

ORIGINAL

SARIMA models for power evolution in photovoltaic systems

Modelos SARIMA para la evolución de la Potencia en sistemas fotovoltaicos

Christian Paul Reyes Orozco¹  , Diana Carolina Campaña Días¹  , Elsa Amalia Basantes Arias¹  , Jesús Rodríguez²  , Juan Ennis Espinoza González³  , Chasiluisa Yanchatuña Sandra Marisol³  

¹Escuela Superior Politécnica de Chimborazo (ESPOCH)- Riobamba, Ecuador.

²Universidad UTE. Quito, Ecuador.

³Investigador Independiente. Ecuador.

Citar como: Reyes Orozco CP, Campaña Días DC, Basantes Arias EA, Rodríguez J, Espinoza González JE, Sandra Marisol CY. SARIMA models for power evolution in photovoltaic systems. Salud, Ciencia y Tecnología. 2025; 5:1971. <https://doi.org/10.56294/saludcyt20251971>

Enviado: 09-02-2025

Revisado: 27-05-2025

Aceptado: 10-08-2025

Publicado: 11-08-2025

Editor: Prof. Dr. William Castillo-González 

Autor para la correspondencia: Juan Ennis Espinoza González 

ABSTRACT

Introduction: the increasing use of renewable energy in power generation systems has highlighted the need for efficient schemes to predict model parameters. In particular, photovoltaic systems require accurate tools to model and forecast solar energy generation behavior.

Objective: to formulate SARIMA models with high accuracy in fitting, explanation, and prediction of energy yields in solar photovoltaic systems, specifically focused on the plant located at Plaza del Duque de Béjar, Spain.

Method: a fitting strategy based on genetic algorithms was adopted to accelerate the estimation of the SARIMA model using hourly solar photovoltaic generation data. The auto.arima package in RStudio was employed as a methodological tool, enabling automatic selection and optimization of the best model parameters. **Results:** the selected model was SARIMA (5,0,0)(2,1,0)242424, characterized by a stationary stochastic process with a clear seasonal component. The model showed remarkable estimation accuracy, with low standard errors in the autoregressive coefficients. Additionally, the model residuals were well-adjusted, displaying independence and absence of serial autocorrelation.

Conclusions: the proposed model demonstrated excellent predictive performance, supported by training error metrics (ME (Mean Error)= -1,344268 and MASE (Mean Absolute Scaled Error)= 0,7048786). Its sound mathematical structure and strong fit make it a reliable tool for forecasting photovoltaic solar energy in systems with similar characteristics.

Keywords: Power; Generation Systems; Photovoltaic; SARIMA Models; Genetic Algorithms; Rstudio.

RESUMEN

Introducción: el creciente uso de energías renovables en los sistemas de generación eléctrica ha puesto de manifiesto la necesidad de contar con esquemas eficientes para la predicción de parámetros de modelos asociados a la producción energética. En particular, los sistemas fotovoltaicos requieren herramientas precisas que permitan modelar y anticipar el comportamiento de la generación de energía solar.

Objetivo: formular modelos SARIMA altamente precisos en cuanto a ajuste, explicación y predicción de los rendimientos energéticos en sistemas solares fotovoltaicos, centrando el estudio en la planta ubicada en la Plaza del Duque de Béjar, España.

Método: se adoptó una estrategia de ajuste basada en algoritmos genéticos para acelerar la estimación del modelo SARIMA a partir de datos horarios de generación solar fotovoltaica. Se utilizó el paquete auto.arima de RStudio como herramienta metodológica, lo que permitió realizar una búsqueda automática de los mejores parámetros del modelo y su correspondiente optimización.

Resultados: el modelo SARIMA seleccionado fue el (5,0,0)(2,1,0)242424, caracterizado por un proceso estocástico estacionario con una marcada componente estacional. La precisión del modelo fue notable, evidenciada por errores estándar reducidos en los coeficientes autorregresivos. Además, los residuos del modelo se ajustaron adecuadamente, mostrando independencia y ausencia de autocorrelación serial.

Conclusiones: el modelo propuesto presentó un excelente desempeño predictivo, respaldado por errores cuantificados en la fase de entrenamiento (ME (Error Medio) = -1,344268 y MASE (Error Absoluto Medio Escalado) = 0,7048786). Su adecuada estructura matemática y buen ajuste lo convierten en una herramienta confiable para la predicción de energía solar fotovoltaica en sistemas con características similares.

Palabras clave: Potencia; Sistemas de Generación; Fotovoltaicos; Modelos SARIMA; Algoritmo Genéticos, RStudio.

INTRODUCCIÓN

La energía solar es un recurso viable para usos industriales, comerciales y domésticos.⁽¹⁾ En los últimos años, la energía solar fotovoltaica (que se basa en sistemas fotovoltaicos) ha ganado popularidad debido a que es abundante, suficiente, limpia y respetuosa con el medio ambiente.⁽²⁾ Dado que el clima puede tener un gran impacto en la generación de electricidad y reducir la integración de sistemas fotovoltaicos en la red, es esencial contar con pronósticos precisos a corto plazo para optimizar la potencia generada en las plantas de energía solar fotovoltaica.⁽³⁾

Una previsión precisa de la potencia en el sistema de energía solar permite contribuir a gestionar eficazmente la generación y el almacenamiento de energía a diario o por horas,⁽⁴⁾ y representa un recurso fundamental para que los sistemas fotovoltaicos participen en el mercado energético, lográndose planificar estrategias de forma eficiente.⁽⁵⁾ En este punto, una exhaustiva revisión en la literatura han descrito diversos métodos para lograr referida precisión, incluidos enfoques estadísticos para la predicción de series temporales según metodología ARIMA (Modelo Autorregresivo Integrado de Media Móvil),⁽⁶⁾ técnicas de aprendizaje automático para la construcción de modelos de redes neuronales artificiales,⁽⁷⁾ modelo de predicción adaptativo de la energía recolectada para redes de sensores de recolección de energía solar,⁽⁸⁾ modelos físicos basados en predicciones meteorológicas numéricas e imágenes satelitales⁽⁹⁾ y enfoques híbridos que combinan los tres primeros métodos.⁽¹⁰⁾ Estos han sido las metodologías que ayudan a las plantas fotovoltaicas a participar de manera más eficaz en el mercado energético.

En concreto, las estimaciones de la potencia generada en watts para sistemas de energía solar se tienden a modelizar mediante ajuste con diversas metodologías que conducen a tipologías para modelos específicos de predicción. Siendo los avances más recientes en este campo, los centrados en las contribuciones derivadas de la implementación de inteligencia artificial en la rama de las redes neuronales, en específico, máquinas de vectores de soporte (SVM) o redes neuronales recurrentes de memoria a corto plazo y largo plazo (LSTM),^(11,12) redes neuronales recurrentes bajo el paradigma de retro propagación,⁽¹³⁾ redes neuronales autorregresivas,^(14,15) modelos autorregresivos de redes neuronales con entrada exógena,⁽¹⁶⁾ inteligencia de enjambre,^(17,18) y usos del aprendizaje automático, incluido el aprendizaje profundo.

Con énfasis en los sistemas de energía solar, los modelos de aprendizaje profundo se han estudiado en profundidad para su posible aplicación en la predicción de series temporales. Si bien, tienen en cuenta las limitaciones y los factores particulares de cada arquitectura, como la eficacia del entrenamiento, la interpretabilidad y el manejo de datos, estos modelos se concentran en las dependencias temporales, la gestión secuencial de datos y la mejora de la precisión de las predicciones.⁽¹⁹⁾

Los argumentos anteriores, se encuentra sujetos al tipo de datos manejados y el enfoque aplicado dentro del paradigma de investigación para la obtención de una adecuada modelización. Estas metodologías para la construcción de modelos se dividen según de interés subyacente, es decir, el análisis del paradigma físico, que considere el uso de variables predictoras asociadas con registros meteorológicos e impactos de la irradiación solar. O en su defecto, manejar un enfoque estadístico con el objeto de predecir comportamiento del fenómeno que orienta la obtención de una potencia en la generación de energía solar en relación con una matriz de datos histórica.^(20,21,22)

Por otro lado, la caracterización de una arraigada condición estacional en los datos temporales, entre dentro de la recomendación según la revisión de la literatura, las estimaciones de parámetros en modelos autorregresivos integrados de medias móviles estacionales (SARIMA), cuyos resultados han sido exitosos para predecir un fenómeno en particular.^(5,21,23,24,25) En esencia, la extensa literatura existente desarrolla de forma analítica y comparativa entre los diferentes modelos para la predicción de la potencia en la generación de energía por sistemas fotovoltaicos, enfocándose en desarrollos de modelos híbridos con loable fortalezas y debilidades implícitas.^(10,26)

La ventaja significativa con el uso de la modelización SARIMA radica en su simplicidad y ésta suele ajustarse con facilidad en procesos estocásticos estacionarios con marcada componente estacional.⁽⁶⁾ Por lo tanto, cuando se está en presencia de una serie que exhibe un proceso no estacionario, se procede a realizar transformaciones para convertir la serie objeto de análisis en estacionaria y ajustar adecuadamente los parámetros del modelo ARIMA. Este modelo se puede construir mediante implementación de técnicas estadísticas sofisticadas.⁽²⁷⁾ Donde el último enfoque aplicable, implica optimizar la selección mediante el uso de estadísticos y determinación de Criterios de decisión, tal como el Criterio de Información de Akaike (AIC) y la suma residual de cuadrados (SSE).

En el contexto actual de transición energética y aprovechamiento de fuentes renovables, los sistemas de generación fotovoltaica han adquirido un rol preponderante. En este sentido, la predicción precisa de la potencia generada constituye un eje estratégico para la planificación operativa y el desarrollo sostenible de dichas tecnologías. De acuerdo con la literatura especializada, los modelos autorregresivos integrados de medias móviles estacionales (SARIMA) han demostrado ser altamente efectivos en la predicción de fenómenos con componentes estacionales, como lo evidencian diversos estudios empíricos exitosos.

A pesar del auge de modelos híbridos que integran técnicas de inteligencia artificial y aprendizaje automático, las aproximaciones SARIMA siguen ofreciendo ventajas significativas, especialmente por su simplicidad, robustez estadística y su capacidad de ajuste a procesos estocásticos estacionarios con fuerte estacionalidad.⁽⁶⁾ Esto resulta particularmente valioso cuando la serie de tiempo presenta comportamiento no estacionario, permitiendo, mediante transformaciones adecuadas, convertirla en estacionaria y ajustar correctamente los parámetros del modelo. Además, la construcción de modelos SARIMA se potencia mediante el uso de herramientas estadísticas avanzadas y la implementación de criterios de decisión como el Criterio de Información de Akaike (AIC) y la Suma Residual de Cuadrados (SSE), que permiten una selección óptima del modelo.⁽²⁷⁾

En este marco, el presente estudio se justifica en la necesidad de modelar la serie de generación de potencia eléctrica en un sistema solar fotovoltaico localizado en la Plaza del Duque de Béjar, Villanueva del Duque, España. Esta serie se caracteriza por su periodicidad diaria (cada 24 horas) durante un periodo continuo de cuatro años (2019-2022), lo cual revela una clara componente estacional. Así, la investigación se propone como una primera etapa hacia la modelización estocástica de dicha serie, estableciendo una base sólida para pronósticos energéticos futuros que mejoren la gestión de recursos renovables.

Según los antecedentes plasmados, el desarrollo de este artículo científico comprende una primera etapa definida por la descripción de la componente estacional presente en la serie potencia generada en el sistema fotovoltaico ubicado en la Plaza del Duque de Béjar, Villanueva del Duque, España, debido a la naturaleza de la recolección de información por cada 24 horas los 365 días del año, durante la cobertura 2019 hasta 2022. En segundo lugar, se procedió a la estimación de los parámetros del modelo SARIMA, incorporando una estrategia de optimización basada en algoritmos genéticos para mejorar la precisión en la predicción de la potencia generada por sistemas de energía renovable. Este enfoque permitió acelerar el proceso de ajuste del modelo a partir de datos horarios de generación solar fotovoltaica. Finalmente, se discuten los resultados obtenidos y se formulan las conclusiones derivadas del análisis, destacando la eficacia del modelo propuesto en contextos similares.

MÉTODO

El presente estudio adopta un diseño cuantitativo, observacional y retrospectivo, basado en datos horarios históricos de generación de energía solar fotovoltaica registrados entre 2019 y 2022 en la planta ubicada en la Plaza del Duque de Béjar, España. Dado el comportamiento estacional claramente identificado en la serie temporal, se seleccionó como técnica de modelado el modelo SARIMA (Modelo Autorregresivo Integrado de Media Móvil Estacional), por su idoneidad para capturar las dinámicas estocásticas con componente estacional.

Con el objetivo de mejorar el proceso de estimación de parámetros, se implementó una estrategia de ajuste basada en algoritmos genéticos, los cuales permitieron optimizar la selección del modelo y acelerar la convergencia hacia una estructura de mejor ajuste. El proceso de modelado se llevó a cabo mediante la función `auto.arima()` del paquete de pronósticos en RStudio, que facilita la selección automática y refinamiento de los parámetros del modelo en función de criterios como el Criterio de Información de Akaike (AIC) y la suma residual de cuadrados (SSE).

El procedimiento metodológico se desarrolló en las siguientes etapas:

1. Preprocesamiento y transformación de los datos para lograr la estacionariedad de la serie.
2. Aplicación del algoritmo `auto.arima` para identificar los parámetros óptimos del modelo SARIMA.
3. Evaluación del rendimiento del modelo mediante indicadores estadísticos como el ME (Error Medio) y el MASE (Error Absoluto Medio Escalado).
4. Diagnóstico de residuos para verificar la independencia y ausencia de autocorrelación serial.

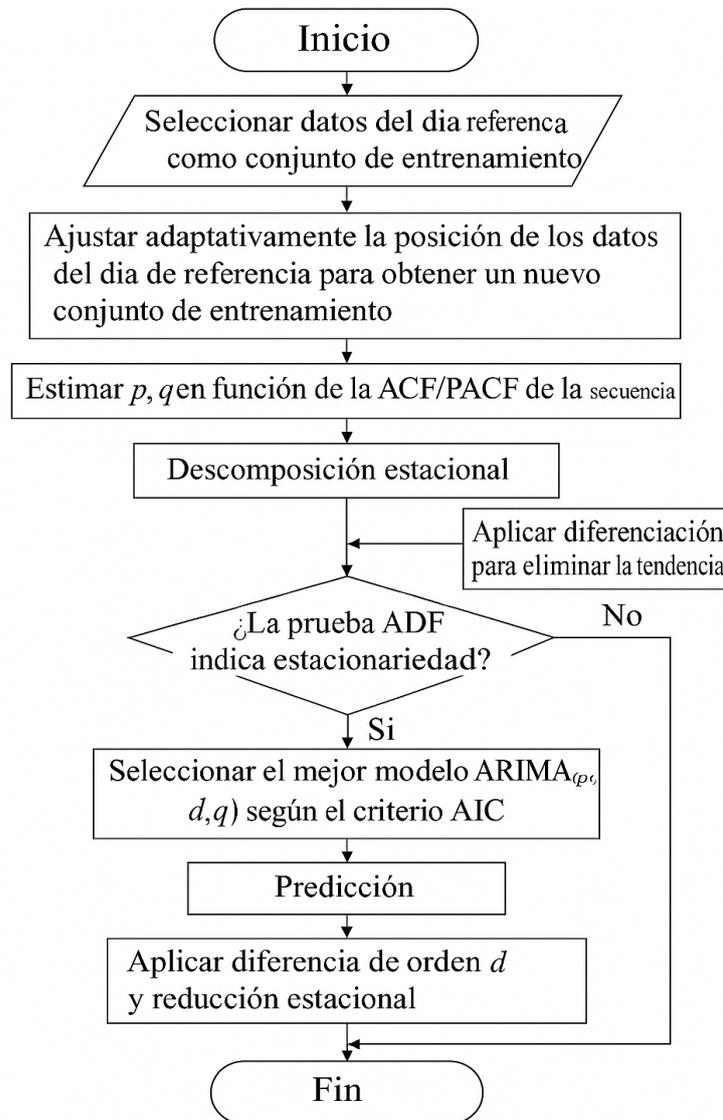
Esta metodología proporcionó un marco robusto y eficiente para el modelado y la predicción de la generación de energía solar, en concordancia con las mejores prácticas en análisis de series temporales.⁽⁴⁾

Descripción de los datos

La construcción de modelos estadísticos con base en datos referidos a la radiación solar superficial (SARAH3) provenientes de los registros del Sistema de Información Geográfica Fotovoltaica (PVGIS). Este conjunto de datos proporciona registros en la ubicación con Latitud: 38,393 y Longitud: -5,000, en grados decimales, para la potencia generada (W) por un sistema fotovoltaico centrado en la Plaza del Duque de Béjar, Villanueva del Duque, España. Esta ubicación, confiere un manejo en función a las normativas de la Unión Europea y otros países.⁽²⁸⁾ La ventana para la generación fotovoltaica está orientada de forma que se logre optimizar la radiación solar incidente durante la cobertura horaria comprendida desde las 08:10 am hasta las 04:10 pm, de forma que la radiación solar incidente es máxima. La base de datos se encuentra registrada en intervalos de 24 horas, específicamente después de los 10 minutos determinados directamente a partir de cada hora, desde el año 2019 hasta el año 2021.

Protocolo para la construcción de Modelos SARIMA

Este apartado es necesario para evaluar la potencia (W) de energía solar generada por sistemas fotovoltaicos, con el objetivo de construir modelos SARIMA; que en esencia requiere asumir un proceso sistemático para el análisis de datos paso a paso, resaltando la importancia del uso de técnicas y pruebas estadísticas para una modelización adecuada de la estructura subyacente en los datos registrados para la potencia generada en sistemas fotovoltaicos (figura 1).



Fuente: Kushwaha et al.⁽²⁹⁾

Figura 1. Diagrama de flujo metodología SARIMA

La figura 1 ilustra un diagrama de cómo debe ser el proceso de predicción de energía solar mediante el método SARIMA. Tal proceso, se inicia con la identificación a partir de datos históricos con un ajuste de las

medidas descriptivas y estadísticos requeridos para una estimación de los parámetros p y q , la realización de un análisis estadístico, con la eliminación de tendencias y la confirmación de la estabilidad de la serie. Como paso clave, la selección del mejor modelo ARIMA para realizar predicciones utilizando la estructura identificada. Finalmente, se obtienen las predicciones finales de la potencia (W) en la generación de energía solar en sistemas fotovoltaicos.

Metodología de procesos ARIMA estacionales (SARIMA)

Si la serie $\{Y_t\}$ presenta una marcada componente estacional de periodo s , éste puede eliminarse aplicando el operador de diferencia estacional con un rezago de orden s , equivalente a los periodos de recolección de datos, y así obtener una serie $\{Z_t\}$ con estructura de proceso ARMA.

En el contexto de una serie temporal que presenta cierto nivel de estacionariedad, cuando la media y la varianza de la serie temporal no presentan oscilaciones notables, se dice que la serie temporal es estacionaria.⁽²³⁾ Esto significa que la curva ajustada por la secuencia de muestra puede continuar en el futuro en función de las circunstancias existentes. En un escenario distinto, se debe aplicar el operador diferencia regular para estabilizar el valor medio de la serie temporal.⁽³⁰⁾ Por otro lado, cuando la serie exhibe características estacionales, se debe considerar corregir esa estacionalidad inherente a la recolección de los datos con el operador diferencia estacional para asegurar eliminar esa componente que distorsiona la correcta identificación de los parámetros en el modelo.⁽¹²⁾ La clave consiste en identificar esos periodos de estacionalidad en el tiempo e identificar con una denotación de s periodos.

Por consiguiente, un proceso temporal que exhibe una tendencia regular y un marcado componente estacional logra definir un parámetro de integración de la forma d y D . En adición, considerándose d y D números enteros no negativos, entonces $\{Z_t\}$ es un proceso estacional multiplicativo ARIMA(p,d,q) \times (P,D,Q) $_s$ con periodo s , si la serie diferenciada en su componente regular y estacional es denotada por $Z_t = (1 - B)^d(1 - B^s)^D Y_t$, representa un proceso ARMA causal definido por:

$$\phi_p(B)\Phi_p(B^s)Z_t = \theta_q(B)\Theta_q(B^s)\varepsilon_t, \{\varepsilon_t\} \sim \varepsilon_t N(0, \sigma^2) \quad (1)$$

Este tipo de modelo, se denomina SARIMA(p,d,q) \times (P,D,Q) $_s$. En esencia, según⁽⁹⁾ el modelo ARIMA estacional incluyen términos autorregresivos y de promedio móvil con rezagos de orden s .

La metodología de procesos ARIMA estacionales se utiliza para pronosticar el futuro de una serie de tiempo que presenta un patrón estacional ($s= 24$ horas). Para ello, se utilizan datos históricos de una serie de tiempo para estimar los parámetros del modelo ARIMA estacional. Una vez estimados los parámetros y validado los supuestos del modelo, este puede ser utilizado para pronosticar el futuro de la serie de tiempo.⁽³⁰⁾

En la etapa de verificación que la serie temporal sea estacionaria, se aplica la prueba Dickey Fuller Aumentada (ADF) que consiste en determinar si una secuencia contiene una raíz unitaria; si es así, la serie temporal se considera no estacionaria; si no, se considera estacionaria.⁽⁶⁾ En este estudio se utiliza las estadísticas ADF sobre la potencia en la generación de energía solar recopiladas en sistemas fotovoltaicos en la Plaza del Duque de Béjar, España. La ausencia de raíces unitarias en la serie temporal indica que la serie temporal recolectada resulta ser estable en media y en variabilidad; situación apropiada para construir modelos SARIMA.

Los modelos obtenidos de un proceso ARIMA estacional se presentan con múltiples aplicaciones para las variables temporales a pronosticar, las cuales deben implementarse para procesos que exhiban una variedad de patrones estacionales, incluyendo patrones estacionales aditivos y multiplicativos. También se pueden utilizar para pronosticar series temporales que presentan una variedad de patrones no estacionarios, incluidos patrones no estacionarios lineales y no lineales.⁽³¹⁾ La ventaja robusta de utilizar este tipo de modelos radica en la precisión avanzada de los resultados de predicción a corto plazo.⁽³²⁾

En términos de interpretación, los modelos autorregresivos y de promedio móvil se combinan para crear el modelo ARMA. Los parámetros medias móviles (MA) simula de cerca los valores de la muestra para explicar los efectos de perturbación en periodos de tiempos previos; mientras que, los parámetros autorregresivos (AR) supone que los valores de muestra en tiempos actuales para la serie de tiempo están relacionados con los valores registrados en tiempos anteriores. El promedio ponderado de los valores anteriores que asume la serie de tiempo representa los comportamientos de los valores de muestra en los tiempos actuales.

RESULTADOS

Pruebas de estacionariedad en la serie que define la potencia en sistemas fotovoltaicos

Implementación del estadístico de Dickey Fuller aumentada según el comando adf_{test} (tabla 1), aplicado en R-Studio para determinar la existencia de raíces unitarias en la serie observada constituida por la potencia en sistemas fotovoltaicos en la Plaza del Duque de Béjar, España. La formulación de hipótesis que se definen, se centra en una hipótesis nula (H_0) que postula la existencia de una raíz unitaria, indicando que la serie es no

estacionaria, mientras que la hipótesis alterna (H_1) sugiere que no hay raíz unitaria, lo que implica que la serie es estacionaria.

| Augmented Dickey-Fuller Test | Valor |
|------------------------------|---------|
| Dickey-Fuller | -15,808 |
| Lag order | 29 |
| p-value | 0,01 |

En función de los hallazgos expuestos en la tabla 1, se puede argumentar que el p-valor=0,01 representa un valor menor al error mínimo permisible ($\alpha=0,05$), situación determinante para concluir que la serie temporal constituida por la potencia registrada en sistemas fotovoltaicos se presenta con condición de estacionariedad.

Ajuste del modelo SARIMA mediante algoritmos genéticos

El modelo utilizado para la visualización se obtiene mediante algoritmos genéticos para acelerar la estimación. El objetivo principal no es obtener un pronóstico perfecto, sino obtener una buena visualización de los resultados.

| | |
|--|--|
| Fitting models using approximations to speed things up... | |
| ARIMA(2,0,2)(1,1,1)[24] with drift | : Inf |
| ARIMA(0,0,0)(0,1,0)[24] with drift | : 683269,9 |
| ARIMA(1,0,0)(1,1,0)[24] with drift | : 661046,7 |
| ARIMA(0,0,1)(0,1,1)[24] with drift | : Inf |
| ARIMA(0,0,0)(0,1,0)[24] | : 683267,9 |
| ARIMA(1,0,0)(0,1,0)[24] with drift | : 668133,7 |
| ARIMA(1,0,0)(2,1,0)[24] with drift | : 658551,4 |
| ARIMA(1,0,0)(2,1,1)[24] with drift | : Inf |
| ARIMA(1,0,0)(1,1,1)[24] with drift | : Inf |
| ARIMA(0,0,0)(2,1,0)[24] with drift | : 676584,4 |
| ARIMA(2,0,0)(2,1,0)[24] with drift | : 658430,5 |
| ARIMA(2,0,0)(1,1,0)[24] with drift | : 660897,2 |
| ARIMA(2,0,0)(2,1,1)[24] with drift | : Inf |
| ARIMA(2,0,0)(1,1,1)[24] with drift | : Inf |
| ARIMA(3,0,0)(2,1,0)[24] with drift | : 658400,5 |
| ARIMA(3,0,0)(1,1,0)[24] with drift | : 660882,5 |
| ARIMA(3,0,0)(2,1,1)[24] with drift | : Inf |
| ARIMA(3,0,0)(1,1,1)[24] with drift | : Inf |
| ARIMA(4,0,0)(2,1,0)[24] with drift | : 658361,5 |
| ARIMA(4,0,0)(1,1,0)[24] with drift | : 660851,7 |
| ARIMA(4,0,0)(2,1,1)[24] with drift | : Inf |
| ARIMA(4,0,0)(1,1,1)[24] with drift | : Inf |
| ARIMA(5,0,0)(2,1,0)[24] with drift | : 658333,6 |
| ARIMA(5,0,0)(1,1,0)[24] with drift | : 660820,9 |
| ARIMA(5,0,0)(2,1,1)[24] with drift | : Inf |
| ARIMA(5,0,0)(1,1,1)[24] with drift | : Inf |
| ARIMA(5,0,1)(2,1,0)[24] with drift | : 658340,1 |
| ARIMA(4,0,1)(2,1,0)[24] with drift | : 658337 |
| ARIMA(5,0,0)(2,1,0)[24] | : 658331,6 |
| ARIMA(5,0,0)(1,1,0)[24] | : 660818,9 |
| ARIMA(5,0,0)(2,1,1)[24] | : Inf |
| ARIMA(5,0,0)(1,1,1)[24] | : Inf |
| ARIMA(4,0,0)(2,1,0)[24] | : 658359,5 |
| ARIMA(5,0,1)(2,1,0)[24] | : 658338,1 |
| ARIMA(4,0,1)(2,1,0)[24] | : 658335 |
| Now re-fitting the best model(s) without approximations... | |
| ARIMA(5,0,0)(2,1,0)[24] | : 658847,1 Best model: ARIMA(5,0,0)(2,1,0)[24] |

El uso de algoritmos genéticos para la estimación y selección de parámetros en modelos SARIMA ha demostrado ser una estrategia eficaz para acelerar el proceso de ajuste en entornos de alta dimensionalidad y fuerte estacionalidad, como es el caso de la serie de generación de potencia en sistemas solares. En la tabla 2, se aprecia una evaluación comparativa entre múltiples combinaciones de parámetros $(p,d,q)(P,D,Q)[s](p,d,q)$

(P,D,Q)[s](p,d,q)(P,D,Q)[s], donde se emplea como métrica de selección el valor del log-verosimilitud negativo o alguna función relacionada (presumiblemente el AIC o el log-likelihood).

Durante el proceso inicial, que utiliza aproximaciones para acelerar los cálculos, se exploran diversos modelos. Algunos modelos son descartados inmediatamente por generar resultados infinitos (“Inf”), lo que indica problemas de convergencia, sobreajuste o inestabilidad numérica. Por otro lado, modelos viables como el ARIMA(5,0,0)(2,1,0)[24], muestran consistentemente los valores más bajos de error, lo cual los posiciona como candidatos óptimos dentro del espacio de búsqueda.

El modelo ARIMA(5,0,0)(2,1,0)[24] se distingue finalmente como el mejor ajustado tras una segunda fase de refinamiento “sin aproximaciones”, confirmando su robustez. Este modelo captura adecuadamente la estructura estacional diaria (frecuencia 24) de la serie, y representa un proceso estocástico estacionario en diferencias, con una estructura autoregresiva dominante tanto a nivel regular como estacional.

Es importante destacar que, si bien el objetivo no era obtener un pronóstico perfecto, la calidad del ajuste logrado refuerza la validez del modelo para visualización y análisis exploratorio, así como su posible aplicación futura en escenarios predictivos. Además, el uso de algoritmos genéticos permite superar limitaciones inherentes a métodos de búsqueda tradicional como grid search, al evitar quedar atrapado en mínimos locales.

Estimación de los parámetros del modelo SARIMA

En la serie potencia en sistema fotovoltaico usada para estimar los parámetros del modelo SARIMA, mediante el uso del paquete auto.arima en RStudio, que permita realizar una búsqueda automática de los mejores parámetros y optimizar el modelo.

Según la tabla 3, se puede observar los valores de los parámetros estimados en el modelo ARIMA(5,0,0)(2,1,0)₂₄ identificado. En adición, se destaca la precisión en las estimaciones de los parámetros en el modelo, debido a los pequeños valores para el error estándar en los coeficientes autorregresivos; tanto para la parte regular de la serie como para la parte estacional.

Por otro lado, la varianza del modelo, según los registros de potencia en los sistemas fotovoltaicos determinan que = 4,519e+09 y la calidad en el ajuste del modelo (log likelihood = -329415,5), revela que existe una cuantificación muy alta, lo que representa un buen ajuste según los parámetros definidos en la estructura matemática.

| Tabla 3. Estimación de parámetros SARIMA | | | | | | | |
|--|--------|--------|---------|---------|---------|---------|---------|
| Parámetros | ar1 | ar2 | ar3 | ar4 | ar5 | sar1 | sar2 |
| Coefficientes | 0,6608 | 0,0955 | -0,0057 | -0,0173 | -0,0343 | -0,6414 | -0,3023 |
| Error estándar (s.e.) | 0,0062 | 0,0074 | 0,0074 | 0,0074 | 0,0062 | 0,0059 | 0,0059 |

El software RStudio adapta series a los modelos ARIMA(p,d,q)x(P,D,Q)s, es decir, las partes autorregresivas y de promedio móvil para los componentes regular y estacional en la serie en estudio. En este sentido, se muestra la estimación de los parámetros en el modelo que se determinó a través del paquete (tabla 2), en este caso un modelo multiplicativo:

$$ARIMA(5,0,0) \times ARIMA(2,1,0)_{24} \rightarrow \phi_5(L) \cdot \Phi_2(L^{24}) \cdot Y_t = \varepsilon_t \quad (2)$$

La representación matemática se formula de forma específica mediante la siguiente expresión:

$$(1 - 0,6608L - 0,0955L^2 + 0,0057L^3 + 0,0173L^4 + 0,0343L^5)(1 + 0,6414L^{24} + 0,3023L^{24}) \cdot Y_t = (1 + 0,6414L^{24} + 0,3023L^{24}) \cdot \varepsilon_t \quad (3)$$

Simplificando la estructura anterior, se obtiene:

$$Y_t = \mu + 0,6608 \cdot Y_{t-1} + 0,0955 \cdot Y_{t-2} - 0,0057 \cdot Y_{t-3} - 0,0173 \cdot Y_{t-4} - 0,0343 \cdot Y_{t-5} - 0,6414 \cdot Y_{t-24} - 0,3023 \cdot Y_{t-24} + \varepsilon_t \quad (4)$$

Este modelo, se utiliza para analizar procesos estocásticos estacionarios con marcada componente estacional de periodo 24 horas en registros diarios para la potencia en sistemas fotovoltaicos. En este contexto, se puede destacar tal representación gráfica en la figura 1, por lo cual representa un proceso determinístico definido por parámetros autorregresivos dentro de los cinco (05 horas anteriores) rezagos en el tiempo. Y con respecto a la componente aleatoria, se caracteriza en la estructura estacional de la serie por una marcada componente

definida por dos parámetros autorregresivos anteriores, en un marco temporal constituidos cada 24 horas de registros diarios.

Una vez lograda la ecuación matemática, se procede a determinar errores en el grupo de datos para entrenamiento, esta se refiere a la determinación de discrepancia entre los valores de salida predichos por un modelo y los valores reales objetivos en el conjunto de datos de entrenamiento. Entre las medidas utilizadas para evaluar el rendimiento del modelo en el conjunto de entrenamiento se encuentran el Error Medio (ME), el Error Cuadrático Medio (RMSE) y el Error Absoluto Medio (MAE).

La tabla 4 resume los principales estadísticos de error utilizados para evaluar el desempeño del modelo SARIMA(5,0,0)(2,1,0)₂₄ durante la fase de entrenamiento. Estas métricas permiten cuantificar la precisión y robustez del ajuste del modelo a los datos históricos, aportando evidencia cuantitativa sobre su capacidad para modelar adecuadamente la generación horaria de energía solar fotovoltaica.

- ME (Error Medio = -1,344268): Indica un leve sesgo negativo en las predicciones, lo que sugiere que el modelo tiende a subestimar ligeramente los valores reales. Sin embargo, dado que el valor es cercano a cero, el sesgo es mínimo y aceptable.
- RMSE (Raíz del Error Cuadrático Medio = 67185,63): Proporciona una medida de la magnitud media del error de predicción, penalizando fuertemente los errores grandes. Aunque el valor es relativamente alto en términos absolutos, debe ser interpretado en función de la escala de la potencia generada, y se considera coherente con la variabilidad inherente a la serie.
- MAE (Error Absoluto Medio = 29122,39): Refleja el promedio del valor absoluto de los errores, lo cual da una idea clara de la magnitud media del error sin considerar su dirección. Su valor es coherente con el RMSE, indicando estabilidad en el modelo sin presencia de grandes errores atípicos.
- MASE (Error Absoluto Medio Escalado = 0,7048786): Es una métrica escalada que permite comparar la precisión del modelo frente a un modelo naïve (sin capacidad predictiva). Un valor menor que 1 indica que el modelo SARIMA ofrece mejor desempeño predictivo que un modelo base, validando su idoneidad para el análisis y la visualización de la serie.
- ACF1 (Autocorrelación del residuo en el rezago 1 = -0,00019): Evalúa si los residuos del modelo presentan correlación serial inmediata. El valor es prácticamente cero, lo que confirma que los residuos son independientes y aleatorios, cumpliendo con un supuesto clave de validez para los modelos SARIMA.

En conjunto, estos indicadores respaldan la consistencia estadística, el bajo sesgo, la capacidad de generalización y el buen ajuste del modelo propuesto, por lo que se justifica plenamente su selección como estructura base para análisis y predicciones futuras en contextos energéticos similares.

| Estadísticos de error | Medidas en conjunto de entrenamiento |
|---|--------------------------------------|
| ME (Error Medio) | -1,344268 |
| RMSE (Raíz del Error Cuadrático Medio) | 67185,63 |
| MAE (Error Absoluto Medio) | 29122,39 |
| MASE (Error Absoluto Medio Escalado) | 0,7048786 |
| ACF1 (Autocorrelación del residuo en el rezago 1) | -0,0001895767 |

Un bajo error de entrenamiento indica que el modelo ha aprendido bien de los datos en el proceso de entrenamiento, pero un error muy bajo también puede ser una señal de sobreajuste. Un valor negativo en el error medio (ME=-1,344268) indica que, en promedio, las predicciones son menores que los valores observados. Otro estadístico, representado por la raíz del error cuadrático medio (RMSE=67185,63), la cual expresa la magnitud del error en las predicciones; un valor más bajo indica un mejor ajuste. El error medio absoluto (MAE=29122,39), proporciona una medida clara de la precisión del modelo y complementario el valor cercano a cero para ACF1 sugiere que no hay correlación significativa en los errores de las predicciones (tabla 4).

Validación del modelo SARIMA

En proceso de validación de un modelo identificado y estimado representa una tarea esencial para comprobar si el modelo SARIMA(5,0,0)(2,1,0)₂₄ se ajusta adecuadamente a los datos, revisando parámetros como la estabilidad y la estacionalidad. La primera premisa, ha de consistir en evaluar si los residuos representan procesos de ruido blanco gaussiano, por consiguiente, se debe realizar gráfico de residuales para verificar que el valor medio esperado para los residuos sea cero para todos los valores en el tiempo (figura 2).

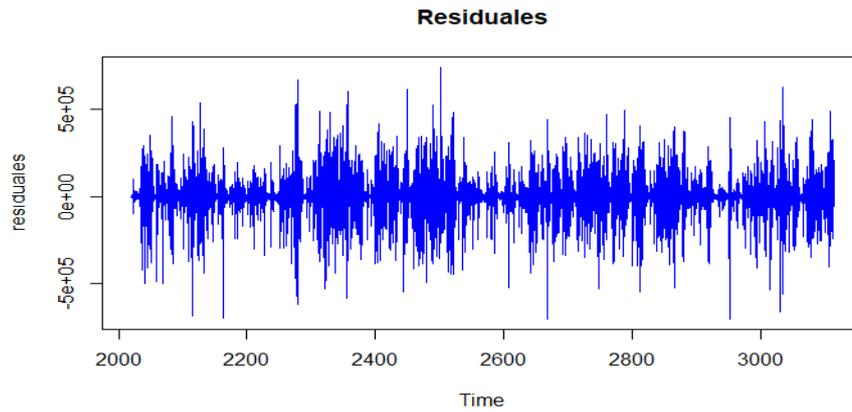


Figura 2. Gráfico de los residuales en el Modelo SARIMA(5,0,0)(2,1,0)₂₄

La conjetura sobre la figura 2, radica en establecer por medio del cálculo del estadístico Dickey-Fuller aumentada= -26,123, orden de retardo = 29, p-valor = 0,01, si los residuales son estacionarios. Por ende este modelo SARIMA(5,0,0)(2,1,0)₂₄, representa una buena estructura que proporciona un buen ajuste. Aunado al resultado arrojado por la prueba Box-Ljung-Box= 0,00094546, con un valor probabilístico (p-valor) = 0,9755, esto sugiere que los residuales no presentan autocorrelación serial significativa en los residuales y por consiguiente resultan ser independientes. Dado que el p-valor es mucho mayor que el nivel de significancia estándar ($\alpha=0,05$), se puede concluir que los residuales obtenidos por el modelo SARIMA son apropiados para definir un ajuste óptimo en términos de la independencia. Referidas aseveraciones se consolidan al realizar un examen visual a los gráficos plasmados en la figura 3.

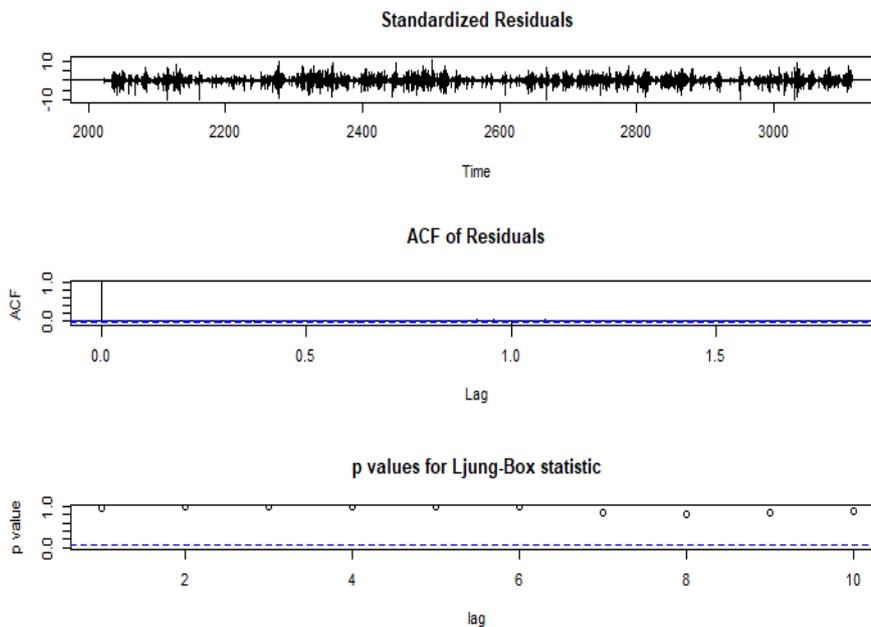


Figura 3. Gráficos de diagnóstico para los residuos en el modelo SARIMA(5,0,0)(2,1,0)₂₄

La figura 3, muestra tres representaciones graficas para la variable residual. En primer lugar, se representan los residuales estandarizados del modelo en función del tiempo centrados alrededor del valor medio cero y han de estar distribuidos de manera aleatoria, sin patrones evidentes. En segundo lugar, el correlograma ACF de los residuos que mantiene rezagos dentro de los límites de confianza, lo que indica que los residuos son independientes y no presentan autocorrelación serial estadísticamente significativa. Por último, se muestra los p-valores del estadístico de Ljung-Box para diferentes rezagos en una representación gráfica, cuya interpretación implica que todos los p-valores son mayores que el nivel de significancia estándar (, lo que sugiere que no hay evidencia de autocorrelación serial en los residuales arrojados por el modelo SARIMA(5,0,0)(2,1,0)₂₄. En definitiva, estos hallazgos sobre el diagnóstico general indican que el modelo SARIMA ajustado a partir de los datos de potencia en sistema fotovoltaicos muestra ser adecuado, ya que los residuales cumplen con los supuestos de independencia y ausencia de autocorrelación. Esto es un buen indicio de que el modelo está bien especificado y puede ser utilizado para realizar pronósticos en periodos futuros.

DISCUSIÓN

La contribución en este estudio ha sido notable en cuanto a los hallazgos, bajo un nivel comparativo de los resultados obtenidos mediante modelización numérica, con los datos registrados en el sistema fotovoltaico en la Plaza del Duque de Béjar, España. Así como un análisis comparativo de eficiencia con los posibles modelos obtenidos con algoritmo genético implementado en paquetería específicas dentro del programa RStudio 2024.09.1. Para realizar el ajuste de los parámetros del modelo SARIMA se ha realizado con un conjunto de 26 304 datos relacionados con los registros de la potencia fotovoltaica registrada en el sistema, desde la fecha 01-01-2019 hasta 31-12-2021, considerando manejar la optimización de los parámetros en el sistema fotovoltaico para, garantizar el incremento de la eficiencia energética y la mejora de su eficiencia global.^(33,34,35)

Por ende, una adecuada estimación de modelos SARIMA, permite emplear la estructura matemática yacente para explicar el fenómeno bajo estudio y predecir posibles escenarios futuros sobre la potencia (W) generada por la radiación solar superficial (SSR) mediante un conjunto de datos recopilados en sistemas fotovoltaicos en la Plaza del Duque de Béjar, España. En este punto, los estadísticos más significativos están en torno a la estructura matemática del modelo SARIMA $(5,0,0)(2,1,0)_{(24)}$, razón por la cual se eligió esta variante después de realizar 35 iteraciones para conformar este tipo de modelos autorregresivos integrados de media móvil con una marcada componente regular y estacional. En un contexto general, para selección final de la mejor estructura, se debe evaluar la adecuación en la identificación y estimación de la referida ecuación matemática SARIMA; por medio de la aplicación de la prueba estadística de Ljung-Box utilizada para identificar el modelo más significativo⁽³⁶⁾ en términos a la no significancia estadística de las autocorrelaciones residuales en los diferentes rezagos temporales.

El modelo SARIMA $(5,0,0)(2,1,0)_{(24)}$, identificado mediante algoritmos genéticos, ha demostrado un desempeño estadísticamente sólido en el conjunto de entrenamiento, evidenciado por un valor de MASE de 0,7048786, lo que indica que su capacidad predictiva supera de forma significativa a la de un modelo naïve. Este resultado valida la eficiencia del modelo para capturar la estructura estacional y estocástica de la serie de generación de potencia solar. Además, al presentar el menor valor de error entre todas las configuraciones evaluadas por el algoritmo genético, se establece como la opción óptima de modelado para el contexto de estudio.

Este hallazgo es coherente con investigaciones previas que han respaldado el uso de modelos SARIMA en contextos similares. Por ejemplo, ⁽³⁹⁾ demostraron que los modelos SARIMA permiten realizar pronósticos precisos de radiación solar diaria, facilitando la planificación energética. De igual forma, ⁽⁴⁰⁾ aplicaron SARIMA a datos horarios de potencia solar en Japón, logrando predicciones robustas en presencia de estacionalidad. Asimismo, ⁽⁴¹⁾ destacaron que, si bien los modelos híbridos ofrecen mejoras marginales en algunos casos, los modelos SARIMA continúan siendo herramientas confiables, especialmente por su estructura interpretable y su baja complejidad computacional.

Bajo este enfoque, resulta pertinente reafirmar la utilidad de aplicar modelos SARIMA optimizados mediante algoritmos evolutivos en sistemas fotovoltaicos con registros horarios extensos, como el de la planta en la Plaza del Duque de Béjar, España. Esta metodología no solo mejora la eficiencia en la estimación de parámetros, sino que proporciona una base metodológica sólida para fortalecer los procesos de pronóstico, gestión operativa y planificación energética en el contexto de la transición hacia fuentes renovables.

El requerimiento de un pronóstico preciso sobre la generación de energía solar fotovoltaica dentro de un periodo de corto tiempo, resulta fundamental para garantizar un funcionamiento seguro y económico,⁽³⁷⁾ así como también para una gestión de la energía en la red. Este pronóstico debe ser con la adopción de un modelo de Media Móvil Integrada Autorregresiva Estacional (SARIMA) para el pronóstico de varios periodos (resolución de 1 hora con 10 minutos) de la generación solar fotovoltaica.⁽²⁹⁾ Si bien, los pronósticos a largo y mediano plazo maximizan los beneficios de la generación de energía fotovoltaica y su penetración en el mercado, los horizontes de pronóstico a corto plazo estiman las tasas de aumento de la potencia fotovoltaica.⁽³⁸⁾ Según Husein et al.⁽⁴⁾ sus aportes fueron determinantes al demostrar que los pronósticos de la radiación solar e incidencia en potencia de generación en sistemas fotovoltaicos con un día de anticipación pueden reducir los gastos anuales de energía de las operaciones de micro-redes en edificios comerciales.

Una alternativa viable para futuras desarrollos de esta línea de investigación, lo representa la creciente preferencia por los enfoques de aprendizaje profundo en la predicción de series temporales para la potencia de generación de energía solar. Esta afirmación se ha respaldado con los hallazgos de Kim et al.⁽³⁾, asumiendo un estudio de precisión en la predicción para la generación de energía fotovoltaica utilizando siete modelos en las ciudades de Ansan y Suwon. Este estudio Descubriendo una baja para los pronósticos estacionales y meteorológicos. O en su defecto, poder usar modelos híbridos de aprendizaje profundo para la previsión de series temporales de energía solar.⁽¹⁰⁾

CONCLUSIONES

La estimación de modelos SARIMA ha demostrado ser una herramienta eficaz para la predicción a corto plazo de la potencia generada por sistemas fotovoltaicos, permitiendo evaluar su desempeño mediante métricas

estadísticas aplicadas al conjunto de datos de entrenamiento. En este estudio, el modelo SARIMA(5,0,0)(2,1,0) fue identificado como la estructura matemática más adecuada, tras múltiples iteraciones realizadas mediante algoritmos genéticos, los cuales facilitaron una optimización eficiente del proceso de ajuste.

La validación del modelo se llevó a cabo mediante el estadístico de Ljung-Box, utilizado para analizar la independencia de los residuos. El resultado (p -valor $> 0,05$) confirmó la ausencia de autocorrelación serial significativa, respaldando así la consistencia del modelo y su capacidad para capturar correctamente la dinámica estacional y estocástica de la serie.

Como proyección futura, se plantea la exploración de enfoques híbridos y avanzados basados en aprendizaje profundo, que permitan abordar la complejidad no lineal inherente a las series temporales de generación fotovoltaica. En particular, se sugiere el desarrollo de una red neuronal polinómica basada en el método GMDH (Group Method of Data Handling), capaz de optimizar de forma inductiva la predicción de la potencia generada. Este tipo de arquitectura resulta prometedora para escenarios donde la irradiación solar presenta cambios bruscos e impredecibles, afectando la fiabilidad y estabilidad del sistema energético.

REFERENCIAS BIBLIOGRÁFICAS

1. Sobri S, Koochi-Kamali S, Rahim NA. Solar photovoltaic generation forecasting methods: A review. *Energy Convers Manag.* 2018; 156:459-97. <http://dx.doi.org/10.1016/j.enconman.2017.11.019>
2. Dada M, Popoola P. Recent advances in solar photovoltaic materials and systems for energy storage applications: a review. *Beni-Suef Univ J Basic Appl Sci.* 2023;12(1). <http://dx.doi.org/10.1186/s43088-023-00405-5>
3. Kim E, Akhtar MS, Yang O-B. Designing solar power generation output forecasting methods using time series algorithms: Global warming affected weather conditions. *SSRN Electron J.* 2022; <http://dx.doi.org/10.2139/ssrn.4170513>
4. Husein M, Chung I-Y. Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach. *Energies.* 2019;12(10):1856. <https://www.mdpi.com/1996-1073/12/10/1856>
5. Aliberti A, Fucini D, Bottaccioli L, Macii E, Acquaviva A, Patti E. Comparative analysis of neural networks techniques to forecast global horizontal irradiance. *IEEE Access.* 2021;9:122829-46. <http://dx.doi.org/10.1109/access.2021.3110167>
6. Atique S, Noureen S, Roy V, Subburaj V, Bayne S, Macfie J. Forecasting of total daily solar energy generation using ARIMA: A case study. En: 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC). IEEE; 2019. p. 0114-9.
7. Runge J, Zmeureanu R. Forecasting energy use in buildings using artificial neural networks: A review. *Energies.* 2019;12(17):3254. <http://dx.doi.org/10.3390/en12173254>
8. Li L, Han C. ASARIMA: An adaptive harvested power prediction model for solar energy harvesting sensor networks. *Electronics (Basel).* 2022;11(18):2934. <http://dx.doi.org/10.3390/electronics11182934>
9. Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy (Oxf).* 2018;148:461-8. <http://dx.doi.org/10.1016/j.energy.2018.01.177>
10. Salman D, Direkoglu C, Kusaf M, Fahrioglu M. Hybrid deep learning models for time series forecasting of solar power. *Neural Comput Appl.* 2024;36(16):9095-112. <http://dx.doi.org/10.1007/s00521-024-09558-5>
11. Junhuathon N, Chayakulkheeree K. Comparative study of short-term photovoltaic power generation forecasting methods. En: 2021 International Conference on Power, Energy and Innovations (ICPEI). IEEE; 2021. p. 159-62.
12. Jiang Y, Zheng L, Ding X. Ultra-short-term prediction of photovoltaic output based on an LSTM-ARMA combined model driven by EEMD. *J Renew Sustain Energy.* 2021;13(4). <http://dx.doi.org/10.1063/5.0056980>
13. Aghmadi A, El Hani S, Mediouni H, Naseri N, El Issaoui F. Hybrid solar forecasting method based on empirical mode decomposition and Back Propagation Neural Network. *E3S Web Conf.* 2021;231:02001. <http://dx.doi.org/10.1051/e3sconf/202123102001>

14. Mughal SN, Sood YR, Jarial RK. Design and optimization of photovoltaic system with a week ahead power forecast using autoregressive artificial neural networks. *Mater Today*. 2022;52:834-41. <http://dx.doi.org/10.1016/j.matpr.2021.10.223>
15. Rogier JK, Mohamudally N. Forecasting photovoltaic power generation via an IoT network using nonlinear autoregressive neural network. *Procedia Comput Sci*. 2019;151:643-50. <http://dx.doi.org/10.1016/j.procs.2019.04.086>
16. Sultan Mohd MR, Johari J, Ruslan FA, Abdul Razak N, Ahmad S, Mohd Shah AS. Analysis on parameter effect for solar radiation prediction modeling using NNARX. En: 2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS). IEEE; 2021. p. 69-74.
17. Boubaker S, Kamel S, Kolsi L, Kahouli O. Forecasting of one-day-ahead global horizontal irradiation using block-oriented models combined with a swarm intelligence approach. *Nat Resour Res*. 2021;30(1):1-26. <http://dx.doi.org/10.1007/s11053-020-09761-w>
18. Zou L, Munir MS, Kim K, Hong CS. Day-ahead energy sharing schedule for the P2P prosumer community using LSTM and swarm intelligence. En: 2020 International Conference on Information Networking (ICOIN). IEEE; 2020. p. 396-401.
19. Benti NE, Chaka MD, Semie AG. Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects. *Sustainability*. 2023;15(9):7087. <http://dx.doi.org/10.3390/su15097087>
20. Basmadjian R, Shaafieyoun A, Julka S. Day-ahead forecasting of the percentage of renewables based on time-series statistical methods. *Energies*. 2021;14(21):7443. <http://dx.doi.org/10.3390/en14217443>
21. Natarajan VA, Karatampati P. Survey on renewable energy forecasting using different techniques. En: 2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC). IEEE; 2019. p. 349-54.
22. Singh B, Pozo D. A guide to solar power forecasting using ARMA models. En: 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe). IEEE; 2019. p. 1-4.
23. Fara L, Diaconu A, Craciunescu D, Fara S. Forecasting of energy production for photovoltaic systems based on ARIMA and ANN advanced models. *Int J Photoenergy*. 2021;2021:1-19. <http://dx.doi.org/10.1155/2021/6777488>
24. Marikkar U, Hassan ASJ, Maithripala MS, Godaliyadda RI, Ekanayake PB, Ekanayake JB. Modified Auto Regressive technique for univariate time series prediction of solar irradiance [Internet]. arXiv [eess.SP]. 2020. Disponible en: <http://arxiv.org/abs/2012.03215>
25. Reikard G, Hansen C. Forecasting solar irradiance at short horizons: Frequency and time domain models. *Renew Energy*. 2019;135:1270-90. <http://dx.doi.org/10.1016/j.renene.2018.08.081>
26. Seyedmahmoudian M, Jamei E, Thirunavukkarasu G, Soon T, Mortimer M, Horan B, et al. Short-term forecasting of the output power of a building-integrated photovoltaic system using a metaheuristic approach. *Energies*. 2018;11(5):1260. <http://dx.doi.org/10.3390/en11051260>
27. Das UK, Tey KS, Seyedmahmoudian M, Mekhilef S, Idris MYI, Van Deventer W, et al. Forecasting of photovoltaic power generation and model optimization: A review. *Renew Sustain Energy Rev*. 2018;81:912-28. <http://dx.doi.org/10.1016/j.rser.2017.08.017>
28. European Commission. PVGIS-SARAH3: Photovoltaic Geographical Information System. Obtenido de EU Science Hub <https://ec.europa.eu/jrc/en/pvgis>. 2024 may.
29. Kushwaha V, Pindoriya NM. Very short-term solar PV generation forecast using SARIMA model: A case study. En: 2017 7th International Conference on Power Systems (ICPS). IEEE; 2017. p. 430-5.
30. Kushwaha V, Pindoriya NM. A SARIMA-RVFL hybrid model assisted by wavelet decomposition for very short-term solar PV power generation forecast. *Renew Energy*. 2019;140:124-39. <http://dx.doi.org/10.1016/j.renene.2019.03.020>

31. Rajagukguk RA, Ramadhan RAA, Lee H-J. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies*. 2020;13(24):6623. <http://dx.doi.org/10.3390/en13246623>
32. Zhang X, Wu X, Zhu G, Lu X, Wang K. A seasonal ARIMA model based on the gravitational search algorithm (GSA) for runoff prediction. *Water Sci Technol Water Supply*. 2022;22(8):6959-77. <http://dx.doi.org/10.2166/ws.2022.263>
33. Moeeni H, Bonakdari H, Ebtehaj I. Monthly reservoir inflow forecasting using a new hybrid SARIMA genetic programming approach. *J Earth Syst Sci*. 2017;126(2). <http://dx.doi.org/10.1007/s12040-017-0798-y>
34. Fashae OA, Olusola AO, Ndubuisi I, Udomboso CG. Comparing ANN and ARIMA model in predicting the discharge of River Opeki from 2010 to 2020. *River Res Appl*. 2019;35(2):169-77. <http://dx.doi.org/10.1002/rra.3391>
35. Harrou F, Taghezouit B, Sun Y. Robust and flexible strategy for fault detection in grid-connected photovoltaic systems. *Energy Convers Manag*. 2019;180:1153-66. <http://dx.doi.org/10.1016/j.enconman.2018.11.022>
36. Taghezouit B, Harrou F, Larbes C, Sun Y, Semaoui S, Arab A, et al. Intelligent monitoring of photovoltaic systems via simplicial empirical models and performance loss rate evaluation under LabVIEW: A case study. *Energies*. 2022;15(21):7955. <http://dx.doi.org/10.3390/en15217955>
37. Yesildal F, Ozakin AN, Yakut K. Optimization of operational parameters for a photovoltaic panel cooled by spray cooling. *Eng Sci Technol Int J*. 2022;25(100983):100983. <http://dx.doi.org/10.1016/j.jestch.2021.04.002>
38. Belghiti H, Kandoussi K, Chellakhi A, Mchaouar Y, El Otmani R, Sadek EM. Performance optimization of photovoltaic system under real climatic conditions using a novel MPPT approach. *Energy Sources Recovery Util Environ Eff*. 2024;46(1):2474-92.
39. Kumar, M., & Kumar, Y. Solar radiation forecasting using SARIMA model for Patiala city, Punjab, India. *Materials Today: Proceedings*. 2013; 33, 3739-3744. <https://doi.org/10.1016/j.matpr.2020.08.413>
40. Yona, A., Senjyu, T., Saber, A. Y., Urasaki, N., & Funabashi, T. Application of recurrent neural network to short-term-ahead generating power forecasting for photovoltaic system. *Energy*. 2013; 30(11-12), 2191-2204. <https://doi.org/10.1016/j.energy.2004.03.001>
41. Benmouiza, K., & Cheknane, A. Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models. *Energy Conversion and Management*. 2013; 75, 561-569. <https://doi.org/10.1016/j.enconman.2013.08.027>

FINANCIACIÓN

Ninguna.

CONFLICTO DE INTERESES

Los autores declaran que no existe conflicto de intereses.

CONTRIBUCIÓN DE AUTORÍA

Conceptualización: Jesús Rodríguez.

Curación de datos: Christian Reyes.

Análisis formal: Diana Campaña.

Investigación: Elsa Basantes.

Metodología: Juan Espinoza.

Administración del proyecto: Sandra Chasiluisa.

Recursos: Jesús Rodríguez.

Software: Christian Reyes.

Supervisión: Diana Campaña.

Validación: Elsa Basantes.

Visualización: Juan Espinoza.

Redacción - borrador original: Sandra Chasiluisa.

Redacción - revisión y edición: Jesús Rodríguez.