

## Desarrollo de modelos de predicción de radiación solar mediante técnicas de machine learning

Abad-Alcaraz, V.<sup>a,b</sup>, Castilla, M.M.<sup>a,b,\*</sup>, Álvarez, J.D.<sup>a,b</sup>, Carballo, J.A.<sup>a,c</sup>, Bonilla, J.<sup>a,c</sup>

<sup>a</sup>CIESOL, Centro de Investigaciones en Energía Solar. Centro mixto UAL-CIEMAT, Universidad de Almería - ceiA3 - Crta. Sacramento, s/n, 04120, La Cañada de San Urbano, Almería, España

<sup>b</sup>Departamento de Informática, Universidad de Almería - ceiA3, Crta. Sacramento, s/n, 04120, La Cañada de San Urbano, Almería, España.

<sup>c</sup>CIEMAT-Plataforma Solar de Almería, Crta. de Senés, s/n, 04200, Tabernas, España.

**To cite this article:** Abad-Alcaraz, V., Castilla, M.M., Álvarez, J.D., Carballo, J.A., Bonilla, J. 2023. Development of solar radiation prediction models using machine learning techniques. XLIV Jornadas de Automática, 382-387. <https://doi.org/10.17979/spudc.9788497498609.382>

### Resumen

La creciente demanda de paneles solares en los últimos años ha hecho que cada vez más empresas se dediquen por completo a la instalación de paneles fotovoltaicos o captadores solares con el fin de aprovechar al máximo la energía solar. Ya sea para un huerto solar o para conseguir edificios más eficientes energéticamente, el disponer de predicciones de radiación solar y, por tanto, de la producción de electricidad o de calor, facilita la toma de decisiones y permite realizar una gestión óptima de la producción y demanda de energía. Sin embargo, la predicción de la radiación solar es un desafío importante debido, entre otros factores, a su alta variabilidad espacial y temporal y a la influencia de las condiciones atmosféricas. Este trabajo presenta una comparación de modelos de predicción de radiación solar desarrollados mediante técnicas de aprendizaje automático utilizando datos históricos de la Plataforma Solar de Almería. Los resultados obtenidos resultan prometedores ya que, en la mayoría de los casos, se han obtenido errores inferiores al 12 %.

*Palabras clave:* Modelado, Modelado e integración de sistemas de energías renovables, Modelado de series temporales, Predicción, Aprendizaje automático.

### Development of solar radiation prediction models using machine learning techniques

#### Abstract

In recent years, the growing demand for solar panels has led an increasing number of companies to focus on installing both photovoltaic panels and solar collectors in order to take advantage of solar energy. Whether it is a solar farm or an energy-efficient building, the availability of predictions of solar radiation, and therefore of electricity and heat production, facilitates decision making and allows the development of optimal management strategies for energy production and demand. However, the prediction of solar radiation is a major challenge due to, among other factors, its high spatial and temporal variability and the influence of atmospheric conditions. This paper presents a comparison of solar radiation prediction models developed using machine learning techniques and historical data from the Solar Platform of Almería. The results obtained are promising since, in most cases, an error less than 12 % has been obtained.

*Keywords:* Modelling, Renewable Energy System Modeling and Integration, Time series modelling, Forecasting, Machine learning.

\*Autor para correspondencia: [mcastilla@ual.es](mailto:mcastilla@ual.es)  
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

## 1. Introducción

Actualmente, la ascendente preocupación sobre el cambio climático está ocasionando la búsqueda de diferentes alternativas que permitan disminuir las emisiones de gases de efecto invernadero, reducir la dependencia de los combustibles fósiles y disponer de diversas opciones de suministro eléctrico. Más concretamente, en la Unión Europea, la Directiva sobre Fuentes de Energía Renovables y el Pacto Verde Europeo Amanatidis (2022); Ciucci (2023) pretenden establecer una cuota de energías renovables con respecto al consumo final de energía igual al 42,5% para el año 2030. En los últimos años, el uso de la energía solar para la producción de energía eléctrica o térmica no deja de crecer. Los principales motivos de su auge están estrechamente relacionados con ser una fuente de energía renovable inagotable, su pequeño impacto en el medio ambiente, la continua mejora de las tecnologías de almacenamiento de energía y su bajada de precio (Esteban Amaro et al., 2023). Asimismo, la integración de la energía solar en edificios, ya sea con el fin producir electricidad mediante paneles fotovoltaicos o generar agua caliente con el uso de captadores solares planos, resulta de gran interés para conseguir que la edificación del futuro sea energéticamente eficiente.

Por otro lado, la producción de los paneles fotovoltaicos o los captadores solares está directamente relacionada con la radiación solar que incide sobre ellos. Sin embargo, estas ondas electromagnéticas emitidas por el sol son de naturaleza intermitente y dependientes de las condiciones meteorológicas, siendo, por tanto, de gran importancia conocer de forma anticipada la radiación solar de la que se dispondrá para garantizar un suministro óptimo y estable de energía eléctrica. En la literatura, se pueden encontrar diversos tipos de modelos para predecir la radiación solar Khatib et al. (2012). Por ejemplo, se pueden utilizar modelos matemáticos para predecir la radiación solar (Plasencia et al., 2007) que tienen en cuenta, entre otras variables, el día del año, la hora solar, la localización geográfica, la declinación solar, el factor de corrección, el ángulo solar, etc. y que consiguen determinar aproximadamente la radiación solar ideal en un cierto instante de tiempo. No obstante, factores externos, como las nubes, resultan complicados de parametrizar, y en consecuencia de predecir, por lo que estos modelos no resultan ser suficientes o no proporcionan buenos resultados bajo ciertas condiciones atmosféricas. Por ello, otra alternativa es el desarrollo de modelos de caja negra utilizando técnicas de aprendizaje automático y datos históricos.

La selección de la técnica de aprendizaje automático más conveniente dependerá de los datos disponibles, la localización sobre la que se realice el estudio, el número de predictores y los objetivos específicos de la predicción. Además, la estrecha relación existente entre la estadística y las predicciones permite obtener resultados aceptables utilizando modelos de regresión entre los cuales destacan la regresión lineal, los árboles de regresión, las máquinas de soporte vectorial (*SVM* del inglés *Support Vector Machine*) o los *Random Forest* (Loayza-Pizarro et al., 2022). En (Voyant et al., 2017) se puede encontrar una amplia revisión bibliográfica de la utilización de técnicas de aprendizaje automático para la obtención de modelos de predicción de la radiación solar. Por otro lado, otra metodología ampliamente utilizada para predecir la radiación solar es mediante el uso de redes neuronales artificiales (*ANN*, por sus siglas

en inglés de *Artificial Neural Networks*) Barrera et al. (2020); Notton et al. (2019); Shboul et al. (2021); Yadav and Chandel (2014). Su funcionamiento mediante diferentes capas de neuronas que procesan la información ofrece la posibilidad de utilizarlas para predecir el comportamiento de entornos complejos y dinámicos. Más concretamente, las redes neuronales tipo *NARX* (por sus siglas en inglés *Nonlinear AutoRegressive with eXogenous input*) es un tipo de arquitectura de redes neuronales que se utilizan para modelar y predecir el comportamiento de series temporales, ya que, son capaces de identificar las relaciones no lineales y las dependencias existentes entre las series temporales y, además, pueden tener en cuenta también variables exógenas. Asimismo, las redes *NARX* se pueden entrenar utilizando diversos algoritmos de aprendizaje como *Levenberg-Marquardt (LM)*, *Bayesian Regularization (BR)* y *Scaled Conjugate Gradient (SCG)* (Carrillo Andrade, 2022). En este trabajo, se muestra la metodología seguida para el desarrollo de modelos de predicción de radiación solar obtenidos mediante técnicas de aprendizaje automático a partir de datos reales registrados en la Plataforma Solar de Almería (*PSA*). Además, se incluye un análisis de los resultados obtenidos en base a distintos índices estadísticos como el error absoluto medio (*MAE*, del inglés *Mean Absolute Error*).

El resto de este documento se organiza en una breve descripción del marco en el que se ha desarrollado este trabajo, en la sección 2. En la sección 3, se realiza una explicación de la metodología seguida para la obtención de los distintos modelos de predicción. Los resultados obtenidos y una comparación entre los mejores modelos de predicción desarrollados se muestra en la sección 4. Finalmente, en la sección 5 se muestra un resumen de las principales conclusiones y futuros trabajos a realizar.

## 2. Marco de investigación

Como se ha mencionado anteriormente, la integración de la energía solar en los edificios es un punto clave para alcanzar los objetivos marcados por las distintas directivas aprobadas por la Unión Europea y, para ello, es vital disponer de modelos de predicción de la radiación solar que nos permitan realizar una gestión óptima de los recursos disponibles. Este trabajo se ha desarrollado en el marco del Centro de Investigación en Energía Solar, el edificio CIESOL (<https://ciesol.com/>), situado en el campus de la Universidad de Almería, en el sureste de España. Más concretamente, el edificio CIESOL es un centro mixto de investigación entre la Universidad de Almería y la Plataforma Solar de Almería, que fue construido utilizando criterios de arquitectura bioclimática con el principal objetivo de aprovechar las ventajas ofrecidas por el clima característico de Almería para obtener un edificio de consumo casi nulo. Para ello, entre otras medidas, el edificio CIESOL cuenta con un campo de paneles fotovoltaicos para producir electricidad y un campo de captadores solares que producen agua caliente para la instalación de climatización basada en frío solar, véase Figura 1(a).

Tras la reciente remodelación integral del sistema de supervisión del edificio CIESOL para su adaptación al paradigma del Internet de las cosas (*IoT*, del inglés *Internet of Things*), no se disponía de una cantidad de datos históricos recientes suficientes que reflejasen las distintas condiciones atmosféricas que se

pueden registrar a lo largo de un año. Por ese motivo, se decidió utilizar los datos meteorológicos registrados por la Plataforma Solar de Almería la cual se sitúa a una distancia inferior a 30 km del centro CIESOL.

La Plataforma Solar de Almería, dependiente del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) se encuentra situada en Tabernas, en la provincia de Almería, y es uno de los mayores centros de investigación energía solar a nivel mundial, véase Figura 1(b). Este centro dispone de una amplia gama de instalaciones y equipos especializados que incluyen instalaciones basadas en tecnología termosolar, de almacenamiento térmico, de receptor solar, de tratamiento de aguas, etc. Además, para este trabajo, se ha utilizado una instalación meteorológica situada en la planta solar de receptor central (CESA-I) que se encarga de registrar la radiación solar (radiación global, directa y difusa), imágenes del cielo, así como, otras variables meteorológicas (temperatura, velocidad y dirección de viento, etc.) con un tiempo de muestreo de 10 minutos. Más concretamente, esta estación incorpora un pirheliómetro para medir la radiación solar directa y piranómetros para la radiación solar difusa y global (resultante de la suma de la radiación solar directa y difusa). Los piranómetros presentan un rango de 0 a 3000  $W/m^2$  y una sensibilidad nominal de  $15 \times 10^{-6} V/(W/m^2)$ .



(a) Edificio CIESOL



(b) Plataforma Solar de Almería

Figura 1: Marco de investigación

### 3. Desarrollo de los modelos de predicción de radiación solar

Se puede definir la predicción como el proceso de predecir acontecimientos futuros en función de datos pasados y presentes. Concretamente, la predicción basada en una sucesión de datos medidos en determinados momentos y ordenados cronológicamente se denomina predicción de series temporales, es decir, es una técnica de modelado que utiliza datos históricos y actuales para predecir valores a lo largo de un periodo de tiempo. Un modelo de predicción de series temporales suele modelarse como una secuencia de variables aleatorias dependientes del tiempo. No obstante, debido a las dependencias temporales de los datos de las series temporales, se necesitan técnicas de validación diferentes a las del modelado tradicional, por ejemplo, la validación cruzada o la validación cruzada dividida en series temporales. En esta sección, se muestra la metodología seguida para el desarrollo de varios modelos de predicción de series temporales utilizando técnicas de aprendizaje automático supervisado.

#### 3.1. Adquisición de los datos

Como se ha mencionado anteriormente, en este trabajo se han utilizado datos registrados en la planta CESA-I de la Plataforma Solar de Almería. Los datos meteorológicos aportados se componen de la humedad ambiente [%], presión atmosférica [mb], radiación solar directa [ $W/m^2$ ], temperatura ambiente [ $^{\circ}C$ ], velocidad del viento [km/h], radiación solar global [ $W/m^2$ ] y radiación solar difusa [ $W/m^2$ ]. Más concretamente, se ha hecho uso de un año de datos que comprende desde el 26 de marzo de 2022 hasta el 25 de marzo de 2023 con un tiempo de muestreo de 10 minutos. Estos datos se encontraban en los metadatos de las imágenes del cielo por lo que el primer paso ha sido la extracción y filtrado de los mismos.

#### 3.2. División de los datos

La radiación solar que incide sobre la superficie terrestre no es constante a lo largo de todo el año, ya que ésta sufre fluctuaciones en función de la estación del año debido a la distancia y ángulos de incidencia que existen entre la Tierra y el Sol. Se distinguen por tanto cuatro periodos claves: solsticio de verano (21 de junio), solsticio de invierno (21 de diciembre) y los equinoccios. En el caso de los solsticios, es cuando el sol presenta su mayor declinación astronómica y, en el caso de los equinoccios, es cuando el sol está situado en el plano del ecuador celeste. Por ello, se ha decidido realizar diferentes modelos: uno para el solsticio de invierno, uno para el solsticio de verano y otro para los equinoccios. Para la división del año en base a lo anterior, se ha establecido que el periodo sea  $\pm 45$  días en torno al solsticio, mientras que los días restantes pertenecerán a los equinoccios.

Otro aspecto clave en esta clasificación es distinguir entre el tipo de día: nublado o despejado. Para ello, se ha hecho uso del denominado *Índice de Claridad (IC)*, que establece la relación entre la radiación global medida y la radiación solar extraterrestre ideal calculada en base a variables como la declinación solar, la latitud, la hora solar, etc. En la Figura 2 se pueden apreciar cinco días pertenecientes al periodo de los equinoccios con diferentes condiciones atmosféricas y el IC asociado con

cada uno de ellos. Finalmente, tras diversos estudios (Meizoso-López, 2012), y considerando el clima de Almería, se ha determinado que aquellos días que tengan un valor de  $IC$  mayor de 0,38 sean clasificados como días despejados.

Por lo tanto, con la clasificación descrita anteriormente se obtienen un total de seis casos de estudio: equinoccios para días despejados, equinoccios para días nublados, solsticio de verano para días despejados, solsticio de verano para días nublados, solsticio de invierno para días despejados y solsticio de invierno para días nublados. Sin embargo, en este trabajo no ha sido posible realizar una predicción de la radiación solar para el solsticio de verano para días nublados debido a la falta de datos, ya que, al caracterizarse Almería por tener un clima subdesértico, mediterráneo, cálido y seco, durante ese periodo del año 2022 no se tuvieron días suficientemente nublados.

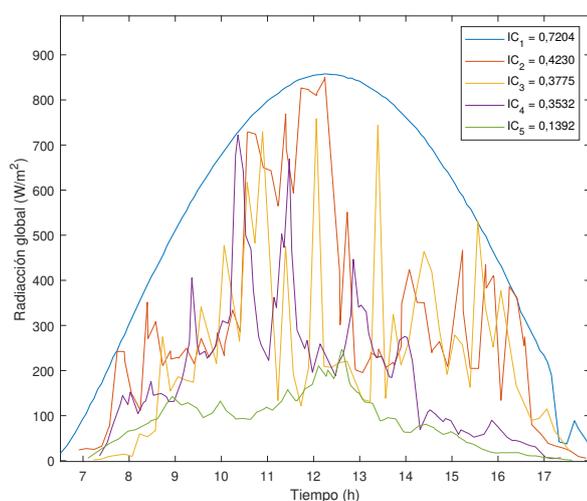


Figura 2: Análisis del Índice de Claridad para la clasificación entre día despejado y día nublado

### 3.3. Selección de entradas y salidas

Seguidamente, se realizó una preselección de variables para entrenar y validar los distintos modelos teniendo en cuenta los seis casos de estudio determinados en la sección anterior. Tal y como se ha mencionado anteriormente, uno de los pasos fundamentales al desarrollar este tipo de modelos consiste en entender el proceso y decidir cuáles serán las entradas y salidas del modelo. Por tanto, tras un análisis detallado de los datos disponibles y de los calculados anteriormente se han seleccionado como entradas:

- La hora.
- La radiación global extraterrestre teórica.
- El índice de claridad diario medio obtenido con los datos temporalmente anteriores.
- La derivada de la radiación global estimada en función de los dos datos temporalmente anteriores.

Asimismo, como salida de los modelos de predicción se ha seleccionado la radiación global.

### 3.4. Entrenamiento y validación de los modelos de predicción

Para el desarrollo de los modelos de predicción es necesario dividir el conjunto total de los datos disponibles para cada uno de los seis casos de estudio en tres subconjuntos: entrenamiento, validación y prueba. Esta división se ha realizado de forma aleatoria pero respetando que fuesen días completos debido a que se está trabajando con series temporales. Respecto al porcentaje de datos que le corresponde a cada categoría, tras diversas pruebas, se ha decidido establecer la siguiente división: 70 % para entrenamiento, 20 % para validación y 10 % para el subconjunto de prueba.

Inicialmente, se desarrollaron modelos de predicción basados en modelos de regresión mediante aprendizaje automático supervisado. Más concretamente, se utilizaron modelos de regresión lineal, varias opciones de árboles de regresión (*Fine Tree*, *Medium Tree* y *Coarse Tree*), máquinas de soporte vectorial y *Random forest*. Todos ellos se validaron utilizando la técnica de validación cruzada con 15 divisiones con el principal objetivo de evitar el sobreajuste.

Posteriormente, se desarrollaron varios modelos de predicción basados en ANNs. La selección de los parámetros de entrenamiento es muy importante para conseguir una alta precisión durante los procesos de entrenamiento y validación. En este caso, se han hecho diferentes pruebas utilizando varios algoritmos de entrenamiento: *Levenberg-Marquardt (LM)*, *Bayesian Regularization (BR)* y *Scaled Conjugate Gradient (SCG)* y se ha utilizado para evaluar el desempeño durante la fase de entrenamiento el índice del error cuadrático medio (*MSE* del inglés, *Mean Squared Error*). Además, se ha seleccionado un tipo de arquitectura denominada NARX que se caracteriza por tener como entradas dos líneas de señales retardadas: una para las entradas y otra para las salidas, y una capa oculta con un cierto número de neuronas que debe seleccionarse de forma cuidadosa. Para ello, se ha realizado un análisis de la variación del desempeño del índice *MSE* de distintos modelos obtenidos aumentando de forma progresiva el número de neuronas en la capa oculta. Finalmente, los modelos desarrollados tienen cuatro señales de entrada y una señal de salida, tal y como se indicaba en la sección 3.3, diez neuronas en la capa oculta y permiten realizar una predicción de la radiación a un paso, es decir, diez minutos.

## 4. Resultados y discusión

En esta sección se realiza un resumen de los principales resultados obtenidos al desarrollar los modelos mencionados en la sección anterior. Más concretamente, se han obtenido un total de nueve modelos diferentes: seis modelos de regresión y tres basados en ANNs. Además, con el objetivo de analizar el ajuste de cada modelo, para cada uno de ellos se ha realizado un análisis estadístico en el que se han considerado los siguientes índices: el *Rango* del conjunto de datos en [ $W/m^2$ ], el número de muestras ( $N$ ), la raíz del error cuadrático medio normalizado (*NRMS E*) en [%], el error absoluto medio (*MAE*) en [ $W/m^2$ ] y el coeficiente de correlación ( $R$ ). Asimismo, se han incluido los resultados gráficos obtenidos para un día de los conjuntos de prueba para cada uno de los seis casos de estudio considerados en la sección 3.2.

Tabla 1: Resumen de los resultados obtenidos para el análisis de los modelos de predicción de radiación obtenidos

<b>Caso 1. Equinoccios para día despejado</b>					
Tipo de modelo	Radiación máxima [W/m <sup>2</sup> ]	N	NRMSE [%]	MAE [W/m <sup>2</sup> ]	R
Modelo regresión <i>Random Forest</i>	1055,688	606	11,606	74,370	0,894
ANN <i>Levenberg-Marquardt</i>	1055,688	606	6,234	44,320	0,972
ANN <i>Bayesian Regularization</i>	1055,688	606	6,359	44,853	0,970
<b>Caso 2. Equinoccios para día nublado</b>					
Tipo de modelo	Radiación máxima [W/m <sup>2</sup> ]	N	NRMSE [%]	MAE [W/m <sup>2</sup> ]	R
Modelo regresión <i>Random Forest</i>	858,427	202	15,651	94,755	0,536
ANN <i>Levenberg-Marquardt</i>	858,427	202	6,111	25,268	0,946
ANN <i>Bayesian Regularization</i>	858,427	202	7,940	33,172	0,900
<b>Caso 3. Solsticio de verano para día despejado</b>					
Tipo de modelo	Radiación máxima [W/m <sup>2</sup> ]	N	NRMSE [%]	MAE [W/m <sup>2</sup> ]	R
Modelo regresión <i>Random Forest</i>	1029,990	445	10,663	57,535	0,942
ANN <i>Levenberg-Marquardt</i>	1029,990	445	5,205	40,072	0,987
ANN <i>Bayesian Regularization</i>	1029,990	445	5,994	42,532	0,982
<b>Caso 4. Solsticio de invierno para día despejado</b>					
Tipo de modelo	Radiación máxima [W/m <sup>2</sup> ]	N	NRMSE [%]	MAE [W/m <sup>2</sup> ]	R
Modelo regresión <i>Random Forest</i>	656,901	425	4,501	14,976	0,982
ANN <i>Levenberg-Marquardt</i>	656,901	425	2,787	13,348	0,993
ANN <i>Bayesian Regularization</i>	656,901	425	3,290	16,451	0,990
<b>Caso 5. Solsticio de invierno para día nublado</b>					
Tipo de modelo	Radiación máxima [W/m <sup>2</sup> ]	N	NRMSE [%]	MAE [W/m <sup>2</sup> ]	R
Modelo regresión <i>Random Forest</i>	347,602	44	14,280	32,900	0,700
ANN <i>Levenberg-Marquardt</i>	347,602	44	25,025	65,205	0,487
ANN <i>Bayesian Regularization</i>	347,602	44	14,549	37,624	0,970

No obstante, debido a la falta de espacio, no se han podido incluir los resultados obtenidos para cada uno de los modelos. En su lugar, se han incluido los resultados para el mejor modelo de regresión y dos modelos basados en redes neuronales. Específicamente, tras un profundo análisis de los modelos obtenidos, los que peores resultados ofrecían eran los basados en regresiones lineales, máquinas de soporte vectorial y el tipo de árbol de regresión denominado *fine tree* debido a que resultan adecuados para problemas más sencillos y lineales. Por lo tanto, entre los modelos de regresión, destacan el *Random forest* y los árboles de regresión del tipo *Coarse Tree* y *Medium Tree*. Con estos dos últimos se obtiene un comportamiento parecido, sin embargo el modelo de predicción del tipo *Random forest* sí que presenta, con algo más de diferencia respecto a los anteriores, mejores resultados. Por otro lado, en el caso de los modelos que han usado ANNs, en general, estos consiguen un mejor ajuste en comparación con los modelos de regresión. Finalmente, en este trabajo, se incluyen los resultados proporcionados por los siguientes modelos de predicción de la radiación solar: (i) Modelo de regresión del tipo *Random Forest*; (ii) Modelo ANN entrenado con el algoritmo *Levenberg-Marquardt*; y (iii) Modelo ANN entrenado con el algoritmo *Bayesian Regularization*.

En la Tabla 1 se muestran los resultados obtenidos del análisis estadístico realizado para los tres modelos seleccionados en los cinco casos de estudio considerados, utilizando la totalidad del conjunto de prueba. En este punto, hay que enfatizar que, como se mencionó anteriormente, no se disponen de datos para el caso de estudio asociado al solsticio de verano para día nublado. Asimismo, en la Figura 4 se muestra la evolución de los diferentes modelos para cada uno de los casos de estudio

utilizando un día del conjunto de datos de prueba.

A partir de esos resultados, se puede observar que las predicciones obtenidas para días despejados son las que presentan un mejor ajuste (casos 1, 3 y 4) ya que estos son periodos normalmente soleados y sin prácticamente nubes en el cielo. En todos los casos, se ha obtenido un *NRMSE* inferior al 12%. Por el contrario, el ajuste para los días nublados es bastante peor que en el caso del día despejado debido a la influencia de la nubes y la alta variabilidad de las mismas en ese tipo de días. Además, se puede observar que los peores resultados se asocian, en todos los casos, con las predicciones proporcionados por los modelos del solsticio de invierno para día nublado debido fundamentalmente a la escasez de datos, ya que, solo estaban disponibles ocho días completos para entrenamiento, validación y prueba.

Finalmente, a partir de la Tabla 1 y la Figura 4 se puede inferir que el modelo de predicción que, en general, se comporta mejor para todos los casos es el obtenido mediante una ANN entrenada con el algoritmo *Levenberg-Marquardt* con excepción del caso 5. Sin considerar el caso 5 (solsticio de invierno para día nublado), este modelo de predicción proporciona un *NRMSE* de 6,2% para el peor de los casos (equinoccio para día despejado) y un *NRMSE* de 2,8% para el mejor de los casos (solsticio de invierno para día despejado).

## 5. Conclusiones y futuros trabajos

Debido a la creciente preocupación por el cambio climático, la integración de energías renovables en la vida diaria está en auge. Una de las más representativas es la energía solar que se trata de una fuente de energía inagotable. Sin embargo, para garantizar una gestión óptima de los recursos disponibles y

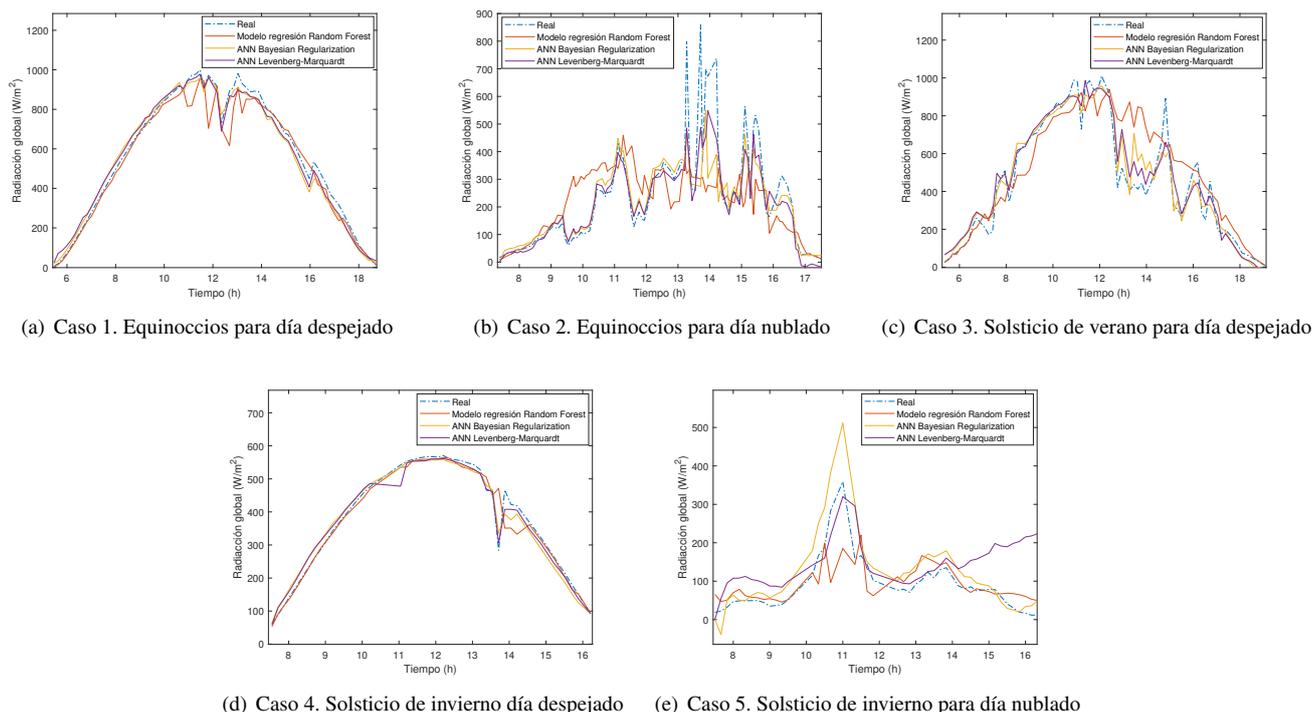


Figura 3: Resultados obtenidos para el conjunto de prueba de los mejores modelos de predicción de radiación solar

garantizar la estabilidad de la red es necesario disponer de predicciones de la radiación solar. En este trabajo, se han desarrollado diferentes modelos utilizando técnicas de aprendizaje automático supervisado y datos recogidos en la Plataforma Solar de Almería. Además, se ha hecho una clasificación en diferentes casos de estudio en base a la época del año y el tipo de día: despejado o nublado. Los resultados obtenidos son bastante prometedores ya que en la mayoría de los casos se ha obtenido un *NRMSSE* inferior al 12% con excepción del caso asociado al solsticio de invierno para día nublado del cual había muy pocos datos disponibles. Por lo tanto, todavía existe un margen de mejora que será analizado en un futuro.

Como futuros trabajos, para mejorar el desempeño de los modelos obtenidos, se podrían utilizar imágenes del cielo para la detección de nubes y utilizar esa información para predecir la influencia de las mismas sobre la radiación solar. Asimismo, también se podría recolectar un mayor volumen de datos asociados con días nublados para mejorar los resultados obtenidos, desarrollar modelos considerando un tercer tipo de día en función del índice de claridad o predecir un día completo.

### Agradecimientos

Este trabajo es parte del proyecto de I+D+i TED2021-131655B-I00, y ha sido realizado gracias a la financiación por AEI/10.13039/501100011033/ y “Unión Europea NextGenerationEU”.

### Referencias

Amanatidis, G., 2022. Fichas temáticas sobre la Unión Europea. La lucha contra el cambio climático. <https://www.europarl.europa.eu/factsheets/es/sheet/72/1a-lucha-contra-el-cambio-climatico> Última fecha de acceso: 30 de mayo 2023.

- Barrera, J. M., Reina, A., Maté, A., Trujillo, J. C., 2020. Solar energy prediction model based on artificial neural networks and open data. *Sustainability* 12 (17), 6915. DOI: 10.3390/su12176915
- Carrillo Andrade, F. A., 2022. Pronóstico del recurso solar a corto plazo para distritos industriales basado en redes neuronales artificiales. B.S. thesis.
- Ciucci, M., 2023. Fichas temáticas sobre la Unión Europea. La Energía Renovable. <https://www.europarl.europa.eu/factsheets/es/sheet/70/1a-energia-renovable> Última fecha de acceso: 30 de mayo 2023.
- Esteban Amaro, R., Lengua Lengua, I., Estelles Miguel, S., 2023. Ayer y hoy de la energía fotovoltaica en España. DOI: 10.4995/INN2022.2022.15780
- Khatib, T., Mohamed, A., Sopian, K., 2012. A review of solar energy modeling techniques. *Renewable and Sustainable Energy Reviews* 16 (5), 2864–2869. DOI: 10.1016/j.rser.2012.01.064
- Loayza-Pizarro, F. G., Nuñez-Medrano, Y., et al., 2022. Assembled methods for the prediction of the incident ghi over the city of puno in tilted solar panels. DOI: 10.18687/LACCEI2022.1.1.529
- Meizoso-López, M.-C., 2012. Distribución de la irradiación solar global, directa y difusa, en la ciudad de A Coruña.
- Notton, G., Voyant, C., Fouilloy, A., Duchaud, J. L., Nivet, M. L., 2019. Some applications of ann to solar radiation estimation and forecasting for energy applications. *Applied Sciences* 9 (1), 209. DOI: 10.3390/app9010209
- Plasencia, E., Matos, L., Posadas, A., Cabrera, C., 2007. Estimación horaria de la irradiancia solar total extraterrestre. *Revista del Instituto de investigación de la Facultad de minas, metalurgia y ciencias geográficas* 10 (19), 72–77. DOI: 10.15381/iigeo.v10i19.540
- Shboul, B., Ismail, A.-A., Michailos, S., Ingham, D., Ma, L., Hughes, K. J., Pourkashanian, M., 2021. A new ANN model for hourly solar radiation and wind speed prediction: A case study over the north & south of the arabian peninsula. *Sustainable Energy Technologies and Assessments* 46, 101248. DOI: 10.1016/j.seta.2021.101248
- 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. DOI: 10.1016/j.renene.2016.12.095
- Yadav, A. K., Chandell, S., 2014. Solar radiation prediction using artificial neural network techniques: A review. *Renewable and Sustainable Energy Reviews* 33, 772–781. DOI: 10.1016/j.rser.2013.08.055