, C., Anh, L. H., Management, S., Do, Q. H., y Technology, T. (2020). *Forecasting Air Passenger Demand: A Comparison of LSTM and SARIMA*. <https://doi.org/10.3844/jcssp.2020.1063>
- Dos Santos, T. R., y Franco, G. C. (2019). Bootstrap for correcting the mean square error of prediction and smoothed estimates in structural models. *Brazilian Journal of Probability and Statistics*, 33(1), 139-160. <https://doi.org/10.1214/17-BJPS381>
- Durstewitz, D., Koppe, G., y Meyer-Lindenberg, A. (2019). Deep neural networks in psychiatry. *Molecular Psychiatry*, 24(11), 1583-1598. <https://doi.org/10.1038/s41380-019-0365-9>
- El Peruano. (2023). *Ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país*. Ley N° 31814. <https://busquedas.elperuano.pe/dispositivo/NL/2192926-1>
- ENEL. (2021). *Informe de Sostenibilidad Enel Americas 2021*. [https://www.enelamericas.com/content/dam/enel-americas/investor/sostenibilidad/reporte\\_de\\_sostenibilidad/2021/ISAmericas2021.pdf](https://www.enelamericas.com/content/dam/enel-americas/investor/sostenibilidad/reporte_de_sostenibilidad/2021/ISAmericas2021.pdf)
- Farsi, B., Amayri, M., Bouguila, N., y Eicker, U. (2021). On short-term load forecasting using machine learning techniques and a novel parallel deep LSTM-CNN approach. *IEEE Access*, 9, 31191-31212. <https://doi.org/10.1109/ACCESS.2021.3060290>
- Feng, C., Liu, Y., y Zhang, J. (2021). A taxonomical review on recent artificial intelligence applications to PV integration into power grids. *International Journal of Electrical Power and Energy Systems*, 132(April), 107176. <https://doi.org/10.1016/j.ijepes.2021.107176>
- Fouilloy, A., Voyant, C., Notton, G., Motte, F., Paoli, C., Nivet, M. L., Guillot, E., y Duchaud, J. L. (2018). Solar irradiation prediction with machine learning: Forecasting models selection method depending on weather variability. *Energy*, 165, 620-629. <https://doi.org/10.1016/j.energy.2018.09.116>
- Gan, L. K., Shek, J. K. H., y Mueller, M. A. (2015). Hybrid wind-photovoltaic-diesel-battery system sizing tool development using empirical approach, life-cycle cost and performance analysis: A case study in Scotland. *Energy Conversion and*

- Management*, 106, 479-494. <https://doi.org/10.1016/j.enconman.2015.09.029>
- Gasparin, A., Lukovic, S., y Alippi, C. (2022). Deep learning for time series forecasting: The electric load case. *CAAI Transactions on Intelligence Technology*, 7(1), 1-25. <https://doi.org/10.1049/cit2.12060>
- Ghenai, C., y Bettayeb, M. (2017). Optimized design and control of an off grid solar PV/hydrogen fuel cell power system for green buildings. *IOP Conference Series: Earth and Environmental Science*, 93(1). <https://doi.org/10.1088/1755-1315/93/1/012073>
- Hammad, M. A., Jereb, B., Rosi, B., y Dragan, D. (2020). Methods and Models for Electric Load Forecasting: A Comprehensive Review. *Logistics & Sustainable Transport*, 11(1), 51-76. <https://doi.org/10.2478/jlst-2020-0004>
- Hanaa Elgohari, M. A. & A. E. (2018). Application Sarima Models on Time Series to Forecast the Number of death in hospital. *International Journal of Applied Mathematics Statistical Sciences (IJAMSS)*, 7(4), 9-18. [http://www.iaset.us/view\\_archives.php](http://www.iaset.us/view_archives.php)
- Hanifi, S., Liu, X., Lin, Z., y Lotfian, S. (2020). A Critical Review of Wind Power Forecasting. *Energies*, 13(15), 1-24.
- Hassan, Q. (2021). Evaluation and optimization of off-grid and on-grid photovoltaic power system for typical household electrification. *Renewable Energy*, 164, 375-390. <https://doi.org/10.1016/j.renene.2020.09.008>
- Hatami, N., Gavet, Y., Debayle, J., Hatami, N., Gavet, Y., Debayle, J., Using, T. I., Hatami, N., Gavet, Y., y Debayle, J. (2018). *Classification of Time-Series Images Using Deep Convolutional Neural Networks To cite this version : HAL Id : hal-01743695 Classification of Time-Series Images Using Deep Convolutional Neural Networks.*
- Hidalgo, R., Amoroso, F., Urquizo, J., Villavicencio, V., Torres, M., Singh, P., y Soriano, G. (2022). Feasibility Study for Off-Grid Hybrid Power Systems Considering an Energy Efficiency Initiative for an Island in Ecuador. *Energies*, 15(5). <https://doi.org/10.3390/en15051776>
- Hong, T., Wang, P., y Willis, H. L. (2011). A naïve multiple linear regression benchmark for short term load forecasting. *IEEE Power and Energy Society General Meeting*, 2, 1-6. <https://doi.org/10.1109/PES.2011.6038881>
- Hove, T., y Tazvinga, H. (2012). A techno-economic model for optimising component sizing and energy dispatch strategy for PV-diesel-battery hybrid power systems.

- Journal of Energy in Southern Africa*, 23(4), 18-28. <https://doi.org/10.17159/2413-3051/2012/v23i4a3175>
- Hsu, D., y Kang, L. (2014). Dispatch analysis of off-grid diesel generator-battery power systems. *International Journal of Emerging Electric Power Systems*, 15(2), 161-170. <https://doi.org/10.1515/ijeeps-2013-0134>
- Huang, C. M. T., Huang, Y. C., y Huang, K. Y. (2014). A hybrid method for one-day ahead hourly forecasting of PV power output. *Proceedings of the 2014 9th IEEE Conference on Industrial Electronics and Applications, ICIEA 2014*, 5(3), 526-531. <https://doi.org/10.1109/ICIEA.2014.6931220>
- Huneke, F., Henkel, J., González, J. A. B., y Erdmann, G. (2012). Optimisation of hybrid off-grid energy systems by linear programming. *Energy, Sustainability and Society*, 2(1), 1-19. <https://doi.org/10.1186/2192-0567-2-7>
- IEA. (2024). *Energía Renovable Progreso Tracker*. <https://www.iea.org/data-and-statistics/data-tools/renewable-energy-progress-tracker>
- INGEMMET. (2019). *EVALUACIÓN DE PELIGROS VOLCÁNICOS DEL SECTOR DE SAN JUAN DE TARUCANI*. [https://repositorio.ingemmet.gob.pe/bitstream/20.500.12544/2332/1/A6932-Evaluacion\\_peligros\\_San\\_Juan\\_de\\_Tarucani-Arequipa.pdf](https://repositorio.ingemmet.gob.pe/bitstream/20.500.12544/2332/1/A6932-Evaluacion_peligros_San_Juan_de_Tarucani-Arequipa.pdf)
- IRENA. (2020a). Costos de generación de energía renovable en 2020: Resumen ejecutivo. *International Renewable Energy Agency*, 1-12.
- IRENA. (2020b). Renewable Capacity Statistics 2020. En *International Renewable Energy Agency*. [https://irena.org/-/media/Files/IRENA/Agency/Publication/2020/Mar/IRENA\\_RE\\_Capacity\\_Statistics\\_2020.pdf](https://irena.org/-/media/Files/IRENA/Agency/Publication/2020/Mar/IRENA_RE_Capacity_Statistics_2020.pdf)
- IRENA. (2022). *Renewable Capacity Statistics 2023*. [https://mc-cd8320d4-36a1-40ac-83cc-3389-cdn-endpoint.azureedge.net/-/media/Files/IRENA/Agency/Publication/2023/Mar/IRENA\\_RE\\_Capacity\\_Statistics\\_2023.pdf?rev=d2949151ee6a4625b65c82881403c2a7](https://mc-cd8320d4-36a1-40ac-83cc-3389-cdn-endpoint.azureedge.net/-/media/Files/IRENA/Agency/Publication/2023/Mar/IRENA_RE_Capacity_Statistics_2023.pdf?rev=d2949151ee6a4625b65c82881403c2a7)
- IRENA. (2023). Las renovables registran un crecimiento récord del 9,6 % pese a la crisis energética. *Internacional Renewable Energy Agency*. <https://www.irena.org/News/pressreleases/2023/Mar/-/media/BB677D59011546F49898C92D05266D59.ashx>
- Islam, M. K., Hassan, N. M. S., Rasul, M. G., Emami, K., y Chowdhury, A. A. (2023). Forecasting of Solar and Wind Resources for Power Generation. *Energies*, 16(17).

<https://doi.org/10.3390/en16176247>

- Jung, T. Y., Kim, D., Lim, S., y Moon, J. (2019). *Evaluation criteria of independent hybrid energy systems*. 1-7. <https://doi.org/10.1093/ijlct/ctz036>
- Khanarsa, P., y Sinapiromsaran, K. (2020). *Automatic SARIMA Order Identification Convolutional Neural Network*. 10(5). <https://doi.org/10.18178/ijmlc.2020.10.5.988>
- Khandelwal, I., Adhikari, R., y Verma, G. (2015). Time series forecasting using hybrid arima and ann models based on DWT Decomposition. *Procedia Computer Science*, 48(C), 173-179. <https://doi.org/10.1016/j.procs.2015.04.167>
- Kharrich, M., Mohammed, O. H., Kamel, S., Selim, A., Sultan, H. M., Akherraz, M., y Jurado, F. (2020). Development and implementation of a novel optimization algorithm for reliable and economic grid-independent hybrid power system. *Applied Sciences (Switzerland)*, 10(18). <https://doi.org/10.3390/APP10186604>
- Kim, H., Park, S., y Kim, S. (2022). Solar radiation forecasting using boosting decision tree and recurrent neural networks. *Communications for Statistical Applications and Methods*, 29(6), 709-719. <https://doi.org/10.29220/CSAM.2022.29.6.709>
- Kurukuru, S. B., y Mohamed, A. K. (2019). *Predictive Control of Energy Management System for Fuel Cell Assisted Photo Voltaic Hybrid Power System*. 698, 581-589. [https://doi.org/https://doi.org/10.1007/978-981-13-1819-1\\_24](https://doi.org/https://doi.org/10.1007/978-981-13-1819-1_24)
- Liu, Y., Esan, O. C., Pan, Z., y An, L. (2021). Machine learning for advanced energy materials. *Energy and AI*, 3, 100049. <https://doi.org/10.1016/j.egyai.2021.100049>
- Lou, Q., Lyu, Q., Na, Z., Ma, D., y Ma, X. (2020). Short-term electric power demand forecasting using a hybrid model of SARIMA and SVR. *IOP Conference Series: Earth and Environmental Science*, 619(1). <https://doi.org/10.1088/1755-1315/619/1/012035>
- Lu, H. J., y Chang, G. W. (2018). A Hybrid Approach for Day-Ahead Forecast of PV Power Generation. *IFAC-PapersOnLine*, 51(28), 634-638. <https://doi.org/10.1016/j.ifacol.2018.11.774>
- Ma, J., y Ma, X. (2018). A review of forecasting algorithms and energy management strategies for microgrids. *Systems Science and Control Engineering*, 6(1), 237-248. <https://doi.org/10.1080/21642583.2018.1480979>
- Mahmoud, F. S., Diab, A. A. Z., Ali, Z. M., El-Sayed, A. H. M., Alquthami, T., Ahmed, M., y Ramadan, H. A. (2022). Optimal sizing of smart hybrid renewable energy system using different optimization algorithms. *Energy Reports*, 8(April), 4935-4956. <https://doi.org/10.1016/j.egy.2022.03.197>

- Masih, A., y Verma, H. K. (2020). Optimization and Reliability Evaluation of Hybrid Solar-Wind Energy Systems for Remote Areas. *International Journal of Renewable Energy Research*, 10(4), 1696-1707. <https://doi.org/10.20508/ijrer.v10i4.11396.g8055>
- Medina, A. A., y Cárdenas-Barrón, L. E. (2022). Optimal Design of Hybrid Renewable Energy Systems Considering Weather Forecasting Using Recurrent Neural Networks. *Energies*, 15(23). <https://doi.org/10.3390/en15239045>
- Medina, A. A., y Cárdenas, L. E. (2022). Optimal Design of Hybrid Renewable Energy Systems Considering Weather Forecasting Using Recurrent Neural Networks. *Energies*, 15(23). <https://doi.org/10.3390/en15239045>
- Mohammed, O. H., Amirat, Y., Benbouzid, M., y Elbaset, A. A. (2014). Optimal design of a PV/fuel cell hybrid power system for the city of Brest in France. *2014 1st International Conference on Green Energy, ICGE 2014*, 119-123. <https://doi.org/10.1109/ICGE.2014.6835408>
- Municipalidad Provincial de Arequipa. (2021). *Plan De Desarrollo Local De Arequipa 2016-2021.41*. <https://www.muniarequipa.gob.pe/descargas/transparencia/pdlc/PDLC.pdf>
- Muthukumar, R., y Balamurugan, P. (2019). A novel power optimized hybrid renewable energy system using neural computing and bee algorithm. *Automatika*, 60(3), 332-339. <https://doi.org/10.1080/00051144.2019.1637173>
- Naciones Unidas-CEPAL. (2019). La Agenda 2030 y los Objetivos de Desarrollo Sostenible: una oportunidad para América Latina y el Caribe. Objetivos, metas e indicadores mundiales. En *Publicación de las Naciones Unidas*.
- NASA. (2020). *NASA Prediction Of Worldwide Energy Resources*. <https://power.larc.nasa.gov/>
- Nasir, M., Khan, H. A., Hussain, A., Mateen, L., y Zaffar, N. A. (2018). Solar PV-based scalable DC microgrid for rural electrification in developing regions. *IEEE Transactions on Sustainable Energy*, 9(1), 390-399. <https://doi.org/10.1109/TSTE.2017.2736160>
- Nasir, M., Zaffar, N. A., y Khan, H. A. (2016). Analysis on central and distributed architectures of solar powered DC microgrids. *Clemson University Power Systems Conference, PSC 2016*. <https://doi.org/10.1109/PSC.2016.7462817>
- Notton, G., y Voyant, C. (2018). Forecasting of Intermittent Solar Energy Resource. En *Advances in Renewable Energies and Power Technologies* (Vol. 1). Elsevier Inc.

<https://doi.org/10.1016/B978-0-12-812959-3.00003-4>

- Oficina de Prensa de Enel Perú. (2022). *Enel X Perú Instalará El Primer Sistema De Almacenamiento De Energía “Detrás Del Medidor” En Perú.* <https://www.enel.pe/es/conoce-enel/prensa/press/d202105-enel-x-peru-instalara-el-primer-sistema-de-almacenamiento-de-ene.html>
- Ohsawa, Y., Emura, S., y Arai, K. (1993). Optimal operation of photovoltaic/diesel power generation system by neural network. *Proceedings of the 2nd International Forum on Applications of Neural Networks to Power Systems, ANNPS 1993*, 99-103. <https://doi.org/10.1109/ANN.1993.264342>
- Oladigbolu, J. O., Ramli, M. A. M., y Al-Turki, Y. A. (2019). Techno-economic and sensitivity analyses for an optimal hybrid power system which is adaptable and effective for rural electrification: A case study of Nigeria. *Sustainability (Switzerland)*, 11(18). <https://doi.org/10.3390/su11184959>
- Olatomiwa, L., Mekhilef, S., Huda, A. S. N., y Sanusi, K. (2015). Techno-economic analysis of hybrid PV–diesel–battery and PV–wind–diesel–battery power systems for mobile BTS: The way forward for rural development. *Energy Science and Engineering*, 3(4), 271-285. <https://doi.org/10.1002/ese3.71>
- OSINERGMIN. (2022). *Osinergmin fue reconocido con el Premio Creatividad Empresarial por usar Inteligencia Artificial para reducir desconexiones eléctricas.* <https://www.gob.pe/institucion/osinergmin/noticias/674984-osinergmin-fue-reconocido-con-el-premio-creatividad-empresarial-por-usar-inteligencia-artificial-para-reducir-desconexiones-electricas>
- Paoli, C., Voyant, C., Muselli, M., y Nivet, M. L. (2010). Forecasting of preprocessed daily solar radiation time series using neural networks. *Solar Energy*, 84(12), 2146-2160. <https://doi.org/10.1016/j.solener.2010.08.011>
- Park, J. S., Katagi, T., Yamamoto, S., y Hashimoto, T. (2001). Operation control of photovoltaic/diesel hybrid generating system considering fluctuation of solar radiation. *Solar Energy Materials and Solar Cells*, 67(1-4), 535-542. [https://doi.org/10.1016/S0927-0248\(00\)00325-1](https://doi.org/10.1016/S0927-0248(00)00325-1)
- Pillai, G. G., Putrus, G. A., y Pearsall, N. M. (2014). Generation of synthetic benchmark electrical load profiles using publicly available load and weather data. *International Journal of Electrical Power and Energy Systems*, 61, 1-10. <https://doi.org/10.1016/j.ijepes.2014.03.005>
- Pombeiro, H., Santos, R., Carreira, P., Silva, C., y Sousa, J. M. C. (2017). Comparative

- assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks. *Energy and Buildings*, 146, 141-151. <https://doi.org/10.1016/j.enbuild.2017.04.032>
- Qin, A. K., Huang, V. L., y Suganthan, P. N. (2009). Differential evolution algorithm with strategy adaptation for global numerical optimization. *IEEE Transactions on Evolutionary Computation*, 13(2), 398-417. <https://doi.org/10.1109/TEVC.2008.927706>
- Qiu, Z. (2013). Electricity Consumption Prediction based on Data Mining Techniques with Particle Swarm Optimization. *International Journal of Database Theory and Application*, 6(5), 153-164. <https://doi.org/10.14257/ijdta.2013.6.5.14>
- Quispe, J. C., Obispo, A. E., y Alcantara, F. J. (2023). Economic feasibility assessment of microgrids with renewable energy sources in Peruvian rural areas. *Clean Technologies and Environmental Policy*. <https://doi.org/10.1007/s10098-023-02463-w>
- Ramli, M. A. M., Bouchekara, H. R. E. H., y Alghamdi, A. S. (2018). Optimal sizing of PV/wind/diesel hybrid microgrid system using multi-objective self-adaptive differential evolution algorithm. *Renewable Energy*, 121, 400-411. <https://doi.org/10.1016/j.renene.2018.01.058>
- Rao, S. N. V. B., Yellapragada, V. P. K., Padma, K., Pradeep, D. J., Reddy, C. P., Amir, M., y Refaat, S. S. (2022). Day-Ahead Load Demand Forecasting in Urban Community Cluster Microgrids Using Machine Learning Methods. *Energies*, 15(17). <https://doi.org/10.3390/en15176124>
- Raza, M. Q., y Khosravi, A. (2015). A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews*, 50, 1352-1372. <https://doi.org/10.1016/j.rser.2015.04.065>
- Razmjoo, A., Shirmohammadi, R., Davarpanah, A., Pourfayaz, F., y Aslani, A. (2019). Stand-alone hybrid energy systems for remote area power generation. *Energy Reports*, 5, 231-241. <https://doi.org/10.1016/j.egyr.2019.01.010>
- RedDavia Solar Power. (2018). *Solutions On-grid or off-grid, find the perfect cost-efficient solution*. <https://www.redviasolar.com/solutions/>
- Rehman, S. (2021). Hybrid power systems – Sizes, efficiencies, and economics. *Energy Exploration and Exploitation*, 39(1), 3-43. <https://doi.org/10.1177/0144598720965022>
- Rinaldi, F., Moghaddampoor, F., Najafi, B., y Marchesi, R. (2021). Economic feasibility

- analysis and optimization of hybrid renewable energy systems for rural electrification in Peru. *Clean Technologies and Environmental Policy*, 23(3), 731-748. <https://doi.org/10.1007/s10098-020-01906-y>
- Sadeghi, H. R., Ghaderi, F., Azade, A., y Razmi, J. (2009). Forecasting electricity consumption by clustering data in order to decline the periodic variable's affects and simplification the pattern. *Energy Conversion and Management*, 50(3), 829-836. <https://doi.org/10.1016/j.enconman.2008.09.036>
- Salman, A. G., Heryadi, Y., Abdurahman, E., y Suparta, W. (2018). Single Layer & Multi-layer Long Short-Term Memory (LSTM) Model with Intermediate Variables for Weather Forecasting. *Procedia Computer Science*, 135, 89-98. <https://doi.org/10.1016/j.procs.2018.08.153>
- Sawle, Y., Gupta, S. C., y Bohre, A. K. (2018). A novel methodology for scrutiny of autonomous hybrid renewable energy system. *International Journal of Energy Research*, 42(2), 570-586. <https://doi.org/10.1002/er.3841>
- Schmerler, D., Velarde, J. C., Rodríguez, A., y Solís, B. (2019). *Energías renovables: Experiencia y perspectivas en la ruta del Perú hacia la transición energética*. OSIDERMIN.
- Science, C., Veeramachaneni, K., Supervisor, T., Lacurts, K., Dunning, T., Friedman, E., Kondarasaiah, M. H., Ananda, S., Varun Chandola, Arindam Banerjee, and V. K., Ariyaluran Habeeb, R. A., Nasaruddin, F., Gani, A., Targio Hashem, I. A., Ahmed, E., Imran, M., Martí, L., Sanchez-Pi, N., Molina, J. M., Garcia, A. C. B., ... Casale, G. (2019). Machine Learning for Time Series Anomaly Detection Ihssan Tinawi. *Chance Encounters: Probability in Education*, 38(2), 2774-2797. <https://doi.org/10.1007/s10586-019-02998-y>  
<https://aws.amazon.com/sagemaker%0Ahttp://dx.doi.org/10.1016/j.eswa.2010.12.014>
- Scioletti, M. S., Newman, A. M., Goodman, J. K., Zolan, A. J., y Leyffer, S. (2017). Optimal design and dispatch of a system of diesel generators, photovoltaics and batteries for remote locations. *Optimization and Engineering*, 18(3), 755-792. <https://doi.org/10.1007/s11081-017-9355-4>
- Shezan, S. A. (2019). Optimization and assessment of an off-grid photovoltaic–diesel–battery hybrid sustainable energy system for remote residential applications. *Environmental Progress & Sustainable Energy*, August. <https://doi.org/10.1002/ep.13340>

- Shiva, B., y Sudhakar, K. (2015). Performance evaluation of 10 MW grid connected solar photovoltaic power plant in India. *Energy Reports*, 1, 184-192. <https://doi.org/10.1016/j.egy.2015.10.001>
- Sima, S., Tavakoli, N., y Siami, A. (2019). The Performance of LSTM and BiLSTM in Forecasting Time Series. *IEEE International Conference on Big Data (Big Data)*, 3285-3292. <https://doi.org/10.1109/BigData47090.2019.9005997>
- Singh, P., Dwivedi, P., y Kant, V. (2019). A hybrid method based on neural network and improved environmental adaptation method using Controlled Gaussian Mutation with real parameter for short-term load forecasting. *Energy*, 174, 460-477. <https://doi.org/10.1016/j.energy.2019.02.141>
- Sinha, S., y Chandel, S. S. (2015). Prospects of solar photovoltaic-micro-wind based hybrid power systems in western Himalayan state of Himachal Pradesh in India. *Energy Conversion and Management*, 105, 1340-1351. <https://doi.org/10.1016/j.enconman.2015.08.078>
- Skoplaki, E., Boudouvis, A. G., y Palyvos, J. A. (2008). A simple correlation for the operating temperature of photovoltaic modules of arbitrary mounting. *Solar Energy Materials and Solar Cells*, 92(11), 1393-1402. <https://doi.org/10.1016/j.solmat.2008.05.016>
- Soares de Araujo, J. M. (2020). Performance comparison of solar radiation forecasting between WRF and LSTM in Gifu, Japan. *Environmental Research Communications*, 2(4), 045002. <https://doi.org/10.1088/2515-7620/ab7366>
- Sorkun, M. C., Durmaz Incel, Ö., y Paoli, C. (2020). Time series forecasting on multivariate solar radiation data using deep learning (LSTM). *Turkish Journal of Electrical Engineering and Computer Sciences*, 28(1), 211-223. <https://doi.org/10.3906/elk-1907-218>
- Ssekulima, E. B., Anwar, M. B., Al Hinai, A., y El Moursi, M. S. (2016). Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: A review. *IET Renewable Power Generation*, 10(7), 885-898. <https://doi.org/10.1049/iet-rpg.2015.0477>
- Storn, R., y Price, K. (1995). Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11(4), 341-359. <https://doi.org/10.1023/A:1008202821328>
- Sultan, H. M., Menesy, A. S., Kamel, S., Korashy, A., Almohaimeed, S. A., y Abdel-Akher, M. (2021). An improved artificial ecosystem optimization algorithm for

- optimal configuration of a hybrid PV/WT/FC energy system. *Alexandria Engineering Journal*, 60(1), 1001-1025. <https://doi.org/10.1016/j.aej.2020.10.027>
- Sumorek, M., y Idzkowski, A. (2023). Time Series Forecasting for Energy Production in Stand-Alone and Tracking Photovoltaic Systems Based on Historical Measurement Data. *Energies*, 16(17). <https://doi.org/10.3390/en16176367>
- Sundén, B. (2019). Thermal management of batteries. *Hydrogen, Batteries and Fuel Cells*, 93-110. <https://doi.org/10.1016/b978-0-12-816950-6.00006-3>
- Tazay, A. (2020). Techno-Economic Feasibility Analysis of a Hybrid Renewable Energy Supply Options for University Buildings in Saudi Arabia. *Open Engineering*, 11(1), 39-55. <https://doi.org/10.1515/eng-2021-0005>
- Tealab, A., Hefny, H., y Badr, A. (2017). Forecasting of nonlinear time series using ANN Ahmed. *Future Computing and Informatics Journal*, 2(1), 39-47. <https://doi.org/10.1016/j.fcij.2017.05.001>
- Teixeira, L. A., de Souza, R. M., de Menezes, M. L., Cassiano, K. M., Pessanha, J. F. M., y Souza, R. C. (2015). Artificial neural network and wavelet decomposition in the forecast of global horizontal solar radiation. *Pesquisa Operacional*, 35(1), 73-90. <https://doi.org/10.1590/0101-7438.2015.035.01.0073>
- Tsoku, J. T., Phukuntsi, N., y Metsileng, L. D. (2017). Gold sales forecasting: The Box-Jenkins methodology. *Risk Governance and Control: Financial Markets and Institutions*, 7(1), 54-60. <https://doi.org/10.22495/rgcv7i1art7>
- Ulleberg, O. (1998). Stand-alone power systems for the future: optimal design, operation & control of solar-hydrogen energy systems. *NTNU, Trondheim, Norvège, December*, 225. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:STAND-ALONE+POWER+SYSTEMS+FOR+THE+FUTURE+:+OPTIMAL+DESIGN+,+OPERATION+&+CONTROL+OF+SOLAR-HYDROGEN+ENERGY+SYSTEMS#0%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Stand-alone+p>
- Ulleberg, O. (1999). *Stand-alone power systems for the future: Optimal design, operation and control of solar-hydrogen energy systems*. [Norges teknisk-naturvitenskapelige universitet, Trondheim (Norway)]. <https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/DE99751232.xhtml>
- US-European Commission. (2022). *the Impact of Artificial Intelligence on the Future of Workforces in the European Union and the United States of America*. 1-55.

- Vafaeipour, M., Rahbari, O., Rosen, M. A., Fazelpour, F., y Ansarirad, P. (2014). Application of sliding window technique for prediction of wind velocity time series. *International Journal of Energy and Environmental Engineering*, 5(2-3), 1-7. <https://doi.org/10.1007/s40095-014-0105-5>
- Vanneschi, L., y Castelli, M. (2018). Multilayer perceptrons. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 1-3, 612-620. <https://doi.org/10.1016/B978-0-12-809633-8.20339-7>
- Voyant, C., Muselli, M., Paoli, C., y Nivet, M. L. (2013). Hybrid methodology for hourly global radiation forecasting in Mediterranean area. *Renewable Energy*, 53, 1-11. <https://doi.org/10.1016/j.renene.2012.10.049>
- Wamba, S. L., Fosso Wamba, S., Kala Kamdjoug, J. R., y Tchatchouang Wanko, C. E. (2020). Influence of artificial intelligence (AI) on firm performance: the business value of AI-based transformation projects. *Business Process Management Journal*, 26(7), 1893-1924. <https://doi.org/10.1108/BPMJ-10-2019-0411>
- Wang, F., Yu, Y., Zhang, Z., Li, J., Zhen, Z., y Li, K. (2018). Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting. *Applied Sciences (Switzerland)*, 8(8), 1-29. <https://doi.org/10.3390/app8081286>
- Wang, H., Alattas, K. A., Mohammadzadeh, A., Sabzalian, M. H., Aly, A. A., y Mosavi, A. (2022). Comprehensive review of load forecasting with emphasis on intelligent computing approaches. *Energy Reports*, 8, 13189-13198. <https://doi.org/10.1016/j.egy.2022.10.016>
- Yang, J., Wang, Y., y Li, X. (2022). Prediction of stock price direction using the LASSO-LSTM model combines technical indicators and financial sentiment analysis. *PeerJ Computer Science*, 8. <https://doi.org/10.7717/PEERJ-CS.1148>
- Yin, J., Ngiam, K. Y., y Teo, H. H. (2021). Role of artificial intelligence applications in real-life clinical practice: Systematic review. *Journal of Medical Internet Research*, 23(4). <https://doi.org/10.2196/25759>
- Zahraee, S. M., Khalaji Assadi, M., y Saidur, R. (2016). Application of Artificial Intelligence Methods for Hybrid Energy System Optimization. *Renewable and Sustainable Energy Reviews*, 66, 617-630. <https://doi.org/10.1016/j.rser.2016.08.028>
- Zhang, J. (2021). Application of Artificial Intelligence in Renewable Energy and Decarbonization. *ES Energy and Environment*, 14, 1-2.

<https://doi.org/10.30919/esee8c550>

Zhao, Y., Li, T., Zhang, X., y Zhang, C. (2019). Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. *Renewable and Sustainable Energy Reviews*, 109(November 2018), 85-101. <https://doi.org/10.1016/j.rser.2019.04.021>

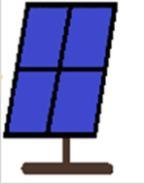
Zhong, X., y Cheng, P. (2020). An Improved Differential Evolution Algorithm Based on Dual-Strategy. *Mathematical Problems in Engineering*, 2020. <https://doi.org/10.1155/2020/9767282>

# ANEXO

## Anexo 1: Consumo energético de la comunidad rural

	Units	Power (w)	Quantity		Duration (Daily) (h)	Energy consumed (Daily) (wh)	Total Power (w)	
			(unit)	(h)				
Households	Ratio	15	1	3		45	15	
	TV	120	1	2		240	120	
	Lightbulbs	15	4	6		360	60	
	Cellphone-Charging	5	4	2		40	20	
	<b>Total watt hour per day for one households</b>						685	215
<b>Total watt hour per day for 200 households</b>						137000	43000	
School	Flurescent lighting	30	6	4		720	180	
	Ratio	15	1	2		30	15	
	Computer	310	3	4		3720	930	
	Microphone	20	1	1		20	20	
	<b>Total watt hour per day for School</b>						4490	1145
Police station	TV	120	1	3		360	120	
	Radio Walkie Talkie	3	8	2		48	24	
	Cellphone-charging	5	8	2		80	40	
	Radio	15	2	3		90	30	
	Flurescent lighting	30	6	4		720	180	
	Computer	310	3	3		2790	930	
	<b>Total watt hour per day for Police Station</b>						4088	1324
Health post	Vaccine store room lighting	30	1	7		210	30	
	Vaccine refrigerator	37	1	10		370	37	
	TV	120	1	6		720	120	
	Toilet Lighting	15	1	4		60	15	
	radio	5	1	4		20	5	
	Delivery room lighting	30	1	4		120	30	
	Patient waiting room lighting	30	1	5		150	30	
	Tratament room lighting	30	1	5		150	30	
	Outdoor lighting	30	1	6		180	30	
	<b>Total watt hour per day for Health post</b>						1980	327
	Church	Light bulbs	15	4	4		240	60
Streets of zone	Security lights	30	50	12		18000	1500	
<b>Total consumption for day in rural zone</b>						165798	47356	
						wh	W	

## Anexo 2: Ficha técnica de los componentes del sistema

MÓDULO		BATERÍA										
	Wp	450 W	Capacidad	150 Ah								
	Voc	50 V	Voltaje	12 V								
	Isc	10.39 A	DOD	60%								
	Vmpp	41.6 V	Corriente max	30 A								
	<table border="1"> <tr> <td>Dólar</td> <td>3.91</td> </tr> <tr> <td>precio Soles</td> <td>S/ 983.33</td> </tr> <tr> <td>Precio Dolares</td> <td>\$ 251.49</td> </tr> <tr> <td>Costo/kW</td> <td>558.869 \$/kW</td> </tr> </table>		Dólar	3.91	precio Soles	S/ 983.33	Precio Dolares	\$ 251.49	Costo/kW	558.869 \$/kW		Temp operaci6n
Dólar	3.91											
precio Soles	S/ 983.33											
Precio Dolares	\$ 251.49											
Costo/kW	558.869 \$/kW											
<table border="1"> <tr> <td>Dólar</td> <td>3.91</td> </tr> <tr> <td>precio Soles</td> <td>S/ 1.375.01</td> </tr> <tr> <td>Precio Dolares</td> <td>\$ 351.66</td> </tr> <tr> <td>Costo/kW</td> <td>\$ 195.37 \$/kWh</td> </tr> </table>		Dólar	3.91	precio Soles	S/ 1.375.01	Precio Dolares	\$ 351.66	Costo/kW	\$ 195.37 \$/kWh	Peso	30 kg	
Dólar	3.91											
precio Soles	S/ 1.375.01											
Precio Dolares	\$ 351.66											
Costo/kW	\$ 195.37 \$/kWh											
<p><b>S/.983,33</b></p>		potencia	1800 Wh									
<p>Marca AMERISOLAR</p> <p><a href="https://autosolar.pe/paneles-solares-24v/panel-solar-ja-solar-450w-24v-monocrystalino-perc">https://autosolar.pe/paneles-solares-24v/panel-solar-ja-solar-450w-24v-monocrystalino-perc</a></p>		<p><a href="https://autosolar.pe/pdf/datasheet-tensite-batteries-GEL-12-250.pdf">https://autosolar.pe/pdf/datasheet-tensite-batteries-GEL-12-250.pdf</a></p> <p><a href="https://autosolar.pe/baterias-de-gel-12v/bateria-ritar-12v-150ah-dg-gel">https://autosolar.pe/baterias-de-gel-12v/bateria-ritar-12v-150ah-dg-gel</a></p>										
<p><b>Inversor Victron Phoenix 24V 1200VA</b></p> <p><a href="https://autosolar.pe/inversores-48v/inversor-victron-phoenix-48v-3000va">https://autosolar.pe/inversores-48v/inversor-victron-phoenix-48v-3000va</a></p>		<p><b>Regulador</b></p> <p><a href="https://autosolar.pe/pdf/Ficha_tecnica_regulador_MPPT_SRNE_sr_mc.pdf">https://autosolar.pe/pdf/Ficha_tecnica_regulador_MPPT_SRNE_sr_mc.pdf</a></p>										
	<p><b>S/.5.984,29</b></p>			<p><b>S/.333,02</b></p>								
	potencia_inv	4000 W		potencia_inv	#iRE							
	Dólar	3.91		Dólar	3.91							
	precio Soles	S/ 5,984.29		precio Soles	S/ 333.02							
	Precio Dolares	\$ 1,530.51		Precio Dolares	\$ 85.17							
Costo/kW	\$ 382.63 \$/kW	Costo/kW	#jREF1 \$/kW									

[Panel Solar JA SOLAR 450W 24V Monocrystalino PERC | AutoSolar](https://autosolar.pe/paneles-solares-24v/panel-solar-ja-solar-450w-24v-monocrystalino-perc)  
[Inversor Victron Phoenix 48V 3000VA | al Mejor Precio en el Perú \(autosolar.pe\)](https://autosolar.pe/inversores-48v/inversor-victron-phoenix-48v-3000va)  
<https://autosolar.pe/pdf/datasheet-tensite-batteries-GEL-12-250.pdf>  
[autosolar.pe/pdf/Ficha tecnica regulador MPPT SRNE sr mc.pdf](https://autosolar.pe/pdf/Ficha_tecnica_regulador_MPPT_SRNE_sr_mc.pdf)

### Anexo 3: Código en Python del Modelo SARIMA

```
## Libreria Sarimax
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import matplotlib as mpl

import requests
import statsmodels.api as sm
import io
import os
import pylab
import scipy.stats as stats
from scipy.stats import kstest
from statsmodels.tsa.api import SimpleExpSmoothing, Holt,
ExponentialSmoothing
from sklearn.model_selection import TimeSeriesSplit
from sklearn.metrics import mean_squared_error
from statsmodels.tsa.stattools import acf, pacf
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from statsmodels.tsa.statespace.sarimax import SARIMAX
from random import seed
from pandas.plotting import lag_plot

plt.rcParams.update({'figure.figsize': (10, 7), 'figure.dpi': 120})
```

```
p = range(0, 3)
d = range(1, 2)
q = range(0, 3)

print(list(p))
print(list(d))
print(list(q))
```

```
df_sales = pd.read_csv("LoadPrimary.csv",
                      sep = ';', encoding='latin1')
#df_sales.head()
#lag_plot(df_sales)
```

```
df_monthsales=df_sales
```

```
from statsmodels.tsa.seasonal import seasonal_decompose
from dateutil.parser import parse
```

```

result_add = seasonal_decompose(df_monthsales['Primaryload'],model
= 'additive',freq = 12)
plt.rcParams.update({'figure.figsize': (20, 10), 'figure.dpi':
120})
result_add.plot().suptitle('Aditive Decompose',fontsize = 15)
plt.show()

```

```

# SEPARACION DE DATA

```

```

df_monthsales.drop(['Time'], axis='columns', inplace=True)
df_monthsales

```

```

n_splits= int(df_monthsales.shape[0]/876)-1#separe o corte
rADIATION TIENE [8760/13=673.84 (0.067%) data test]

```

```

n_splits

```

```

tss = TimeSeriesSplit(n_splits)

```

```

# 10 ---> X_train =667 X_test=67 794

```

```

# 11 ---> X_train =687 X_test=62 749

```

```

# 12 ---> X_train =687 X_test=57 744

```

```

X = df_monthsales.drop(labels=['Primaryload'], axis=1)

```

```

y = df_monthsales['Primaryload']

```

```

#from sklearn.model_selection import train_test_split

```

```

#X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.33, random_state=42)

```

```

#1 print(X)

```

```

#1 print(y)

```

```

for train_index, test_index in tss.split(X):

```

```

    print("TRAIN:",train_index,"TEST:", test_index)

```

```

    X_train, X_test = X.iloc[train_index, :], X.iloc[test_index,:]

```

```

    y_train, y_test = y.iloc[train_index], y.iloc[test_index]

```

```

X_train.shape, X_test.shape

```

```

y_train.shape, y_test.shape

```

```

# Graficamos los periodos train y test

```

```

plt.figure(figsize=(16,5), dpi=120)

```

```

plt.title('Timeserie of Global Radiation',fontsize =18)

```

```

plt.plot(pd.DataFrame(y_train).index,

```

```

pd.DataFrame(y_train).Primaryload,color='tab:green')

```

```

plt.plot(pd.DataFrame(y_test).index,

```

```

pd.DataFrame(y_test).Primaryload,color='tab:red')

```

```

plt.xticks(fontsize=13)

```

```

plt.yticks(fontsize=13)

```

```

plt.xlabel('Date', fontsize=16)

```

```

plt.ylabel('Globalsolar', fontsize=16)

```

```

plt.show()

```

```

df1 = pd.DataFrame()

```

```

df1.insert(0,"Train",y_train)

```

```

df1.insert(1,"Test",None)
df2 = pd.DataFrame()
df2.insert(0,"Test",y_test)
df2.insert(0,"Train",None)
df3 = pd.concat([df1,df2])
df3.to_excel('excel_out.xlsx')
# MODELO SARIMA ÓPTIMO

model_sarima = SARIMAX(y_train, order= (1,0,2),
seasonal_order=(1,0,1,12))
results_sarima = model_sarima.fit()
print(results_sarima.summary())
results_sarima.resid.plot()
print(results_sarima.resid.describe())
pred = results_sarima.get_prediction(start = 7884, end = 8759,
dynamic=False)    #36 dias

        # 1 semana 8593 hasta 8759 con 167 punto para
prediccion

        # 2 dias y medio 8699, end = 8759

        # 1 semana 8591, 8759
#pred = results_sarima.get_prediction(start = 8134, end = 8759,
dynamic=False) # 691, end = 743
pred_ci = pred.conf_int()
pred =abs(pred.predicted_mean)
#pred = results_sarima.get_prediction(step=-84,
dynamic=False).predicted_mean #.rename('forecast sarimax') #
predicción horizonte futuro

#pred_ci.head()
pred

#mte_forecast = pred.predicted_mean
#mte_forecast = pred
#mte_truth = y_test
plt.fill_between(pred_ci.index,
                pred_ci.iloc[:, 0],
                pred_ci.iloc[:, 1], color='gray', alpha=.07)
plt.plot(df_monthsales.Primaryload[7880:])    # y_test inicia en
8134
plt.plot(pred)

plt.show()
results_sarima.plot_diagnostics()
plt.show()
plt.savefig('staticaldianostig_sarima_213.png')

```

```

y_test

pred.shape, y_test.shape
from sklearn.metrics import mean_squared_error,
mean_squared_log_error, mean_absolute_error, r2_score
print("MSE:"+str(mean_squared_error(y_test, pred)))
print("RMSE:"+str(np.sqrt(mean_squared_error(y_test, pred))))
#print("MSLE:"+str(mean_squared_log_error(mte_forecast,
mte_truth)))
#print("RMSLE:"+str(np.sqrt(mean_squared_log_error(mte_forecast,
mte_truth))))
print("MAE:"+str(mean_absolute_error(y_test, pred)))

#from sklearn.metrics import r2_score
#y_true = [3, -0.5, 2, 7]
#y_pred = [2.5, 0.0, 2, 8]
print("R2:"+str(r2_score(y_test, pred)))

print(results_sarima.summary())
M
plt.plot(pred,y_test, 'o', label='Y_predicho vs Y_prueba')
#plt.plot(trainPredict,Y_train, 'x', label='Y_predicho vs
Y_entrena')
plt.legend()
plt.show()
mte_pred_concat = pd.concat([y_test, pred])
mte_pred_concat = pd.concat([y_test, pred])
forecast.head()
forecast_ci1.head()
forecast_ci2
ax = df_monthsales.Primaryload[8000:].plot(label='observed',
figsize=(20, 15))
# forecast.plot(ax=ax, label='Forecast')
ax.plot(forecast, label='Forecast')
ax.fill_between(forecast_ci1.index,
                forecast_ci1.iloc[:, 0],
                forecast_ci1.iloc[:, 1], color='black', alpha=.05)
ax.fill_between(forecast_ci2.index,
                forecast_ci2.iloc[:, 0],
                forecast_ci2.iloc[:, 1], color='blue', alpha=.5)
ax.set_xlabel('Time (year)')
ax.set_ylabel('NG CO2 Emission level')

plt.legend()
plt.show()
df1 = pd.DataFrame()
df1.insert(0,"Primaryload",df_monthsales.Primaryload)
df1.insert(1,"Predictive",None)
df1.insert(2,"Forecast", None)

```

```

df2 = pd.DataFrame()
df2.insert(0, "Predictive", pred)
df2.insert(0, "Primaryload", None)
df2.insert(0, "Forecast", None)
df3 = pd.DataFrame()
df3.insert(0, "Forecast", forecast)
df3.insert(0, "Primaryload", None)
df3.insert(0, "Predictive", None)
df4 = pd.concat([df1, df2, df3])
df4.to_excel('AllData&predictive1.xlsx')
df4.head()
plt.plot(forecast)
plt.show()
plt.figure(figsize=(8, 8), dpi=80)

from pandas.plotting import lag_plot
lag_plot(df_monthsales)
plt.show()
plt.figure(figsize=(8, 8), dpi=80)

from pandas.plotting import autocorrelation_plot
autocorrelation_plot(df_monthsales)
plt.show()
plt.figure(figsize=(8, 8), dpi=80)

from statsmodels.graphics.tsaplots import plot_acf
plot_acf(df_monthsales, lags=876)
plt.show()
plt.figure(figsize=(8, 8), dpi=80)

# Predecimos valores futuros
prediction_sarima=results_sarima.get_prediction(start = -
53).predicted_mean.rename('forecast sarima')
prediction_sarima
#Calculamos el ECM
error = mean_squared_error(y_test, forecast_sarima)
print('Test MSE: %.4f' % error)
# plot
plt.figure(figsize=(16,5), dpi=80)
plt.plot(df_monthsales)
prediction_sarima.plot(color = 'green', legend = True)
y_test.plot(color='red', legend = True)
#pred_cis=forecast_sarima.conf_int(alpha=0.5)
#plt.fill_between(lower_limits, upper_limits, color='pink')

plt.show()
#plt.figure(figsize=(16,5), dpi=100)

y_test.shape, prediction_sarima.shape

```

```
results_sarima.plot_diagnostics()
plt.show()
plt.plot(forecast_sarima,y_test, 'o', label='Y_predicho vs
Y_prueba')
#plt.plot(trainPredict,Y_train, 'x', label='Y_predicho vs
Y_entrena')
plt.legend()
plt.show()

plt.plot(forecast_sarima,y_test, 'o', label='Y_predicho vs
Y_prueba')
#plt.plot(trainPredict,Y_train, 'x', label='Y_predicho vs
Y_entrena')
plt.legend()
plt.show()

forecast_sarima =
results_sarima.get_forecast(step=100).predicted_mean.rename('foreca
st')
forecast_sarima
```

## Anexo 4: Código en Python del Modelo ANN

```
# ANN Librerias
from tensorflow.keras.optimizers import RMSprop
import numpy as np
import pandas as pd
from keras.layers.core import Dense, Dropout
#from keras.optimizers import RMSprop
from keras.models import Sequential
import matplotlib.pyplot as plt
start_time = time.time()
# SARIMAX librerias
!pip install seaborn
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
#from numpy.random import seed
#from tensorflow import set_random_seed
#!setting the seed
#!seed(0)
#!set_random_seed(0)
# import matplotlib as mpl
plt.rcParams.update({'figure.figsize': (10, 7), 'figure.dpi': 120})

#df = pd.read_csv("Load&radiation to
forecast_soloradiacion_onemonth.csv",
#                 sep = ';',encoding='latin1')
#, parse_dates=['date'], index_col='date'
df = pd.read_csv("Global_radiation.csv",
                 sep = ';',encoding='latin1')
df

df.drop(['Time'], axis=1, inplace=True)
df.dtypes
# Plot timeserie
plt.figure(figsize=(30,6))
plt.plot(df, color='blue')
plt.show()

# Data test and train

len(df) # definir data train y test (aumentar data test)
train, test = df[:-876], df[-876:]
# data train 684 y data test 60 (con poca data)
# 876 data de prueba
train.shape, test.shape
```

```

print(train)

# Scalando data entre (0,1)

from sklearn.preprocessing import MinMaxScaler
Scaler = MinMaxScaler()
Scaler.fit(train)
Scaler.data_min_
Scaler.data_max_
train = Scaler.transform(train)
test = Scaler.transform(test)
train=train[:,0]
test=test[:,0]

train, test

# Tratamiento Data

from keras.preprocessing.sequence import TimeseriesGenerator
# define generator
n_input = 876 # agregar mas puntos para aumentar mas puntos de
pronostico
generator = TimeseriesGenerator(train, train, length=n_input,
batch_size=1)
len(generator)
print(generator[0])

print('Samples: %d' % len(generator))
# print each sample
X=[]
Y=[]
for i in range(len(generator)):
    x, y = generator[i]
    X.append(x)
    Y.append(y)
    print('%s => %s' % (x, y))

generator.data
#generator.data.shape

#MODELO ANN
from keras import metrics
#from keras.optimizers import SGD,Adam,RMSprop
from tensorflow.keras.optimizers import SGD, Adam
model = Sequential()
model.add(Dense(100, activation='relu', input_dim= n_input)) #
MODIFICAR CADA PRUEBA

```

```

model.add(Dropout(0.2))
model.add(Dense(1))
model.summary()
opt=Adam(lr=0.001)
model.compile(optimizer=opt, loss='mse',
metrics=[metrics.MeanSquaredError(name='mse'),
metrics.RootMeanSquaredError(name='rmse'),
metrics.MeanAbsolutePercentageError(name='mape') ])

```

```

#Entrenamiento del modelo
#model.compile(loss='mse',optimizer='adam')
h= model.fit(generator, epochs=80, validation_data=(generator)) ##
GENERADOR : (X_TRAIN, Y_TRAIN) #FIJAR EPOCH 80

```

```

pd1=pd.DataFrame()
pd1.insert(0,'loss',h.history['mse'])
pd1.insert(1,'val_loss',h.history['val_loss'])
#pd2=pd.DataFrame()
pd1.to_excel('SD_cruvadeentrenamiento_17.xlsx') #MODIFICAR PARA
CADA PRUEBA
#pd3=concat([pd1,pd2])
pd1.head()
plt.plot(h.history['mse'])
plt.plot(h.history['val_mse'], color = "r")
plt.show()

```

```

pred_list = []
n_features=1
batch = train[-n_input:].reshape((1,n_input,n_features))[:, :, 0]
df_test = df[-876:]
print(batch)
print(train[681], ' ', train[682], ' ', train[683])
print('prediciendo')
for i in range(876):
    pred_list.append(model.predict(batch)[0])
    batch = np.append(batch[:,1:], [pred_list[i]],axis=1)
    print(pred_list[i])
    print(batch)
ventas_predict = pd.DataFrame(Scaler.inverse_transform(pred_list),
index=df_test[-876:].index, columns=['Predictions']) #168 ---#867
ventas_predict.to_excel('SDRADIATIONPredict_ANN__17.xlsx') #
MODIFICAR CADA PRUEBA
ventas_predict

```

```

ventas_test =pd.concat([df_test, ventas_predict], axis=1)

```

```
ventas_test.head()
```

```
#ventas_test =pd.concat([df_test, ventas_predict], axis=1)
#ventas_test.head(10)
plt.figure(figsize=(12,5))
plt.plot(df.index[8000:], df.Globalsolar[8000:], color='grey')
plt.plot(ventas_test.index, ventas_test['Globalsolar'])
plt.plot(ventas_test.index, ventas_test['Predictions'], color='r')
plt.legend(['Observated', 'Prediction'], loc='upper left')

plt.show()
```

```
plt.figure(figsize=(12,5))
#plt.plot(df)
plt.plot(ventas_test.index, ventas_test['Globalsolar'])
plt.plot(ventas_test.index, ventas_test['Predictions'], color='r')
plt.legend(['Observated', 'Prediction'], loc='upper left')

plt.show()
```

```
from sklearn.metrics import mean_squared_error,
mean_squared_log_error, mean_absolute_error, r2_score

#mte_forecast = ventas_predict
#ventas_test['Predictions'][-876:]
#mte_truth = df_test
#ventas_test['Globalsolar'][-876:]

#df_test, ventas_predict
print("MSE:"+str(mean_squared_error(df_test, ventas_predict)))
print("RMSE:"+str(np.sqrt(mean_squared_error(df_test,
ventas_predict))))
#print("MSLE:"+str(mean_squared_log_error(mte_forecast,
mte_truth)))
#print("RMSLE:"+str(np.sqrt(mean_squared_log_error(mte_forecast,
mte_truth))))
print("MAE:"+str(mean_absolute_error(df_test, ventas_predict)))
print("R2:"+str(r2_score(df_test, ventas_predict)))
#from sklearn.metrics import r2_score
#y_true = [3, -0.5, 2, 7]
#y_pred = [2.5, 0.0, 2, 8]

#limport math
```

```

#1from sklearn.metrics import mean_squared_error
#from sklearn.metrics import mean_absolute_percentage_error

#1mse = mean_squared_error(ventas_test['Globalsolar'][-876:],
ventas_test['Predictions'][-876:])
#1rmse = math.sqrt(mse)
#mape = mean_absolute_percentage_error(ventas_test['radiation'][-60:],
ventas_test['Predictions'][-60:])
#1print('mse', mse)
#1print('rmse:', rmse)
#print('mape:', mape)
# error de Test

```

```

print(model.layers[0].get_config())
print(model.layers[1].get_config())
#print(model.layers[2].get_config())

```

```

model.summary()
#print(model.layers[0].get_weights()) #pesos y peso capa 1
print(model.layers[0].get_weights()[0]) #pesos 1
print(model.layers[0].get_weights()[1]) #bias 1

```

```

model.optimizer
model.save('SDradiation_model_17.h5')

```

```

from keras.models import load_model
new_model =load_model('SDradiation_model_17.h5') # MODIFICAR CADA
PRUEBA
new_model.summary()
new_model.get_weights()
#print(model.layers[0].get_weights()) #pesos y peso capa 1
print(new_model.layers[0].get_weights()[0]) #pesos 1
print(new_model.layers[0].get_weights()[1]) #bias 1
new_model.optimizer

```

```

train1 =df
Scaler.fit(train1)
train2=Scaler.transform(train1)
train3=train2[:,0]

n_input= 876 #876
n_feature =1
generator = TimeseriesGenerator(train3, train3, length= n_input,
batch_size=1)
new_model.fit_generator(generator,epochs=60) #50
pred_list = []

```

```

batch = train3[-n_input:].reshape((1, n_input, n_features))[:, :, 0]

for i in range (n_input):
    pred_list.append(new_model.predict(batch) [0])
    batch=np.append(batch[:,1:], [pred_list[i]], axis=1)

from pandas.tseries.offsets import DateOffset

df.tail(876)    #168): #867
[df.index[-1]+ df.index[x] for x in range(1,877)] # 744 ....
from pandas.tseries.offsets import DateOffset
add_dates = [df.index[-1] + df.index[x] for x in range(1,878)] #
dateoffset () ~# 744 (inicia con 0) hasta 803(61)
future_dates = pd.DataFrame(index =add_dates[1:],
columns=df.columns)
ventas_predict=pd.DataFrame(Scaler.inverse_transform(pred_list),
index=future_dates[-n_input:].index, columns=['Prediction'])

ventas_predict.shape #
ventas_predict_meses=ventas_predict[0:876] # 6 puntos
ventas_predict_meses.to_excel('SDRADIATIONForecats_ANN__17.xlsx')
# MODIFICAR CADA PRUEBA
ventas_proj=pd.concat([df, ventas_predict_meses], axis=1)
plt.figure(figsize=(16,5), dpi=100)
plt.title('Serie de tiempo de ventas', fontsize=18)
plt.plot(df.index, df.Globalsolar, color='mediumblue' )
plt.xticks(fontsize=13)
plt.yticks(fontsize=13)
plt.xlabel('Fecha', fontsize=18)
plt.ylabel('Globalsolar', fontsize=18)
plt.show()

plt.figure(figsize=(18,8), dpi=80)
plt.plot(ventas_proj.index, ventas_proj['Globalsolar'],
color='mediumblue', label='Real')
plt.plot(ventas_proj.index, ventas_proj['Prediction'], color='r',
label='Pronostico')
plt.title('Prediccion de las ventas en el siguiente mes',
fontsize=18)
plt.xlabel('Fecha', fontsize=15)
plt.ylabel('Radiacion total', fontsize=15)
plt.xticks(fontsize=13)
plt.yticks(fontsize=13)
plt.legend()
plt.show()

plt.figure(figsize=(18,8), dpi=50)

```

```
plt.plot(ventas_proj.index, ventas_proj['Prediction'],
color='darkblue')
plt.title('prediciion de la radiacion en el mes', fontsize=18)
plt.xlabel('Fecha', fontsize=15)
plt.ylabel('Globalsolar', fontsize=15)
plt.xticks(fontsize=13)
plt.yticks(fontsize=13)
plt.legend()
plt.show();
```

## Anexo 5: Código en Python del Modelo LSTM

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.preprocessing import MinMaxScaler
import seaborn as sns
from pylab import rcParams
from sklearn.metrics import mean_absolute_error,
mean_squared_error, r2_score
from numpy import sqrt
import tensorflow as tf

# Data comes from:
csv_path = "Global_radiation_uni.csv"

dfx = pd.read_csv(csv_path, sep=';')
dfx

df=dfx.copy()

plt.figure(figsize=(20,20))
dfx.plot(x='Time', y='Globalsolar', kind='line')

# Normalization:

scaler = MinMaxScaler()
close_price = dfx['Globalsolar'].values.reshape(-1, 1)
scaled_close = scaler.fit_transform(close_price)
scaled_close.shape
np.isnan(scaled_close).any()
scaled_close = scaled_close[~np.isnan(scaled_close)]
scaled_close = scaled_close.reshape(-1, 1)
np.isnan(scaled_close).any()

# Normalization:
SEQ_LEN = 100

def to_sequences(data, seq_len):
    d = []
    for index in range(len(data) - seq_len):
        d.append(data[index: index + seq_len])
    return np.array(d)

def preprocess(data_raw, seq_len, train_split):
```

```

data = to_sequences(data_raw, seq_len)
num_train = int(train_split * data.shape[0])

X_train = data[:num_train, :-1, :]
y_train = data[:num_train, -1, :]

X_test = data[num_train:, :-1, :]
y_test = data[num_train:, -1, :]

return X_train, y_train, X_test, y_test

```

```

X_train, y_train, X_test, y_test = preprocess(scaled_close,
SEQ_LEN, train_split = 0.9)

```

```

X_train.shape

```

```

X_test.shape

```

```

# Model:
model = tf.keras.Sequential()
# adding 1st LSTM layer and some dropout regularization
model.add(tf.keras.layers.LSTM(units=50,
input_shape=(X_train.shape[1], 1), return_sequences=True,
activation = 'relu'))
model.add(tf.keras.layers.Dropout(0.2))
# adding 2nd LSTM layer and some dropout regularization
model.add(tf.keras.layers.LSTM(units=50, return_sequences=True))
model.add(tf.keras.layers.Dropout(0.2))
# adding 3rd LSTM layer and some dropout regularization
model.add(tf.keras.layers.LSTM(units=50, return_sequences=True))
model.add(tf.keras.layers.Dropout(0.2))
# adding 4th LSTM layer and some dropout regularization
model.add(tf.keras.layers.LSTM(units=50))
model.add(tf.keras.layers.Dropout(0.2))
# adding output layer
model.add(tf.keras.layers.Dense(units=1))
#compiling RNN
model.compile(
    loss='mean_squared_error', metrics=['accuracy'],
    optimizer='adam'
)
model.summary()

# Training:

BATCH_SIZE = 60

```

```

history = model.fit(
    X_train,
    y_train,
    epochs=100,
    batch_size=BATCH_SIZE,
    shuffle=False,
    validation_split=0.1
)
model.save('my_model_h.h5') # creates a HDF5 file 'my_model.h5'

model.evaluate(X_test, y_test)

plt.figure(figsize=(10,6))
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.show()

# Prediction:
y_hat = model.predict(X_test)

y_test_inverse = scaler.inverse_transform(y_test)
y_hat_inverse = scaler.inverse_transform(y_hat)

plt.figure(figsize=(15,7))

plt.plot(dfx.index[-len(X_test):],y_test_inverse, label="Actual",
color='green')
#plt.plot(y_test_inverse, label="Actual", color='green')
plt.plot(dfx.index[-len(X_test):],y_hat_inverse, label="Predicted",
color='red')

plt.title('Solar Power Forecasting LSTM')
plt.xlabel('Time [hours]')
plt.ylabel(r'Global Solar $(kW-hr/m^2/day)$')
plt.legend(loc='best')

plt.show();

print('Mean squared error: %.2f'%
mean_squared_error(y_test_inverse,y_hat_inverse))
print('Root Mean squared error: %.2f'%
sqrt(mean_squared_error(y_test_inverse,y_hat_inverse)))

```

```

print('Coefficient of determination: %.2f'%
r2_score(y_test_inverse,y_hat_inverse))

# Global Radiation Grid:

def metricas_sk(y_test_inverse,y_hat_inverse):
    m1=mean_squared_error(y_test_inverse,y_hat_inverse)
    m2=sqrt(mean_squared_error(y_test_inverse,y_hat_inverse))
    m3=r2_score(y_test_inverse,y_hat_inverse)

    return m1,m2,m3
def modelo(act,opt):
    model = tf.keras.Sequential()
    # adding 1st LSTM layer and some dropout regularization
    model.add(tf.keras.layers.LSTM(units=50,
input_shape=(X_train.shape[1], 1), return_sequences=True,
activation = act))
    model.add(tf.keras.layers.Dropout(0.2))
    # adding 2nd LSTM layer and some dropout regularization
    model.add(tf.keras.layers.LSTM(units=50, return_sequences=True))
    model.add(tf.keras.layers.Dropout(0.2))
    # adding 3rd LSTM layer and some dropout regularization
    model.add(tf.keras.layers.LSTM(units=50, return_sequences=True))
    model.add(tf.keras.layers.Dropout(0.2))
    # adding 4th LSTM layer and some dropout regularization
    model.add(tf.keras.layers.LSTM(units=50))
    model.add(tf.keras.layers.Dropout(0.2))
    # adding output layer
    model.add(tf.keras.layers.Dense(units=1))
    #compiling RNN
    model.compile(
        loss='mean_squared_error', metrics=['accuracy'],
        optimizer=opt
    )

    return model

def predicciones(X_test,y_test):
    y_hat = model.predict(X_test)
    y_test_inverse = scaler.inverse_transform(y_test)
    y_hat_inverse = scaler.inverse_transform(y_hat)

    return y_test_inverse,y_hat_inverse

def predicciones_new(date):
    y_hat = model.predict(date)
    y_hat_inverse = scaler.inverse_transform(y_hat)

    return y_hat_inverse

```

```

activation=['relu', 'tanh', 'sigmoid']
#activation=['tanh', 'sigmoid']
optimizer = ['rmsprop', 'adam']

print('Initializing')
for act in activation:
    for opt in optimizer:
        #reiniciando dict
        predicx = {}
        losses = {}
        metricas = {}
        #=====
        print(act+'_'+opt)
        model=modelo(act,opt)
        ## history
        BATCH_SIZE = 60
        history = model.fit(
            X_train,
            y_train,
            epochs=90,
            batch_size=BATCH_SIZE,
            shuffle=False,
            validation_split=0.1
        )
        #data history
        lossx=history.history['loss']
        val_lossx=history.history['val_loss']
        #predicts
        y_test_inverse,y_hat_inverse=predicciones(X_test,y_test)
        #new pred
        #date = pd.date_range(start='2020-01-01', end='2020-01-31',
freq='H')
        #new_pred=predicciones_new(date)
        #metrics
        mse,rmse,r2=metricas_sk(y_test_inverse,y_hat_inverse)
        #dictionaries
        time=df['Time'][-len(X_test):]
        predicx =
{'time':time,'true':y_hat_inverse.flatten(),'test':y_test_inverse.f
latten()}
        losses = {'loss':lossx, 'val_loss':val_lossx}
        metricas = {'mse':mse,'rmse':rmse,'r2':r2}
        #new_pred={'date'=date,'new_pred'=new_pred.flatten()}
        #pandas
        df_losses = pd.DataFrame.from_dict(losses)
        df_pred = pd.DataFrame.from_dict(predicx)
        df_metricas = pd.DataFrame.from_dict(metricas, orient='index')
        #df_new_pred = pd.DataFrame.from_dict(predicx)

```

```

#save data
name = act+'_'+opt+'.xlsx'
writer = pd.ExcelWriter(name)
df_losses.to_excel(writer, 'Sheet1')
df_pred.to_excel(writer, 'Sheet2')
df_metricas.to_excel(writer, 'Sheet3')
#df_new_pred.to_excel(writer, 'Sheet4')
writer.save()
#save keras model
namemod=act+'_'+opt+'.h5'
model.save(namemod)

#Predict NEW tIME
from keras.models import load_model

# load model
new_model = load_model('tanh_adam.h5')
# summarize model.
new_model.summary()

from keras.preprocessing.sequence import TimeseriesGenerator
from sklearn.preprocessing import MinMaxScaler
Scaler = MinMaxScaler()
dfx.drop(['Time'], axis=1, inplace=True)
dfx.dtypes

train1=dfx
Scaler.fit(train1)
train2=Scaler.transform(train1)
train3=train2[:,0]

n_input=99 #876
n_feature = 1
generator=TimeseriesGenerator(train3, train3,
length=n_input,batch_size=1)
print(generator)
new_model.fit_generator(generator, epochs=1)
pred_list = []
batch = train3[-n_input:].reshape((1,n_input,n_feature))[:, :,0] #
(,,,,) 99 columnas
#for i in range (n_input): #predecir 99 puntos
forecast_value=168 # pronosticar 168 puntos futuros fuera de la
data
for i in range (forecast_value): #predecir 99 puntos
    pred_list.append(new_model.predict(batch) [0])
    batch=np.append(batch[:,1:], [pred_list[i]], axis=1)

```

```

from pandas.tseries.offsets import DateOffset
df=dfx
len(pred_list)
df.index[-1]
from pandas.tseries.offsets import DateOffset
add_dates = [df.index[-1] + df.index[x] for x in range(1,169)] #
dateoffset () ~# 744 (inicia con 0) hasta 803(61)
future_dates = pd.DataFrame(index =add_dates[0:],
columns=df.columns) #0 inicia
print(future_dates)

ventas_predict=pd.DataFrame(Scaler.inverse_transform(pred_list),
index=future_dates[-forecast_value:].index, columns=['Prediction'])
print(ventas_predict)
ventas_predict.shape #
ventas_predict_meses=ventas_predict[0:168] # 6 puntos
print(ventas_predict)
ventas_predict_meses.to_excel('SDRADIATIONForecats_ANN___17.xlsx')
# MODIFICAR CADA PRUEBA

ventas_proj=pd.concat([df, ventas_predict_meses], axis=1)
print(ventas_proj)

plt.figure(figsize=(18,8), dpi=80)
plt.plot(ventas_proj.index, ventas_proj['Globalsolar'],
color='mediumblue', label='Real')
plt.plot(ventas_proj.index, ventas_proj['Prediction'], color='r',
label='Pronostico')
plt.title('Prediccion de las ventas en el siguiente mes',
fontsize=18)
plt.xlabel('Fecha', fontsize=15)
plt.ylabel('Radiacion total', fontsize=15)
plt.xticks(fontsize=13)
plt.yticks(fontsize=13)
plt.xlim(8690,8930)
plt.legend()
plt.show()
plt.figure(figsize=(18,8),dpi=50)
plt.plot(ventas_proj.index, ventas_proj['Prediction'],
color='darkblue')
plt.title('prediccion de la radiacion en el mes', fontsize=18)
plt.xlabel('Fecha', fontsize=15)
plt.ylabel('Globalsolar', fontsize=15)
plt.xticks(fontsize=13)
plt.yticks(fontsize=13)
plt.legend()
plt.show();

```

## Anexo 6: Código en Python de la operatividad del sistema y el algoritmo de optimización Evolución Diferencial

```

import pandas as pd

data = pd.read_csv('DATA_NORMAL_W_UNI_3_CORTADA.csv', sep=";",
encoding= 'unicode_escape')
data['T']=data['T(°C)'] # [°C]
data['Gh']=data['Gh(w/m2)'] # [W/m2]
data['load1']=data['load1(W)'] # [W]
NOCT=45 # [°C]
FF= (Impp*Vmpp) / (Isc*Voc)

GNOCT=800 # [w/m2]
TaNOC=20 # [°C]
Tstc=25 # [°C]
Ki=0.0005 # [A/°C] 0.05%/°C
Kv=-0.0028 # [V/°C] -0.28%/°C
Isc=10.39
Pfv=450
## MODULO FOTOVOLTAICO
Pfv=450 # [W/unid]
Voc=50 # [V]
Isc=10.39 # [A]
Vmpp=41.6 # [V]
Impp=10.812

Tc = np.zeros(data.shape[0])
Voc_ = np.zeros(data.shape[0])
Isc_=np.zeros(data.shape[0])
Ppv=np.zeros(data.shape[0])
Ppv_output=np.zeros(data.shape[0])
print(FF)
print(Pfv)
print(n_inv)
for t in range(data.shape[0]):
    Tc[t]=data['T'] [t]+((NOCT-TaNOCT)/GNOCT)*data['Gh(w/m2)'] [t]
# [°C]=[°C]+([°C*m2/w]*[w/m2])
    Voc_[t]=Voc+Kv*(Tc[t]-Tstc) #
[V]=[V]+[V/°C]*([°C-°C])
    Isc_[t]=(Isc+Ki*(Tc[t]-Tstc))*data['Gh(w/m2)'] [t]/Gref #
[A]=[A]+[A/°C]*[°C-°C]*[w/m2*m2/w]
    #Ppv_output[t] = Voc_[t]*Isc_[t]*FF
    Ppv[t]=Voc_[t]*Isc_[t]*n_inv*FF #
[W]=[V]*[A]*[cte]*[cte] # POTENCIA PARA 1 MODULO
    # n_con = mean(Ppv_output[t])/Ppv_output[t]
    # Ppv[t]=Ppv_output[t]*n_inv*n_con
print(Ppv)
#df1 = pd.DataFrame(Ppv)
#print(df1)

```

```

#df.to_excel('output.xlsx')
print(Tc)
print(Voc_)
print(Isc_)
print(Ppv.shape)
#hora=data['t(hours)'].shape(1,8760)
fig, ax = plt.subplots() #axs=plt.subplots(2, sharex=True) #,
sharey=True)
fig.suptitle('Comparation Power FV of one module vs Radiation for
m2')
ax.plot(data['t(hours)'],Ppv, color='yellow', label='Power FV(W)')
# axs[0], axs[1]
ax.plot(data['t(hours)'],data['Gh(w/m2)'], color='red',
label='Radiation(w/m2)')
ax.plot(data['t(hours)'], data['load1(W)'], label='Consumtion (w)')
plt.xlim(0, 500)
plt.ylim(0, 60000)
plt.legend()#frameon=False,fontsize='xx-small')
ax.set_ylabel('Power(W)')
ax.set_xlabel('Time(Hours)')
ax.set_title('Potencia vs Tiempo')
#axs[2].plot(x, y, '+')

```

```

def charge(Pp,Eb,Ebmax,Pl,t,Ech,Edch,n_bat,Edump,E_deficit): #
(((( ( 2.1 )))))
    uconv = 0.9 # [%]
    univ = 0.9 # [%]
    Edch[t] = 0 # [Wh]
    Pch = Pp[t]-Pl[t]/univ # [W]= [W] -[W]/[%]

    Ech[t] = np.multiply(Pch,n_bat) # [Wh]= [W]*[%][1h]
    if Ech[t]<=Ebmax-Eb[t-1]: # [ ! ! ]
        #Ebmin Eb(t-1) Ebmax
        Eb[t] = Eb[t-1]+Ech[t] # [Wh]
        Edump[t]=0.0
        #print("Ebateria",Eb)
        #print("Edeficit_1",E_deficit)
        E_deficit[t]=0.0
        #print("Edeficit_2",E_deficit)
    else:
        Eb[t] = Ebmax # [Wh]
        Edump[t] = Ech[t]-(Ebmax-Eb[t-1]) # [Wh]
        Ech[t] = Ebmax-Eb[t-1] # [Wh]
        E_deficit[t]=0.0
    return Eb,Ech,Edch,Edump,E_deficit

```

```

def
discharge(Pp,Eb,Ebmax,Pl,t,Ebmin,Edch,Ech,NS,Edump,E_deficit): #
(((((( 2.2 )))))
    uconv = 0.9 # [%]
    uinv = 0.9 # [%]
    Pdch = (Pl[t]-(Pp[t]*uinv))/uconv # [W] =[W]-
[W]*[%]/[%] [Pdch]-> [contr] ->uconv*Pdch=Pl[t]-(Pp[t]*uinv)
    Edch[t] = Pdch*1 # [Wh]=[W]*[1h]

    if(Eb[t-1]-Ebmin)>=Edch[t]: ##(40-20)>=15
        Eb[t] = Eb[t-1]-Edch[t] # [Wh] =[Wh]-[Wh]
        Edump[t] = 0.0 # [Wh]
        E_deficit[t]=0.0
    else: # (40-20)< 25 # (20% mini de bateria la cual no se puede
llegar a descargar) [ ' ]
        ## if(Eb[t-1]-Edch[t])>=0: # (40-25)>0 # [Wh] =[Wh]-
[Wh]
        E_deficit[t]=Edch[t]-(Eb[t-1]-Ebmin) # [wh]-([Wh]-
[wh]) { ! . }
        # LO Q
TENGO # 20 50 100
        Eb[t]=Ebmin
        Edump[t] = 0.0
        ##Eb[t] = Eb[t-1]-Edch[t] # [Wh] =[Wh]-[Wh]
        ##E_deficit[t]=0
    ## else:
        #Eb[t]=Ebmin
        #Edump=Edch[t]-(Eb[t-1]-Ebmin)
        ## Eb[t] = Eb[t-1] # [Wh] =[Wh]
        ## Edch[t] = 0 # [Wh]
        # E_deficit[t]=Pdch[t]
        NS = 1
        print('se esta violando la restriccion Edch<Socmin
en',t,'EnergiaBat',Eb)
        print('Constrained voilated at:')
        print(t)

    return Eb,Edch,NS,Ech,Edump, E_deficit

```

```

from pandas._libs.tslibs.tzconversion import tz_localize_to_utc
#def economic_cost(x):
def economic_cost(NSol,Nbat,Ncontr,Ninv,Expected_Life_Battery):
#Nreg, Nnveror Expected_Life_Battery

```

```

print('ingresa economic_cost')
#print('Corroborar Best',Best)
# initial cost
#Csol = 350.91*1.22 # costo fv

#bateray unit cost
#BAT_C = 333.55*1.22
#inverter cost
#INV_C = 127

#economic index
#REAL_INTREST = 6
#life time
#PRJ_LF = 25
#PV_L = 25
#Bate_L = 20
#ir = 6/100
### Parametros economicos###
inter=0.08 # tasa de
interes                                     #[cte]
rp=0.015 # 1.5% Porcentaje del costo de OM con respecto al Cost
capital                                     #[cte]
rr=1 # porcentaje (100%) del costo de reemplazo con respecto al
capital                                     #[cte]
#### Tiempo de Operacion y Mantenimiento de las tecnologias
tOM=1 # [año] tiempo de operacion en
año                                         #[Año]
tremBat=Expected_Life_Battery #tremBat=15 # [años] tiempo de
reemplazo de Bateria                       #[Año]
tremInv_contr= 10 # [años] tiempo de reemplazo de Controlador y
inversor                                    #[Año]
t_l=20 # tiempo de vida del
proyecto[años]                             #[Año]

###COST
##CAPEX##
Cfv=558.87 # [USD]/[KW]
CBat=195.37 # [USD/KWh]
Cinv= 382.63 # [USD/Kw]
Ccontr=29.57 # [USD/Kw]
#####

## OPEX ##
Cfv_OM=rp*Cfv # [USD/Kwp-año]
CBat_OM=rp*CBat # [USD/Kwh-año]
Cinv_OM=rp*Cinv # [USD/Kwp-año]
C_contr_OM=rp*Ccontr # [USD/Kwp-año]

```

```

# Costo de Reemplazo #
CBat_reemp=rr*CBat      # [USD/KWh]
Cinv_reemp=rr*Cinv      # [USD/KW]
Ccontr_reemp=rr*Ccontr  # [USD/KW]
#####
#### SALVAGE #####
rProject = t_l          #[años] : Años del Proyecto
rCom_Bat = Expected_Life_Battery #[años] : tiempo de vida de
la BATERIA
rCom_Con = tremInv_contr #[años] : tiempo de vida de
la CONTROLADOR
rCom_Inv = tremInv_contr #[años] : tiempo de vida de
la INVERSOR
#####
## Numero de componentes
#N_pfv=x[0]      #[Unid] 10
#N_pbat=x[1]     #[Unid] 10
# Ncontr=18 #[Unid] X[2]
# Ninv=12      #[Unid] X[3]
#####
## Potencia de cada componente
pfv= 450        #[W/unid]
Ebat=1800       #[Wh/unid]
pcontr= 2880    #[W/unid]
pinv= 4000     #[W/unid]

## Potencia Total de cada componente
Pfv_T=Nsol*(pfv/1000)      #[kW]   X1  ##NSOL= NUMERO DE
FOTOVOLTAICO A OPTIMIZAR
Pbat_T=Nbat*(Ebat/1000)    #[kWh]   X2  ##NBAT= NUMERO DE
BATERIAS A OPTIMIZAR
Pcontr_T=Ncontr*(pcontr/1000) #[kW]   X3
Pinv_T=Ninv*(pinv/1000)    #[kW]   X4
print('numeros de dispositivos', Nsol,Nbat,Ncontr,Ninv)
#####
sumal=0
P1=[]
for i in range(1, t_l+1,tOM):
    k = (1/((1+inter)**i))
    sumal=sumal+k
    P1.append(k)
    #print('a(',i,')',k)
    #print('sumal=',sumal)
## suma 2
#P2=[P1[j] for j in range(tremBat-1,t_l-4,tremBat)] # el valor de
4 es para no reemplazar años cercanos al proyecto
#suma2=sum(P2)          # 15-1 = 14, 20-4 = 16, 15
suma2 =0
for i in range (int(tremBat*10),(t_l+1)*10, int(tremBat*10)):

```

```

    r = i/10
    k = (1/((1+inter)**r))
    suma2 += k
#print("suma2=", suma2)
#suma 3
P3=[P1[ij] for ij in range(tremInv_contr-1,t_l-4,tremInv_contr)]
suma3=sum(P3)
#print("suma3=", suma3)
#####

#Battery_cost =
np.multiply(Nbat, [1+BAT_C*(ir*(1+ir)**(Bate_L))/(((1+ir)**Bate_L)-
1)])
#PV_cost =
np.multiply(Nsol, [(1*315/1000)+Csol*(ir*(1+ir)**(PV_L))/(((1+ir)**P
V_L)-1)])
#Inverter_C =
np.multiply(np.multiply(Nsol, [(1)+INV_C*(ir*(1+ir)**(20))/(((1+ir)*
*20)-1)]), (315/1000))
PV_cost =
np.multiply(Pfv_T, Cfv+suma1*Cfv_OM) #
[USD]=[kw]*[USD/kw]
Battery_cost
=np.multiply(Pbat_T, CBat+suma1*CBat_OM+suma2*CBat_reemp) #
[USD]=[kwh]*[USD/kwh]
Inverter_Cost
=np.multiply(Pinv_T, Cinv+suma1*Cinv_OM+suma3*Cinv_reemp) #
[USD]=[kw]*[USD/kw]
Reg_cost =
np.multiply(Pcontr_T, Ccontr+suma1*C_contr_OM+suma3*Ccontr_reemp) #
[USD]=[kw]*[USD/kw]
#Sol_cost = Nsol*Csol;
#Bat_cost = BAT_C*Nbat;
#Inv_cost = INV_C;
#####print('costos total de
dispositivos', PV_cost, Battery_cost, Inverter_Cost, Reg_cost)
#####
##### C_Salvage#####
##### Bateria Salvage #####
rCom_Bat =Expected_Life_Battery
rRep_Bat = rCom_Bat*int(rProject/rCom_Bat)
rRem_Bat = rCom_Bat-(rProject - rRep_Bat)
C_Salvage_Bat = CBat_reemp*rRem_Bat/rCom_Bat
##### Controlador Salvage #####
rRep_Con = rCom_Con*int(rProject/rCom_Con)
rRem_Con = rCom_Con-(rProject - rRep_Con)
C_Salvage_Con =Ccontr_reemp*rRem_Con/rCom_Con
##### inversor Salvage #####
rRep_Inv = rCom_Inv*int(rProject/rCom_Inv)

```

```

rRem_Inv = rCom_Inv-(rProject - rRep_Inv)
C_Salvage_Inv =Cinv_reemp*rRem_Inv/rCom_Inv
#####
#####Costo Salvage Total#####
C_Salvage_Total = C_Salvage_Bat + C_Salvage_Con + C_Salvage_Inv
#####
##### Imprimir costo Csalvage Anualizado #####
C_Salvage_Total = C_Salvage_Total*(1/(1.08))**25
#####
#####
#####CALCULO DEL LCC = NPC#####
#####
LCC = PV_cost+Battery_cost+Inverter_Cost+Reg_cost -
C_Salvage_Total
#LCC = PV_cost+Battery_cost+Inverter_Cost+Reg_cost # USD por
vida del proyecto
print('LCC',LCC)
#print("PV_cost=",PV_cost)
#print("Battery_cost=",Battery_cost)
#print("Inverter_Cost=",Inverter_Cost)
#print("Reg_cost=",Reg_cost)

initial_cost=LCC
Anual_cost = initial_cost#[0]
#LCOE=Anual_cost/sum(data['load1(W)']/1000)
print('costo anualizadoooo',Anual_cost)
print('Fin_economic_cost')
if NS==1:
    print('NS_eleccion_economic_cost')
    Annual_cost=5e5

return Annual_cost #,Sol_cost,Bat_cost,Inv_cost

```

```

def economic_analysis(x): #(((((( 2 ))))))
    print('ingresa economic analysis')
    print('datos de entrada X',x)
    ASC = []
    n_bat = 0.90
    #data = pd.read_csv('Data_normal_W.csv')
    data = pd.read_csv('DATA_NORMAL_W_UNI_3_CORTADA.csv', sep=";",
encoding= 'unicode_escape')

```

```

data['T']=data['T(°C)'] # [°C]
data['Gh']=data['Gh(w/m2)'] # [W/m2]
data['load1']=data['load1(W)'] # [W]
NS = 0

Floss = np.zeros(data.shape[0])
Psol = np.zeros(data.shape[0])
T_cell=np.zeros(data.shape[0])
# SOLAR SYSTEM
NOCT=43.6 # [°C]
Pr = 0.315;
Gs = 1000; # [w/m2]
gamma=-0.41 #[%/°C]
INV_ef=0.95
INV_Reg=0.8
pfv= 450 #[W/unid]
## modeling Photovoltaic ####
Tc = np.zeros(data.shape[0])
Voc_ = np.zeros(data.shape[0])
Isc_=np.zeros(data.shape[0])
Psol=np.zeros(data.shape[0])
#print(FF)
#print(Pfv)
#print(n_inv)
for t in range(data.shape[0]):
    Tc[t]=data['T'][t]+((NOCT-TaNOCT)/GNOCT)*data['Gh(w/m2)'][t]
# [°C]=[°C]+([°C*m2/w]*[w/m2])
    Voc_[t]=Voc+Kv*(Tc[t]-
Tstc) # [V]=[V]+[V/°C]*([°C-°C])
    Isc_[t]=(Isc+Ki*(Tc[t]-
Tstc))*data['Gh(w/m2)'][t]/Gref # [A]=[A]+[A/°C]*[°C-
°C]*[w/m2*m2/w]

    Psol[t]=Voc_[t]*Isc_[t]*n_inv*FF #
[W]=[V]*[A]*[cte]*[cte] # POTENCIA PARA 1 MODULO

#for t in range(data.shape[0]):
    #Floss[t] = 1 # [Cte]
    #if data['T'][t]>25:
    #    Floss[t] = (1-(data['T'][t]-25)*0.0038) #
[cte]=[cte-[°C-°C]*[%/°C]]
    #Psol[t] = Floss[t]*pfv*Pr*(data['Gh'][t]/Gs); #
[W]=[cte**[W/m2]/[W/m2]]*[W] * 1 modulo FV (UN POTENCIA
FOTOVOLTAICO en el tiempo t)

```

```

        #T_cell[t]=(data['T(°C)'][t]+((NOCT-
20/0.8)*data['Gh(w/m2)'][t])) # [°C]=[°C]+([°C]-[°C])*[W/m2]*?
        #Psol[t]=pfv*((data['Gh(w/m2)'][t]/Gs)-gamma*(T_cell[t]-
25))*INV_ef*INV_Reg
        #print(Psol)
        #if Psol[t]<0:
        #    Psol[t]=0
        #else:
        #    Psol[t]=Psol[t]
##
##Fotovoltaico
Nsol = x[0];    ## 1° Variable
Pp = np.multiply(Nsol,Psol)    # [W]
#####

# BATTERY
Nbat = x[1]                # [Cte]  2° Variable
P11 = data['load1']        # [W]
uin = 0.90                 # [%]
dod = 0.8                  # [%]
#####
#
#Ninvermin=round((Nfvmin*Pfv*1.11/(Pinver)))
#Ncontr=x[0]*Pfv/(Pcontr)
#Controlador
Ncontr = x[2]    ## 3° Variable
#####
#Inversor
#Ninv=x[0]*Pfv*1.11/(Pinver)
Ninv = x[3]    ## 4° Variable
#####
# MAIN PROGRAM
data_length = data.shape[0]
Ebmax = Nbat*12*150    #
[Wh]=[cte]*[V]*[Ah]    ###Etot_bat=Nbat*Vbat_nom*Cbat_nom ###
Etotal_bat=nbattery*Vsis*CapTotal_battery,
CapTotal_battery=Capbat*nbat_paralelo , Vist=Vnom*Nbatser
Ebmin = Ebmax*(1-dod)    # [Wh]
Eb = np.zeros((data_length))
fuel = np.zeros((data_length))
Edump = np.zeros((data_length))
Edch = np.zeros((data_length))
Ech = np.zeros((data_length))
E_deficit=np.zeros((data_length)) #data.shape[0]
#E_deficit_suma=[]
Eb[0] = Ebmax
print('numero de NPV,Nbat,Ncon,Ninv',Nsol,Nbat,Ncontr,Ninv)
# START
Pl = P11                # [W]

```

```

    univ = 0.90 # [Cte]
#E_deficit_total=np.zeros((data_length))
    for t in range(1,data_length):
        if (Pp[t]*univ)>=(Pl[t]): # [W]-[W]
            if (Pp[t]*univ)>Pl[t]:
                Eb,Ech,Edch,Edump, E_deficit=
charge(Pp,Eb,Ebmax,Pl,t,Ech,Edch,n_bat,Edump,E_deficit) #
(((((( 2.1 ))))))
                E_deficit[t]=0
            else:
                Eb[t] = Eb[t-1] # [Wh]
                E_deficit[t]=0
        else:
            Eb,Edch,NS,Ech,Edump,E_deficit =
discharge(Pp,Eb,Ebmax,Pl,t,Ebmin,Edch,Ech,NS,Edump,E_deficit) #
(((((( 2.2 ))))))

        if NS==1:
            break;
#E_deficit_suma=sum(E_deficit)
#E_deficit_total.append(E_deficit_suma)
#print("Deficit suma=",E_deficit_suma)
E_deficit_total.append(sum(E_deficit))
print("Deficit suma=",E_deficit_total)
Vsis=48
V_bat=12 # V
Cap_Bat=150 # Ah
Cap_nom_bat=V_bat*Cap_Bat # [wh]
n_rontrip=0.98
Ncycle=740
Throuthput_Bat=Cap_nom_bat*n_rontrip*DOD*Ncycle

Expected_Life_Battery=x[1]*(Throuthput_Bat/1000)/(sum(Edch)/100
0)
print('Expected Life Battery=',Expected_Life_Battery,'y')
###(COMENTARIO)price_electricity =
economic_cost(Nsol,Nbat,Ncontr,Ninv,NS) # #NReg
Nbat extraido ,Sol_cost,Bat_cost,Inv_cost # Price
electricity puede ser anual cost
#####price_electricity=LCC=Anual_cost ### es igual
#####(COMENTARIO) print('precio de la
electricidad',price_electricity)
print('fin economic analysis ')
if(NS==1):
    price_electricity = 5e5
else:
    price_electricity =
economic_cost(Nsol,Nbat,Ncontr,Ninv,Expected_Life_Battery)
print('precio de la electricidad',price_electricity)

```

```

ASC.append(price_electricity)

return price_electricity

## Problem Definition
#####
NS=0
Nfvmin=500 #50      F=300 - 1500 / 300 - 1500
Nfvmax=1500
Pfv=450

N_Bat_min=100      # 100 - 1500 / 100 - 1500
N_Bat_max=1500

Pcontr=60*48
Pinver=4000
Ncontrmin=round(Nfvmin*Pfv/(Pcontr))
Ncontrmax=round(Nfvmax*Pfv/(Pcontr))
Ninvermin=round((Nfvmin*Pfv*1.11/(Pinver)))
Ninvermax=round((Nfvmax*Pfv*1.11/(Pinver)))
print('Numero de contr mini,max; num invmin inv
max',Ncontrmin,Ncontrmax,Ncontrmin,Ninvermax)

# Problem Statement
Npar = 4  ## # problem dimension (4 dmensions)
## S.T:
### LIMITE DE LAS VARIABLES
VarLow = [Nfvmin,
N_Bat_min,Ncontrmin,Ninvermin]          #[300,300]          [50 , 10,
10, 10 ]
VarHigh =
[Nfvmax,N_Bat_max,Ncontrmax,Ninvermax]          #[1400,1400]          [300,
400 , 500, 600]

# parameters
N = 20          # number of population members
MaxIter = 50   # Maximun number of iteration (generations) #
generacion de prueba 2
F = 0.85       # DE-scaling factor
CR = 0.7       # crossover probability
Pmu = 0.5      # mutation probability

#initialize a random value as default best value
Best =
np.add(np.multiply(np.random.random((Npar)), np.subtract(VarHigh,Var
Low)),VarLow)
###
Best[2]=Best[0]*Pfv/Pcontr

```

```

Best[3]=Best[0]*Pfv*1.11/(Pinver)
E_deficit_total=[]

bestVal = economic_analysis(Best)
GB = bestVal
GBV = np.asarray([Best])

t = time.time()

#intialize a random population and memory
pop = np.zeros((N, Npar))
fitness = np.zeros((N))
print(fitness)
E_deficit_total.clear()
for ii in range(N):

    # initialize with a random solution
    pop[ii] = np.add(np.multiply(np.random.random((1,
Npar)), np.subtract(VarHigh, VarLow)), VarLow)
    ##### RESTRICCIÓN: pop[ii]=[pp[0],pop[1],pop[2],pop[3]]
; (pop[2]=pop[0]*pfv/pcontr, pop[3]=pop[0]*pop[n,0]*Pfv*1.11/Pinver)
    for n in range(N):
        pop[n,2]=pop[n,0]*Pfv/(Pcontr)
        pop[n,3]=pop[n,0]*Pfv*1.11/(Pinver)
    #####
    # calculate the fitness of the solution

    fitness[ii] = economic_analysis(pop[ii])

# Evaluate the best member
if fitness[ii]<bestVal:
    Best = pop[ii]
    bestVal = fitness[ii]

rot = np.arange(N)

E_deficit_total=[]
#Main loop
#np.zeros((MaxIter-1))
#MEJOR=[]
MIRA=[]
for ii in range(MaxIter-1):

    popold = pop

```

```

ind = np.random.permutation(4)
a1 = np.random.permutation(N)
rt = np remainder(np.add(rot, ind[0]), N)
a2 = np.array([a1[i] for i in rt])
rt = np remainder(np.add(rot, ind[1]), N)
a3 = np.array([a2[i] for i in rt])
rt = np remainder(np.add(rot, ind[2]), N)
a4 = np.array([a3[i] for i in rt])
rt = np remainder(np.add(rot, ind[3]), N)
a5 = np.array([a4[i] for i in rt])

pm1 = np.array([popold[i] for i in a1])
pm2 = np.array([popold[i] for i in a2])
pm3 = np.array([popold[i] for i in a3])
pm4 = np.array([popold[i] for i in a4])
pm5 = np.array([popold[i] for i in a5])

bm = []
for k in range(N):
    bm.append(Best)
## Mutando
mui = np.array([[1 if j<CR else 0 for j in i] for i in
np.random.random((N, Npar))])
mpo = np.logical_not(mui).astype(int)

ui = popold + F*(bm-popold) + F*(pm1 - pm2)
#### RESTRICCION: pop[ii]=[pp[0],pop[1],pop[2],pop[3]]
; (pop[2]=pop[0]*pfv/pcontr, pop[3]=pop[0]*pop[n,0]*Pfv*1.11/Pinver)
for m in range(N):
    ui[m,2]=ui[m,0]*Pfv/(Pcontr)
    ui[m,3]=ui[m,0]*Pfv*1.11/(Pinver)
#####
ui = np.multiply(popold,mpo) + np.multiply(ui,mui)
print('Vector MUTADO sin filtrado nPV,Nbat,Pcon,Pinv', ui)
print('Valor minimo y max de Ninv', VarLow, VarHigh)
### Filtrando

for k in range(N):
    ui[k] = limiter1(ui[k], VarHigh, VarLow)
#### RESTRICCION: pop[ii]=[pp[0],pop[1],pop[2],pop[3]]
; (pop[2]=pop[0]*pfv/pcontr, pop[3]=pop[0]*pop[n,0]*Pfv*1.11/Pinver)
for p in range(N):
    ui[p,2]=ui[p,0]*Pfv/(Pcontr)
    ui[p,3]=ui[p,0]*Pfv*1.11/(Pinver)
print('vector mutado_Valor minimo y max de
Ninv', ui[3], ui[3])
#####
#E_deficit_suma=[]

```

```

        tempval = economic_analysis(ui[k])
## Evaluando
        if tempval<fitness[k]:
            pop[k] = ui[k]
            fitness[k] = tempval

            if tempval<bestVal:
                bestVal = tempval
                Best = ui[k]

print('Vector MUTADO con filtrado nPV,Nbat,Pcon,Pinv', ui)
print('NUMERO DE DISPOSITIVOS FINALES Y LCC',GB,bestVal)
#E_deficit_total=np.append(E_deficit_total)
print('DEFICIT TOTAL PARA CADA ITERACION',E_deficit_total)
# store the best value in each iteration
GB = np.append(GB,bestVal)
#E_deficit_total.append(E_deficit_suma)
#E_deficit_total.append(E_deficit_suma)
GBV = np.append(GBV, [Best], axis=0)
print('NDISPOSI AND LCC', GB)
MIRA.append([bestVal,Best[0],Best[1]])

t1 = time.time()

print('The time taken is %3.2f seconds' % (t1-t))
print('The best value is :')

print('Best: ',Best)
print('bestVal: ',bestVal)
x = Best;

```

```

data = pd.read_csv('DATA_NORMAL_W_UNI_3_CORTADA.csv', sep=";",
encoding= 'unicode_escape')
data['T']=data['T(°C)'] # [°C]
data['Gh']=data['Gh(w/m2)'] # [w/m2]
data['load1']=data['load1(W)'] # [w]
NS = 0

Floss = np.zeros(data.shape[0]) # [cte]
Psol = np.zeros(data.shape[0])
T_cell=np.zeros(data.shape[0])
# SOLAR SYSTEM
NOCT=43.6 # [°C]
Pr = 0.85; # [Cte]
Gs = 1000; # [w/m2]
gamma=-0.41 # [%/°C]

```

```

INV_ef=0.95
INV_Reg=0.8
pfv=450 # [W/unid]

uinvs = 0.90 # [Cte]
dod = 0.8 # [Cte]
n_bat = 0.90
###MODELING PHOTOVOLTAIC ###
Floss = np.zeros(data.shape[0])
Psol = np.zeros(data.shape[0])
T_cell=np.zeros(data.shape[0])
# SOLAR SYSTEM
NOCT=43.6 # [°C]
Pr = 0.315;
Gs = 1000; # [w/m2]
gamma=-0.41 # [%/°C]
INV_ef=0.95
INV_Reg=0.8
pfv= 450 # [W/unid]
## modeling Photovoltaic ###
Tc = np.zeros(data.shape[0])
Voc_ = np.zeros(data.shape[0])
Isc_=np.zeros(data.shape[0])
Psol=np.zeros(data.shape[0])
print (FF)
print (Pfv)
print (n_inv)
for t in range(data.shape[0]):
    Tc[t]=data['T'][t]+((NOCT-TaNOCT)/GNOCT)*data['Gh(w/m2)'][t]
# [°C]=[°C]+([°C*m2/w]*[w/m2])
    Voc_[t]=Voc+Kv*(Tc[t]-Tstc) #
[V]=[V]+[V/°C]*([°C-°C])
    Isc_[t]=(Isc+Ki*(Tc[t]-Tstc))*data['Gh(w/m2)'][t]/Gref #
[A]=[A]+[A/°C]*[°C-°C]*[w/m2*m2/w]
    Psol[t]=Voc_[t]*Isc_[t]*n_inv*FF #
[W]=[V]*[A]*[cte]*[cte] # POTENCIA PARA 1 MODULO

#####
##Fotovoltaico
x=[120,180,20,20]
Nsol = x[0];
Pp = np.multiply(Nsol,Psol) # [W]
#####
# BATTERY
Nbat =x[1] # [Cte]
Pl1 = data['load1'] # [W]
uinvs = 0.90 # [%]
dod = 0.8 # [%]
#####

```

```

## MAIN PROGRAM
data_length = data.shape[0]
Ebmax = Nbat*12*150 # [Wh]=[cte]*[v]*[Ah]
Ebmin = Ebmax*(1-dod) # [Wh]
Eb = np.zeros((data_length))
fuel = np.zeros((data_length))
Edump = np.zeros((data_length))
Edch = np.zeros((data_length))
Ech = np.zeros((data_length))
Eb[0] = Ebmax # [Wh]

#
#Ninvermin=round((Nfvmin*Pfv*1.11/(Pinver)))
#Ncontr=x[0]*Pfv/(Pcontr)
#Controlador
Ncontr = x[2]
#####
#Inversor
#Ninv=x[0]*Pfv*1.11/(Pinver)
Ninv = x[3]
#####
# MAIN PROGRAM
data_length = data.shape[0]
Ebmax = Nbat*12*150 #
[Wh]=[cte]*[V]*[Ah] ###Etot_bat=Nbat*Vbat_nom*Cbat_nom ###
Etotal_bat=nbattery*Vsis*CapTotal_battery,
CapTotal_battery=Capbat*nbat_paralelo , Vist=Vnom*Nbatser
Ebmin = Ebmax*(1-dod) # [Wh]
Eb = np.zeros((data_length))
fuel = np.zeros((data_length))
Edump = np.zeros((data_length))
Edch = np.zeros((data_length))
Ech = np.zeros((data_length))
E_deficit=np.zeros((data_length))
Eb[0] = Ebmax
print('numero de NPV,Nbat,Ncon,Ninv',Nsol,Nbat,Ncontr,Ninv)
# START
Pl = P11 # [W]
univ = 0.90 # [Cte]

for t in range(1,data_length):
    if (Pp[t]*univ)>=(Pl[t]): # [W]-[W]
        if (Pp[t]*univ)>Pl[t]:
            Eb,Ech,Edch,Edump,E_deficit =
charge(Pp,Eb,Ebmax,Pl,t,Ech,Edch,n_bat,Edump,E_deficit) #
(((((( 2.1 ))))))
        else:

```

```

Eb[t] = Eb[t-1]          # [Wh]
else:
    Eb,Edch,NS,Ech,Edump, E_deficit=
discharge(Pp,Eb,Ebmax,Pl,t,Ebmin,Edch,Ech,NS,Edump,E_deficit) #
(((( ( 2.2 )))))
    #print(Edump.shape)

##### EXPECT LIFE BATTERY
Vsis=48
V_bat=12 # V
Cap_Bat=150 # Ah
Cap_nom_bat=V_bat*Cap_Bat # [wh]
n_rountrip=0.98
Ncycle=740
Throuthput_Bat=Cap_nom_bat*n_rountrip*DOD*Ncycle

Expected_Life_Battery=x[1]*(Throuthput_Bat/1000)/(sum(Edch)/100
0)
print('Expected Life Battery=',Expected_Life_Battery,'y')
#####

price_electricity =
economic_cost(Nsol,Nbat,Ncontr,Ninv,Expected_Life_Battery) # #NReg
Nbat      extraido ,Sol_cost,Bat_cost,Inv_cost      # Price
electricity puede ser anual cost
print('fin economic analysis ')

if(NS==1):
    price_electricity = 5e5
#####
#Nsol = x[0]
print('N fv optimo',Nsol)
Pp = np.multiply(Nsol,Psol)          # [W]=[Cte]*[W]      (Potencia
total FV= NSOL_OPTMO*Pp(t) [w])

m1 = 4880
m2 = 5048
t1 = np.arange(168)

ax.plot(data['t(hours)'], data['load1(W)'], label='Consumtion (w)')

#figure 8
ax1 = plt.subplot()
ax1.set_ylabel('Potencia (kW)',fontsize=24,fontname = 'Times New
Roman')
#ax1.plot(t1,data['load1'][m1:m2],color = 'red')          # [W]*1h

```

```

ax1.plot(data['t(hours)'],data['load1']/1000,color = 'blue',
label='Potencia de consumo (kW)' #[W]*1h
ax1.plot(data['t(hours)'],np.multiply(Pp,uinv)/1000,color =
'yellow',zorder=1, label='Potencia FV (kW)')
ax1.fill_between(data['t(hours)'],np.multiply(Pp,uinv)/1000,
color='yellow',alpha=0.8,zorder=1)
ax1.plot(data['t(hours)'],Edch/1000,color = 'green',zorder=2,
label='Descarga de la batería (kW)')
ax1.fill_between(data['t(hours)'],Edch/1000,
color='green',alpha=0.8,zorder=2)
#ax1.plot(t1,np.multiply(Pp[m1:m2],uinv),color = 'blue',zorder=1)
#ax1.fill_between(t1,np.multiply(Pp[m1:m2],uinv),
color='blue',alpha=0.8,zorder=1)
#ax1.plot(t1,Edch[m1:m2],color = 'green',zorder=2)
#ax1.fill_between(t1,Edch[m1:m2], color='green',alpha=0.8,zorder=2)

plt.ylim(0,70)
plt.xlim(0, 72)
ax2 = ax1.twinx()
ax2.set_ylabel('SOC(%)',fontsize=24,fontname = 'Times New Roman')
#ax2.plot(t1,Eb[m1:m2]*100/Ebmax,color = 'black')
ax2.plot(data['t(hours)'],Eb*100/Ebmax,'.:',color = 'olivedrab',
markersize=6, label='Estado de la batería (SOC%)')
ax1.legend(loc='upper left')
ax2.legend(loc=1)
#ax1.legend(loc="upper right")
#ax2.legend(loc="upper right")
#ax2.legend(loc="upper right")
ax1.set_title('Operación del sistema de generación (FV,
Batería)',fontsize=24,pad=20,fontname = 'Times New Roman')#,
fontsize=20, fontname = 'Times New Roman')
plt.ylim(0,120)
##ax1.set_xticks(x)
##ax1.set_xticklabels(x, fontsize=20)
#ax1.set_xticks(y)
#ax1.set_xticklabels(y, fontsize=20)
#ax1.set_ysticks(fontsize=15)
ax1.set_xlabel('Horas',fontsize=24,fontname = 'Times New Roman')
plt.savefig("Figura8.png")
plt.show()

```

### Global solar radiation time series forecasting using different architectures of the multilayer perceptron model

J J Peñalva<sup>1\*</sup>, D A Lozano<sup>2</sup>, J C Murillo<sup>2</sup> and F M Ortega<sup>3</sup>

<sup>1</sup>Faculty of Mechanical Engineering, Universidad Nacional de Ingeniería, Lima, Peru

<sup>2</sup>Faculty of Informatics Engineering, Pontificia Universidad Católica del Perú, Lima, Peru

<sup>3</sup>Faculty of Chemical Engineering, Universidad Nacional del Callao, Lima, Peru

\*E-mail: jpenalvas@uni.pe

**Abstract.** In this work, the multilayer perceptron model was used to forecast the time series of global solar radiation for a near future about a week. Different architectures of this model were built through varying its different hyperparameters such as optimizers, activation functions, number of neurons and neuron dropout in which their performance was evaluated using error metrics. It was found that the architectures (60, SGD, Sigmoid), (10, Adam, Relu) and (60, SGD, Sigmoid) presented an  $R^2$  around 0.877, 0.873 and 0.872, respectively. The architecture with neuron dropout (150, SGD, Sigmoid, 0.2) presented a higher performance among all the architectures evaluated and its  $R^2$  value was 0.884. Architectures with higher performance are used to predict future values of solar radiation.

#### 1. Introduction

In recent years, energy demand has increased due to industrial development, the growth of the economy and population [1]. This high energy demand produces a large increase in the burning of fossil fuels that releases large amounts of carbon dioxide and toxic gases into the atmosphere causing climate change and global warming, making it a global concern because of change on normal conditions of planet [2]. To mitigate and solve this problem, the use of clean energy such as renewable energy was increased to decarbonize the conventional electricity grid [3].

Solar energy is one of the most promising sources due to its availability and energy potential [4]. For this reason, the generation of photovoltaic energy has been increasing its installation both for systems connected and isolated from the electricity grid [5]. However, renewable energy such as solar and wind present an intermittent generation due to its variable nature caused by climatic factors [6], making it necessary to enter other generators and storage systems such as diesel and batteries to satisfy the energy demand of the place in a reliable way [7, 8].

In order to ensure the precise balance between production and electricity consumption using photovoltaic energy with greater penetration in the electrical grid, the precise forecasting of the solar resource is important to solve the problem of its uncertainty [9]. This solution allows the grid operator to guarantee a precise balance with a greater share of photovoltaic generation in a safe and reliable way [10-12].

In the literature, there are a lot of ways for forecasting time series such as classic models of the type Exponential Smoothing, Holt-Winters, Box-Jenking [13-14] and intelligent models such multilayer



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Published under licence by IOP Publishing Ltd

perceptron neural network [15]. Multilayer perceptron performance (MLP) is dependent on many hyperparameters, its evaluating is important to get the best model predictive [16].

This study aims to evaluate the combination of these hyperparameters and have good model performances. The work is organized as follows: Section 1 shows the introduction. Section 2 describes the methodology. Section 3 presents the results and discussions of different architectures of the MLP model. In section 4, we conclude and suggest future works.

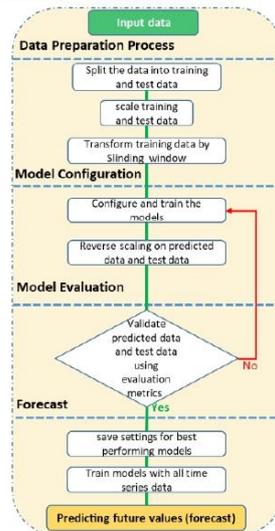
## 2. Methodology

In this work, the solar radiation time series is used to predict future values using different configurations of the multilayer perceptron model. Time series data of global solar radiation is obtained NASA' website for a study location, which is used to train a supervised learning model and predict its future behavior.

The flow diagram of the data processing, configuration, evaluation and forecasting of the MLP model is shown in figure 1. The first stage is about the data preparation process, where the data is divided into training and test data. Training data is normalized or scaled between values from 0 to 1 using the MinMax method [17]. Within this stage, an important sub-process is the transformation of the training data into input data for the MLP model, this transformation is carried out using the sliding window method that allows generating data series displaced at a later time [18-19].

The second stage is the configuration and training of the MLP model. In this stage, the model is defined by different architectures due to its hyperparameters. Each MLP architecture uses data to train and find a minimum error in its predictions.

The evaluation of the MLP model is the third stage, where the predictions of each architecture are evaluated with test data through error metrics, which indicates the performance of each architecture. High performance architectures continue to the next stage, otherwise they are trained again. Finally, the last stage is the prediction of future values or forecast. At this stage, the highest performing architectures are saved and used to predict future values of time series.

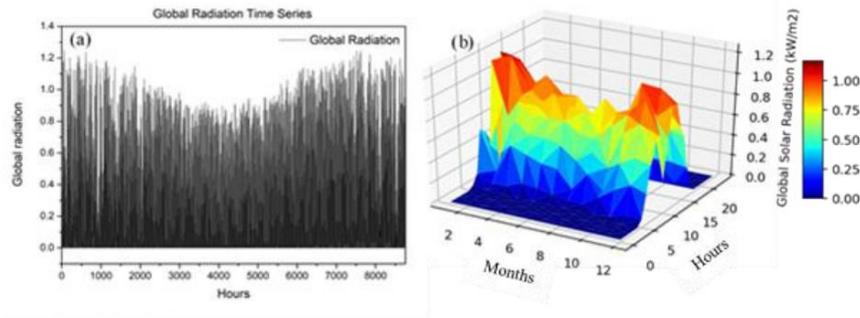


**Figure 1.** Flow chart to obtain future values: Data preparation process, configuration, evaluation and forecast of the model.

### 2.1. Study area and data of solar radiation

The study place was chosen in the community called San Juan de Tarucani which is located in Peru, Arequipa region with coordinates  $16^{\circ} 11' S$  and  $71^{\circ} 3.6' E$  of latitude and longitude, respectively.

Time series data of global solar radiation obtained at these coordinates is shown in figure 2a. On the other hand, in figure 2b, the radiation curve in hours and months is shown in a 3-dimensional graph; This curve shows greater radiation in the summer months, reaching values of  $1.2 \text{ kW/m}^2$ .



**Figure 2.** (a) Time series of solar radiation from the study site throughout the year. (b) Graph of global solar radiation in 3D for months and hours.

### 2.2. Data preparation process: Data partitioning, scaling and transformation

The time series data is partitioned into training and test data with 90% and 10%, respectively. The training data allows to find the optimal weights of the MLP model to ensure efficient learning and high performance in the predictions on training and test data.

Intelligent models such as MLP improve their performance when their input values are scaled in a standard range. The MinMax method is used to scale data values between the range of 0 and 1. The training data is scaled to train MLP model through equation 1, and the predicted values by the model are rescaled to their normal values through the equation 2:

$$x'_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (1)$$

$$x_i = x'_i(\max(x_i) - \min(x_i)) + \min(x_i) \quad (2)$$

Where  $\min(x_i)$  and  $\max(x_i)$  are the minimum and maximum values that are present in the training data, respectively,  $x'_i$  is the scaled value of  $x_i$  which is between values of 0 and 1.

Univariate time series are a unique set of observable values in a time order that can be predicted using supervised learning models such as the MLP model. However, this set of time series requires being prepared as input and output data to train the model. This preparation consists of generating samples with observable values shifted in previous time steps as input variable and an observable value of next time step as output variable. This procedure is called the sliding window method, where the number of previous time steps is called the window width, and each window shift predicts an observable value.

Figure 3 graphically shows data preparation, training, and prediction of the MLP model. The value of  $(t + N)$  and  $M$  represent the number of observable values for training and test data, respectively. When preparing data, the observable values of the time series were divided into  $(N + M + 1)$  samples with  $t$  number of time steps as input variable and one time step as output variable. The new training data has  $(N + 1)$  sliding windows or samples which are used as input data to train the MLP model and predict the

following values. When the model finds an optimal learning with new training data and a minimum error, the predictions on test data are carried out. These predicted values are obtained when test data samples are entered into the model; As an example, when entering the sample with label (N + 2) to the model, an observable value  $x_{(t+N+2)}$  is obtained. These predictions are evaluated with values of test data using error metrics.

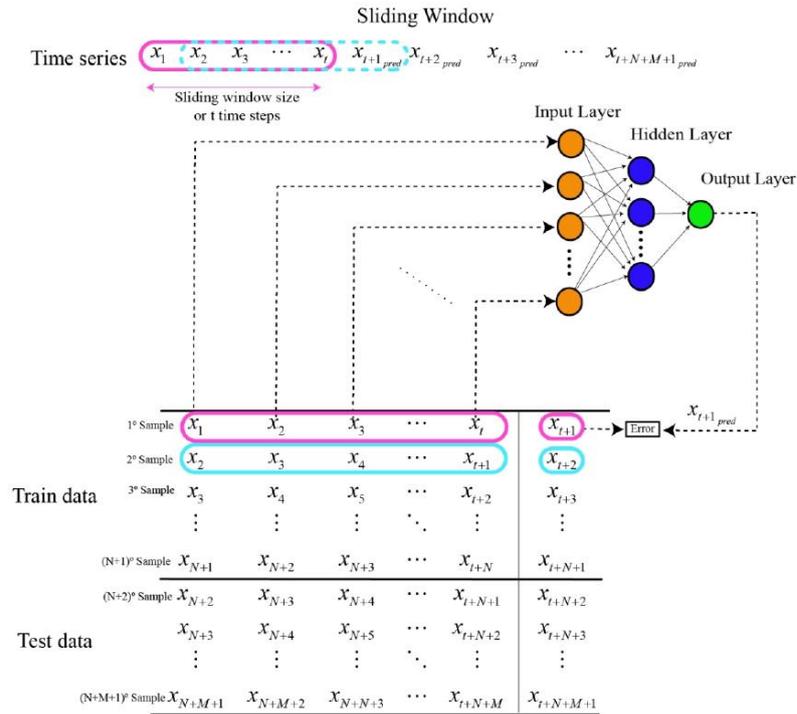


Figure 3. Data preparation, training and prediction of the MLP model.

2.3. The multilayer perceptron (MLP) model

Within the artificial neural network, the MLP model is one of the most used architectures because it is an efficient and practical learning algorithm, which allows a wide application in various fields of science and industry [20-22].

Multilayer perceptron neural network is comprised of three types of layers, which are input layer, hidden layer and output layer where each one has a certain number of neurons which are interconnected. The data processing of the MLP model has several stages [23-24]: In the first stage, this process occurs in the input layer neurons and hidden layer neurons. The observable values assigned to each neuron in the input layer are multiplied by the assigned weights, then added by bias and stored in each neuron in the hidden layer. This process is mathematically expressed in equation (3):

$$h_l = \sum_{i=1}^I w_{li} x_i + \beta_l, \forall l \in \{1, 2, 3, \dots, j\} \tag{3}$$

Where  $l$  represents the number of input neurons,  $x_i$  represents the values of each neuron in the input layer and  $w_{il}$  is the weight assigned to connect the input neuron  $x_i$  and the hidden neuron  $l$ . The second term is the bias  $\beta_l$  which is assigned to each neuron in the hidden layer. The sum of the product between input values and their assigned weights, plus the bias of each hidden neuron are represented by  $h_l$ .

The second stage occurs at the output of hidden layer neurons where the output values of each hidden neuron are calculated as follows:

$$H_l = f(h_l), \forall l \in \{1, 2, 3, \dots, j\} \quad (4)$$

Where  $H_l$  is the output value of each hidden neuron and  $f$  is the activation function that transforms the values of  $h_l$ .

The activation functions are varied, being the most used the types of activation functions Sigmoid, Tanh and Relu [25], which are represented by equations (5), (6) and (7), respectively:

$$f(h_l) = \frac{1}{1 + e^{-h_l}} \quad (5)$$

$$f(h_l) = \frac{(e^{h_l} - e^{-h_l})}{(e^{h_l} + e^{-h_l})} \quad (6)$$

$$f(h_l) = \begin{cases} 0, & \text{if } h_l < 0 \\ h_l, & \text{if } h_l \geq 0 \end{cases} \quad (7)$$

The last stage occurs in the output layer neurons, where the output values of hidden layer neurons are multiplied by new assigned weights and added by a value of bias to each output neuron. That is represented by:

$$O_k = f\left(\sum_{l=1}^j w_{lk} H_l + \beta_k\right), \forall k \in \{1, 2, 3, \dots, n\} \quad (8)$$

Where  $\beta_k$  is the bias of each output neuron,  $w_{lk}$  is the weight assigned to the output values of each hidden neuron which is the connection between the hidden neurons and output neurons,  $f$  is the activation function of the output layer neurons and  $O_k$  is the output value of the multilayer perceptron.

The main objective of the MLP model is to predict values closer to the observable values. For this reason, it is important to minimize the error of the MLP predictions. The cost or error function ( $E$ ) is represented in equation (9), which has been considered for a single neuron in the output layer making  $k = 1$  and  $O_1 = y_{pred}$  derived from the equation (8):

$$E = \frac{1}{2} \sum_{i=1}^n (y_i - y_{pred})^2 \quad (9)$$

Where  $y_{pred}$  is the predicted value by the model and  $y_i$  is the observable value of the time series with  $n$  observable quantities.

This error function is iteratively minimized by optimization algorithms, which aims to find optimal weights that improve the learning process of the model and minimize the error between predicted and observable values. The MLP model uses several types of optimization algorithms or also called learning algorithms such as Gradient Descending (SGD), RMSprop and Adam, among others [26].

The performance of the MLP model is evaluated using error metrics. These metrics are the mean square error (MSE), mean square error (RMSE), mean absolute error (MAE) and the coefficient of determination ( $R^2$ ). These metrics indicate how good is the model performs when the MSE, RMSE, and MAE values are close to zero and  $R^2$  is close to 1 [27].

### 3. Experimental setup of MLP models

To obtain an adequate MLP model, it is important to study its hyperparameters and the performance of its predictions. For this, different hyperparameters are evaluated which generate diverse configurations of the model, among them we have a learning algorithm, activation function, number of neurons in the hidden layer and neuron dropout. There are other hyperparameters such as input windows or input neurons, learning rate, number of output neurons and other parameters which are kept fixed in this study.

The solar radiation time series data is one year of record in units of kW/m<sup>2</sup> per hour and has 8760 observable values that are divided into training and test data in a relationship of 90% and 10%, respectively. In the training data, the minimum and maximum values are 0 and 1.2 kW/m<sup>2</sup> to be used in scaling and rescaling. For data preparation, the observable values of the time series were divided into 7884 samples with 876 number of time steps as input variable and one time step as output variable. These input and output data are used to train the model and make predictions.

Table 1 shows the different values for these hyperparameters used in the MLP configuration, which generate 216 configurations to be evaluated. These configurations are represented as architectures of the type (number of neurons, optimizer, activation function, with and without neuron dropout, among others). The architectures are trained with training data at a learning rate of 0.001 and with 80 training epochs for each optimizer.

The predictions generated by the MLP model architectures are evaluated with the test data using error metrics (MSE, RMSE, MAE, R<sup>2</sup>) to indicate how good their performance is. Subsequently, the MLP architectures are saved to later train them with the entire time series and predict future values outside the series.

**Table 1.** Hyperparameters and their values for various architectures of MLP model.

Hyperparameters	Possible values	Count
Number of time steps or input neurons	876	1
Learning algorithm or optimizer	SGD, Adam, RMSprop	3
Learning rate	0.001	1
Activation function	Sigmoid, Relu and Tanh	3
Hidden layer neurons number	2, 3, 4, 5,6, 10, 20, 40, 60, 80, 100, 150	12
dropping out units	0% (dropout=0) and 20% (dropout=0.2)	2
Number of output time steps or output neurons	1	1
Total number of hyperparameter combinations		216

### 4. Results and discussion

In this section, the training results, predictions and performance evaluation of the MLP model with different architectures will be given.

#### 4.1. MLP training process for various architectures

The training processes for different MLP architectures are shown in figure 4. These graphs show the learning (solid lines) and validation (dotted lines) convergence curve for each architecture.

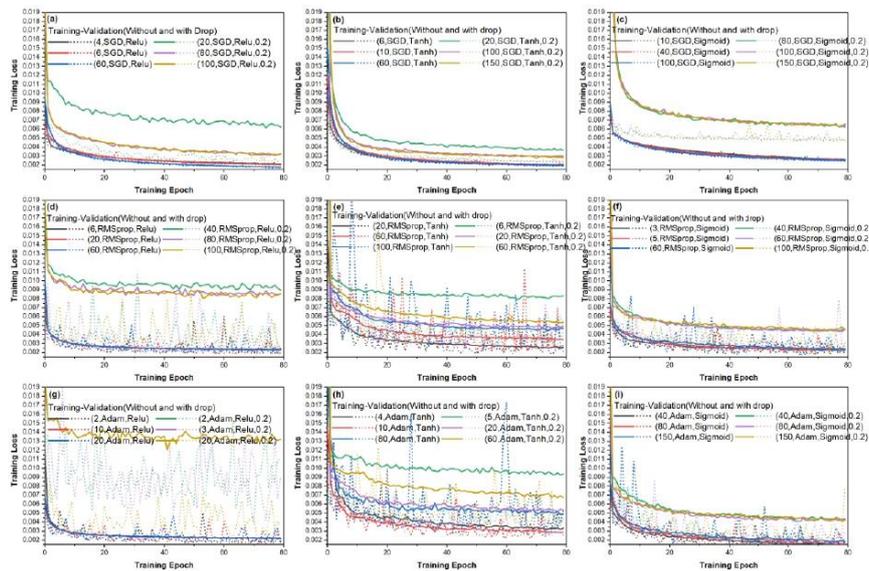
Figures 4a - 4c show the overlap of the training and validation curves when neuron dropout is not used (dropout = 0). However, when using neural dropout (dropout = 20%), there is a separation distance between the training and validation curves. Oneto *et al.* [28] indicate that the use of neuron dropout is important to reduce the overfitting generated in training and to improve the performance of the model. Furthermore, in figure 4a and figure 4b, it is observed that the use of neuron dropout and the increase in

the number of hidden neurons decreases the distance between the training and test curve, reaching a minimum error.

Furthermore, these figures show that when neurons in the hidden layer increase for SGD optimizer, activation function (Relu, Tanh, and sigmoid), and neuron dropout, there is an overlap of training curves. The opposite occurs when decreased neurons, these training curves are separated, being observed in figure 4a and figure 4b. It is also observed that the training and validation curves reach a minimum of training loss with the Relu activation function and the SGD optimizer than the other activation functions.

Figure 4d - 4f, the training and validation curves are shown in the same range as the previous curves. These curves converge when neurons are increased except for RMSprop and Tanh, which present a different behavior. It is also observed many pronounced peaks appear as noise in the validation curve being RMSprop and Sigmoid a little more stable.

Figure 4g shows the training and validation curves for the combination of Adam and Relu. Without neuron dropout, curves converge and overlap with little noise; however, when neuron dropout is added, these curves tend to separate with pronounced peaks. On the other hand, in figure 4h and figure 4i, the activation functions (Tanh, sigmoid) with the Adam optimizer have a similar behavior to the curves generated by Tanh and Sigmoid with the RMSprop optimizer.



**Figure 4.** Training and validation curves of the MLP model for different architectures with activation functions (Relu, Tanh, Sigmoid) and the optimizers (SGD, RMSprop, Adam).

4.2. Predictions of different architectures on test data of solar radiation

Figure 5 shows the predictions of global solar radiation for each MLP architecture concerning to its activation function (Relu, Tanh and Sigmoid), optimizer (SGD, RMSprop, Adam), with neuron dropout (0.2) and without neuron dropout (0). The architectures without neuron dropout are shown with solid lines, while the architectures with neuron dropout are represented by dotted lines.

Figure 5a – 5c show the predictions for the architectures (SGD, activation function, dropout=0) with activation functions Relu, Tanh, Sigmoid and different numbers of neurons. These figures show high

and low values of predictions on the peaks of solar radiation profile. It is also observed that the highs and lows of the predictions occur when solar radiation takes zero value.

Figure 5d shows the architecture predictions (RMSprop, Relu, dropout = 0.2, neurons). This figure shows how the predictions of these architectures have good precision with initial observable values and then lose precision. In figure 5 (e) and (f), the predictions improve when considering more neurons.

Figure 5g- 5i, it is observed that the architectures (Adam, Relu, neurons) have better predictions of solar radiation including values close to zero. At these architectures when neuron dropout is considered, the predictions decay over time.

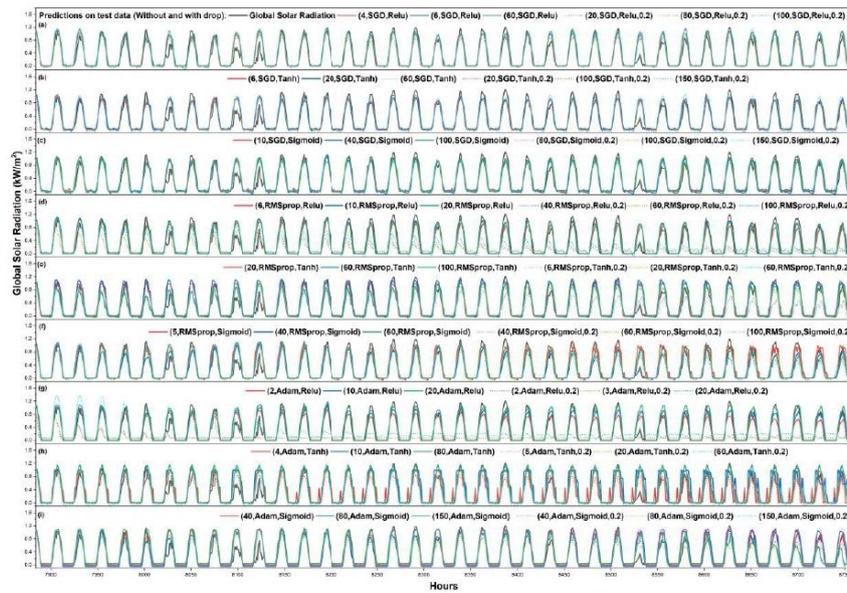


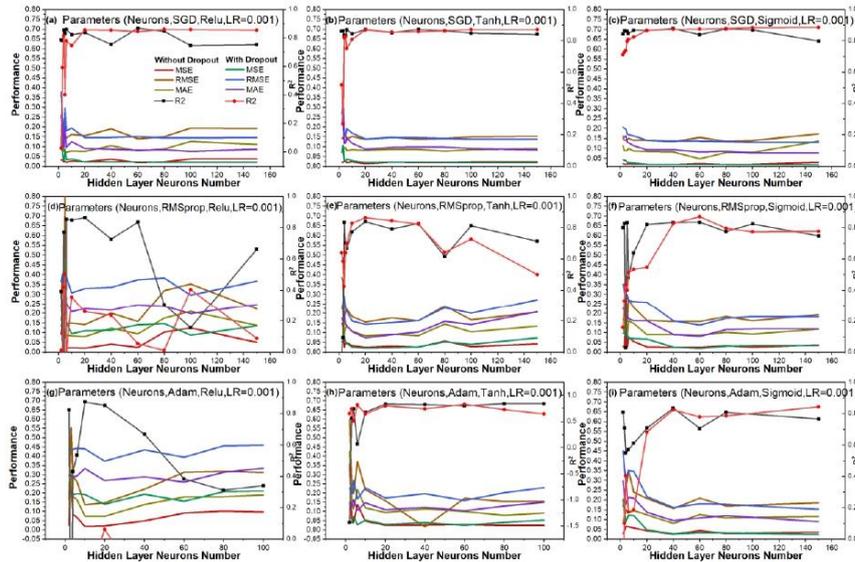
Figure 5. Predictions of the different architectures on the solar radiation test data.

4.3. Error metrics for different MLP architectures

Figure 6 shows the performances of MLP model architectures as a function of neurons number and learning rate ( $Lr = 0.001$ ), these performances are obtained by means error metrics of type MSE, RMSE, MAE and  $R^2$ .

The architectures with SGD optimizer and different activation functions (Relu, Tanh and Sigmoid) have better performance than with the other optimizers, being shown in figure 6a – 6c. On the other hand, in figure 6d - 6f, the architectures with RMSprop optimized and different activation functions have low performance when neurons are increased for cases with and without neuron dropout. It is also observed that the RMSprop optimizer with Tanh and Sigmoid have lower performance than the SGD optimizer and the activation functions studied.

In figure 6g, the architectures with Adam and Relu have similar performances to the RMSprop optimizer and Relu function; however, with the Tanh and Sigmoid activation functions these performances tend to increase when the neurons increase as shown in the figure 6h and figure 6i.



**Figure 6.** Error metrics for different architectures with SGD, RMSprop, Adam optimizer and Relu, Tanh and Sigmoid activation functions to different cases with and without neuron dropout.

4.4. Solar radiation forecast

Table 2 and table 3 show error metrics values of architectures evaluated with and without neuron dropout.

Table 2 shows that the SGD optimizers with different activation functions have low MSE, RMSE and MAE, and a high value of  $R^2$  compared to the other architectures. It is also observed that SGD and Sigmoid present an  $R^2$  with a value of 0.87761 greater than SGD using Tanh and Relu functions.

Table 3 shows different architectures with neuron dropout. These have high  $R^2$  when neurons are increased. It is also observed that the architecture (150, SGD, Sigmoid, 0.2) have the highest  $R^2$  among all architectures evaluated with and without neuron dropout.

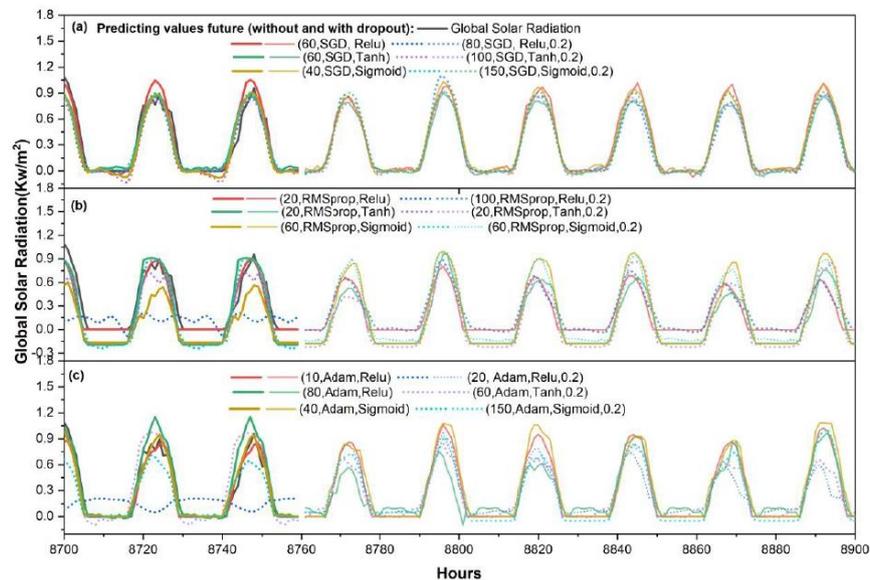
**Table 2.** Error metrics for different architectures without neuron dropout.

MLP architectures	MSE	RMSE	MAE	R <sup>2</sup>
(60, SGD, Relu)	0.01948	0.13956	0.07768	0.86782
(60, SGD, Tanh)	0.01885	0.13729	0.08561	0.87209
(40, SGD, Sigmoid)	0.01804	0.13429	0.08184	0.87761
(20, RMSprop, Relu)	0.02006	0.14162	0.07905	0.86388
(20, RMSprop, Tanh)	0.02378	0.15419	0.08369	0.83865
(60, RMSprop, Sigmoid)	0.02439	0.15618	0.08599	0.83447
(10, Adam, Relu)	0.01859	0.13636	0.07527	0.87382
(80, Adam, Tanh)	0.02371	0.15399	0.07855	0.83908
(40, Adam, Sigmoid)	0.02443	0.15631	0.08010	0.83419

**Table 3.** Error metrics for the different architectures with neuron dropout.

MLP architectures	MSE	RMSE	MAE	R <sup>2</sup>
(80, SGD, Relu, 0.2)	0.02063	0.14364	0.08479	0.85998
(100, SGD, Tanh, 0.2)	0.01892	0.13755	0.09041	0.87161
(150, SGD, Sigmoid, 0.2)	0.01697	0.13026	0.07744	0.88484
(100, RMSprop, Relu, 0.2)	0.08752	0.29584	0.19751	0.40602
(20, RMSprop, Tanh, 0.2)	0.02015	0.14196	0.07285	0.86324
(60, RMSprop, Sigmoid, 0.2)	0.01907	0.13808	0.08115	0.87061
(20, Adam, Relu, 0.2)	0.13837	0.37198	0.26685	0.06097
(60, Adam, Tanh, 0.2)	0.02527	0.15898	0.10314	0.82848
(150, Adam, Sigmoid, 0.2)	0.02311	0.15200	0.08946	0.84320

The best architectures of the model are used to train with the entire time series of solar radiation and make forecasts for the next 6 days, this is shown in figure 7. These architectures are taken from table 1 and 2, which show the highest performance among all the architectures evaluated. The thick lines of figure 7 show the architectures predictions on the test data without neuron dropout and the thin lines of the same color show the architectures predictions for the next 6 days. Similarly, the dotted lines describe the architectures predictions with neuron dropout and the dotted lines of the same faint color are the forecast for the same following days.



**Figure 7.** Forecast of global solar radiation for different architectures with SGD, RMSprop and Adam optimizers and their activation functions (Relu, Tanh and Sigmoid).

## 5. Conclusion

Solar radiation forecasting allows photovoltaic generation to be safe and reliable for an isolated or grid-connected hybrid generation system. Knowing the solar radiation in the future will allow the operators of these systems to give a greater percentage of participation to photovoltaic generation, reducing the operating costs, technical and environmental effects of the system.

In this article, the behavior of MLP was studied for its different architectures in relation to its 4 hyperparameters (optimizer, activation function, number of neurons in the hidden layer, neuron dropout) to predict the global solar radiation time series and their future values.

In the training and validation curves were found that when neuron dropout is added (dropout = 0.2) in the hidden layer, these curves tend to separate and their separation distances decrease or increase when their hyperparameters vary (number of neurons, type activation function and optimizer). The importance of introducing neuron dropout is to avoid overfitting the model and to have better performance. It was also found in the validation curves that the SGD optimizer with the three activation functions had very little peaks or noises, which indicates a good learning of the model. This good performance of the MLP model using the SGD optimizer can be seen in error metrics, predictions on the time series and its realized forecast.

In the performance of architectures without neuron dropout, it was found that the architectures (60, SGD, Sigmoid), (10, Adam, Relu) and (60, SGD, Sigmoid) have an  $R^2$  around 0.877, 0.873 and 0.872, respectively. For other cases, where neuron dropout was considered, the architecture (150, SGD, Sigmoid, 0.2) had a higher performance among all the architectures evaluated, giving an  $R^2$  value of 0.884.

The architectures with the SGD optimizer and three activation functions had more stable predictions in the future predictive values than other architectures.

As future work, it is proposed to use intelligent optimizers that allow to make a great combination of more hyperparameters, find the most optimal MLP architecture, and obtain a higher performance than we found. It is also suggested to work with greater time series data for the analysis, as well as to study their confidence intervals for future values.

#### Acknowledgments

This work is supported by the Doctoral Program in Sciences with mention in Energy from the National University of Engineering and the Peruvian National Council for Science and Technology (CONCYTEC) through Contract No. 207-2015-Fondecyt-UNI.

#### References

- [1] Zohuri B 2020 *Nuclear fuel cycle and decommissioning* (Albuquerque, NM, United States: Elsevier Ltd.)
- [2] Das U K, Tey K S, Seyedmahmoudian M, Mekhilef S, Idris M Y I, Van Deventer W, Horan B and Stojcevski A 2018 Forecasting of photovoltaic power generation and model optimization: A review *Renew. Sustain. Energy Rev.* **81** 912–28
- [3] Papadis E and Tsatsaronis G 2020 Challenges in the decarbonization of the energy sector *Energy* **205** 118025
- [4] Kannan N and Vakeesan D 2016 Solar energy for future world: - A review *Renew. Sustain. Energy Rev.* **62** 1092–105
- [5] IRENA 2019 *Future of solar photovoltaic: Deployment, investment, technology, grid integration and socio-economic aspects* vol November
- [6] Yin J, Molini A and Porporato A 2020 Impacts of solar intermittency on future photovoltaic reliability *Nat. Commun.* **11** 1–9
- [7] Blanco H and Faaij A 2018 A review at the role of storage in energy systems with a focus on Power to Gas and long-term storage *Renew. Sustain. Energy Rev.* **81** 1049–86
- [8] Trevizan R D, Headley A J, Geer R, Atcity S and Gyuk I 2021 Integration of energy storage with diesel generation in remote communities *MRS Energy Sustain.* **8** 57–74
- [9] Sengupta M, Habte A, Stoffel T, Perez R, Myers D, Gueymard C and Philippe Blanc 6 and Stefan Wilbert 2017 *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications: Second Edition*
- [10] Shadab A, Said S and Ahmad S 2019 Box–Jenkins multiplicative ARIMA modeling for prediction of solar radiation: a case study *Int. J. Energy Water Resour.* **3** 305–18
- [11] Nespoli A, Ogliari E, Leva S, Pavan A M, Mellit A, Lughì V and Dolara A 2019 Day-ahead photovoltaic forecasting: A comparison of the most effective techniques *Energies* **12** 1–15
- [12] Inman R H, Pedro H T C and Coimbra C F M 2013 Solar forecasting methods for renewable energy integration *Prog. Energy Combust. Sci.* **39** 535–76
- [13] Ristow D C M, Henning E, Kalbusch A and Petersen C E 2021 Models for forecasting water demand using time series analysis: a case study in Southern Brazil *J. Water, Sanit. Hyg. Dev.* **11** 231–40
- [14] Lima S, Gonçalves A M and Costa M 2019 Time series forecasting using Holt-Winters exponential smoothing: An application to economic data *AIP Conf. Proc.* **2186**
- [15] Voyant C, Notton G, Darras C, Fouilloy A and Motte F 2017 Uncertainties in global radiation time series forecasting using machine learning: The multilayer perceptron case *Energy* **125** 248–57
- [16] Meenu Sreedharan, Khedr A M and El Bannany M 2021 A Multi-Layer Perceptron Approach to Financial Distress Prediction with Genetic Algorithm *Autom. Control Comput. Sci.* **2020** 546 54 475–82
- [17] Jain Y K and Bhandare S K 2013 Min Max Normalization Based Data Perturbation Method for Privacy Protection *Int. J. Comput. Commun. Technol.* **4** 233–8
- [18] Vafaeipour M, Rahbari O, Rosen M A, Fazelpour F and Ansarirad P 2014 Application of sliding

- window technique for prediction of wind velocity time series *Int. J. Energy Environ. Eng.* **5** 1–7
- [19] Sugiartawan P, Pulungan R and Kartika A 2017 Prediction by a Hybrid of Wavelet Transform and Long-Short-Term-Memory Neural Network *Int. J. Adv. Comput. Sci. Appl.* **8**
- [20] Serin G, Sener B, Ugur Gudelek M, Ozbayoglu A M and Unver H O 2020 Deep multi-layer perceptron based prediction of energy efficiency and surface quality for milling in the era of sustainability and big data *Procedia Manuf.* **51** 1166–77
- [21] Matindife L, Sun Y and Wang Z 2020 A Machine-Learning Based Nonintrusive Smart Home Appliance Status Recognition *Math. Probl. Eng.* **2020**
- [22] Gardner M W and Dorling S R 1998 Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences *Atmos. Environ.* **32** 2627–36
- [23] Alboaneen D A, Tianfield H and Zhang Y 2017 Glowworm swarm optimisation for training multi-layer perceptrons *BDCAT 2017 - Proc. 4th IEEE/ACM Int. Conf. Big Data Comput. Appl. Technol.* 131–8
- [24] Mas J F and Flores J J 2008 The application of artificial neural networks to the analysis of remotely sensed data *Int. J. Remote Sens.* **29** 617–63
- [25] Banerjee C, Mukherjee T and Pasilio E 2020 The Multi-phase ReLU Activation Function *ACMSE 2020 - Proc. 2020 ACM Southeast Conf.* 239–42
- [26] Nwankpa C E 2020 Advances in optimisation algorithms and techniques for deep learning *Adv. Sci. Technol. Eng. Syst.* **5** 563–77
- [27] Khan M and Noor S 2019 Performance Analysis of Regression-Machine Learning Algorithms for Predication of Runoff Time *Agrotechnology* **08** 1–12
- [28] Watt N and du Plessis M C 2020 Dropout for Recurrent Neural Networks (Springer, Cham) pp 38–47