

**Modelo de Estimación de la Potencia Generada por un Sistema Fotovoltaico a partir
de Variables Provenientes de Bases de Datos de Recurso Energético Solar**

Edisson David Rosas Santacruz

Asesor:

Roberto Mauricio Cárdenas Cárdenas

Universidad Nacional Abierta y a Distancia – UNAD

Escuela de Ciencias Básicas Tecnología e Ingeniería – ECBTI

Maestría en Gestión de Tecnología de Información

2025

Nota de Aceptación

Presidente del jurado

Jurado

Jurado

Bogotá D.C., septiembre de 2025

Dedicatoria

En primer lugar, doy gracias a Dios y a María santísima por las bendiciones concedidas a lo largo de este proceso académico. Sin Su amor y fortaleza, nada de esto habría sido posible.

Agradezco a mis padres Margoth S y Vicente R por su constante aliento, comprensión y paciencia, los cuales han sido fundamentales para mi éxito académico. Ellos han sido mi principal fuente de inspiración y motivación a lo largo de mi vida.

Quiero expresar mi sincero agradecimiento a mi novia Lucero B, por su amor incondicional, paciencia infinita y apoyo constante a lo largo de este proceso académico.

Agradecimientos

Quiero agradecer al asesor Roberto Mauricio Cárdenas Cárdenas, por su acompañamiento y compromiso, por dedicar gran parte de su tiempo para guiarme en la construcción de este trabajo de grado.

Al Coasesor, Javier Revelo Fuelagan, quien desde su experiencia, conocimiento y motivación orientó el desarrollo de la investigación, por su grandiosa asesoría académica que sin duda fue el pilar fundamental en el planteamiento y culminación esta investigación.

Además, un especial agradecimiento al proyecto Campus Verde Udenar por facilitar el acceso y uso de los datos utilizados en la presente investigación, de igual manera, al Grupo de Investigación en Ingeniería Eléctrica y Electrónica – GIIEE- de la Universidad de Nariño.

Resumen

Este trabajo de investigación presenta el desarrollo de un modelo para estimar la potencia generada por un sistema fotovoltaico conectado a la red, empleando bases de datos de recurso solar y registros reales de producción fotovoltaica. Se selecciona la base de datos de SOLCAST de acuerdo con criterios de resolución temporal y espacial, cobertura geográfica y disponibilidad de variables climáticas. Con base en esta información, se propone un modelo de estimación basado en una red neuronal recurrente de tipo Long Short-Term Memory (LSTM), diseñada para capturar dependencias temporales a corto y largo plazo presentes en series temporales de irradiancia global horizontal (GHI), nubosidad y temperatura. La red LSTM utiliza como variable objetivo la potencia generada por un sistema fotovoltaico on-grid de 16,2 kWp. El modelo se entrena y valida utilizando datos correspondientes al año 2023, alcanzando un error porcentual promedio inferior al 2,74 %. Finalmente, el modelo propuesto se plantea como una herramienta de apoyo para procesos de planificación, diseño y evaluación de soluciones de energía solar para autoconsumo y generación distribuida.

Palabras clave: sistema fotovoltaico, potencia estimada, irradiancia, redes neuronales, generador distribuido

Abstract

This research presents the development of a model to estimate the power generated by a grid-connected photovoltaic system using solar resource databases and real power generation records. The SOLCAST database was selected based on criteria such as temporal and spatial resolution, geographical coverage, and availability of climate variables. Based on this information, an estimation model based on a Long Short-Term Memory (LSTM) recurrent neural network is proposed. The LSTM architecture is specifically designed to capture short- and long-term temporal dependencies in time series of global horizontal irradiance (GHI), cloud cover, and air temperature. The target variable corresponds to the power generated by a 16.2 kWp grid-connected photovoltaic system. The model is trained and validated using data from the year 2023, achieving an average percentage error below 2.74%. The proposed approach is presented as a useful tool for the planning and design of photovoltaic systems for self-consumption and distributed generation.

Keywords: photovoltaic system, estimated power, irradiance, neural networks, distributed generator

Tabla de Contenido

Introducción	12
Justificación.....	14
Objetivos	16
Objetivo General.....	16
Objetivos Específicos.....	16
Marco Conceptual y Teórico.....	17
Bases de Datos y Modelos de Estimación de Generación de Energía Solar.....	21
Tecnologías para Capturar Información del Recurso Solar.....	21
Bases de Datos de Recurso Solar.....	23
Análisis Comparativo de Bases de Datos de Recurso Solar	26
Modelos Basados en Datos para Estimar el Potencial de Generación Solar	29
Estimación de la Potencia Generada con Redes Neuronales Artificiales a partir de Series Temporales.....	30
Descripción y Funcionamiento del Modelo de Estimación de Producción Fotovoltaica	31
Resultados	35
Caso de Estudio del Sistema Fotovoltaico Instalado en la Universidad de Nariño.....	35
Análisis de las Series Temporales Provenientes SOLCAST.....	43
Análisis de la Estimación del Recurso Solar Haciendo Uso del Modelo Propuesto	51
Entrenamiento del Modelo.....	53
Análisis de los Resultados Obtenidos con el Modelo	58
Conclusiones y Trabajos Futuros	62

Referencias Bibliográficas.....64

Lista de Tablas

Tabla 1 <i>Criterios de Selección de las Bases de Datos de Recurso Solar</i>	26
Tabla 2 <i>Descripción de Variables del Inversor FRONIUS Symo 15kW</i>	38
Tabla 3 <i>Variables Usadas en la Investigación Provenientes de la Base de Datos de SOLCAST</i>	44
Tabla 4 <i>Parámetros Técnicos del Sistema Fotovoltaico del Caso de Estudio</i>	48
Tabla 5 <i>Métricas de Rendimiento del Modelo</i>	55
Tabla 6 <i>Predicción de Generación Fotovoltaica vs la Producción Fotovoltaica Real Registrada por el Inversor Fronius</i>	58
Tabla 7 <i>Producción Fotovoltaica Entregada por el Modelo Predictivo</i>	61

Lista de Figuras

Figura 1 <i>Arquitectura del Modelo de Estimación de la Producción Fotovoltaica</i>	32
Figura 2 <i>Sistema Fotovoltaico Instalado en la Universidad De Nariño</i>	36
Figura 3 <i>Inversor Fronius On-Grid de 15 kW</i>	37
Figura 4 <i>Proceso de Transformación y Limpieza de Producción Fotovoltaica</i>	40
Figura 5 <i>Potencia Generada el 1 de Julio del 2023</i>	41
Figura 6 <i>Potencia Generada por el Sistema Fotovoltaico</i>	42
Figura 7 <i>Ubicación Geográfica de la Instalación Fotovoltaica del Caso de Estudio</i>	43
Figura 8 <i>Flujo para la Limpieza de los Datos Originales de la Base de Datos SOLCAST</i>	
Figura 9 <i>Análisis del Comportamiento de las Variables Provenientes de SOLCAST para el 19 de Enero de 2023</i>	46
Figura 10 <i>Potencia Calculada a partir de la Irradiancia de SOLCAST vs la Potencia Medida</i>	49
Figura 11 <i>Potencia Calculada a partir de los Datos de GHI de SOLCAST vs la Producción Fotovoltaica Registrada por el Inversor Fronius Symo</i>	50
Figura 12 <i>Diagrama de Flujo de Datos para la Red Neuronal</i>	51
Figura 13 <i>Flujo para Crear Tensores</i>	52
Figura 14 <i>Flujo para el Entrenamiento y Validación del Modelo a partir de los Tensores de Entrada</i>	53
Figura 15 <i>Progreso del Entrenamiento de la Red Neuronal: Pérdida (MSE) y Precisión (MAE) en Función de las Épocas</i>	55

Figura 16 *Graficas de Dispersión: Conjunto de Datos de Prueba y Conjunto de Datos**Completo*57**Figura 17** *Comparación Entre la Energía Medida con la Energía Estimada por el**Modelo*.....59**Figura 18** *Producción Fotovoltaica Obtenida por el Modelo vs Producción Fotovoltaica**Real*60

Introducción

Con el creciente interés en las fuentes no convencionales de energía renovable (FNCR) en Colombia, especialmente en Nariño, es fundamental comprender la relación entre el recurso energético solar y el rendimiento de los sistemas fotovoltaicos. La irradiancia solar es la principal variable que determina el potencial energético del recurso solar, y a partir de sus medidas se puede estimar la eficiencia de conversión de energía solar a electricidad que presentan los equipos de los sistemas fotovoltaicos. De igual manera, los cambios en la variable de irradiancia ya sean por variaciones estacionales, condiciones climáticas o fenómenos atmosféricos, impactan directamente la producción de energía solar. Por otra parte, al momento de diseñar soluciones de energía se requiere estimar con gran exactitud el potencial de energía para definir con precisión los tiempos de retorno de inversión y generar más confianza en los inversionistas, de igual manera, cuando los sistemas fotovoltaicos se encuentran en operación, un modelo de estimación de producción fotovoltaica ayuda a determinar si se encuentra operando en el punto de mayor producción, con el fin de plantear estrategias de mantenimiento y reconfiguración de equipos.

Por lo anterior, en este trabajo de investigación se presenta el desarrollo de un modelo para estimar la potencia generada por un sistema fotovoltaico a partir de variables provenientes de bases de datos de recurso energético solar y registros de producción real. El modelo propuesto está basado en una red neuronal recurrente de tipo Long Short-Term Memory (LSTM), la cual permite modelar explícitamente la dependencia temporal de las series de datos solares. La arquitectura LSTM es especialmente adecuada para este tipo de aplicación, ya que incorpora mecanismos de memoria a corto y largo plazo que permiten capturar la dinámica temporal asociada a la variabilidad de la irradiancia solar y las condiciones climáticas. El modelo se

entrena y valida con datos de un sistema fotovoltaico real, demostrando su capacidad para estimar la producción de energía con buen desempeño.

Este trabajo no solo aborda una necesidad práctica en el ámbito de la energía renovable, sino que también contribuye al avance de la ciencia, el análisis de datos y la tecnología en el campo de la energía solar. Finalmente, este modelo puede ser utilizado en procesos de planeación cuando se requiere diseñar sistemas fotovoltaicos configurados como soluciones para autoconsumo o generadores distribuidos, además puede ser replicado y adaptado en otros escenarios de generación y ubicación.

Justificación

Actualmente, la demanda energética mundial continúa en aumento debido al crecimiento de los procesos productivos. Esta situación ha impulsado la implementación de soluciones basadas en energías renovables. En particular, los sistemas fotovoltaicos destacan por ser los de mayor crecimiento, gracias a su facilidad de adaptación a distintos espacios y al aprovechamiento del recurso solar. Para diseñar una solución energética con un retorno de inversión adecuado, resulta indispensable contar con modelos que permitan estimar la potencia generada por un sistema fotovoltaico. Esto facilita una planeación más objetiva y contribuye a garantizar la sostenibilidad del proyecto. En este sentido, el desarrollo de la presente investigación se justifica por la necesidad creciente de optimizar la generación de energía solar y mejorar la precisión en la estimación de su potencial energético.

En el escenario nacional e internacional, la adopción de energías renovables es fundamental para reducir la dependencia de combustibles fósiles y mitigar el cambio climático, en línea con el Objetivo 7 de los Objetivos de Desarrollo Sostenible. Sin embargo, la variabilidad inherente del recurso solar plantea desafíos en la planificación y estimación del potencial de generación de estos sistemas.

Esta investigación aborda una necesidad técnica y científica que contribuye al avance de la sostenibilidad energética y la transición hacia fuentes de energía más limpias y eficientes. La implementación de un modelo de predicción basado en algoritmos especializados para la estimación del potencial energético solar permitirá que los proyectos mejoren la previsión y gestión de la generación de energía, beneficiando tanto a consumidores como a proveedores de energía y aportando a lo establecido en el marco regulatorio nacional en transición energética.

La radiación solar constituye un factor determinante en la eficiencia de la conversión de energía solar en electricidad. Los cambios en la radiación solar ya sean atribuibles a variaciones estacionales, condiciones climáticas o fenómenos atmosféricos, inciden directamente en la producción de energía en los sistemas fotovoltaicos. Por lo tanto, investigar estas relaciones se torna esencial mejorar el rendimiento y la eficiencia de estos sistemas de generación de energía.

Además, el aprovechamiento de los recursos energéticos de la generación de energía con FNCER no solo fomentará la sostenibilidad, reduciendo la dependencia de fuentes no renovables, sino que también promoverá el aprovechamiento de fuentes no convencionales de energía, conforme a la transición energética contemplada en la ley 2099 de 2021. Es por esto, que el desarrollo de modelos de predicción de la producción fotovoltaica permitirá la aplicación de estrategias de gestión de energía teniendo en cuenta la variabilidad del recurso energético, mejorando la estabilidad y confiabilidad de los sistemas fotovoltaicos.

De este modo, la investigación propuesta no solo aborda una necesidad práctica en el ámbito de la energía renovable, sino que también contribuye al progreso de la ciencia, el análisis de datos y la tecnología en el sector de las energías renovables, en particular la energía solar.

Adicionalmente, el análisis y procesamiento de los datos del recurso energético solar permite determinar el potencial de generación de energía que puede ser comparado con la generación de sistemas fotovoltaicos reales, aspecto relevante que muy pocas investigaciones lo presentan. Con estos resultados se logra determinar si el sistema fotovoltaico está operando en su mejor configuración o si es necesario realizar actividades de mantenimiento prioritarias.

Objetivos

Objetivo General

Desarrollar un modelo de estimación de la potencia generada por un sistema fotovoltaico a partir de variables provenientes de bases de datos de recurso energético solar que permita la planificación en el diseño de soluciones fotovoltaicas

Objetivos Específicos

Identificar la base de datos de recurso solar que permita estimar la potencia de generación fotovoltaica en una determinada ubicación.

Proponer un modelo de estimación de la potencia generada por un sistema fotovoltaico mediante el uso de variables provenientes de bases de datos de recurso solar.

Analizar los resultados del modelo de estimación de potencia generada por un sistema fotovoltaico on-grid.

Marco Conceptual y Teórico

En el contexto actual, las energías renovables desempeñan un papel fundamental en la búsqueda de alternativas más sostenibles y amigables con el medio ambiente. Estas fuentes de energía se derivan de recursos disponibles en los territorios, que mediante la utilización de tecnología se puede construir generadores para autoconsumo o generadores distribuidos, convirtiendo a las fuentes renovables en opciones viables y prometedoras para abordar el agotamiento de los recursos y la mitigación de la crisis provocada por el cambio climático.

Entre las diversas formas de energías renovables se destaca la energía solar, que está presente en la mayor parte de la superficie terrestre y puede obtenerse aun con la presencia de nubes. Es de resaltar que, la velocidad a la que la Tierra intercepta la energía solar es aproximadamente 10.000 veces superior a la velocidad con la que la humanidad consume la energía.

En el proceso de conversión, los rayos solares, compuestos por fotones cargados de energía, inciden en una celda fotovoltaica, donde pueden ser reflejados o absorbidos. Cuando los fotones son absorbidos, transfieren su energía activando el flujo de electrones en el material semiconductor de la celda fotovoltaica y permitiendo que escapen de su posición inicial y generan una corriente eléctrica en un circuito conectado (Droege, 2011). Esta tecnología presenta numerosas ventajas, como su capacidad para generar electricidad de forma silenciosa y sin emisiones contaminantes, así como su modularidad y escalabilidad.

La tecnología fotovoltaica desempeña un papel fundamental en la transición energética hacia un modelo más sostenible, en particular, el fomento del uso de fuentes no convencionales de energía en Colombia se enmarca en la ley 1715 de 2014, que es considerada como la ley de energía renovables.

A pesar de que, según la Asociación Colombiana de Generadores, en 2021 Colombia se posicionó como la sexta matriz energética más limpia a nivel mundial, el administrador del mercado eléctrico colombiano, XM, revela que el Sistema Interconectado Nacional emitió alrededor de 9 millones de toneladas de CO₂ equivalente (tonCO₂eq) ese mismo año. Estos datos sugieren la necesidad de investigar y aprovechar las fuentes no convencionales de energías renovables (FNCER), en concordancia con el Plan Minero-Energético hacia el 2050.

La implementación de tecnologías como la fotovoltaica se alinea con la dinamización del mercado de energía y la reactivación económica del país, según lo establecido en la Ley de transición energética. Este enfoque, a su vez, contribuye al cumplimiento del acceso a una energía asequible y menos contaminante, como lo establecen los Objetivos de Desarrollo Sostenible de las Naciones Unidas.

Con respecto a las bases de datos del potencial energético solar, estas contienen principalmente información sobre la radiación solar de dos componentes principales: la irradiancia normal directa (DNI), que es el haz directo en un plano perpendicular a los rayos del sol, y la irradiancia difusa, generalmente medida en un plano horizontal (DHI - irradiancia horizontal difusa). La irradiancia total sobre una superficie horizontal se conoce como irradiancia horizontal global (GHI) expresada en W/m². En el caso de superficies inclinadas, como paneles fotovoltaicos o colectores solares térmicos, se debe tener en cuenta un tercer componente, que es la irradiancia solar reflejada en el suelo y en los objetos circundantes. Por otro lado, en el caso de sistemas de concentración solar, el componente de interés es el DNI, incluida la irradiancia circumsolar (Pereira et al., 2022).

Las mediciones terrestres de la irradiancia solar se realizan comúnmente mediante piranómetros para los componentes globales o difusos y mediante pirheliómetros para el

componente normal directo. Estas mediciones son útiles para evaluar el recurso solar en lugares específicos o regiones, utilizando observaciones desde diferentes ubicaciones (Feng et al., 2021).

Además de las observaciones terrestres, se han empleado datos satelitales (Polo et al., 2020) y modelos numéricos de predicción meteorológica (PNT) (Perdigão et al., 2017) para estudiar el recurso solar en regiones más extensas. Adicionalmente, el año meteorológico típico (TMY) consiste en series de datos a largo plazo para una ubicación determinada, concatenando meses meteorológicos típicos de mediciones reales que reproducen las estadísticas a largo plazo. Estos datos son útiles para la simulación energética de plantas de energía solar en edificios, ya sea para determinar su respuesta térmica bajo diferentes condiciones ambientales o para diseñar sistemas integrados de energía renovable. Un TMY generado para la evaluación del recurso solar está compuesto por la radiación solar y otras variables meteorológicas durante un período de un año (Lhendup & Lhundup, 2007), que generalmente incluye valores medios y extremos de temperatura, viento y humedad relativa.

Por otra parte, los modelos de estimación abarcan diversos enfoques metodológicos que han evolucionado a lo largo del tiempo. Entre ellos se encuentran los modelos determinísticos, basados en principios físicos y matemáticos, que calculan una magnitud a partir de ecuaciones exactas y bien definidas. También destacan los modelos estadísticos, que formulan relaciones entre variables mediante distribuciones de probabilidad o regresiones, ofreciendo aproximaciones útiles en contextos de variabilidad moderada.

En las últimas décadas, el avance del aprendizaje automático ha impulsado una nueva generación de modelos de estimación. A diferencia de los enfoques tradicionales, los algoritmos de aprendizaje automático no requieren definir explícitamente la relación entre las variables, sino

que aprenden directamente de los datos, capturando dependencias no lineales y patrones complejos que resultan difíciles de modelar con técnicas convencionales (Cerqueira et al., 2022)

En cuanto a los modelos para la estimación del recurso solar, estos se utilizan para predecir la cantidad de radiación solar que llega a la superficie terrestre en un lugar específico y durante un período de tiempo determinado. Estos modelos son fundamentales para la planificación y diseño de instalaciones solares, como plantas fotovoltaicas y térmicas, ya que proporcionan información sobre la disponibilidad y variabilidad de la radiación solar en una ubicación particular.

Existen varios tipos de modelos para la estimación del recurso solar, y su elección depende de la disponibilidad de datos, la complejidad del modelo y el propósito específico de la estimación (Teke et al., 2015). En particular, los modelos basados en ciencia de datos para la estimación del recurso solar son herramientas que utilizan técnicas y algoritmos de aprendizaje automático y estadísticos para predecir la radiación solar en una ubicación específica en función de la información contenida en bases de datos de recurso solar (Sehrawat et al., 2023). Entre estos enfoques destacan las redes neuronales recurrentes de tipo *LSTM (Long Short-Term Memory)*, diseñadas para el aprendizaje de secuencias y dependencias temporales mediante el uso de memoria a corto y largo plazo, lo que las convierte en una alternativa especialmente adecuada para modelar series temporales asociadas a la variabilidad de la radiación solar (Lan et al., 2021).

Finalmente, estos algoritmos, al trabajar de manera conjunta o individual, permiten desarrollar modelos de estimación del recurso solar más precisos, los cuales sirven de apoyo en procesos de planificación de proyectos de energía renovable, así como también en procesos de

operación y mantenimiento de centrales de generación, de esta manera se contribuye a la transición energética en el territorio nacional.

Bases de Datos y Modelos de Estimación de Generación de Energía Solar

En esta sección, se analizan las bases de datos de recurso solar requeridas para el desarrollo de la presente investigación. Además, se revisan los diferentes modelos de predicción presentados en la literatura que puedan ser útiles para realizar la predicción de generación energética a partir de series temporales. Estas series, provenientes de la base de datos identificada y seleccionada, se configuran como parámetros de entrada en un modelo de estimación de la potencia generada por un sistema fotovoltaico en un punto geográfico de la superficie terrestre.

Tecnologías para Capturar Información del Recurso Solar

La adquisición de datos precisos sobre el recurso solar es de gran importancia para la implementación de las tecnologías de energía solar. La tecnología fotovoltaica (PV) solar, que convierte directamente la radiación solar en electricidad, es un método destacado para aprovechar la energía solar (Rufangura & Sabah, 2018)

Se han desarrollado diversas tecnologías para hacer un uso efectivo de la energía solar, incluyendo la tecnología de iluminación solar, tecnología de energía térmica solar, tecnología fotovoltaica (PV) solar y tecnología de producción de hidrógeno solar (Zhang et al., 2018). Además, se ha explorado la integración de la tecnología PV solar con vehículos eléctricos e híbridos, lo que indica la versatilidad de las aplicaciones de la energía solar (Waseem et al., 2019). También han analizado la viabilidad y las implicaciones de usar tecnologías térmicas solares para satisfacer los requisitos energéticos de captura de CO₂, destacando el potencial de la energía solar para abordar desafíos ambientales (Deng et al., 2014).

La importancia de datos precisos del recurso solar es evidente en el desarrollo de un "Atlas Solar de Nueva Zelanda" usando imágenes satelitales, lo que enfatiza la necesidad de información sobre la irradiancia solar para la utilización efectiva de la energía solar (Brent et al., 2020). Además, el análisis de adaptabilidad regional de las tecnologías de utilización de techos solares en china demuestra la importancia de evaluar las tecnologías de utilización solar en regiones específicas para optimizar su efectividad (Li et al., 2022).

Los datos de recurso solar se capturan en satélites mediante diversos métodos y tecnologías. Los satélites geoestacionarios y en órbita polar se han utilizado para estimar la irradiancia solar a nivel del suelo, proporcionando conjuntos de datos que cubren casi la totalidad de la Tierra con una resolución espacial y temporal variable (Amillo et al., 2014) El uso de la energía solar depende del conocimiento sobre el comportamiento de la irradiancia solar en ubicaciones geográficas específicas, y numerosos estudios se han centrado en la viabilidad del recurso solar basado en datos obtenidos de estaciones de monitoreo y satélites meteorológicos (Ordoñez Palacios et al., 2022). Los modelos basados en satélites se han convertido en una herramienta para estimar el recurso solar con alta resolución espacial y temporal en grandes áreas geográficas (Amillo et al., 2018). Las tecnologías de última generación, incluyendo satélites meteorológicos y software avanzado, permiten la estimación del recurso solar con alta precisión y exactitud, considerando tanto las características climáticas como la topografía de una región (A. Pettazzi & S. Salsón, 2024).

La información de imágenes satelitales se procesa utilizando técnicas de computación paralela para estimaciones del recurso solar, demostrando la aplicación de tecnologías de procesamiento de imágenes en la captura de datos de recurso solar (Suárez & Nesmachnow, 2012). Además, la teledetección satelital de la radiación solar ofrece una mejor cobertura

espacial, y se han presentado métodos para el análisis de la irradiancia solar de imágenes satelitales y validarla con datos terrestres (Kaskaoutis & Polo, 2019). Además, se han desarrollado y modificado modelos de radiación solar basados en satélites para el cálculo de la irradiancia solar global superficial a partir de datos satelitales, mostrando el avance continuo en este campo (Janjai et al., 2013).

De esta manera, la captura de datos del recurso solar en satélites implica el uso de satélites geoestacionarios y en órbita polar, software avanzado y técnicas de procesamiento de imágenes. Estos métodos permiten la estimación del recurso solar con alta precisión y exactitud en grandes áreas geográficas, contribuyendo al avance de la utilización de la energía solar.

Bases de Datos de Recurso Solar

Las series temporales son secuencias de datos recopilados o registrados en intervalos regulares durante un período determinado. En el contexto del recurso solar, las bases de datos de series temporales permiten evaluar el potencial de la energía solar y su variabilidad a lo largo del tiempo. La disponibilidad de bases de datos de series temporales de alta calidad es importante para una evaluación del rendimiento de los proyectos fotovoltaicos solares, y se recomienda el uso de bases de datos de series temporales medidas a largo plazo o de alta resolución para la estimación del rendimiento energético de los proyectos fotovoltaicos (Purohit & Purohit, 2018). Además, técnicas de imputación de datos se han utilizado en situaciones donde hay datos faltantes de series temporales, que son requeridas para la predicción a largo plazo de la radiación solar (Jiménez-Valero et al., 2022).

La relevancia de estas bases de datos se enfatiza aún más por el hecho de que han llegado a ser el tipo de base de datos de más rápido crecimiento, indicando su creciente importancia y aplicabilidad (Calatrava et al., 2021). Estas bases de datos se utilizan para varios análisis, como

el análisis de series temporales de radiación solar registradas en diferentes estaciones, lo que puede proporcionar información valiosa sobre la variabilidad del recurso solar (Fortuna et al., 2016). Las bases de datos compuestas que combinan diferentes fuentes de datos de irradiancia solar permiten el acceso a datos de recurso solar para varias aplicaciones (Lefevre et al., 2014).

La complementariedad espacial y temporal del recurso solar y eólico es un aspecto clave que puede evaluarse utilizando datos de series temporales. Estudios han empleado series temporales horarias de radiación solar y velocidad del viento para evaluar esta complementariedad, lo cual es fundamental para la integración efectiva de energías renovables y la planificación de recursos (Jurasz et al., 2017). Además, la comparación y validación de bases de datos de recurso solar derivadas de satélites de código abierto destacan la importancia de contar con datos fiables y precisos para regiones geográficas específicas (Harsarapama et al., 2020).

En regiones con alto potencial de energía solar, como la región nororiental de Brasil, la evaluación de la variabilidad estacional y el análisis de tendencias del recurso de energía solar utilizando datos de series temporales, proporciona información complementaria del recurso solar, tanto espacial como estacional (Martins et al., 2019). De manera similar, la evaluación de la complementariedad energética en el tiempo entre diferentes recursos energéticos, como el viento y el solar, requiere el uso de datos de series temporales para un análisis preciso (Borba & Brito, 2017).

Con la base de datos de SOLCAST se han desarrollado diversos estudios orientados a la aplicación de modelos de estimación solar, gracias a su alta resolución temporal y espacial, así como a la disponibilidad de históricos y pronósticos de radiación. Estas características han permitido validar metodologías de predicción de generación fotovoltaica, optimización de

microrredes y análisis de integración de energías renovables en distintos contextos. Asimismo, su capacidad de proporcionar datos confiables y actualizados ha facilitado la construcción de modelos estadísticos y de inteligencia artificial para la estimación del recurso solar, consolidando a SOLCAST como una fuente de referencia en investigaciones académicas y proyectos de aplicación práctica en el sector energético.

Además, la base de datos SolarGIS se ha utilizado para evaluar los recursos locales de energía solar, y se han validado bases de datos derivadas de satélites sobre la radiación solar superficial en Brasil e Indonesia, demostrando su idoneidad para estudiar el recurso solar en regiones específicas (Thomas et al., 2016). La base de datos HelioClim-1 ha proporcionado datos diarios de radiación solar, indicando sus beneficios en el acceso a la información del recurso solar (Lefevre et al., 2014). Además, las mediciones realizadas en campo, mediante estaciones meteorológicas han sido utilizadas para determinar el recurso solar in-situ, que se complementan con datos satelitales. Estas mediciones también tienen otros propósitos que pueden ser aplicados a temas de gestión del riesgo, que dependiendo de las variables utilizadas se analizan modelos flujos de lodos, tormentas eléctricas, eventos sísmicos, entre otros (Eckman & Stackhouse, 2012).

En resumen, la disponibilidad y utilización de bases del recurso solar de series temporales son esenciales para diversas aplicaciones, incluyendo la evaluación del rendimiento de proyectos solares, la predicción a largo plazo del potencial energético, estudios de complementariedad energética y evaluación del potencial de recursos. Estas bases de datos juegan un papel en proporcionar información confiable y precisa para la toma de decisiones en el sector de energía renovable.

Análisis Comparativo de Bases de Datos de Recurso Solar

En el marco de esta investigación, se realizó un análisis comparativo entre diversas bases de datos del recurso solar, específicamente: NASA, SolarGIS, SOLCAST y NREL (Laboratorio Nacional de Energía Renovable). Este análisis se centró en evaluar la cantidad de información proporcionada, la granularidad de la información y la accesibilidad al histórico de datos. A continuación, se profundiza en la comparación según los criterios preestablecidos para determinar la base de datos más adecuada para evaluar el potencial energético solar.

Los criterios de selección utilizados en las bases de datos de recurso solar se agrupan en categorías relevantes para la investigación y se muestran en la Tabla 1.

Tabla 1

Criterios de selección de las bases de datos de recurso solar

Criterios de comparación de las bases de datos		
Categoría	Criterio	Descripción
Resolución	Temporal	Refiere a la frecuencia con la que se actualizan o registran los datos.
Resolución	Espacial	Se refiere al nivel de detalle o escala a la que se representan los datos geográficos.
Cobertura	Geográfica	Extensión territorial que abarcan los datos.
Usabilidad	Facilidad de uso	Facilidad con la que se pueden obtener y utilizar los datos.
Usabilidad	Utilización de los datos	Cómo y para qué se pueden usar los datos.
Mantenimiento	Base de datos	Cuidado y actualización continua de la base de datos.

Nota. Modelo LSTM propuesto para estimar la potencia fotovoltaica a partir de GHI, temperatura y nubosidad. Elaboración propia (2025).

La elección de SOLCAST como base de datos de recurso solar se soporta en la combinación de resolución temporal, resolución espacial, cobertura geográfica, usabilidad y mantenimiento. En términos de resolución temporal, SOLCAST ofrece datos horarios e incluso en intervalos de 5 a 10 minutos, lo que proporciona un nivel de detalle superior para análisis dinámicos de generación solar y para la integración con modelos de predicción energética o sistemas de gestión de microrredes. En la resolución espacial, SOLCAST trabaja con modelos satelitales y algoritmos avanzados que permiten representar la radiación solar a una escala muy detallada, incluso a nivel de proyectos individuales, lo que resulta más útil que las aproximaciones más generales de NASA o NREL.

En cuanto a la cobertura geográfica, aunque bases como NASA y SolarGIS también tienen alcance global, SOLCAST destaca porque ofrece datos adaptados a condiciones locales mediante algoritmos de nowcasting y predicción que permiten personalizar los resultados para ubicaciones específicas. Desde el punto de vista de la usabilidad, SOLCAST dispone de plataformas de acceso rápido vía API, con interfaces modernas y documentación clara, lo que facilita la integración de sus datos en aplicaciones prácticas de investigación, planeación de proyectos y simulaciones energéticas. Respecto a la utilización de los datos, su principal ventaja es la capacidad de generar pronósticos solares a corto y mediano plazo con alta precisión, característica que es esencial para estudios de operación de microrredes, gestión de energía distribuida y evaluación de proyectos fotovoltaicos.

En cuanto al mantenimiento, SOLCAST se caracteriza por una actualización continua de sus bases de datos, con algoritmos de aprendizaje automático que mejoran de forma progresiva la exactitud de las estimaciones, garantizando un recurso confiable y vigente. En síntesis, la elección de SOLCAST se justifica porque combina precisión, detalle temporal y espacial,

facilidad de integración tecnológica y capacidad de predicción, aspectos que lo hacen más robusto y aplicable en proyectos actuales de energía solar respecto a NASA, SolarGIS y NREL.

Después de analizar cada base de datos según los criterios mencionados, se concluyó que la base de datos de SOLCAST es la que mejor se ajusta a los requisitos de este proyecto. Los factores decisivos para esta elección fueron:

1. SOLCAST utiliza una combinación de imágenes de alta resolución de los satélites GOES 16 y 18, Meteosat-09, 10 y 11, e Himawari 9, junto con modelos de predicción numérica del tiempo (NWP), para ofrecer datos de irradiancia solar con una granularidad de 5 minutos. Este enfoque permite un seguimiento de las condiciones de nubosidad y la estimación de la irradiancia solar en una ubicación geográfica en particular.

2. La metodología de SOLCAST incluye el uso de algoritmos propios y modelos estándar para procesar las imágenes satelitales y estimar la nubosidad. Además, los datos se someten a control de calidad por parte de SOLCAST para garantizar su precisión y fiabilidad.

3. Estos datos de alta resolución, junto con la información adicional sobre temperatura y nubosidad, son importantes para validar modelos de predicción de generación de energía solar y simular el comportamiento del recurso solar. La interfaz intuitiva de SOLCAST y su completo soporte técnico facilitan el acceso y análisis de datos, convirtiéndolos en una herramienta ampliamente utilizada entre profesionales de la industria solar.

De esta manera, la base de datos de SOLCAST cumple con todos los criterios propuestos en la investigación, especialmente en términos de resolución temporal y la disponibilidad de datos climáticos adicionales. Estos atributos hacen de la base de datos de SOLCAST, la herramienta más adecuada para la presente investigación, permitiendo realizar análisis detallados y con fundamentos teóricos requeridos para evaluar el potencial energético solar.

Modelos Basados en Datos para Estimar el Potencial de Generación Solar

Para estimar el potencial de generación solar, se han desarrollado y aplicado diversos modelos basados en datos. Estos modelos utilizan algoritmos de aprendizaje automático para estimar el potencial de energía solar. Por ejemplo, en (Golestaneh et al., 2016) emplearon una Máquina de Aprendizaje Extremo (ELM) para el pronóstico de la generación de energía solar, mientras que (Yu et al., 2018) desarrollaron un modelo de aprendizaje automático, DeepSolar, para construir una base de datos de despliegue solar en los Estados Unidos. Además, en (Khodayar et al., 2020) utilizaron algoritmos de aprendizaje automático como aproximadores de la función objetivo para el pronóstico espaciotemporal de la irradiancia solar, y (Aslam et al., 2019) emplearon modelos de aprendizaje profundo para el pronóstico a largo plazo de la radiación solar para estimar los potenciales energéticos influenciados por la tasa de degradación de los paneles fotovoltaicos.

Por otra parte, en investigaciones se presentan un modelo predictivo diseñado para estimar la disposición de instalaciones solares en áreas de estudio con censos demográficos, en (Khatri et al., 2022) utilizaron el Modelo Asesor del Sistema de NREL (SAM) para estimar el potencial geográfico y técnico de la energía solar fotovoltaica. Estos estudios demuestran la diversa gama de modelos basados en datos, incluyendo aprendizaje automático, aprendizaje profundo y pronóstico basado en la teoría de probabilidad, que pueden utilizarse para estimar el potencial de generación solar.

De igual manera, en (Fuentes et al., 2020) propusieron un método para estimar el potencial de energía solar basado en la fotogrametría de vehículos aéreos no tripulados, mientras que en (Kulkarni et al., 2023) utilizaron regresión polinomial, un algoritmo ANN y métodos basados en píxeles para estimar la producción de energía solar sobre superficies de edificios.

Además, en (Park et al., 2021) demostraron el uso de métodos de aprendizaje automático, como cámaras orientadas al cielo con aprendizaje automático, para estimar la producción de energía solar, mientras que en (Chow et al., 2011) enfatizaron la necesidad de un enfoque eficiente para la modelización con un nivel adecuado de detalle y menor tiempo computacional en la estimación del potencial de energía solar con alta resolución espaciotemporal.

En general, la literatura presenta una amplia variedad de modelos basados en datos, incluyendo técnicas de aprendizaje automático, aprendizaje profundo, fotogrametría y técnicas de regresión, que pueden utilizarse para estimar el potencial de generación solar. Sin embargo, se resalta la necesidad de investigar en el desarrollo de algoritmos basados en redes neuronales artificiales que contemplen el uso de datos de sistemas fotovoltaicos implementados.

Estimación de la Potencia Generada con Redes Neuronales Artificiales a partir de Series Temporales

La predicción del recurso solar a partir de series temporales ya sea a través de medidores en campo o datos satelitales, se maneja principalmente mediante redes neuronales artificiales. Por ejemplo, en (Praynlin & Jenson, 2017) aplicaron redes neuronales artificiales para estimar la radiación solar, resaltando el potencial de este enfoque en el ámbito de los sistemas fotovoltaicos. De manera similar, (Husein & Chung, 2019) emplearon Redes Neuronales Recurrentes (RNN) de Memoria a Corto y Largo Plazo para la previsión de irradiancia solar diaria, demostrando cómo el aprendizaje profundo puede ser aplicado. Además, en (Madhiarasan et al., 2021) destacaron el análisis de rendimiento de un perceptrón multicapa para analizar el potencial energético solar.

En (Voyant et al., 2017) proporcionaron una revisión de los métodos de aprendizaje automático para la previsión de radiación solar, destacando la creciente utilización de redes neuronales artificiales en el análisis de datos y la predicción en meteorología.

Varios estudios han demostrado la versatilidad de las redes neuronales artificiales en diferentes aspectos de la predicción de energía solar. Por ejemplo, en (Khatib & Elmenreich, 2015) desarrollaron un modelo para la generación de datos de radiación solar horaria utilizando una red neuronal artificial de regresión generalizada. En (Gopinath & Prince, 2018) aplicaron redes neuronales para la predicción de la evolución futura de los ciclos solares, mostrando las diversas aplicaciones de estas redes en la investigación de la energía solar.

Adicionalmente, para desarrollar un modelo de red neuronal secuencial, es esencial considerar las capacidades y características de diferentes tipos de redes neuronales recurrentes. El modelo Codificador-Decodificador RNN propuesto por (Cho et al., 2014) introduce un enfoque novedoso que utiliza dos RNN para aprender representaciones de frases, con datos de naturaleza secuencial. Esto es especialmente relevante para el procesamiento de datos secuenciales en la predicción solar, donde la precisión y la capacidad de predicción son importantes para realizar pronósticos a largo plazo.

En resumen, las redes neuronales artificiales, incluyendo modelos como RNN y LSTM, se han consolidado como herramientas fundamentales para la predicción del recurso solar. Estas tecnologías permiten abordar la variabilidad relacionada al recurso solar, facilitando así la integración de datos de múltiples fuentes en aplicaciones de energía solar.

Descripción y Funcionamiento del Modelo de Estimación de Producción Fotovoltaica

En esta sección se presenta un modelo de estimación de la producción fotovoltaica basado en una red neuronal recurrente de tipo Long Short-Term Memory (LSTM), cuya

arquitectura se muestra en la Figura 1. El modelo está diseñado para procesar series temporales multivariadas, preservando el orden temporal de los datos y permitiendo el aprendizaje de dependencias dinámicas en el tiempo. Las entradas del modelo corresponden a series temporales de irradiancia global horizontal (GHI), temperatura del aire y nubosidad, obtenidas de la base de datos SOLCAST, mientras que la salida corresponde a la potencia generada por un sistema fotovoltaico real. A diferencia de los modelos densamente conectados, la red LSTM no requiere el aplanamiento de los datos, ya que opera directamente sobre tensores tridimensionales con la forma (muestras, pasos de tiempo, variables).

Figura 1

Arquitectura del Modelo de Estimación de la Producción Fotovoltaica



Nota. El modelo LSTM estima la potencia del sistema fotovoltaico a partir de irradiancia, temperatura y nubosidad obtenidas de SOLCAST. Elaboración propia.

En este proceso, el modelo busca estimar la producción de energía solar mediante el uso de redes neuronales recurrentes de tipo Long Short-Term Memory (LSTM), las cuales están diseñadas para procesar información secuencial y capturar dependencias temporales en series de datos. En este caso, los datos de entrada se estructuran como series temporales multivariadas, conservando explícitamente el orden temporal de las observaciones y evitando transformaciones que eliminen dicha dependencia, como el aplanamiento de los tensores.

Cada muestra de entrada corresponde a una secuencia temporal diaria con una granularidad de 5 minutos, conformada por múltiples pasos de tiempo y variables asociadas al recurso energético solar, tales como irradiancia global horizontal, temperatura y nubosidad. Esta estructura permite que la red LSTM aprenda tanto patrones de corto plazo como tendencias de largo plazo relacionadas con la variabilidad de la producción fotovoltaica.

Una vez organizadas las series temporales, los datos se dividen en dos conjuntos: uno destinado al entrenamiento del modelo y otro para su validación. En esta investigación, se utilizó el 5 % del conjunto total de datos para la validación, garantizando que la mayor parte de la información esté disponible para el aprendizaje de los parámetros del modelo, al tiempo que se evalúa su capacidad de generalización.

La arquitectura del modelo está conformada por dos capas LSTM con 50 unidades cada una, seguidas de una capa densa de salida con activación lineal, las cuales incorporan mecanismos internos de memoria mediante compuertas de entrada, olvido y salida, permitiendo modelar dependencias temporales complejas presentes en las series de datos solares. Posteriormente, se emplea una capa de salida, encargada de transformar la representación temporal aprendida por la red en una estimación de la potencia fotovoltaica generada. Para las capas LSTM se utiliza la función de activación interna característica de esta arquitectura, mientras que la capa de salida emplea una activación lineal, adecuada para problemas de regresión.

El proceso de compilación del modelo se realiza configurando el optimizador Adam con una tasa de aprendizaje de 0.001, y se selecciona el error cuadrático medio como función de pérdida, con el fin de cuantificar la desviación entre las estimaciones del modelo y los valores

reales durante el entrenamiento. Como métrica de evaluación del desempeño se utiliza el error absoluto medio, el cual permite interpretar de forma directa la magnitud del error de predicción.

El entrenamiento de la red se lleva a cabo a través de múltiples iteraciones, conocidas como épocas. En cada época, el modelo ajusta sus parámetros internos mediante retro propagación en el tiempo, con el objetivo de minimizar la función de pérdida. De manera simultánea, se realiza la validación utilizando el conjunto de datos reservado, lo que permite monitorear el proceso de aprendizaje y evaluar la capacidad del modelo para generalizar a datos no vistos.

Finalmente, una vez concluido el entrenamiento, el modelo se evalúa empleando el conjunto de datos de prueba, obteniendo métricas de desempeño que reflejan su capacidad de estimación. Los valores obtenidos para la función de pérdida y el error absoluto medio evidencian que el modelo LSTM entrenado presenta resultados confiables y adecuados para la estimación de la producción fotovoltaica a partir de series temporales de recurso solar.

Resultados

En esta sección, se presentan los resultados de la evaluación del desempeño del modelo de estimación propuesto. Dicho modelo se entrena utilizando series temporales de irradiancia Global Horizontal (GHI), nubosidad y temperatura, conforme a la información suministrada por la base de datos de SOLCAST. Se toma como referencia la potencia generada por un sistema fotovoltaico de 16,2 kWp situado en la Universidad de Nariño. Este sistema es empleado como caso de estudio para verificar la exactitud del modelo.

Caso de Estudio del Sistema Fotovoltaico Instalado en la Universidad de Nariño

El sistema fotovoltaico de la Universidad de Nariño, ubicado en la sede Torobajo de la ciudad de Pasto, Nariño, Colombia, está compuesto por 30 paneles solares fotovoltaicos de 540 W de potencia pico, instalados en la terraza del Bloque 1 Sur. Este bloque cuenta con una amplia terraza y una subestación eléctrica propia que opera con un transformador de distribución de 225 kVA.

Como se muestra en la Figura 2, a través de una perspectiva aérea, los paneles solares se colocaron en tres zonas específicas de la terraza para preservar la estética y funcionalidad del espacio. Se distribuyeron equitativamente, con 20 paneles solares en cada zona. En cuanto a la configuración, los paneles se interconectaron en cuatro cadenas de 15 unidades cada una, concentrando la energía en dos inversores On-Grid de 15 kW, donde cada inversor recibe la energía de dos cadenas de paneles solares.

Figura 2

Sistema Fotovoltaico Instalado en la Universidad De Nariño



Nota. Sistema fotovoltaico instalado en la Universidad de Nariño, sede Torobajo. Elaboración propia (2025).

El elemento central de cada sistema fotovoltaico es el inversor solar on-grid de 15 kW, que se muestra en la Figura 3. Este inversor tiene la capacidad de sobredimensionamiento hasta un 150 %, por lo que puede gestionar una potencia pico de 16,2 kWp, distribuida en dos cadenas de 15 paneles. Estos inversores, junto con los tableros de protección, están instalados en el cuarto técnico situado en la terraza del bloque 1 sur de la Universidad de Nariño.

Para monitorear el flujo de potencia desde las cadenas de paneles solares, se colocó un tablero en corriente continua (DC) que alberga dos barrajes independientes para positivo (PV+) y negativo (PV-), fusibles en DC y dispositivos de protección y desconexión tanto de entrada como de salida en DC. La Figura 3 muestra el inversor on-grid de 15 kW del fabricante Fronius. La potencia procedente del tablero DC es conducida a los puertos de conexión del inversor, el cual transforma la tensión de DC a AC trifásica en nivel de tensión de 208 V a una frecuencia de 60 Hz.

Figura 3

Inversor Fronius On-Grid de 15 kW



Nota. Inversor solar on-grid de 15 kW utilizado en el caso de estudio. Elaboración propia (2025).

La salida de cada inversor se conecta al tablero de protecciones en corriente alterna (AC), que integra los barrajes de las 3 fases, el neutro y la tierra, un interruptor principal, un medidor de energía integrado al inversor, protecciones en AC y dispositivos de protección contra sobretensiones (DPS) en AC.

En cuanto al monitoreo, los dos sistemas fotovoltaicos on-grid son supervisados en tiempo real. Para ello, se instaló una red de fibra óptica con un convertidor de medios, estableciendo una red LAN y WLAN a través de un Access Point. Cada inversor se conecta directamente a este Access Point mediante un cable ethernet, lo que permite una transmisión de datos fiable y en tiempo real de las variables registradas por el inversor On Grid y el medidor bidireccional.

De acuerdo con el manual de usuario, se configuraron todos los parámetros de los inversores para la inyección de potencia en la red y la transmisión de datos a la plataforma web de Fronius, accesible en <https://www.solarweb.com/>.

Para el caso de estudio de esta investigación, se recolectaron datos del inversor 1 a lo largo del año 2023 desde el sitio web del fabricante. Las variables reportadas por el inversor se describen en la Tabla 2 las cuales tienen una granularidad de 5 minutos.

Tabla 2

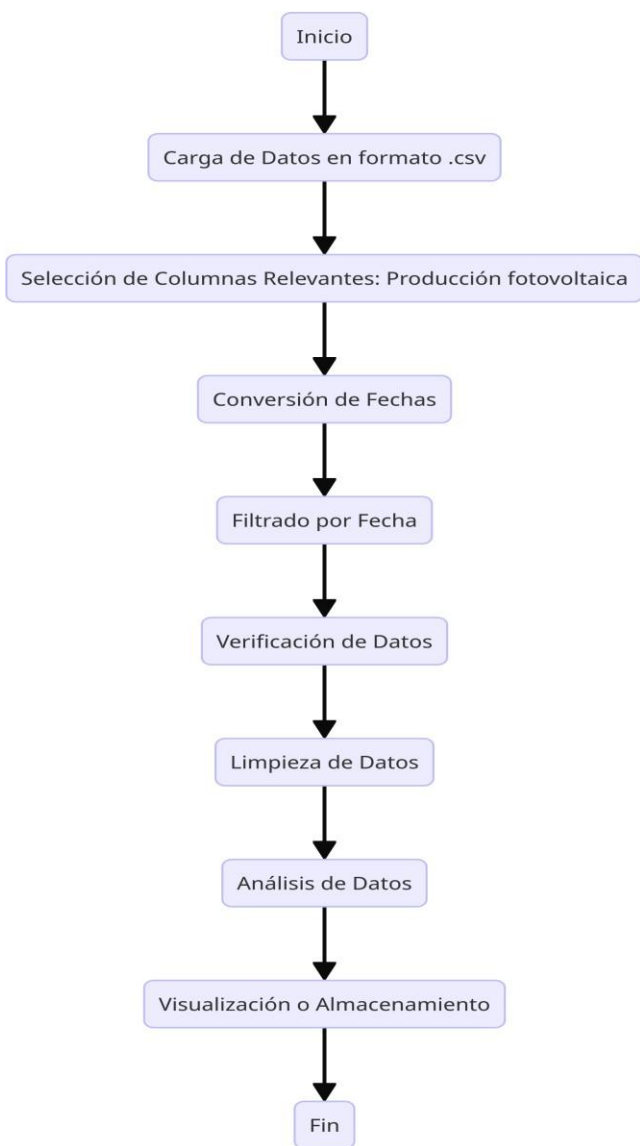
Descripción de Variables del Inversor FRONIUS Symo 15kW

Parámetro	Descripción de variables del inversor FRONIUS Symo 15kW
Fecha y hora	Tiempo específico de registro de datos
Corriente CA L1	Corriente de línea 1 en corriente alterna
Corriente CA L2	Corriente de línea 2 en corriente alterna
Corriente CA L3	Corriente de línea 3 en corriente alterna
Corriente CC MPP1	Corriente del primer punto de máxima potencia en corriente continua
Corriente CC MPP2	Corriente del segundo punto de máxima potencia en corriente continua
Energía	Energía total generada
Energía MPP1	Energía generada por el primer punto de máxima potencia
Energía MPP2	Energía generada por el segundo punto de máxima potencia
Factor de potencia	Relación entre la potencia real y la potencia aparente generada
Potencia aparente	Potencia total generada sin considerar las pérdidas de energía
Potencia reactiva	Potencia usada para crear y mantener los campos eléctricos y magnéticos
Rendimiento específico	Eficiencia de la generación de energía con relación a la capacidad instalada
Tensión CA L1-L2	Tensión entre las líneas 1 y 2 de corriente alterna
Tensión CA L2-L3	Tensión entre las líneas 2 y 3 de corriente alterna
Tensión CA L3-L1	Tensión entre las líneas 3 y 1 de corriente alterna
Tensión CC MPP1	Tensión en el primer punto de máxima potencia en corriente continua
Tensión CC MPP2	Tensión en el segundo punto de máxima potencia en corriente continua
Consumido directamente	Energía consumida directamente sin almacenamiento
Consumo	Energía total consumida
Energía guardada en la batería	Energía almacenada en la batería para uso posterior
Energía obtenida de la batería	Energía consumida proveniente de la batería
Energía obtenida de la red	Energía consumida proveniente de la red eléctrica

Parámetro	Descripción de variables del inversor FRONIUS Symo 15kW
Energía suministrada a la red	Energía entregada de vuelta a la red eléctrica
Producción fotovoltaica	Energía total generada por los paneles solares

Nota. Variables eléctricas registradas por el inversor fotovoltaico con resolución de 5 minutos durante 2023. Elaboración propia (2025).

En la arquitectura propuesta para el modelo de estimación de producción fotovoltaico, se tiene como entrada las series temporales de recurso energético solar de la base de datos SOLCAST y la producción del sistema fotovoltaico registrada por un medidor. Para tener un modelo de estimación de producción más preciso se aplica una fase de limpieza de datos que se muestra en la figura 4.

Figura 4*Proceso de Transformación y Limpieza de Producción Fotovoltaica*

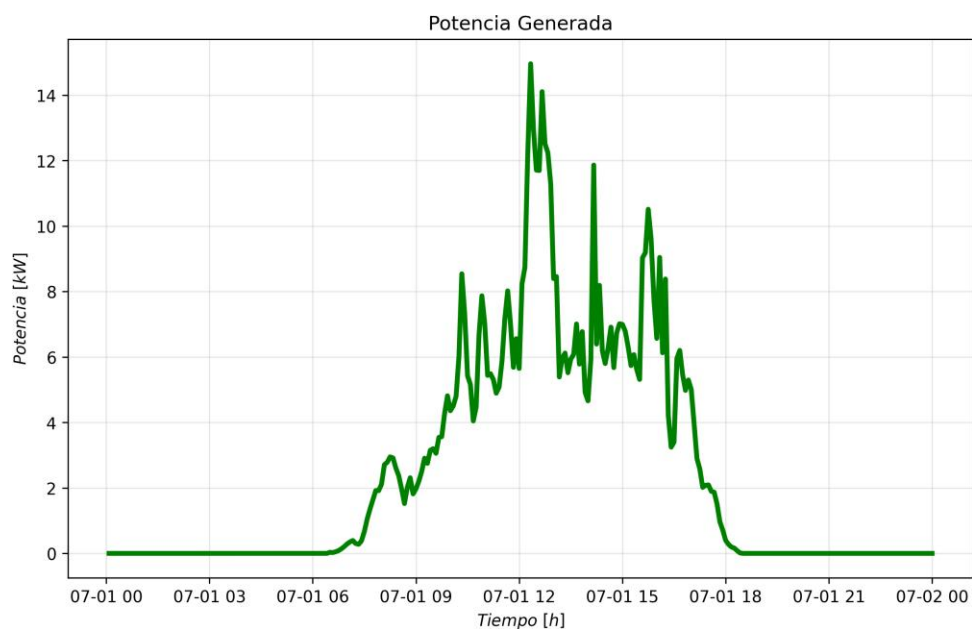
Nota. Proceso de limpieza y transformación de los datos de producción fotovoltaica. Elaboración propia (2025).

Una vez depurados los datos, se implementa un algoritmo en Python para llevar a cabo un análisis visual de la información. La Figura 5 muestra la potencia generada por el inversor

Fronius Symo el primero de julio de 2023, representando en el eje vertical la potencia expresada en [kW] y en el eje horizontal, el tiempo expresado en horas [h] para un día completo.

Figura 5

Potencia Generada el 1 de Julio del 2023



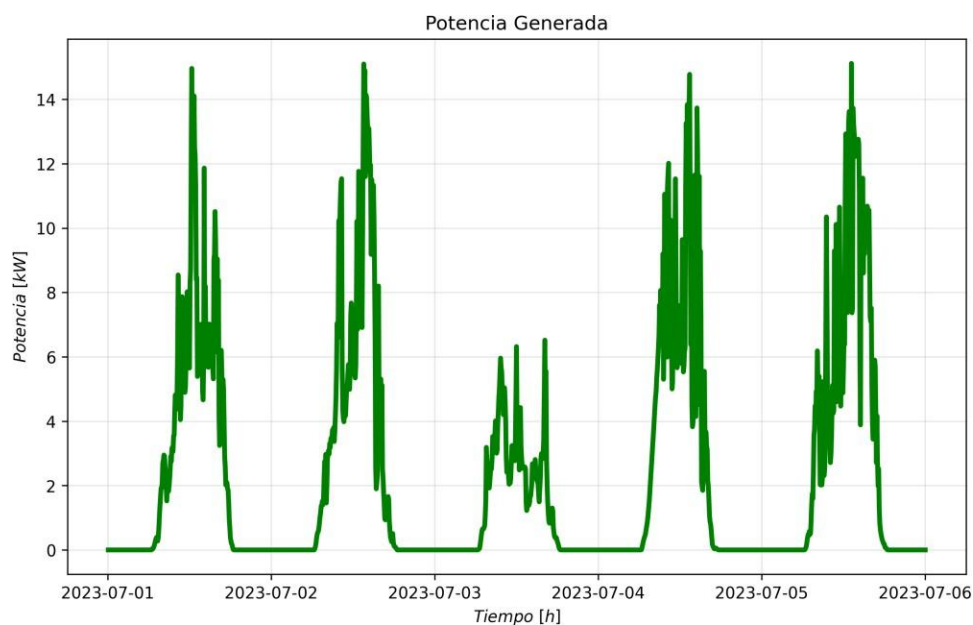
Nota. Perfil diario de potencia fotovoltaica correspondiente al 1 de julio de 2023. Elaboración propia (2025).

La producción de potencia inicia en 0 kW, lo cual es de esperar durante las horas nocturnas antes del amanecer. A medida que transcurre la mañana, se observa un incremento en la producción de energía, alcanzando picos máximos alrededor del mediodía. La presencia de picos pronunciados en el gráfico comprende las variaciones en la generación de energía por la presencia de nubosidad intermitente. Hacia la tarde, la producción disminuye gradualmente, siguiendo la pauta de la intensidad solar decreciente. Al aproximarse la noche, la producción se reduce nuevamente a cero. El patrón observado en la figura es característico de un sistema

fotovoltaico, con su punto más alto de generación energética durante las horas de máxima insolación y sin producción durante la noche.

Figura 6

Potencia Generada por el Sistema Fotovoltaico



Nota. Potencia fotovoltaica registrada durante cinco días consecutivos de julio de 2023.

Elaboración propia (2025).

La Figura 6 presenta la potencia generada para 5 días del mes de julio de 2023. Es de notar que, aunque la producción fluctúa diariamente, existe una tendencia general de altos valores de producción, lo que indica que, en gran parte de los días, se generó una cantidad considerable de energía.

La variabilidad observada en cada día se debe a cambios en la irradiancia solar, influenciada por condiciones climáticas variables como días nublados o soleados. La regularidad en la obtención de valores elevados de producción de energía para la mayor de los días sugiere

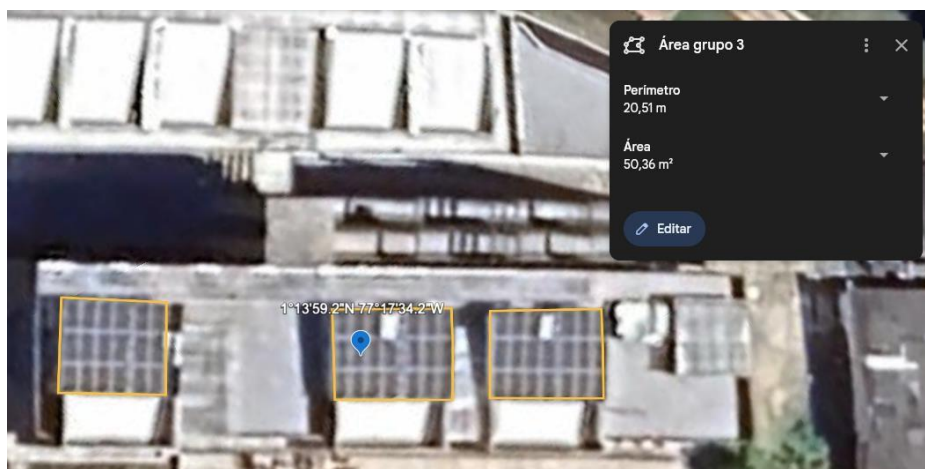
una correcta operación del sistema. Además, indica que el sistema fotovoltaico produce una energía relativamente uniforme, lo cual es característico en meses de estaciones con mayor insolación o en regiones con que el clima es predominantemente despejado.

Análisis de las Series Temporales Provenientes SOLCAST

Después de identificar la base de datos más apropiada para los fines de esta investigación, se descargan los datos del sitio web oficial de SOLCAST (<https://SOLCAST.com/>) correspondientes a la ubicación específica del sistema, como se muestra en la figura 7.

Figura 7

Ubicación Geográfica de la Instalación Fotovoltaica del Caso de Estudio



Nota. Ubicación geográfica del sistema fotovoltaico del caso de estudio. Elaboración propia (2025).

Cabe resaltar que SOLCAST permite descargar un gran número de variables ambientales desde su base de datos, con la posibilidad de configurar la granularidad de la información. Para esta investigación, se seleccionó una granularidad de 5 minutos

Los datos utilizados para entrenar el modelo contienen las variables descritas en la tabla 3, las cuales fueron descargadas de la base de datos seleccionada:

Tabla 3*Variables Usadas en la Investigación Provenientes de la Base de Datos de SOLCAST*

Variable	Descripción
air_temp	Temperatura del aire a 2 metros por encima del nivel de la superficie.
clearsky_ghi	Irradiancia total en una superficie horizontal si no existieran nubes.
clearsky_gti	Irradiación total recibida en una superficie si no hay nubes, con inclinación y azimuth (suma de componentes directos, difusos y reflejados) fijo o de seguimiento.
cloud_opacity	Atenuación de la luz solar debido a la nube.
dewpoint_temp	Temperatura del punto de rocío a 2 metros sobre el nivel de la superficie. Se alcanza cuando el aire se enfría a una temperatura que satura el vapor de agua y forma rocío.
gti	Irradiancia total recibida sobre una superficie con inclinación y azimuth definidos (suma de componentes directos, difusos y reflejados). Fijos o de seguimiento.
dhi	Diffuse Horizontal Irradiance (Irradiancia Horizontal Difusa) en W/m ² : cantidad de radiación solar que recibe una superficie horizontal desde el cielo, excluyendo la radiación solar directa del sol.
dni	Direct Normal Irradiance (Irradiancia Normal Directa) en W/m ² : cantidad de radiación solar que recibe una superficie perpendicular a los rayos del sol, es decir, radiación directa sin reflejos ni difusión.
ghi	Global Horizontal Irradiance (Irradiancia Horizontal Global) en W/m ² : total de radiación solar directa y difusa que llega a una superficie horizontal. Es la suma de DNI proyectado en el plano horizontal más DHI.
Clear GHI	Clear Sky Global Horizontal Irradiance (Irradiancia Horizontal Global de Cielo Despejado) en W/m ² : representa la cantidad de GHI que se esperaría bajo condiciones de cielo claro, sin nubes. Es útil para comparar con el GHI actual para entender el impacto de la cobertura nubosa.
period_end	Momento en el que termina el periodo de medición, generalmente expresado en UTC.
period	Duración del periodo de medición, que en este caso es PT5M, indicando un intervalo de 5 minutos entre cada punto de dato.

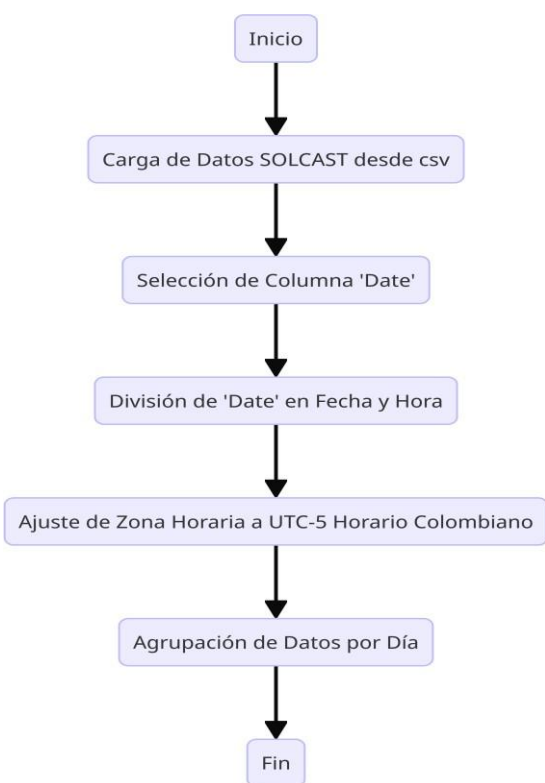
Nota. Variables ambientales descargadas desde SOLCAST con resolución de 5 minutos.

Elaboración propia (2025).

Una vez almacenada la información, se procede a la limpieza de los datos siguiendo el flujo descrito a continuación:

Figura 8

Flujo para la Limpieza de los Datos Originales de la Base de Datos SOLCAST

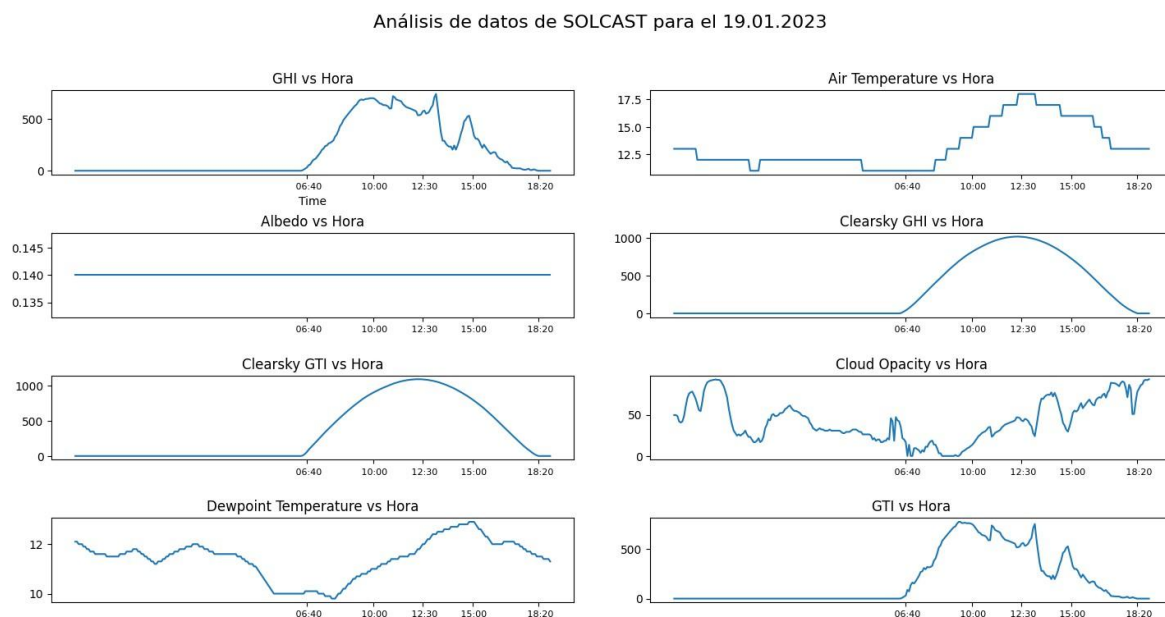


Nota. Flujo de limpieza aplicado a los datos descargados de SOLCAST. Elaboración propia (2025).

Una vez limpiados los datos, se lleva a cabo un análisis visual para examinar el comportamiento temporal de las variables de la base de datos. Para ilustrar este análisis, se utiliza como ejemplo un día de referencia, como se muestra en la Figura 9.

Figura 9

Análisis del Comportamiento de las Variables Provenientes de SOLCAST para el 19 de Enero de 2023



Nota. Comportamiento diario de variables meteorológicas y de irradiancia para el 19 de enero de 2023. Elaboración propia (2025).

Según las series temporales obtenidas de SOLCAST, la imagen presenta un conjunto de gráficos que ilustran diversos aspectos del clima y la radiación solar a lo largo de un día, específicamente el 19 de enero de 2023. El GHI comienza en cero durante las primeras horas de la mañana, aumenta hasta alcanzar un pico al mediodía y luego disminuye hacia la tarde, reflejando el movimiento del sol a través del cielo. La temperatura del aire aumenta de manera constante desde la mañana hasta la tarde. El Clearsky GHI, que muestra la radiación esperada en un día despejado, exhibe un patrón suave y simétrico, con su máximo al mediodía, típico de un día claro. La nubosidad varía a lo largo del día, indicando una presencia fluctuante de nubes que afectan la radiación recibida en la superficie. El Clearsky GTI, ajustado para una superficie

inclinada específica, muestra un perfil similar al de Clearsky GHI, alcanzando el pico al medio día. La temperatura del punto de rocío se mantiene relativamente constante, con ligeras disminuciones que sugieren momentos más secos. Estos datos sirven para comprender las condiciones meteorológicas y la disponibilidad de radiación solar para la aplicación en modelos de estimación de energía solar.

Con la información obtenida de la base de datos de SOLCAST y los registros reales de 2023 del inversor Fronius en el sistema fotovoltaico del caso de estudio, se puede realizar una comparativa. Esta comparación permite evaluar la energía generada a partir de los valores de irradiancia con la producción fotovoltaica registrada por el inversor en campo.

Para calcular la potencia a partir de los datos de irradiación, se debe tener en cuenta la eficiencia de los paneles solares, la cantidad de paneles y el valor de GHI en el sitio. La ecuación de la potencia generada por el sistema fotovoltaico se calcula con

$$P_{\text{total}} = GHI \times \eta \times A \times N, \quad (1)$$

donde:

P_{total} representa la potencia total generada por todos los paneles, medida en vatios (W).

GHI es la Irradiación Horizontal Global, medida en vatios por metro cuadrado $\frac{W}{m^2}$.

η es la eficiencia del panel solar.

A es el área de un panel solar, expresada en metros cuadrados m^2 .

N es la cantidad de paneles fotovoltaicos.

Por lo anterior, la energía producida por los paneles fotovoltaicos a lo largo de un periodo de tiempo se puede calcular integrando la potencia con respecto al tiempo. Si se tiene la irradiancia solar GHI , la eficiencia del panel η , y el área del panel A , se puede calcular la energía E expresada en [Wh] multiplicando la potencia en cada muestra por el delta de tiempo Δt

expresado en fracción de la hora, que corresponde a la granularidad de los datos. La energía es calculada por

$$E_{diaria} [Wh/día] = \sum_{i=1}^n (P_i \Delta t), \quad (2)$$

donde:

E_{diaria} es la energía total generada expresada en [Wh] durante un día.

P_i es la potencia total generada por todos los paneles en un instante dado, medida en vatios [W].

Δt es el tiempo durante el cual la irradiación es constante, medido en horas [h].

Adicionalmente, en la Tabla 4 se detallan los parámetros del sistema fotovoltaico utilizado en el caso de estudio, que serán usados para validar las ecuaciones descritas:

Tabla 4

Parámetros Técnicos del Sistema Fotovoltaico del Caso de Estudio

Parámetro	Valor	Unidades	Descripción
Eficiencia	0.2094	[%]	Eficiencia del panel solar
Cantidad Pv	30	[número de paneles solares]	Cantidad de paneles
Área	2.274 X 1.134	[m^2]	Área de un panel solar

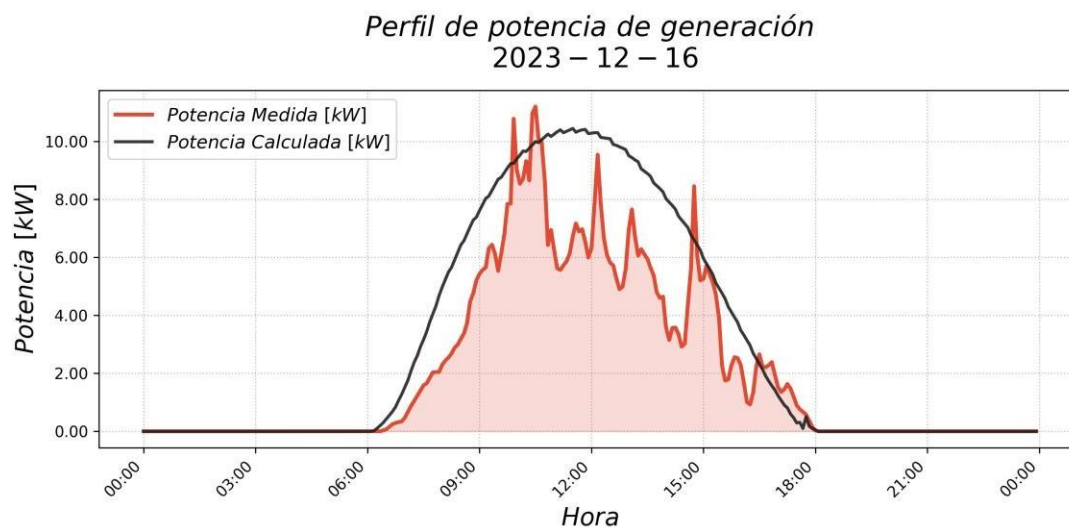
Nota. Parámetros técnicos del sistema fotovoltaico utilizados en el cálculo de potencia y energía.

Elaboración propia (2025).

Una vez relacionada esta información con las ecuaciones descritas, la Figura 10 permite comparar la potencia fotovoltaica medida, registrada por el inversor Fronius, con la potencia calculada a partir de la irradiancia de la base de datos de SOLCAST.

Figura 10

Potencia Calculada a partir de la Irradiancia de SOLCAST vs la Potencia Medida



Nota. Comparación entre potencia medida y potencia calculada a partir de GHI. Elaboración propia (2025).

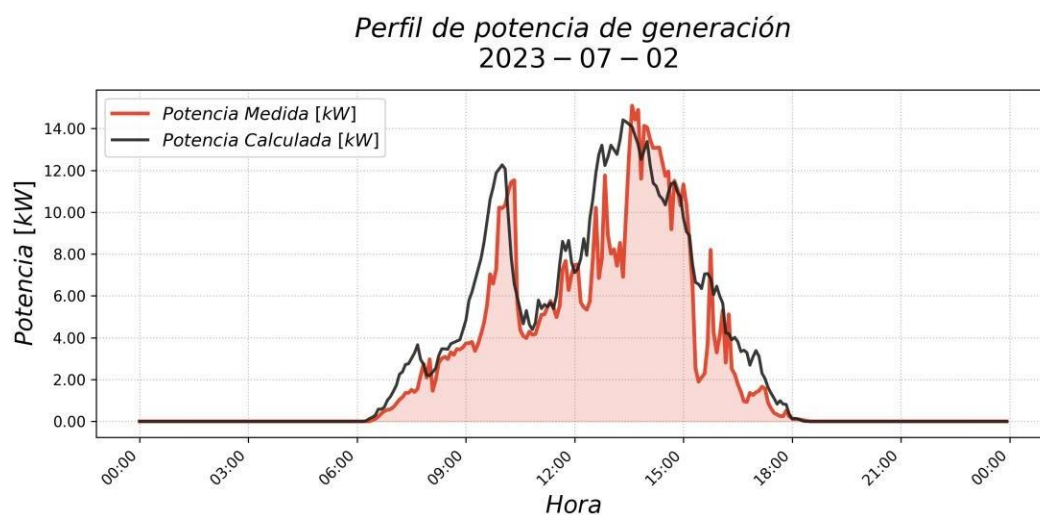
En este análisis se observa que la línea roja, correspondiente a los datos de Fronius, exhibe una notable variabilidad con picos y valles pronunciados, lo cual indica fluctuaciones en la generación de energía. Estas variaciones se deben a factores como la nubosidad intermitente, sombras o el rendimiento del sistema debido a condiciones operativas específicas.

La diferencia entre las dos curvas sugiere que, mientras que la potencia calculada a partir de la irradiancia proporcionado por SOLCAST presenta una curva uniforme, la potencia medida muestra las condiciones reales y la variabilidad propia de un sistema fotovoltaico en funcionamiento. Comprender estas diferencias es importante para interpretar adecuadamente el comportamiento del sistema fotovoltaico bajo circunstancias reales y es fundamental para mejorar la operación de los modelos de predicción.

De igual manera, la Figura 11 muestra la energía calculada a partir de la serie temporal de GHI obtenida de la base de datos de SOLCAST, comparada con los datos registrados por el inversor Fronius en el caso de estudio.

Figura 11

Potencia Calculada a partir de los Datos de GHI de SOLCAST vs la Producción Fotovoltaica Registrada por el Inversor Fronius Symo



Nota. Comparación entre potencia medida y calculada para el 2 de julio de 2023. Elaboración propia (2025).

De manera similar, en la Figura 11 muestra que las dos curvas cuentan con un comportamiento similar, pero demostrando que al usar el Índice de Radiación Global Horizontal (GHI) proveniente de la base de datos de SOLCAST en la ecuación de energía, se puede llegar a un comportamiento muy similar al registrado por las mediciones del inversor Fronius en el caso de estudio real. Las variaciones que se presentan en las curvas sugieren que la nubosidad, temperatura y demás variables están relacionadas con la producción de energía fotovoltaica, y estas deben tenerse en cuenta como parámetros de entrada del modelo de estimación.

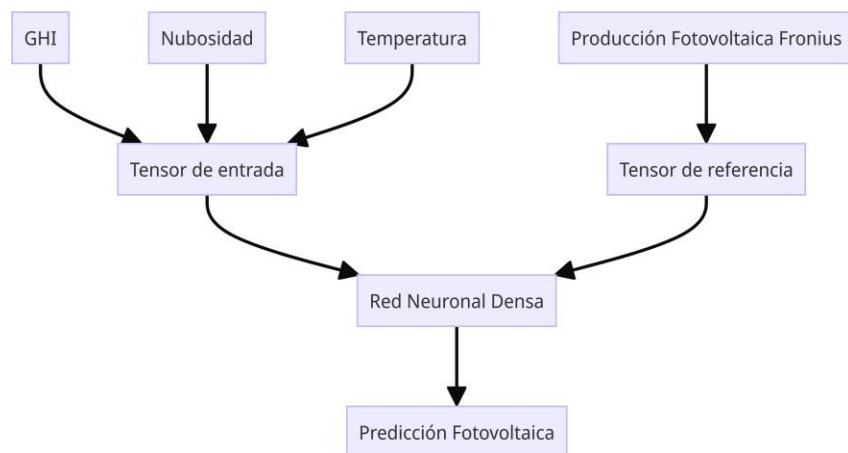
Análisis de la Estimación del Recurso Solar Haciendo Uso del Modelo Propuesto

Para utilizar el modelo propuesto, es necesario ajustar las series temporales provenientes de la base de datos de SOLCAST como parámetros de entrada de la red neuronal. Al mismo tiempo, se deben emplear las series temporales de producción fotovoltaica obtenidas de los registros de mediciones reales del sistema fotovoltaico del inversor Fronius.

Para lograrlo, se debe manipular los datos de manera que las series temporales provenientes de la base de datos de SOLCAST actúen como parámetros de entrada del modelo de red neuronal, mientras que las series temporales de producción fotovoltaica registradas por el inversor Fronius actúen como parámetros de referencia del modelo, como se muestra en el siguiente diagrama.

Figura 12

Diagrama de Flujo de Datos para la Red Neuronal



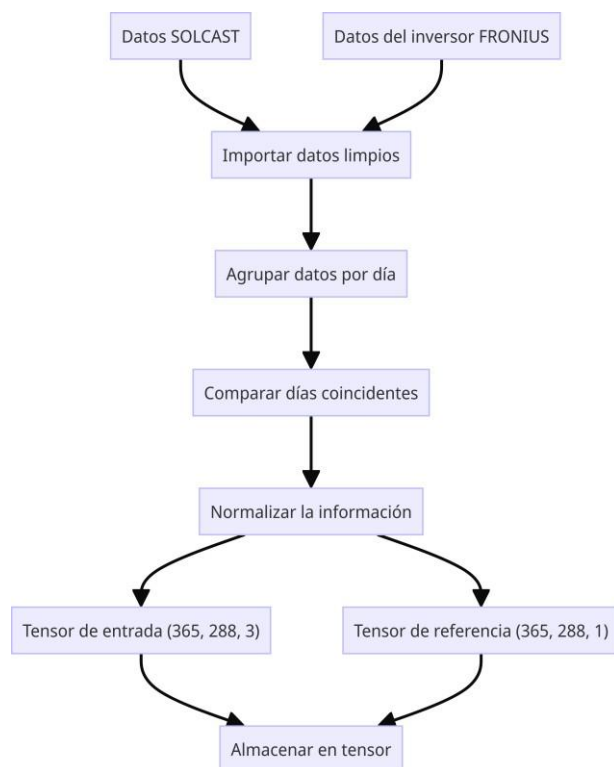
Nota. Diagrama de entrada, modelo y salida de la red neuronal. Elaboración propia (2025).

Para lograr el flujo descrito en la Figura 12, es necesario asegurar que las series temporales provenientes de la base de datos de SOLCAST y la serie temporal de producción fotovoltaica obtenida de los registros de Fronius cumplan con las dimensiones y la forma de los

tensores, de modo que puedan ser interpretadas por el modelo de red neuronal recurrente de tipo LSTM. Esto implica seguir el siguiente flujo para transformar las series temporales en tensores.

Figura 13

Flujo para Crear Tensores



Nota. Proceso de transformación de series temporales en tensores para entrenamiento del modelo. Elaboración propia (2025).

De acuerdo con el diagrama de la Figura 13, se crea un tensor que incluye todos los días del año 2023. Dado que la granularidad de los datos es de 5 minutos, cada serie temporal tendrá una longitud de 288 registros. Además, se observa que el tensor de entrada, proveniente de los

datos de la base de datos de SOLCAST, tiene tres dimensiones relacionadas con el Índice de Radiación Global Horizontal (GHI), temperatura y nubosidad.

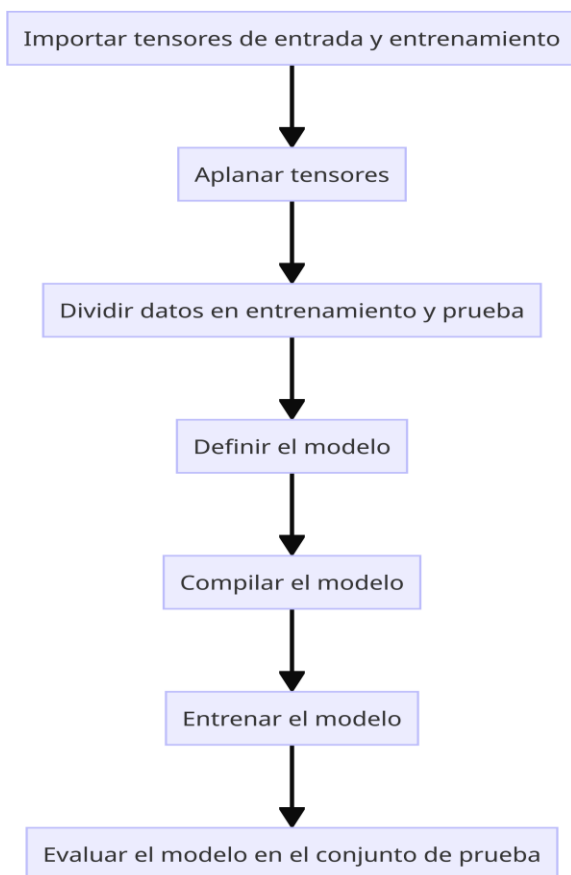
Por otro lado, el tensor de referencia, derivado de los datos capturados en el caso de estudio por el inversor Fronius, consta de 365 registros, correspondientes a cada día del año 2023. Cada uno de estos registros contiene 288 datos, coincidiendo con la granularidad de la información, y tiene una única dimensión relacionada con la producción fotovoltaica.

Entrenamiento del Modelo

Una vez que los tensores están listos para ser entrenados en el modelo, se almacenan en formato ". npy" para alimentar el flujo de entrenamiento y validación descrito en la Figura 14.

Figura 14

Flujo para el Entrenamiento y Validación del Modelo a partir de los Tensores de Entrada



Nota. Flujo del proceso de entrenamiento y validación del modelo LSTM, desde la importación de los tensores hasta la evaluación en el conjunto de prueba. Elaboración propia (2025).

El modelo LSTM se entrena utilizando tensores tridimensionales que preservan la estructura temporal de las series de datos. Cada muestra corresponde a una secuencia temporal diaria con una granularidad de 5 minutos, compuesta por 288 pasos de tiempo y múltiples variables de entrada.

La arquitectura del modelo está conformada por capas LSTM que permiten el aprendizaje de dependencias temporales, seguidas de una capa densa de salida para la estimación de la potencia fotovoltaica. El entrenamiento se realiza utilizando el optimizador Adam y la función de pérdida de error cuadrático medio (MSE), mientras que el desempeño se evalúa mediante el error absoluto medio (MAE).

El modelo se define como una arquitectura secuencial compuesta por capas LSTM, donde la primera capa recibe como entrada un tensor tridimensional con la forma (muestras, pasos de tiempo, variables). Esta configuración permite a la red procesar directamente las secuencias temporales sin necesidad de aplanar los datos, preservando la dependencia temporal inherente a las series de entrada. La salida de la última capa LSTM se conecta a una capa densa final encargada de generar la estimación de la potencia fotovoltaica.

Para el entrenamiento, se utilizaron conjuntos de datos de entrenamiento y validación, ajustando los parámetros del modelo en lotes de 64 muestras por época. La evaluación final del modelo se realizó con un conjunto de datos de prueba, obteniendo métricas de pérdida y precisión que indican su desempeño en tareas de predicción con datos nuevos. Los resultados se resumen en la Tabla 5:

Tabla 5*Métricas de Rendimiento del Modelo*

Métrica	Valor
Pérdida (Loss)	0.026043090969324112
Error Absoluto Medio (Mean Absolute Error)	0.1175408735871315

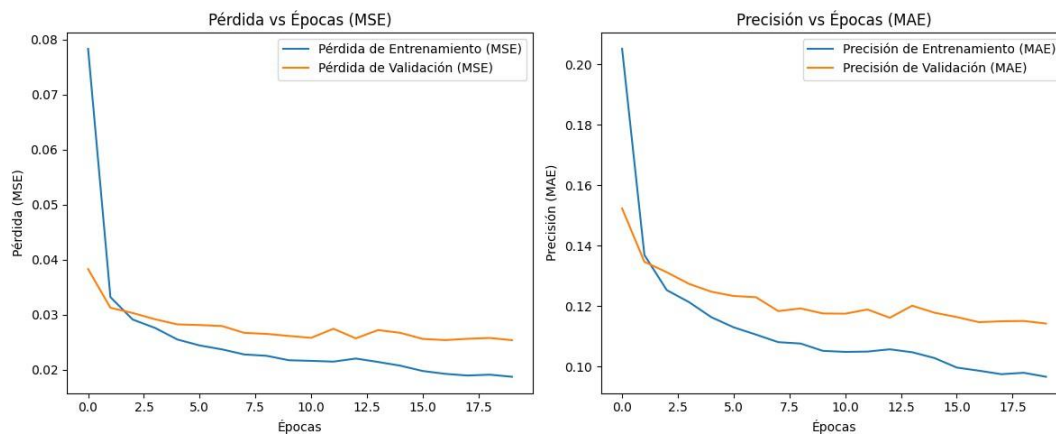
Nota. Métricas de rendimiento del modelo LSTM evaluadas sobre el conjunto de prueba, incluyendo la pérdida (MSE) y el error absoluto medio (MAE). Elaboración propia (2025).

La pérdida, medida por el error cuadrático medio, indica una pequeña divergencia de 0.026 entre las predicciones del modelo y los valores reales. Este valor sugiere que el modelo tiene un buen desempeño en la estimación del recurso fotovoltaico.

El Error Absoluto Medio (MAE) de 0.118 refleja la desviación promedio de las predicciones del modelo respecto a los valores reales. En el ámbito energético, donde cada mejora marginal puede llevar a ahorros significativos, un MAE de esta magnitud indica un modelo confiable.

Figura 15

Progreso del Entrenamiento de la Red Neuronal: Pérdida (MSE) y Precisión (MAE) en Función de las Épocas



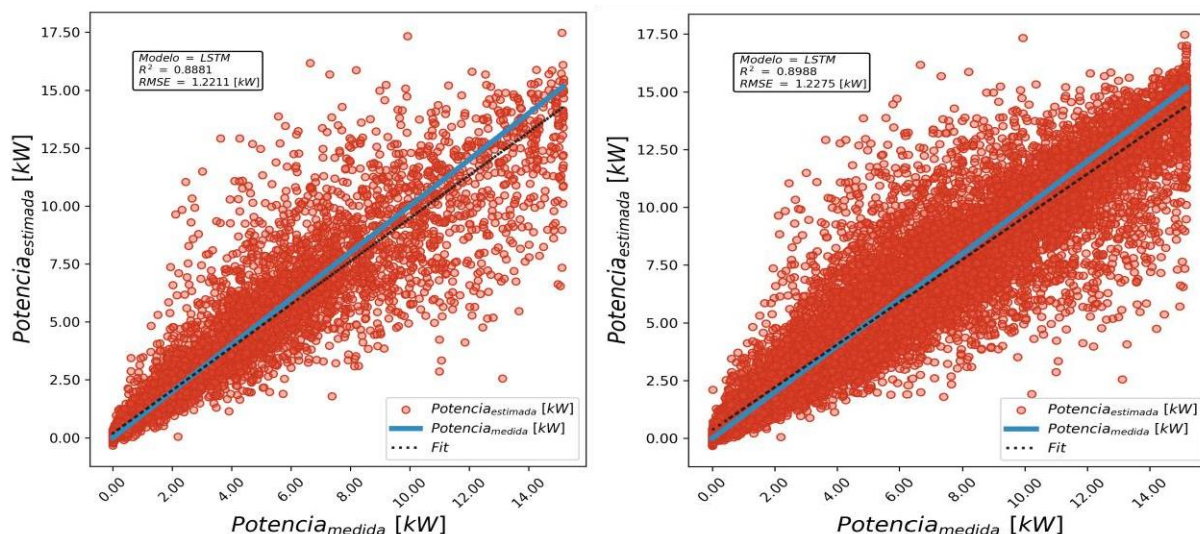
Nota. Evolución de la pérdida (MSE) y el error absoluto medio (MAE) durante el entrenamiento y validación del modelo LSTM en función de las épocas. Elaboración propia (2025).

La Figura 15 muestra la evolución de la pérdida (MSE) y la precisión (MAE) durante el entrenamiento de la red neuronal, tanto en los datos de entrenamiento (azul) como en los datos de validación (naranja). En general, se observa un buen desempeño, con una disminución de la pérdida y un aumento de la precisión a medida que avanzan las épocas. Esto indica que la red está aprendiendo efectivamente a modelar los patrones presentes en los datos.

Hacia el final del entrenamiento, se evidencia una ligera divergencia entre las curvas de entrenamiento y validación, especialmente en la gráfica de pérdida. Este fenómeno sugiere un leve sobreajuste, donde la red comienza a memorizar los datos de entrenamiento en lugar de generalizar a nuevos datos. No obstante, para mitigar este ligero sobreajuste y potencialmente mejorar el rendimiento de la red, se propone incrementar la cantidad de datos de entrenamiento para mejorar la capacidad de generalización de la red. En general, el entrenamiento de la red neuronal ha sido exitoso, logrando un buen equilibrio entre aprendizaje y la generalización de los datos de entrenamiento, el modelo propuesto puede ser utilizado para realizar ajustes precisos en la gestión energética de sistemas fotovoltaicos o en la estimación de la producción fotovoltaica.

Figura 16

Graficas de Dispersión: Conjunto de Datos de Prueba y Conjunto de Datos Completo



Nota. Gráficas de dispersión entre la potencia medida y la potencia estimada por el modelo LSTM para el conjunto de prueba (superior) y el conjunto de datos completo (inferior), incluyendo la recta de ajuste y métricas R^2 y RMSE. Elaboración propia (2025).

En la Figura 16 se presentan los gráficos de dispersión correspondientes al conjunto de datos de prueba (izquierda) y al conjunto de datos completo (derecha). En ambos casos se observa que el modelo LSTM logra una estimación precisa de la potencia generada, evidenciando su capacidad para capturar las dependencias temporales presentes en los datos solares y en la producción fotovoltaica real, lo cual se refleja en los altos valores de R^2 obtenidos. Sin embargo, se evidencia que, a medida que aumenta la potencia medida, también lo hace la dispersión de los valores estimados alrededor de la recta de ajuste. Esto indica que el modelo mantiene un buen desempeño general, aunque su precisión tiende a disminuir en los rangos más altos de potencia.

Análisis de los Resultados Obtenidos con el Modelo

En la evaluación del modelo de predicción de producción fotovoltaica, los resultados demostraron un buen rendimiento con los valores reales observados. En la Tabla 6 se presenta los resultados de la producción de energía:

Tabla 6

Predicción de Generación Fotovoltaica vs la Producción Fotovoltaica Real Registrada por el Inversor Fronius

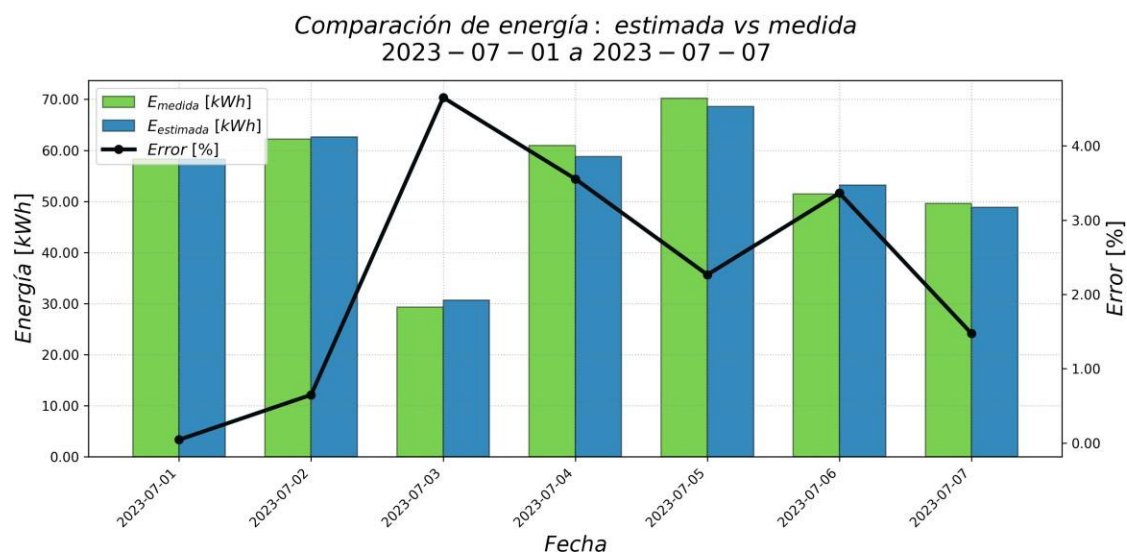
Descripción	Energía Medida (kWh/día)	Energía Estimada (kWh/día)	Error Porcentual (%)
Energía generada para el 1 de julio de 2023.	58,29	58,31	0,04

Nota. Comparación entre la energía fotovoltaica medida y la energía estimada por el modelo para el 1 de julio de 2023, incluyendo el error porcentual de predicción. Elaboración propia (2025).

El error entre la energía medida y la energía estimada, expresado en términos porcentuales de 0,04%, lo que indica que la predicción fue bastante precisa, ya es menor al 1%, pero en promedio el error es del 2,74% del todo el conjunto de días. En la Figura 17 se presenta la comparación entre la potencia medida y la potencia estimada, observando la variabilidad en los diferentes días.

Figura 17

Comparación entre la Energía Medida con la Energía Estimada por el Modelo



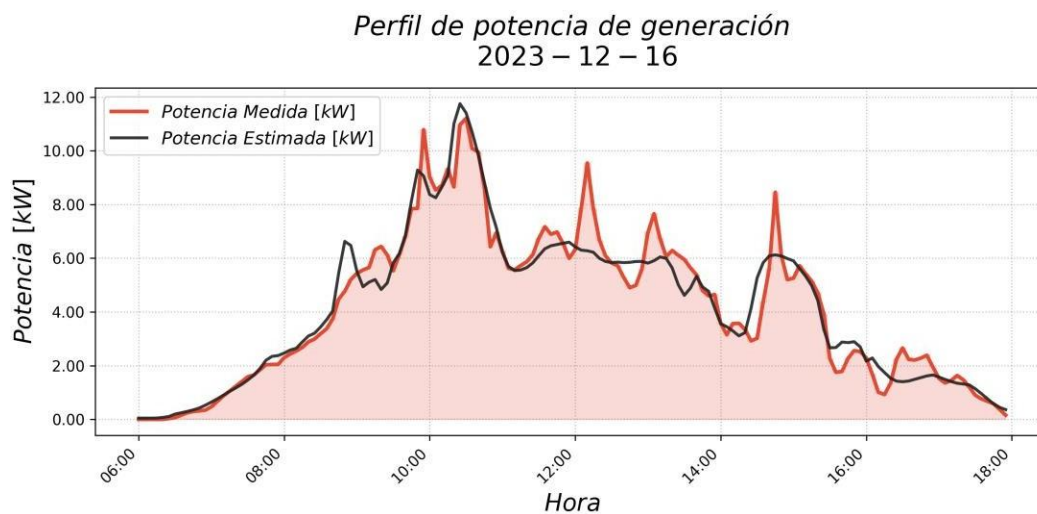
Nota. Comparación diaria entre la energía fotovoltaica medida y la energía estimada por el modelo LSTM del 1 al 7 de julio de 2023, incluyendo el error porcentual de predicción.

Elaboración propia (2025).

De igual manera, en el análisis visual se observa que la línea temporal de producción muestra picos y valles correspondientes a la variabilidad intrínseca en la generación de energía fotovoltaica. Esta variabilidad se debe a factores ambientales transitorios y a las condiciones meteorológicas, así como a la posición solar relativa a lo largo del día. La estimación realizada por el modelo de la Figura 18 (línea de color rojo) sigue de cerca estas fluctuaciones, lo que indica un adecuado seguimiento de la estimación de la producción de energía con el comportamiento de la producción real.

Figura 18

Producción Fotovoltaica Obtenida por el Modelo vs Producción Fotovoltaica Real



Nota. Comparación del perfil diario de potencia medida y potencia estimada por el modelo LSTM para el 16 de diciembre de 2023. Elaboración propia (2025).

La gráfica de la producción demuestra una alineación cercana entre los valores observados y los resultados del modelo construido a partir de la red neuronal, reflejando la habilidad del modelo para representar las dinámicas complejas de la generación fotovoltaica. Estas dinámicas están influenciadas por múltiples factores, incluyendo, pero no limitándose a las condiciones atmosféricas cambiantes y la trayectoria del sol a lo largo del día.

En lo que respecta a las cifras acumuladas, la producción real medida fue de aproximadamente 45,98216 kWh/día, mientras que el modelo proyectó una producción ligeramente superior de 46,10961 kWh/día. Este pequeño margen se traduce en un error porcentual de solo 0,2772%. Un error de menos de 0,3% en el campo de la predicción de la

producción de energía solar es bajo y excede las expectativas convencionales de exactitud para aplicaciones de proyectos de energía solar.

Tabla 7

Producción Fotovoltaica Entregada por el Modelo Predictivo

Producción Real (kWh/día)	Producción Predicha (kWh/día)	Error Porcentual (%)
46	46,10961	0,2772

Nota. Comparación entre la producción fotovoltaica real y la producción predicha por el modelo, incluyendo el error porcentual diario. Elaboración propia (2025).

Estos resultados destacan la utilidad del modelo para la toma de decisiones en el sector de la energía solar. Con un error porcentual de 0,2772%, el modelo ofrece unos resultados fiables para la estimación de producción de energía solar, que pueden ser aplicados para analizar el aprovechamiento energético con la integración de los sistemas fotovoltaicos a la red y el planteamiento de estrategias de gestión de energía con sistema de almacenamiento. Finalmente, esta solución se convierte en una herramienta de apoyo a procesos de planificación para el diseño de sistemas fotovoltaicos, lo que resulta en un avance hacia la integración de las fuentes no convencionales de energía.

Conclusiones y Trabajos Futuros

Se resalta la importancia de contar con datos precisos y confiables sobre el recurso solar para la implementación de tecnologías de energía solar. De esta manera, en esta investigación se seleccionó la base de datos SOLCAST, que combinan imágenes satelitales de alta resolución con modelos de predicción numérica del tiempo (NWP). Para esto, se contó con registros de series temporales de irradiancia, nubosidad y temperatura con una granularidad de 5 minutos.

Se propone un modelo de estimación de producción de energía en sistemas fotovoltaicos on-grid basado en redes neuronales artificiales entrenada a partir de series temporales de datos de recurso solar y de registro de mediciones sistemas fotovoltaicos reales. El modelo propuesto demuestra un buen rendimiento en la predicción de la producción de energía con un error porcentual promedio de 2,74 %.

El modelo de estimación desarrollado tiene aplicaciones en procesos de planificación del diseño de sistemas fotovoltaicos on-grid tanto de autoconsumo como generadores distribuidos. De igual manera, al proporcionar estimaciones precisas de la producción de energía, el modelo ayuda a los operadores de sistemas fotovoltaicos a tomar decisiones informadas sobre el mantenimiento, la operación y la gestión de recursos energéticos.

Esta investigación aporta al fortalecimiento de los esfuerzos orientados a promover el uso de FNCER. Al estimar la potencia generada por un sistema fotovoltaico on-grid de forma más precisa, favorece su integración en la matriz energética, impulsando una transición hacia un modelo más limpio y sostenible. De esta manera, se contribuye tanto a la mitigación del cambio climático como a la disminución de la dependencia de los combustibles fósiles.

En el futuro, se podrían explorar mejoras al modelo mediante la incorporación de variables meteorológicas adicionales, como la velocidad y dirección del viento, la humedad y la

presión atmosférica. También se podría investigar la aplicación de técnicas de aprendizaje profundo más avanzadas, como redes neuronales recurrentes (RNN) o redes neuronales convolucionales (CNN), para capturar patrones más complejos en los datos y mejorar aún más la precisión de las predicciones. Además, se podría ampliar el alcance del estudio para incluir diferentes tipos de sistemas fotovoltaicos y ubicaciones geográficas, con el fin de evaluar la generalización del modelo y su aplicabilidad en diversos contextos.

Referencias Bibliográficas

- A. Pettazzi, & S. Salsón. (2024). Merging remote sensing data with in-situ measurements of global solar radiation: the right path to estimate the solar resource in Galicia. *RE&PQJ*, 10(7).
<https://doi.org/10.24084/repqj10.505>
- Amillo, A. G., Ntsangwane, L., Huld, T., & Trentmann, J. (2018). Comparison of satellite-retrieved high-resolution solar radiation datasets for South Africa. *Journal of Energy in Southern Africa*, 29(2). <https://doi.org/10.17159/2413-3051/2017/v29i2a3376>
- Amillo, A., Huld, T., & Müller, R. (2014). A New Database of Global and Direct Solar Radiation Using the Eastern Meteosat Satellite, Models and Validation. *Remote Sensing*, 6(9), 8165–8189. <https://doi.org/10.3390/rs6098165>
- Aslam, M., Lee, J.-M., Kim, H.-S., Lee, S.-J., & Hong, S. (2019). Deep Learning Models for Long-Term Solar Radiation Forecasting Considering Microgrid Installation: A Comparative Study. *Energies*, 13(1), 147. <https://doi.org/10.3390/en13010147>
- Borba, E. M., & Brito, R. M. (2017). An Index Assessing the Energetic Complementarity in Time between More than Two Energy Resources. *Energy and Power Engineering*, 09(09), 505–514. <https://doi.org/10.4236/epe.2017.99035>
- Brent, A. C., Hinkley, J. (Jim), Burmester, D., & Rayudu, R. (2020). Solar Atlas of New Zealand from satellite imagery. *Journal of the Royal Society of New Zealand*, 50(4), 572–583.
<https://doi.org/10.1080/03036758.2020.1763409>
- Calatrava, C. G., Fontal, Y. B., Cucchietti, F. M., & Cuesta, C. D. (2021). NagareDB: A Resource-Efficient Document-Oriented Time-Series Database. *Data*, 6(8), 91.
<https://doi.org/10.3390/data6080091>

- Cerqueira, V., Torgo, L., & Soares, C. (2022). A case study comparing machine learning with statistical methods for time series forecasting: size matters. *Journal of Intelligent Information Systems*, 59(2), 415–433. <https://doi.org/10.1007/s10844-022-00713-9>
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734. <https://doi.org/10.3115/v1/D14-1179>
- Chow, C. W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., & Washom, B. (2011). Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy*, 85(11), 2881–2893. <https://doi.org/10.1016/j.solener.2011.08.025>
- Deng, S., Tan, Y., Zhao, L., Zhao, R., & Yu, Z. (2014). Energy Efficient Considerations on Carbon Dioxide Capture: Solar Thermal Engineering (Part II). *Energy Procedia*, 61, 2674–2677. <https://doi.org/10.1016/j.egypro.2014.12.273>
- Droege, Peter. (2011). *Urban Energy Transition*.
- Eckman, R. S., & Stackhouse, P. W. (2012). CEOS contributions to informing energy management and policy decision making using space-based Earth observations. *Applied Energy*, 90(1), 206–210. <https://doi.org/10.1016/j.apenergy.2011.03.001>
- Feng, Y., Zhang, X., Jia, Y., Cui, N., Hao, W., Li, H., & Gong, D. (2021). High-resolution assessment of solar radiation and energy potential in China. *Energy Conversion and Management*, 240, 114265. <https://doi.org/10.1016/j.enconman.2021.114265>
- Fortuna, L., Nunnari, G., & Nunnari, S. (2016). *Analysis of Solar Radiation Time Series* (pp. 17–27). https://doi.org/10.1007/978-3-319-38764-2_2

- Fuentes, J. E., Moya, F. D., & Montoya, O. D. (2020). Method for Estimating Solar Energy Potential Based on Photogrammetry from Unmanned Aerial Vehicles. *Electronics*, 9(12), 2144. <https://doi.org/10.3390/electronics9122144>
- Golestaneh, F., Gooi, H. B., & Pinson, P. (2016). Generation and evaluation of space–time trajectories of photovoltaic power. *Applied Energy*, 176, 80–91. <https://doi.org/10.1016/j.apenergy.2016.05.025>
- Gopinath, S., & Prince, P. R. (2018). Prediction of future evolution of solar cycle 24 using machine learning techniques. *Proceedings of the International Astronomical Union*, 13(S340), 317–318. <https://doi.org/10.1017/S1743921318001485>
- Harsarapama, A. P., Aryani, D. R., & Rachmansyah, D. (2020). Open-Source Satellite-Derived Solar Resource Databases Comparison and Validation for Indonesia. *Journal of Renewable Energy*, 2020, 1–14. <https://doi.org/10.1155/2020/2134271>
- Husein, M., & Chung, I.-Y. (2019). Day-Ahead Solar Irradiance Forecasting for Microgrids Using a Long Short-Term Memory Recurrent Neural Network: A Deep Learning Approach. *Energies*, 12(10), 1856. <https://doi.org/10.3390/en12101856>
- Janjai, S., Masiri, I., Pattarapanitchai, S., & Laksanaboonsong, J. (2013). Mapping Global Solar Radiation from Long-Term Satellite Data in the Tropics Using an Improved Model. *International Journal of Photoenergy*, 2013, 1–11. <https://doi.org/10.1155/2013/210159>
- Jiménez-Valero, P., Larrañeta, M., López-García, E., Moreno-Tejera, S., Silva-Pérez, M. A., & Lillo-Bravo, I. (2022). *Synthetic Generation of Plausible Solar Years For Long-Term Forecasting of Solar Radiation*. <https://doi.org/10.21203/rs.3.rs-1350312/v1>

- Jurasz, J., Wdowikowski, M., Kaźmierczak, B., & Dąbek, P. (2017). Temporal and spatial complementarity of wind and solar resources in Lower Silesia (Poland). *E3S Web of Conferences*, 22, 00074. <https://doi.org/10.1051/e3sconf/20172200074>
- Kaskaoutis, D., & Polo, J. (2019). Editorial for the Special Issue “Solar Radiation, Modeling, and Remote Sensing”. *Remote Sensing*, 11(10), 1198. <https://doi.org/10.3390/rs11101198>
- Khatib, T., & Elmenreich, W. (2015). A Model for Hourly Solar Radiation Data Generation from Daily Solar Radiation Data Using a Generalized Regression Artificial Neural Network. *International Journal of Photoenergy*, 2015, 1–13. <https://doi.org/10.1155/2015/968024>
- Khatiri, S. A., Harijan, K., Uqaili, M. A., Shah, S. F., Mirjat, N. H., & Kumar, L. (2022). Solar photovoltaic potential and diffusion assessment for Pakistan. *Energy Science & Engineering*, 10(7), 2452–2474. <https://doi.org/10.1002/ese3.1149>
- Khodayar, M., Mohammadi, S., Khodayar, M. E., Wang, J., & Liu, G. (2020). Convolutional Graph Autoencoder: A Generative Deep Neural Network for Probabilistic Spatio-Temporal Solar Irradiance Forecasting. *IEEE Transactions on Sustainable Energy*, 11(2), 571–583. <https://doi.org/10.1109/TSTE.2019.2897688>
- Kulkarni, R., Kakade, S., Mawande, S., & Sadakale, R. (2023). Estimation of Solar PV Plant output using Polynomial Regression and ANN Algorithm. *Journal of Physics: Conference Series*, 2601(1), 012017. <https://doi.org/10.1088/1742-6596/2601/1/012017>
- Lan, H., Gou, Z., & Xie, X. (2021). A simplified evaluation method of rooftop solar energy potential based on image semantic segmentation of urban streetscapes. *Solar Energy*, 230, 912–924. <https://doi.org/10.1016/j.solener.2021.10.085>
- Lefevre, M., Blanc, P., Espinar, B., Gschwind, B., Menard, L., Ranchin, T., Wald, L., Saboret, L., Thomas, C., & Wey, E. (2014). The HelioClim-1 Database of Daily Solar Radiation at

Earth Surface: An Example of the Benefits of GEOSS Data-CORE. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(5), 1745–1753.

<https://doi.org/10.1109/JSTARS.2013.2283791>

Lhendup, T., & Lhundup, S. (2007). Comparison of methodologies for generating a typical meteorological year (TMY). *Energy for Sustainable Development*, 11(3), 5–10.

[https://doi.org/10.1016/S0973-0826\(08\)60571-2](https://doi.org/10.1016/S0973-0826(08)60571-2)

Li, Q., Fan, J., & Huang, J. (2022). Regional Adaptability Analysis of Solar Roof Utilization Technologies in China. *Applied Sciences*, 12(6), 2792. <https://doi.org/10.3390/app12062792>

Madhiarasan, M., Louzazni, M., & Roy, P. P. (2021). Novel Cooperative Multi-Input Multilayer Perceptron Neural Network Performance Analysis with Application of Solar Irradiance Forecasting. *International Journal of Photoenergy*, 2021, 1–24.

<https://doi.org/10.1155/2021/7238293>

Martins, F. R., Lima, F. J. L. de, Costa, R. S., Gonçalves, A. R., & Pereira, E. B. (2019). Data generated by evaluating the seasonal variability and trend analysis of the solar energy resource in the Northeastern Brazilian region. *Data in Brief*, 26, 104529.

<https://doi.org/10.1016/j.dib.2019.104529>

Ordoñez Palacios, L. E., Bucheli Guerrero, V., & Ordoñez, H. (2022). Machine Learning for Solar Resource Assessment Using Satellite Images. *Energies*, 15(11), 3985.

<https://doi.org/10.3390/en15113985>

Park, S., Kim, Y., Ferrier, N. J., Collis, S. M., Sankaran, R., & Beckman, P. H. (2021). Prediction of Solar Irradiance and Photovoltaic Solar Energy Product Based on Cloud Coverage Estimation Using Machine Learning Methods. *Atmosphere*, 12(3), 395.

<https://doi.org/10.3390/atmos12030395>

- Perdigão, J., Salgado, R., Magarreiro, C., Soares, P. M. M., Costa, M. J., & Dasari, H. P. (2017). An Iberian climatology of solar radiation obtained from WRF regional climate simulations for 1950–2010 period. *Atmospheric Research*, *198*, 151–162.
<https://doi.org/10.1016/j.atmosres.2017.08.016>
- Pereira, S., Abreu, E. F. M., Iakunin, M., Cavaco, A., Salgado, R., & Canhoto, P. (2022). Method for solar resource assessment using numerical weather prediction and artificial neural network models based on typical meteorological data: Application to the south of Portugal. *Solar Energy*, *236*, 225–238. <https://doi.org/10.1016/j.solener.2022.03.003>
- Polo, J., Fernández-Peruchena, C., Salamalikis, V., Mazorra-Aguiar, L., Turpin, M., Martín-Pomares, L., Kazantzidis, A., Blanc, P., & Remund, J. (2020). Benchmarking on improvement and site-adaptation techniques for modeled solar radiation datasets. *Solar Energy*, *201*, 469–479. <https://doi.org/10.1016/j.solener.2020.03.040>
- Praynlin, E., & Jenson, J. I. (2017). Solar radiation forecasting using artificial neural network. *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 1–7.
<https://doi.org/10.1109/IPACT.2017.8244939>
- Purohit, I., & Purohit, P. (2018). Performance assessment of grid-interactive solar photovoltaic projects under India's national solar mission. *Applied Energy*, *222*, 25–41.
<https://doi.org/10.1016/j.apenergy.2018.03.135>
- Rufangura, P., & Sabah, C. (2018). Perfect metamaterial absorber for applications in sustainable and high-efficiency solar cells. *Journal of Nanophotonics*, *12*(02), 1.
<https://doi.org/10.1117/1.JNP.12.026002>

- Sehrawat, N., Vashisht, S., & Singh, A. (2023). Solar irradiance forecasting models using machine learning techniques and digital twin: A case study with comparison. *International Journal of Intelligent Networks*, 4, 90–102. <https://doi.org/10.1016/j.ijin.2023.04.001>
- Suárez, R. A., & Nesmachnow, S. (2012). Parallel Computing Applied to Satellite Images Processing for Solar Resource Estimates. *CLEI Electronic Journal*, 15(3). <https://doi.org/10.19153/cleiej.15.3.4>
- Teke, A., Yıldırım, H. B., & Çelik, Ö. (2015). Evaluation and performance comparison of different models for the estimation of solar radiation. *Renewable and Sustainable Energy Reviews*, 50, 1097–1107. <https://doi.org/10.1016/j.rser.2015.05.049>
- Thomas, C., Wey, E., Blanc, P., & Wald, L. (2016). Validation of three satellite-derived databases of surface solar radiation using measurements performed at 42 stations in Brazil. *Advances in Science and Research*, 13, 81–86. <https://doi.org/10.5194/asr-13-81-2016>
- Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., & Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105, 569–582. <https://doi.org/10.1016/j.renene.2016.12.095>
- Waseem, M., Sherwani, A. F., & Suhaib, M. (2019). Integration of solar energy in electrical, hybrid, autonomous vehicles: a technological review. *SN Applied Sciences*, 1(11), 1459. <https://doi.org/10.1007/s42452-019-1458-4>
- Yu, J., Wang, Z., Majumdar, A., & Rajagopal, R. (2018). DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States. *Joule*, 2(12), 2605–2617. <https://doi.org/10.1016/j.joule.2018.11.021>

Zhang, T., Wang, M., & Yang, H. (2018). A Review of the Energy Performance and Life-Cycle Assessment of Building-Integrated Photovoltaic (BIPV) Systems. *Energies*, *11*(11), 3157.
<https://doi.org/10.3390/en11113157>