

Modelización y predicción espacio-tiempo de la irradiancia solar global a corto plazo mediante redes neuronales artificiales y geoestadística

Space-time short-term solar radiation modeling and forecasting through artificial neural networks and geostatistics

Federico Vladimir Gutiérrez-Corea*

Miguel Ángel Manso-Callejo**

Francisco Serradilla-García***

Recibido 20 de noviembre de 2018; aceptado 03 de enero de 2019

Resumen

Mejorar el conocimiento de la Irradiancia Solar (IS) sobre la superficie terrestre, así como su predicción (pronóstico), cobra especial interés por su importancia para las energías renovables, a como lo son los sistemas basados en Energía Solar (ES), y para distintas aplicaciones industriales o ecológicas. En la presente investigación se ha experimentado con cinco técnicas de estimación espacial de la IS a intervalos de 15 minutos, en el territorio peninsular español, con distintas configuraciones espaciales. Encontrándose que la geoestadística mediante el Kriging con Regresión, usando variables auxiliares —una de ellas la IS estimada a partir de imágenes satelitales— permite estimar espacialmente la IS más allá de los 25 km, identificados en las investigación científicas previas, como límite de distancia máxima al punto de estimación.

* Universidad Politécnica de Madrid, España-Instituto Nicaragüense de Estudios Territoriales, Nicaragua, correos electrónicos: fv.gutierrez@upm.es; vladimir.gutierrez@ineter.gob.ni.
ORCID: <https://orcid.org/0000-0002-5081-2934>.

** Universidad Politécnica de Madrid, España, correo electrónico: m.manso@upm.es.
ORCID: <https://orcid.org/0000-0003-2307-8639>.

*** Universidad Politécnica de Madrid, España, correo electrónico: francisco.serradilla@upm.es.
ORCID: <https://orcid.org/0000-0001-7621-0627>.

Se ha experimentado con el modelado de Redes Neuronales Artificiales (RNA) para la predicción en tiempo futuro (temporal) —a corto plazo— de la IS utilizando observaciones próximas (componentes espaciales) en sus entradas y los resultados son prometedores. Así los niveles de errores disminuyen, en relación a investigaciones relacionadas, bajo las siguientes condiciones: cuando el horizonte temporal de predicción es inferior o igual a tres horas, las estaciones vecinas que se incluyen en los modelos deben encentrarse a una distancia máxima aproximada de 55 km.

Palabras clave: *redes neuronales artificiales, geoestadística, pronóstico.*

Abstract

The enrichment of knowledge about solar irradiance (SI) on the Earth's surface and its prediction (forecast) has a great interest in renewable energy (RE), such as systems based on solar energy (SE) and for different applications industrial and environmental. At the present research it has been investigated five techniques of spatial estimation of the SI in 15 minutes of temporal resolutions for the Spanish mainland, with several spatial configurations. It's been found that the Geostatistics through Regression Kriging, using auxiliary variables —one of this: the SI estimated from Satellite Images— allows spatially estimates the SI beyond the 25 km identified by the related researches as the maximum distance limit to the estimation point. It has been experimented with the Artificial Neural Networks (ANN) modelling for the short-term forecasting of the SI, using close observations (spatial component) as part of its inputs, and the results are promising. In this way the error levels diminish, regarding to the related researches, under the following conditions: when the temporal horizons of the forecast is lower or equal to 3 hours, the neighbors stations to be included as input to the models should be at a 55 km of maximum distance.

Key words: *artificial neural networks, geostatistics, forecast.*

Introducción

En las sociedades modernas la utilización de los Recursos y Energías Renovables (RER) como la energía solar, hidroeléctrica, eólica o geotérmica entre otras, han tomado fuerza como una de las soluciones a largo plazo con mayor potencial, sostenible y más adecuado para ayudar a minimizar el impacto ambiental y ayudar a solucionar los problemas climáticos (Kaya, 2006). La importancia de las fuentes de energía alternativas en el consumo energético mundial, se evidencia en el incremento superior al 100% de la utilización de estas fuentes en las últimas décadas (IEA, 2014). Este valor aún se puede considerar bajo, si se tiene en cuenta la capacidad de las fuentes de energía renovables existente, como es el caso del Sol. Según Moebius (2006), para cubrir la demanda energética actual, bastaría con un área total de celdas solares inferior al 1% de la superficie total del planeta, ya considerando celdas solares con el

10% de efectividad, la rotación de la tierra (día-noche), la interacción de la cobertura nubosa y otras componentes meteorológicas. Conscientes de la necesidad de aprovechar mejor los RER y promover el desarrollo sostenible, las grandes economías mundiales como en Estados Unidos, Unión Eueopea, China, entre otras; han definido sus objetivos de uso de RER a medio-largo plazo (EC, 2013; Toole *et al.*, 2010; EEN, 2010: 11; MITYC, 2009; IDAE, 2010).

El conocimiento y la predicción de los valores de la Irradiancia Solar (IS) desempeñan un papel importante en el desarrollo y la explotación de los sistemas basados en ES como son las plantas termo-solares y las fotovoltaicas (Mellit y Pavan, 2010), para las cuales la información de la irradiancia, y en particular la Irradiancia Solar Global (ISG), es especialmente relevante para el cálculo de numerosos procesos que van desde, la selección de sitios idóneos para la construcción de nuevas plantas de ES, pasando por su diseño, hasta finalmente llegar a la etapa de explotación; requiriéndose datos de la IS para la predicción de la producción energética, para la planificación y administración de la energía misma (Voyant *et al.*, 2014; Hernández *et al.*, 2012; Perpiñan, 2008: 21).

A pesar de la importancia de la información de la IS para investigaciones relacionadas con la ES (así como para otros tipos de estudios y aplicaciones), estos datos no siempre están disponibles en los sitios de interés debido a la ausencia de estaciones meteorológicas y/o sensores de IS (Yadav y Chandel, 2014).

El objetivo principal es investigar la predicción espacial y temporal de la ISG en el territorio peninsular español mediante RNA y geoestadística, con el propósito de generar mejores modelos predictivos que ayuden en la estimación de la producción de energía renovable procedente del Sol, tanto en sitios desde donde se disponen observaciones de esta variables como en emplazamientos donde no existen.

Estado de la cuestión

En esta sección se describen los trabajos relacionados que permiten identificar los aportes científicos a realizar en materias tanto, de la estimación espacial en intervalos de 15 minutos de la IS, como de su predicción temporal a corto plazo.

Disponibilidad de observaciones meteorológicas e IS

Se sabe que los mejores valores de la IS son los observados por los sensores de IS en las Estaciones Meteorológicas (EM) emplazadas en la superficie terrestre (Polo *et al.*, 2008). Los datos de IS son necesarios en la generación y verificación de modelos, inferencia de datos o, asimilación de datos en modelos usados en contextos tales como la meteorología, climatología o hidrología, entre otros campos de estudio. Sin embargo, existe una baja densificación espacial de estos. De acuerdo con Hernández

et al. (2012), Antonanzas-Torres *et al.* (2013) y Badescu (2008), la escasa densificación de EM en superficie o sensores que observen la IS se debe a la inversión necesaria para su instalación y mantenimiento.

La Radiación Solar

La Radiación Solar (RS) es el término utilizado para describir la energía que por fusión nuclear se libera desde el centro del Sol cuando este convierte átomos de hidrógeno en helio, generándose así una pérdida de masa y una liberación de energía (Mayo 2013). La IS que alcanza a la Tierra puede verse en dos etapas. La primera es la IS Extraterrestre (ISE), que puede ser calculada de forma determinista, y la segunda es la ISG que llega a la superficie de la Tierra como resultado de la ISE que se transmite desde la periferia de la atmósfera hasta la superficie y, que como resultado de este viaje, se ve afectada por las condiciones estocásticas atmosféricas. A la vez, la ISG en un plano horizontal está compuesta por la suma de las irradiancias Directa, Difusa y Albedo (Gueymard y Myers, 2008).

Estimación espacial de la ISG intra-horaria

Los valores de la IS en la superficie terrestre pueden obtenerse por observación directa *insitu* o estimada de un modo indirecto. A continuación se presentan brevemente los métodos utilizados en la presente investigación para estimar superficies de ISG en intervalos inferiores a una hora, así como su relación con la literatura científica asociada.

Método Inverse Distance Weighting (IDW)

En este método de interpolación la influencia entre observaciones está definida de modo determinista/analítico. Los pesos que se le asignan a las observaciones de IS de cada estación son inversamente proporcionales a la distancia entre el punto a estimar (interpolar) y las estaciones de referencia. Distintos estudios han utilizado IDW para interpolar la IS (Pons y Ninyerola, 2008; Zelenka *et al.*, 1999; Perez *et al.*, 1997).

Método Ordinary Kriging (OK)

Tanto el método Kriging Ordinario (OK) como el Kriging con Regresión (RK), son métodos geoestadísticos. Para definir objetivamente los pesos a utilizar en la interpolación, la geoestadística modela un semivariograma que, considera tanto la distancia entre los puntos observados como la variación de sus valores. El método OK ha sido utilizado en distintas ocasiones para la estimación de la IS (Bojanowski *et al.*, 2013; Alsamamra *et al.*, 2009).

Método Regression Kriging (RK)

Kriging con Regresión (RK) incluye información exógena a la variable objeto de la interpolación (variables auxiliares explicativas) compensándose así la relativa escasa densidad de distribución espacial de la variable objetivo, implicando que el RK está compuesto por la suma de una parte estocástica y una determinista de la variación espacial (Alsamamra *et al.*, 2009; Bojanowski *et al.*, 2013).

El método RK también ha sido utilizado en distintas ocasiones para estimar la IS (Antonanzas-Torres *et al.*, 2013; Alsamamra *et al.*, 2009; Evrendilek y Ertekin, 2007), así como para interpolar otras variables meteorológicas y climáticas (Hengl *et al.*, 2012; Pebesma, 2006; Hudson, 1993).

Estimación mediante Sensores Remotos

La base para la estimación de la IS mediante sensores remotos es la relación inversa que existe entre, la reflectividad en el exterior de la atmósfera y la IS asociada a la nubosidad; esta relación fue derivada, en la década de los sesenta, a partir de la correlación entre las medidas de IS registrada por EM y la reflectividad obtenida por el satélite TIROS III (Wielicki *et al.*, 1996). A partir de ese momento, varios estudios han usado de imágenes de satélite para estimar la IS (Posselt *et al.*, 2011; Laszlo *et al.*, 2008; Pinker y Laszlo, 1992; Gautier *et al.*, 1980; Tarpley, 1979).

En Europa, a partir de la familia de satélites geoestacionarios Meteosat, se han generado distintas bases de datos de superficies de IS, entre ellas destaca la base de datos HelioClim3 (SoDA-is, 2011), la cual se genera a partir del método HelioSat-2 (Rigollier *et al.*, 2004).

Predicción de la ISG mediante Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) forman parte del área de conocimiento de la Inteligencia Artificial (IA) y del Aprendizaje Automático (*Machine Learning*), emulando mediante programas informáticos, la capacidad humana de aprender, memorizar y encontrar relaciones. Las RNA en particular intentan reproducir de manera extremadamente simple el comportamiento de las redes neuronales biológicas (Hagan *et al.*, 1996). La habilidad para aprender relaciones no lineales y su capacidad de modelar sistemas complejos, las han convertido en una herramienta útil en distintos ámbitos científicos (Yadav y Chandel, 2013; Paoli *et al.*, 2010; Reed y Marks, 1998). La unidad básica de las RNA es la Neurona Artificial (NA), que es una abstracción matemática simplificada del comportamiento de una Neurona Biológica (NB). Las RNA se componen de un gran número de NA agrupadas en capas, y altamente conectadas entre sí, para trabajar conjuntamente en la solución de un problema.

Perceptrón Multicapa

El Perceptrón Multicapa (Multi Layer Perceptron: MLP) es la arquitectura de RNA más popular empleada en la solución de problemas científicos (Tymvios *et al.*, 2008), por su capacidad demostrada para aproximar relaciones no lineales (Reed y Marks, 1998: 37). En la Figura 1 se ilustra una RNA de tipo MLP con “n” valores de entrada, una capa oculta con “Ni” neuronas artificiales y una sola neurona “o” en la capa de salida. En ella también se aprecian los pesos “w” asociados a cada conexión entre neuronas. Los superíndices “L1” o “L2” indican la capa a la cual pertenece el peso; el primer subíndice indica la entrada o neurona desde donde parte la conexión y el segundo subíndice indica la neurona hacia la que se dirige la conexión.

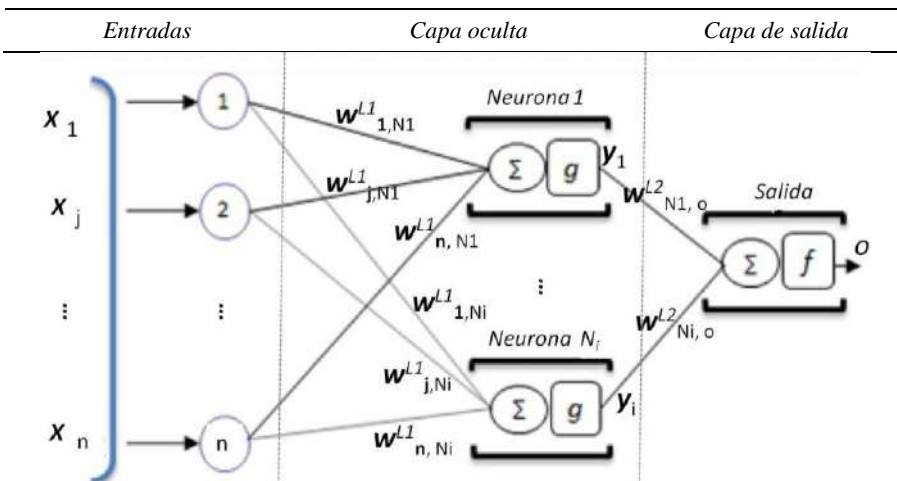


Figura 1. Ilustración de una RNA de tipo Perceptrón Multicapa (MLP).

Modelización de la IS a corto plazo con RNA

Las RNA han sido ampliamente utilizadas en la modelización de la predicción de IS en distintos horizontes de tiempo y en distintos lugares del mundo como, España (Linares-Rodríguez *et al.*, 2011; Bosch *et al.*, 2008; Hontoria *et al.*, 2002), Italia (Mellit y Pavan, 2010), Isla de Córcega en Francia (Voyant *et al.*, 2014; Voyant *et al.*, 2011; Paoli *et al.*, 2010), Turquía (Koca *et al.*, 2011), Arabia Saudita (Mohandes *et al.*, 1998), China (Wang *et al.*, 2012), India (Yadav y Chandel, 2012), Argelia (Mellit *et al.*, 2006), Uganda (Mubiru y Banda, 2008; Mubiru, 2008), entre otros (Mellit *et al.*, 2008; Mellit y Kalogirou, 2008).

Las investigaciones de Voyant *et al.* (2014), Wang *et al.* (2012), y Mellit y Pavan (2010), se relacionan a la presente investigación al focalizarse directamente en la predicción a corto plazo de la IS, utilizando RNA del tipo MLP y utilizando distintos estadísticos: Error Medio Cuadrático Normalizado (nRMSE), Error Medio Absoluto Porcentual (%MAE), Error Medio Cuadrático (RMSE) y Error Medio Cuadrático Porcentual (%RMSE). En la Tabla 1 se resumen las principales características de estas investigaciones.

Tabla 1
Niveles de errores en trabajos relacionados y otras características

<i>Autores</i>	<i>Estadísticos</i>	<i>Periodo de cálculos</i>	<i>Nivel de agregación</i>	<i>Niveles de errores</i>
Voyant <i>et al.</i> (2014)	nRMSE	Estimaciones en las horas de luz	Promedios verano invierno	(27.3%) 21.7% 42.4%
Wang <i>et al.</i> (2012)	RMSE	Estimaciones para las 24 horas	Promedios claros nubosos	63.47 42.29 W/m ² 84.65 W/m ²
Mellit y Pavan (2010)	nRMSE	Estimaciones en las horas de luz	Promedios claros nubosos	(62.5%) 58% (promedio) 67%

Estimación espacial de la ISG

El principal objetivo en esta sección es investigar la bondad de las estimaciones espaciales de la IS en intervalos de 15 minutos, de cinco métodos, en el territorio peninsular español y el archipiélago Balear, aplicado en seis configuraciones espaciales de los datos.

Fuentes de datos

Se utilizaron cuatro fuentes de datos distintas, tres son Redes de Estaciones Meteorológicas (REM) (AEMet, Castilla y León y Meteoclimatic) que registraron observaciones de ISG en la superficie terrestre con periodicidades subhoras, y la última fuente es la base de datos de superficies de ISG HelioClim3v2, estimadas a partir de imágenes del satélite Meteosat.

Se han utilizado datos de dos años de observaciones de la red estaciones de la Agencia Estatal de Meteorología (AEMet), desde noviembre de 2010 hasta octubre de 2012. Durante este periodo, se identificaron 19 estaciones que recolectan información de ISG a intervalos de 10 minutos en el área de estudio. Se utilizaron las observaciones de 50 estaciones que registraron valores de ISG en el año 2011 del Instituto

Tecnológico Agrario de Castilla y León (Estévez y Gavilán, 2008). La tercera fuente de datos REM fueron las observaciones de 206 estaciones la de la Red Meteoclimatic (red de aficionados voluntarios), las cuales solo se muestran como referencia de sus ubicaciones, para posibles futuras consideraciones. Las estimaciones de ISG derivadas a partir de imágenes de satélite utilizadas provienen de la base de datos HelioClim3 versión 2 (HC3v2) (SoDA-is, 2011), creada con el método HelioSat2 (Rigollier *et al.*, 2004). La incertidumbre de la ISG HelioClim3v2 la estiman siguiendo un proceso de benchmarking recomendado por la tarea 36 del Programa de Calentamiento y Enfriamiento Solar (Solar Heating and Cooling Programme: SHC) de la Agencia Internacional de Energía (International Energy Agency: IEA) (IEA-SHC, 2010). Este *benchmark* define el uso de umbrales (0.1; 10; 50 y 200) W/m², resultando en errores cuadráticos medios de 22.9%, 22%, 20.4% y 16.3% respectivamente (SoDa-is, 2013b).

Metodología

Colecta de datos y adaptación temporal

La información de ISG (observaciones) fue descargada de forma periódica mediante robots web programados para el propósito de esta investigación. Todas las observaciones se transformaron a un marco temporal común: tiempo UTC y valores de la ISG en los minutos múltiplos de 15 (minutos: 0, 15, 30 y 45).

Preparación de variables explicativas

Son varios los factores que pueden influir en la cantidad de ISG percibida en una posición de la superficie terrestre. Así diferentes autores utilizan distintas variables explicativas (Bojanowski *et al.*, 2013; Moreno *et al.*, 2011; Alsamamra *et al.*, 2009; Antonanzas-Torres *et al.*, 2009; Kumar *et al.*, 2009; Evrendilek y Ertekin, 2007), entre otros.

Sin embargo ninguna estas investigaciones se focalizó la estimación de valores de ISG a 15 minutos. En este trabajo se seleccionaron tres variables auxiliares. La primera es una abstracción de las horas de luz durante el día y los cambios continuos en los ángulos del Sol, estructurada como una distancia estandarizada calculada para cada observación de la ISG desde su Tiempo Solar Verdadero (True Solar Time: TST) hasta su Medio Día Solar (MDS) con respecto al amanecer y el atardecer, para cada estación meteorológica y para cada día en cuestión. La variable se ha denominado Distancia al Medio Día Solar (DMDS). La segunda variable es la Estimación de la Irradiancia Solar Global (EISG) desde sensores remotos, siguiendo la tendencia de otras investigaciones como es el caso de Hengl *et al.* (2012). La tercera variable utilizada es la latitud, esta se utiliza en los dos métodos de RK y su selección se basó

en la identificación de tendencias con esta variable en Evrendilek y Ertekin (2007), y Kumar *et al.* (2009).

Agrupación de los datos fuente

Se agruparon los datos en seis configuraciones geográficas, a como se ilustra en la Figura 2, bajo distintos criterios: solo estaciones oficiales, romper aglomeraciones, solamente inclusión de estaciones de AEMet, consideración de toda el área de estudio o solo una subárea.

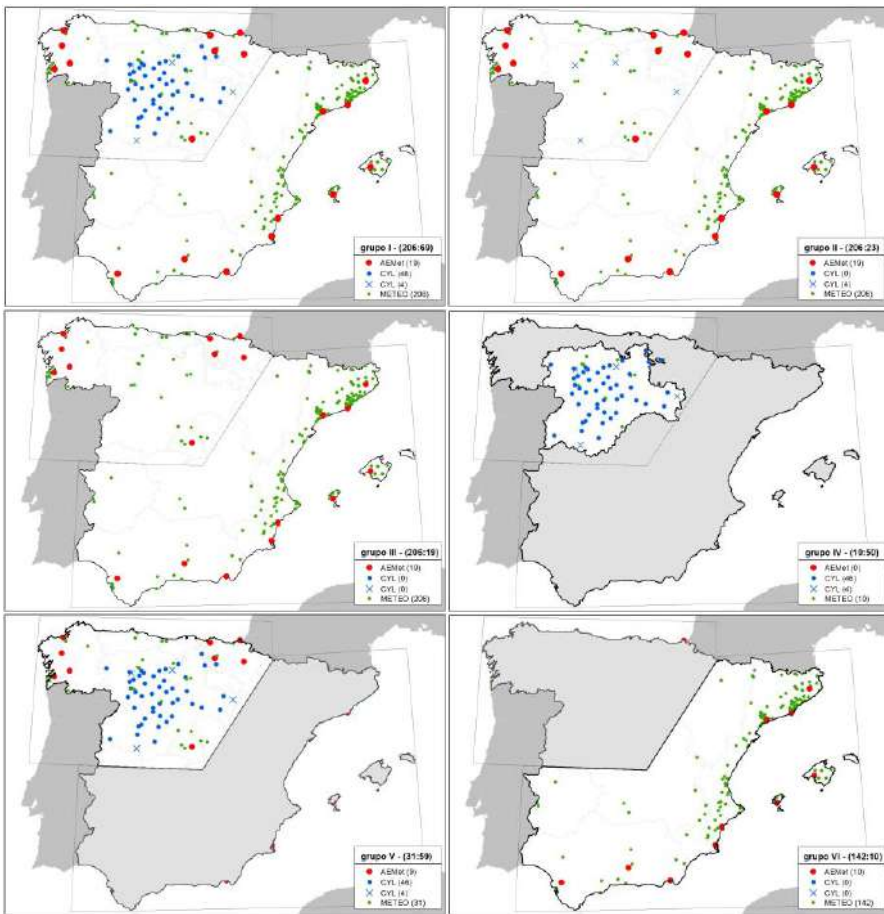


Figura 2. Seis agrupaciones los datos para las estimaciones espaciales. (a) Grupo I (AEMet + CYL); (b) Grupo II (AEMet + 4 CYL); (c) Grupo III (AEMet); (d) Grupo IV (CYL); (e) Grupo V (CYL + 9 AEMet); (f) Grupo VI (AEMet).

Estimación espacial de la ISG para intervalos de 15 minutos

La estimación espacial de la ISG, se ha realizado mediante los métodos descritos en la sección anterior. Se ejecutaron de forma independiente un conjunto de procedimientos para cada una de las seis agrupaciones antes mencionadas. Los procedimientos y cálculos de la metodología se realizaron con el paquete informático R (R Development Core Team, 2012). Para los procesos de interpolación utilizó el paquete geostadístico para R *gstat* (Pebesma, 2004), con excepción de la generación de los semivariogramas, en el que se utilizó el paquete para R *automap* (Automap, 2011).

Detección de valores atípicos

La detección de los valores atípicos (*outliers*) se realizó mediante el método de Box-Plot (Croarkin, 2010), en base a las diferencias entre las observaciones y las estimaciones, denotadas por (Y) e (\hat{Y}) respectivamente. Sus diferencias se agruparon por día y por estación, y se agregaron usando dos estadísticos RMSE (ecuación 1) y MAE (Ecuación 2), a los que se les aplicó el BoxPlot, considerando *outliers* aquellas estaciones identificadas por este método tanto por RMSE como por el MAE.

$$RMSE = \sum_{i=1}^n \sqrt{\frac{1}{n} (Y_i - \hat{Y}_i)^2} \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (2)$$

A continuación se procedió a implementar los métodos IDW, OK y las dos formas de RK. Ambas formas de RK (denominadas RK1 y RK2 respectivamente) se diferencian en una de sus dos variables auxiliares (DMDS en el caso de RK1 y EISG en el caso de RK2), sin embargo ambas comparten como segunda variable auxiliar la “Latitud” geográfica. No se juntaron ambas variables (EISG y DMDS) en un mismo método de RK, por el hecho de que ambas presentan valores de multicolinealidad casi extrema, a pesar de ser ambas variables conceptualmente distintas (Allison, 1998: 141, 150), con valores de correlación de a.m. es 0.7825 y un coeficiente de determinación R2 de 0.6124, y en p.m. el valor de correlación es -0.8102 con R2 de 0.6566.

Validación de los métodos

Para determinar la precisión de los métodos se realizaron validaciones cruzadas. La precisión del modelo HelioSat2 fue obtenida desde la base de datos HC3v2, calculándose las diferencias entre las observaciones en tierra y estos valores.

En todos los casos se calcularon los estadísticos RMSE relativo (%RMSE) y el MAE relativo (%MAE) en base al ratio entre el error estadístico y la media de la variable observada en tanto por 100, como se puede observar en las Ecuaciones (3) y (4) respetivamente:

$$\%RMSE = RMSE / \left(\frac{1}{n} \sum_{i=1}^n (Y_i) \right) \cdot 100 \tag{3}$$

$$\%MAE = \frac{MAE}{\frac{1}{n} \sum_{i=1}^n (Y_i)} \cdot 100 \tag{4}$$

Con el fin último de poder comparar los resultados de los distintos modelos con el benchmarking de referencia en (SoDa-is, 2013b), se calculan los errores para cinco umbrales de la ISG observada en tierra (todo, 0.1, 10, 50 y 200) W/m2, los que también son utilizados por SoDa-is (2013b) para el cálculo de la incertidumbre del modelo HelioSat2.

Resultados y discusión

Puntuación de los métodos de estimación

La precisión y puntuación de los métodos de estimación en cada una de las agrupaciones geográficas, se presentan en la Tabla 2. Se obtuvieron calculando agregaciones diarias de los estadísticos %RMSE y %MAE en base a las diferencias 15-minutales (ISG – EISG) y descartando los días identificados como atípicos por el método BoxPlot.

Tabla 2
Clasificación de los métodos sin umbralizar los valores de irradiancia (agregación diaria) de acuerdo a los 6 grupos de datos: I (AEMet + CYL); II (AEMet + 4 CYL); III (AEMet); IV (CYL); V (CYL + 9 AEMet); VI (10 AEMet)

Puntuación	I		II		III	Estadístico	
1°	RK2	23.75	RK2	27.09	SAT	26.93	
2°	SAT	29.09	SAT	27.49	RK2	28.02	
3°	RK1	30.59	RK1	39.28	IDW	39.86	%RMSE
4°	IDW	30.79	IDW	39.50	RK1	41.17	
5°	OK	31.61	OK	40.56	OK	43.22	

Continuación Tabla 2

Puntuación	IV		V		VI	Estadístico
1°	RK2	21.77	RK2	23.27	SAT	23.12
2°	RK1	25.25	RK1	28.59	RK2	29.82
3°	OK	25.41	IDW	28.86	IDW	34.16
4°	IDW	25.86	OK	28.92	OK	36.75
5°	SAT	28.95	SAT	29.94	RK1	37.91

Desde los valores anteriores se deduce que los métodos RK2 y SAT son siempre los ganadores (por simplicidad solo se muestra %RMSE). Siendo el método RK2 el ganador en las formas de agrupamiento (I, II, IV y V) y SAT en las formas de agrupamiento (III y VI) en estas últimas, solo participan las estaciones de AEMet con el menor número de estaciones por área de trabajo (poca densificación).

Se seleccionó como mejor método el que presenta el valor agregado de la media más bajo para cada tipo de agrupamiento. En la Tabla 3 se presentan los métodos ganadores junto con otros valores de relevancia (por simplicidad solo se muestra %RMSE). Se aprecia que aplicando el método RK2 al grupo IV (CYL) se obtiene el menor valor medio de error (el mejor resultado). También se observa que el grupo VI (10 AEMet) con el método SAT tiene el segundo valor más bajo de media, sin embargo, la unión de las áreas de los dos mejores grupos/métodos (IV y V) no cubren toda el área de estudio. Por lo anterior, también se toma en consideración el Grupo (V) con el método RK2, por presentar el siguiente mejor valor de media.

Tabla 3
El mejor método por grupo y por estadístico, sin discriminar valores de irradiancia (agregación para el número total de días)

Grupo	Mejor método	% RMSE	Stdv (absoluta)	Estadístico
I (CYL + AEMet)	RK2	23.75	11.38	
II (AEMet + 4 CYL)	RK2	29.09	10.02	
III (AEMet)	SAT	26.93	10.03	
IV (CYL)	RK2	21.77	11.57	%RMSE
V (CYL + 9 AEMet)	RK2	23.27	11.63	
VI (10 AEMet)	SAT	23.12	10.53	

Análisis detallado de los mejores métodos

Los análisis presentados a continuación incluyen todos los días que forman parte de los experimentos (es decir, sin descartar los días identificados como atípicos), esto permite comparar los resultados del benchmarking en (SoDa-is, 2013b) y los trabajos relacionados.

- RK2 en la forma de agrupación IV (CYL)

La Tabla 4 presenta los estadísticos para el método RK2 en el agrupamiento (IV-CYL) con seis umbrales de ISG (todo, 0.1, 10, 50, 200 W/m²). Se puede apreciar que los estadísticos porcentuales (%RMSE y %Stdv) van disminuyendo a medida que el umbral aumenta (ISG), además estos valores son más bajos que los respectivos de (69, 72, 77, 83) W/m² y (22.9, 20.2, 20.04, 16.3) % y para umbrales (0.1, 10, 50 y 200) W/m² del benchmarking en (SoDa-is, 2013b).

Tabla 4
Estadísticos para el método RK2 en el grupo IV (CYL)

<i>Umbral</i>	<i>%RMSE (%)</i>	<i>Stdv %RMSE (%)</i>	<i>RMSE (W/m²)</i>	<i>Stdv RMSE (W/m²)</i>	<i>Rango RMSE (W/m²)</i>	<i>%RMSE (All)</i>	<i>Stdv RMSEp (All)</i>
NA	22.46	2.87	68.95	10.36	59-79	247.31	508.06
0.1	22.46	2.87	68.95	10.36	59-79	148.74	362.46
10	21.03	2.75	68.97	10.59	59-81	51.90	25.22
50	17.51	2.45	72.86	11.51	61-84	32.73	9.15
200	11.26	1.35	75.21	12.97	62-88	17.88	3.23

El análisis de la relación entre el %RMSE y la distancia que separa el punto a interpolar de los vecinos, en esta forma de agrupamiento se encontró que, el %RMSE es inferior o igual al 25.5%, independientemente de la distancia que separa el punto a interpolar del vecino más cercano, o independientemente de la distancia media a todos los vecinos, a como se muestra en la Figura 3, esto implica que el método desarrollado no se superan el límite de 25% de %RMSE para observaciones horarias indicado por Zelenka et al. (1999) y Pérez et al. (1997).

- RK2 en la forma de agrupación V (CYL + 9 AEMet)

En la Tabla 5 se presentan los estadísticos resultantes de aplicar el método RK2 en la forma de agrupamiento V (CYL + 9 AEMet) para 6 umbrales de la ISG

(todo, 0.1, 10, 50, 200). De aquí cabe destacar que los valores %RMSE para los umbrales 50 y 200 (18.95 y 12.29) %, son mejores que los respectivos valores de (20.04 y 16.3) % del benchmarking en (SoDa-is, 2013b); mientras que los valores RMSE a partir de los umbrales 10 (70.49, 68.95 y 63.39) W/m², son mejores que los equivalentes del citado benchmarking.

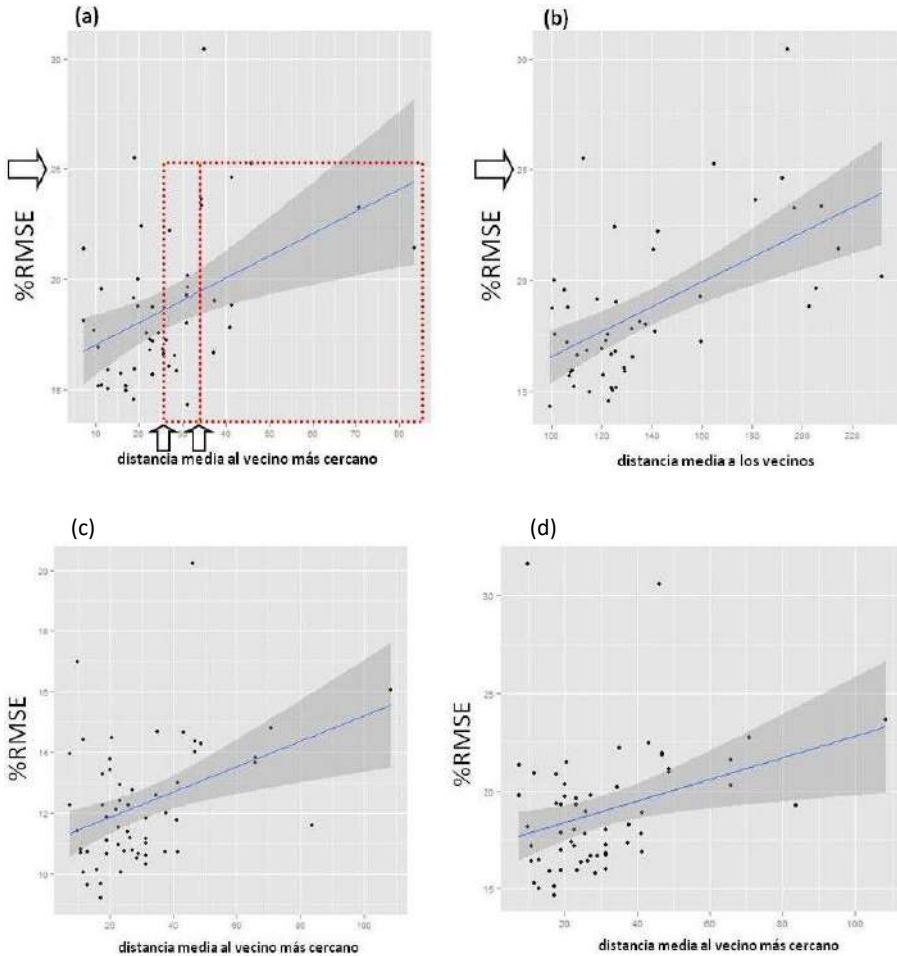


Figura 3. %RMSE de las interpolaciones en relación a los vecinos para el método RK2 en CYL; (a) y (b) muestra el análisis sin umbrales aplicados; (c) y (d) muestra los valores con un umbral aplicado de mayor igual a 50 w/m².

Tabla 5
Estadísticos para el método RK2 en el grupo V (CYL + 9 AEMet)

<i>Umbral</i>	<i>%RMSE (%)</i>	<i>Stdv %RM SE (%)</i>	<i>RMSE: Wm2</i>	<i>Stdv RMSE (Wm2)</i>	<i>Rango RMSE (Wm2)</i>	<i>%RMSE (All)</i>	<i>Stdv %RMSE (All)</i>
NA	24.70	4.26	71.23	23.38	48-95	533.20	1'114.30
0.1	24.69	4.26	71.23	23.38	48-95	357.53	933.02
10	22.90	3.83	70.49	22.96	48-93	63.97	53.70
50	18.95	3.22	68.95	19.11	50-88	37.98	17.61
200	12.29	2.01	63.39	12.25	51-76	22.42	13.76

- SAT en la forma de agrupación VI (10 AEMet):
En la Tabla 6 se presentan los estadísticos para el método HelioSat2 en la forma de agrupar (VI- 10 AEMet) para los seis umbrales de la ISG (todo, 0.1, 10, 50, 200) W/m2. Estos se comparan con al benchmarking en (SoDa-is 2013b), destacando que aunque los % RMSE para los umbrales 0.1 y 10 W/m2 son muy parecidos, los valores de los umbrales 50 y 200 W/m2 son más bajos que los del citado benchmarking y parecidos a los de los casos RK2 de las agrupaciones IV y V.

Tabla 6
Estadísticos para las estimaciones de la ISG derivadas de imágenes de satélite en el grupo VI (10 AEMet)

<i>Umbral</i>	<i>% RMSE</i>	<i>Stdv % RMSE</i>	<i>RMSE (Wm2)</i>	<i>Stdv RMSE (Wm2)</i>	<i>Rango RMSE (Wm2)</i>	<i>% RMSE (All)</i>	<i>Stdv % RMSE (All)</i>
NA	22.85	0.87	105.13	15.67	89-121	713.09	823.02
0.1	22.84	0.87	105.13	15.68	89-121	500.27	476.25
10	21.55	1.03	104.64	15.78	89-120	84.05	33.26
50	18.04	1.27	103.59	15.12	88-119	46.06	8.14
200	11.46	1.25	101.21	15.91	85-117	23.13	2.64

Predicción espacio-tiempo de la ISG

Cuando se trabaja con contextos espaciales reducidos y en el corto plazo de tiempo (pequeña escala espacio-tiempo), la ISG se ve afectada por las condiciones dinámicas

de la atmósfera y a un alto ruido debido a factores como a la nubosidad. Estas condiciones presentan relaciones no lineales que vuelven compleja la predicción a corto plazo de la ISG. Las RNA forman parte de la IA y el ML, las que han demostrado su utilidad para investigar modelos asociados a procesos naturales que aproximan funciones no lineales, como es el caso de la predicción de la ISG a corto plazo.

El objetivo principal de esta sección ha sido investigar científicamente la predicción de la ISG en distintos horizontes temporales a corto plazo, donde este tipo de modelación es compleja, mediante el modelado de RNA, incluyendo la componente espacial como entrada a los mismos.

Ámbito de estudio y datos experimentales

Al igual que los trabajos relacionados (Voyant *et al.*, 2014; Wang *et al.*, 2012; Mellit y Pavan, 2010), se realizó la predicción de la ISG a corto plazo sobre un punto fijo (objetivo de predicción), sin embargo una innovación de esta investigación reside en el incluir la componente espacial en los datos de entrada, mediante el uso de datos en paralelo procedentes de estaciones vecinas al punto objetivo.

Sitio de estudio

La selección de las EM a utilizar ha estado condicionada al uso de REM oficiales, a la distribución geográfica y a la densidad de estaciones con sensores de la ISG en ellas contenidas, quedando los experimentos acotados, al uso de 10 estaciones cercanas al centro de gravedad de Castilla y León.

Datos experimentales

Los datos experimentales utilizados son las observaciones del año 2011 obtenidos desde esta REM de Castilla y León, utilizándose cinco variables meteorológicas: (1) ISG, (2) Temperatura Ambiente (TA), (3) Humedad Relativa del Aire (HRA), (4) Dirección del Viento (DV) y (5) Velocidad del Viento (VV). Adicionalmente se utilizaron 4 variables, relacionadas, calculadas de un modo determinista y referidas al instante de tiempo exacto de la toma de las observaciones meteorológicas, siendo éstas: (6) la Irradiancia Solar Extraterrestre Horizontal (ISEH), (7) el Índice de Claridad Instantáneo (KTi), (8) Distancia al Medio Día Solar (DMDS) y (9) el Ángulo Cenit (AC).

Metodología

Las fases de la metodología utilizada para investigar las predicciones de la ISG a corto plazo mediante RNA se vio influenciadas por el objetivo de experimentar in-

cluyendo un gran número de entradas en estas RNA. Todos los procesos se desarrollaron con el lenguaje de programación y *software* estadístico R.

Normalización y selección de datos (entrenamiento y test)

Con el objeto de lograr convergencia en las redes, los datos se normalizaron entre [0, 1] según la ecuación 5. Donde, $x \in [x_{min}, x_{max}]$ y, $y \in [y_{min}, y_{max}]$; “x” es el dato original y “y” es el su valor normalizado; Asumiendo los valores de $y_{min} = 0$; $y_{max} = 1$.

$$y = \frac{(y_{max} - y_{min})(x - x_{min})}{x - x_{min}} + y_{min} \quad (5)$$

Parámetros de entrenamiento de las RNA

Con el propósito de determinar las mejores configuraciones de RNA aplicadas a una arquitectura de tipo Perceptrón Multicapas (MLP), manteniendo la mayor cantidad de entradas posibles, se estudiaron algunos parámetros que influyen en el aprendizaje y convergencia de la red como lo son: (i) el número de capas ocultas, (ii) la cantidad de neuronas en cada capa, (iii) el intervalo de normalización de los datos, (iv) el ratio de aprendizaje, (v) la cantidad de veces que se entregan todos los ejemplos a la RNA para su entrenamiento (Epochs). Otros parámetros se dejaron fijos, basados en los buenos resultados que estos ya han demostrado en estudios previos: (vi) el algoritmo de entrenamiento utilizado, ha sido BackPropagation, (vi) la función de comparación seleccionada ha sido la “suma de los errores al cuadrado”.

Modelos de RNA y estadísticos para medir la calidad de las predicciones

Los modelos de RNA creados se concentraron en la predicción futura de la ISG en horizontes temporales desde una hasta seis horas, experimentándose de forma independiente con un conjunto de RNA por cada horizonte de tiempo. En todos los casos, los datos de entrada utilizaron información de las distintas variables en instantes de tiempo previos, lo que se conoce como el método de ventana deslizante. Este método consiste en predecir el valor futuro de la ISG (Δt) —el cual define el horizonte temporal de la predicción— utilizando como entradas un conjunto de variables correspondientes a una ventana de tiempo previa.

Los modelos de RNA aquí creados se pueden agrupar en dos tipos que, se diferencian entre sí por la utilización, o no, de datos procedentes de estaciones vecinas. En la figura 4 se representa la arquitectura general simplificada de los métodos basados en RNA desarrolladas. Aquí se puede apreciar cómo llegan a las entradas de la RNA datos de estaciones vecinas, y valores de sus variables para distintos intervalos

de tiempo mediante (método de ventana deslizante), también se aprecia que estas RNA poseen dos capas ocultas de neuronas, y como salida se obtiene el valor de la ISG predicho para un horizonte Δt .

La calidad de las predicciones de las distintas RNA creadas, se llevó a cabo utilizando cuatro estadísticos: RMSE, %RMSE, %MAE, nRMSE. El nRMSE se define en la siguiente ecuación (los otros se definieron previamente). Donde Y_i es el valor observado de la ISG, mientras que \hat{Y}_i es el valor predicho por la RNA. Se seleccionaron estos estadísticos para poder hacer comparaciones con investigaciones relacionadas.

$$nRMSE = \sum_{i=1}^n \sqrt{\frac{(Y_i - \hat{Y}_i)^2}{Y_i^2}} \quad (6)$$

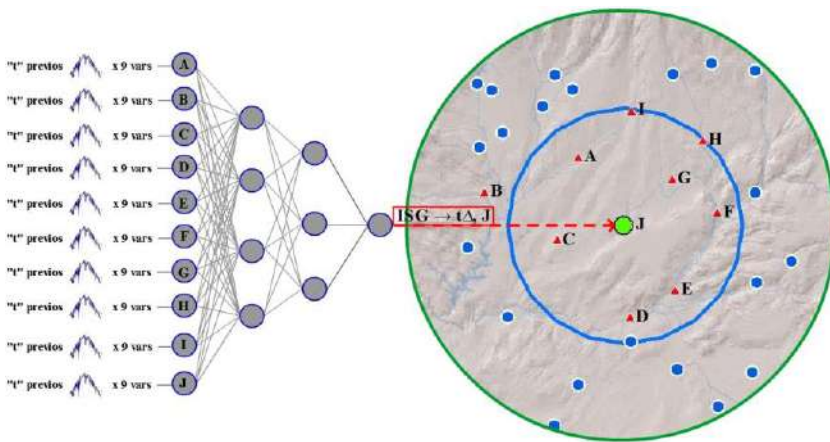


Figura 4. Representación simplificada de una RNA para la predicción la ISG en un horizonte “ Δt ” futuro, en base a datos de distintas estaciones vecinas y de sus variables en diferentes instantes de tiempo.

Resultados y discusión

Parámetros para la selección de las arquitecturas de RNA

El proceso de identificación de los parámetros idóneos para el correcto entrenamiento de las redes fue iterativo, programándose en R experimentos con varias arquitecturas y configuraciones de red. A continuación se presentan los resultados del %RMSE de las validaciones cruzadas (utilizando los datos en test) de una muestra representativa

de algunos de los experimentos, distribuidos en distintas tablas. Los experimentos han permitido identificar los parámetros prometedores para, posteriormente utilizarlos en el entrenamiento de nuevas RNA que optimizaran el resultado de las predicciones.

- **Parámetros y RNA para horizonte temporal de 1 hora.**
En esta tabla 7, las primeras RNA (1-3) realizan la predicción utilizando datos de estaciones vecinas, mientras que las ultimas (4-6) lo hacen con sólo observaciones de instantes previos de la estación objetivo. En ambos casos, se puede apreciar que los mejores niveles de error se encuentran con un MLP de una capa oculta (HL1). También se observa cómo el uso de dos capas ocultas (HL1 y HL2) asemeja los niveles de errores al del Perceptron Simple, sin embargo cabe destacar que tanto el Perceptron Simple como las otras configuraciones de RNA mantienen niveles de errores no muy alejados entre sí, lo cual es un indicio de que para este horizonte temporal el problema tiene una fuerte relación lineal.

Tabla 7
Identificación de parámetros para RNA que predicen la ISG:
Pruebas con horizonte de una hora

<i>Red #</i>	<i>Otros parámetros</i>	<i>Arquitectura de Red</i>	<i>% RMSE en TEST</i>
1	Delta-Deriv = 0.03	PERCEPTRON	24.62%
2	Delta-Deriv = 0.1	10HL1	23.99%
3	Delta-Deriv = 0.05	10HL1-5HL2	25.10%
4	Delta-Deriv = 0.03	PERCEPTRON	26.25%
5	Delta-Deriv = 0.025	30HL1	25.77%
6	Delta-Deriv = 0.025	10HL1-5HL2	25.82%

- **Parámetros y RNA para horizontes de más de una hora.**
A continuación se presenta una muestra de las RNA con las que se experimentó la predicción futura de la ISG para un horizonte temporal de dos horas, mostrándose los resultados en la tabla 8. Se aprecia que las mejores predicciones se alcanzan con MLP de dos capas ocultas, también que los Perceptrones Simples no logran converger ni alcanzar niveles de errores aceptables, lo que implica que la predicción de la ISG para un horizonte de 2 o más horas, deja de ser un problema con relaciones lineales. En esta Tabla, no se muestran resultados de los experimentos para horizontes temporales de 3-6 horas puesto que el patrón encontrado fue similar a lo antes mencionado.

Tabla 8
Identificación de parámetros para RNA que predicen la ISG:
Pruebas con horizonte de dos horas

<i>Red #</i>	<i>Otros parámetros</i>	<i>Arquitectura de Red</i>	<i>%RMSE en TEST</i>
7	Delta-Deriv = 0.03	PERCEPTRON	mayor al 100%
8	Delta-Deriv = 0.05	10HL1	26.97%
9	Delta-Deriv = 0.05	100HL1-10HL2	24.23%
10	Delta-Deriv = 0.035	300HL1	mayor al 100%
11	Delta-Deriv = 0.035	300HL1-25HL2	27.67%
12	Delta-Deriv = 0.055	50HL1-25HL2	No convergió

Predicciones de la ISG a corto plazo

A continuación se presentan los resultados definitivos de modelar la predicción a corto plazo de la ISG mediante RNA. Una vez identificada la arquitectura RNA y sus parámetros de entrada, se generaron más experimentos para la optimización de los modelos de RNA definitivos. Se identificó que para la predicción a corto plazo de la ISG con estaciones vecinas, los mejores resultados se alcanzan con 100 neuronas en la primera capa oculta y 10 neuronas en la segunda. En el caso de no utilizar datos de estaciones vecinas, se encontró que los mejores resultados los genera un MLP con una capa oculta con 30 neuronas, siendo estos los parámetros definitivos utilizados.

Análisis global de los resultados

En la Tabla 9 se presentan los resultados globales (sin desagregar) de las predicciones de la ISG calculados a partir de las validaciones cruzadas (utilizando los datos en test) y agrupadas según los distintos horizontes de predicción (1-6 horas) y por la inclusión de estaciones vecinas en los datos de entrada.

Los resultados mostrados en la Tabla 9 muestran cómo en todos los casos los modelos de RNA creados, logran predecir la ISG a corto plazo con niveles de errores inferiores al 20% de nRMSE. Además desde la tabla se destaca cómo las RNA que predicen la ISG para horizontes temporales de una a tres horas y que incluyen en sus entradas los datos de estaciones vecinas, presentan mejores valores de nRMSE (12.88%, 13.80%, 15.98%), contrario a lo que ocurre con las RNA que sólo usan datos de instantes previos del mismo punto objetivo de predicción (13.38%, 14.84%, 16.04%). También es importante destacar que para los horizontes temporales de entre cuatro a seis horas los resultados se invierten, siendo mejores el segundo tipo de RNA

con valores de nRMSE de (17.36%, 17.27% y 18.27%) contra los valores de (17.94%, 20.07% y 19.62%) del primer tipo de RNA.

Tabla 9
Resumen de resultados de las validaciones cruzadas de las RNA que estiman la ISG a corto plazo para horizontes desde una a

Nombre de Red	Arquitectura de Red	% RMSE	% MAE	RMSE W/m2	nRMSE
<i>RNA que utilizan datos de estaciones vecinas así como valores de instantes de tiempo previos de las distintas variables</i>					
ANN1	100-10 (D-0.05)	22.63	35.55	43.33	12.88
ANN2	100-10 (D-0.035)	24.24	38.15	47.49	13.80
ANN3	100-10 (D-0.05)	28.07	37.52	54.99	15.98
ANN4	100-10 (D-0.05)	31.52	46.96	61.75	17.94
ANN5	100-10 (D-0.05)	35.25	45.32	69.04	20.07
ANN6	100-10 (D-0.05)	34.47	43.95	67.52	19.62
Promedios:		29.36	41.24	57.36	16.72
<i>RNA que sólo utilizan datos de instantes de tiempo previos de la propia estación objetivo de predicción</i>					
ANN1-B	30 (D-0.05)	23.05	26.21	46.04	13.38
ANN2-B	30 (D-0.05)	26.06	44.52	51.05	14.84
ANN3-B	30 (D-0.05)	28.81	33.78	56.43	16.04
ANN4-B	30 (D-0.055)	30.49	44.02	59.74	17.36
ANN5-B	30 (D-0.055)	30.34	46.74	59.43	17.27
ANN6-B	30 (D-0.055)	32.09	47.99	62.86	18.27
Promedios		28.95	40.92	56.69	16.47

Los resultados alcanzados en esta investigación presentan un valor promedio en nRMSE de 16.45%. Este valor es mejor, y a la vez consistente, con los valores alcanzados en los trabajos de Voyant *et al.* (2014) y Paoli *et al.* (2010), en los que obtuvieron valores de nRMSE de 27.8% y ~20% respectivamente.

El valor medio del RMSE resultante es de 57.02 W/m2, este valor es mejor promedio, entre días claros y nuboso, que el valor de 63.47 W/m2 obtenido por Wang *et al.* (2012). La comparación de los resultados con el otro estadístico común (%MAE) y el citado trabajo previo, no se ha podido establecer. Es posible que las diferencias se deban a las distintas localizaciones de las áreas de estudio, estando Wang *et al.*

(2012) ubicado en China, mientras que, tanto la investigación aquí desarrollada como las de Paoli *et al.* (2010) y Voyant *et al.* (2014) están ubicadas en el Sur de Europa. Otra diferencia es el hecho de que, en Wang *et al.* (2012), el análisis de los errores se hace mezclando las predicciones en periodos de día como en periodos nocturnos.

La Figura 5 muestra un ejemplo de pronóstico de serie temporal para un día (20 de agosto de 2011), utilizando los modelos de RNA desarrollados.

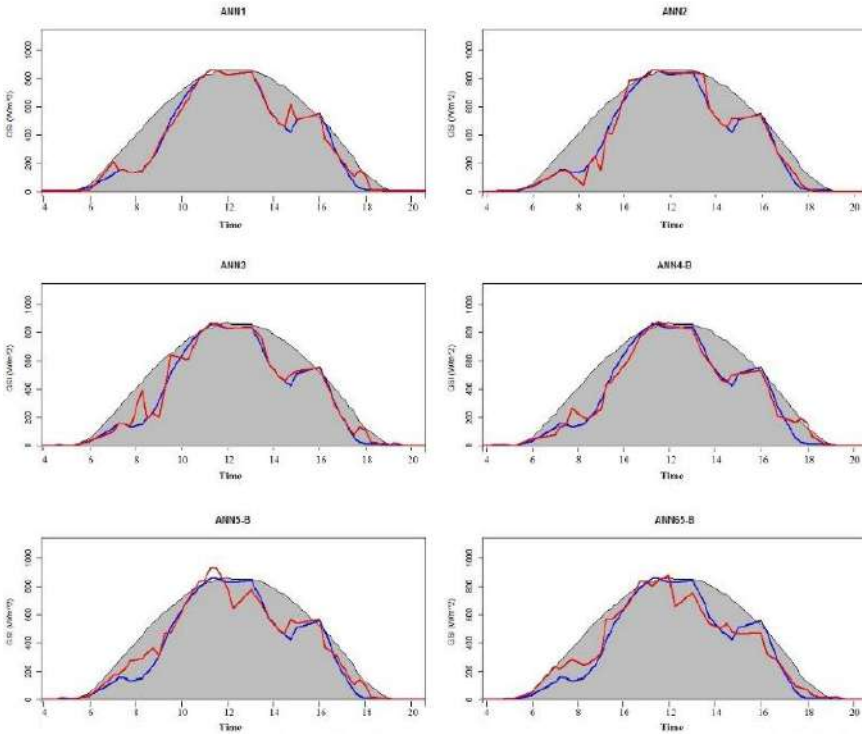


Figura 5. Ejemplo de series temporales basadas en valores pronosticados por las RNA.

Las Figuras 5 (“a”, “b” y “c”) presenta las predicciones para RNA que usan datos de estaciones vecinas (ANN1, ANN2 y ANN3) para horizontes temporales de una, dos y tres horas respectivamente; las figuras 5 (“d”, “e” y “f”) muestra las predicciones de RNA que solo usan observaciones de la propia estación objetivo (ANN4-B, ANN5-B y ANN6-B) para horizontes temporales de cuatro, cinco y seis horas respectivamente. Las líneas azules representan los valores observados por los sensores. Las líneas rojas representan los valores de las predicciones realizados por los modelos de RNA. El área gris representa el pronóstico basado persistencia de tipo *clear-sky*

expectations. Aquí se puede apreciar como la curva de predicción se acerca a la curva de las observaciones, a pesar de que la primera se proyecta horas antes de la segunda.

Conclusiones

El mejor método de estimación espacial de la IS en intervalos de 15 minutos para el territorio español es el Kriging con Regresión (RK), soportado por los valores de la IS estimados desde imágenes de satélite y la latitud geográfica de la localización de la estación como variables auxiliares (KR2).

Se encontró que el método KR2 permite predecir espacialmente (interpolando) la IS en superficie más allá de los 25 km, límite que indica la bibliografía de referencia. El método KR2 ha sido probado hasta los 108 km, distancia para la que el %RMSE nunca alcanza el 25% indicado como límite.

Por lo tanto, el método KR2 desarrollado en esta investigación aprovecha las fortalezas de ambas fuentes de información (observaciones en superficie y valores derivados de imágenes de satélite); esto permite cubrir distancias de interpolación más allá del límite que implica el hecho de sólo utilizar las observaciones de las estaciones en tierra pero con un nivel de error menor al de sólo utilizar valores derivados desde imágenes de satélite.

Lo anterior habilita algunas aplicaciones prácticas. Por ejemplo, la futura generación de superficies de IS en intervalos de 15 minutos con el propósito de estimar el potencial de producción de energía generada por el Sol en lugares donde no hay observaciones directas de la IS y con un error inferior al 25% de RMSE.

Para cualquiera de los métodos (Interpolaciones o Estimación por Satélite), los outliers encontrados son mayores cuando se trata de estimar la IS en franjas de tiempo donde los valores de la IS son bajos (amanecer/anochecer).

La consideración en los modelos de RNA de la mayor cantidad de entradas, mediante la inclusión de la componente espacial, ha implicado experimentar con el modelado de RNA de hasta 900 entradas (altas dimensiones).

Los resultados de esta investigación científica han permitido generar modelos de RNA que predican la ISG a corto plazo con niveles de exactitud inferiores al 20% del nRMSE, en todos los horizontes temporales, lo cual en valores promedios, es una mejora en cuanto a las investigaciones relacionadas.

Los resultados también han permitido identificar el aporte y los límites de la inclusión de estaciones vecinas en los datos de entrada a los modelos de RNA; así se encontró que existe una relación entre la distancia de los vecinos y el aporte que hacen al horizonte temporal de predicción. Para horizontes de predicción de hasta 3 horas, lo mejor es utilizar datos de las estaciones vecinas a distancias inferiores a 55 km respecto a la estación objetivo de predicción, mientras que para horizontes de entre 4 y 6 horas es mejor no utilizar dichas estaciones en ese rango de distancias.

La relación entre las predicciones en distintos horizontes de tiempo y las distancias en que se encuentran las estaciones vecinas, se puede explicar por el traslado del efecto meteorológico local de éstas hacia la estación objetivo de predicción, de tal forma que estaciones a 55 km o menos de lejanía y las condiciones metrológicas temporalmente cercanas (una a tres horas) influirán en la estación que es objetivo de predicción, mientras que los efectos de las observaciones de estas estaciones ya habrán pasado para horizontes de (cuatro a seis) horas, por lo cual, en trabajo futuros se debería investigar otras configuraciones especiales para la inclusión de la componente especial en los modelos RNA para mayores horizontes de tiempo que aporten mejoras en los niveles de errores.

Bibliografía

- Allison, P. (1998). Federico Valdimir *Multiple regression: A primer*, Thousand Oaks, California, USA, Pine Forge Press.
- Alsamamra, H.; Ruiz-Arias, J.A.; Pozo-Vázquez, D.; Tovar-Pescador, J. (2009). "A Comparative Study of Ordinary and Residual Kriging Techniques for Mapping Global Solar Radiation Over Southern Spain", *Agric. For. Meteorol*, 149: 1343-1357. <http://dx.doi.org/10.1016/j.agrformet.2009.03.005>.
- Antonanzas-Torres, F.; Cañizares, F. and Perpiñán, O. (2013). "Comparative Assessment of Global Irradiation from a Satellite Estimate Model (CM SAF) and on-Ground Measurements (SIAR): A Spanish Case Study", *Renewable and Sustainable Energy Reviews*, 21: 248-261. <http://dx.doi.org/10.1016/j.rser.2012.12.033>.
- Automap (2011). Automatic Interpolation Package. Recuperado de <http://cran.r-project.org/web/packages/automap>.
- Badescu, V. (ed.) (2008). *Modeling Solar Radiation at the Earth's Surface*, Springer-Verlag Berlin Heidelberg, pp. vi-xvii.
- Bojanowski, J.S.; Vrieling, A.; Skidmore, A.K. (2013). "Calibration of Solar Radiation Models for Europe using Meteosat Second Generation and Weather Station Data", *Agric. For. Meteorol*, 176: 1-9. <http://dx.doi.org/10.1016/j.agrformet.2013.03.005>.
- Bosch, J.L., Lopez, G. & Batlles, F.J. (2008). "Daily solar irradiation estimation over a mountainous area using artificial neural networks", *Renewable Energy*, 33(7): 1622-1628.
- Croarkin, C. & Tobias, P. (2010). *NIST/SEMATECH e-Handbook of Statistical Methods. The National Institute of Standards and Technology 2010*, Gaithersburg, MD, USA.

- Estévez, J. and Gavilán, P. (2008). “Procedimientos de validación de datos de estaciones meteorológicas automáticas. Aplicación a la Red de Información Agroclimática de Andalucía”, II Jornadas de Agrometeorología 2008.
- EC (European Commission) (2013). *Joint Research Center Scientific and Policy Report. PV Status Report 2013*, Publication office of the European Union 2013, Ispra, Italy. Recuperado de https://helapco.gr/pdf/pv_status_report_2013.pdf.
- EEN (European Economy News) (2010). SMART, Sustainable and Inclusive Growth. EUR-LEX 2010. Recuperado de <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC2020&from=EN>.
- Evrendilek, F. and Ertekin, C. (2007). “Statistical Modeling of Spatio-Temporal Variability in Monthly Average Daily Solar Radiation Over Turkey”, *Sensors*, 7 (11): 2763-2778. <http://dx.doi.org/10.3390/s7112763>.
- Finn, P.G.; Udy, N.S.; Baltais, S.J.; Price, K. and Coles, L. (2010). “Assessing the Quality of Seagrass Data Collected by Community Volunteers in Moreton Bay Marine Park, Australia”, *Environmental Conservation*, 37(1): 83-89.
- Gautier, C.; Diak, G. and Masse, S. (1980). “A Simple Physical Model to Estimate Incident Solar Radiation at the Surface from GOES Satellite Data”, *J. Appl. Meteor*, 19: 1005-1012.
- Gueymard, C.A. and Myers D.R. (2008). Solar radiation measurement: Progress in radiometry for improved modeling. *Modeling Solar Radiation at the Earth Surface*, Springer-Verlag Berlin Heidelberg, pp. 1-27. http://dx.doi.org/10.1007/978-3-540-77455-6_1.
- Hagan, M.T.; Demuth, H. B. and Beale, M.H. (1996). *Neural network design*, vol. 1, Boston: Pws.
- Hartley, B. (2012). “The Internet of Things - Weather Monitoring Too”, *WMO technical conference on meteorological and environmental instruments and methods of observation*, Belgium, Brussels.
- Hengl, T.; Heuvelink, G.B.; Tadić, M.P. and Pebesma, E.J. (2012). “Spatio-Temporal Prediction of Daily Temperatures using Time-Series of MODIS LST Images”, *Theoretical and applied climatology*, 107(1-2): 265-277.
- Hernández, L.; Baladrón, C.; Aguiar, J.M.; Calavia, L.; Carro, B.; Sánchez-Esguevillas, A.; Cook, D.J.; Chinarro, D. and Gómez, J. (2012). “A Study of the Relationship between Weather Variables and Electric Power Demand Inside a Smart grid/smart World Framework”, *Sensors*, 12(9): 11571-11591.
- Hontoria, L.; Aguilera, J. and Zufiria, P. (2002). “Generation of hourly irradiation synthetic series using the neural network multilayer perceptron”, *Solar Energy*, 72(5): 441-446.
- Hudson, G. (1993). “Kriging temperature in Scotland using the external drift method 2”, *Geostatistics Troia '92*, 1: 577-588, Dordrecht, The Netherlands Springer Netherlands. http://dx.doi.org/10.1007/978-94-011-1739-5_45.

- IDAE (2010). *Informe de previsión. Plan de Energías Renovables 2011-2020*. Recuperado de <https://www.idae.es/tecnologias/energias-renovables/plan-de-energias-renovables-2011-2020>.
- IEA (2014). *2014 key world energy statistics*. Recuperado de <http://www.iea.org/publications/freepublications/publication/KeyWorld2014.pdf>.
- IEA-SHC (2010). *Task 36 Objectives*. Recuperado de <http://task36.iea-shc.org/publications>.
- Kaya, D. (2006). "Renewable energy policies in Turkey", *Renewable and Sustainable Energy Reviews*, 10 (2): 152-163.
- Knights, M. (2007). "Web 2.0", *Communications Engineer*, 5(1): 30-35.
- Koca, A., Oztop, H.F., Varol, Y. & Koca, G.O. (2011). "Estimation of solar radiation using artificial neural networks with different input parameters for Mediterranean region of Anatolia in Turkey", *Expert Systems with Applications*, 38(7): 8756-8762.
- Kumar, L.; Skidmore, A.K.; Knowles, E. (2009). "Modelling Topographic Variation in Solar Radiation in a GIS Environment", *Int. J. Geogr. Inf. Sci.*, 11(5):475-497.
- Laszloa, I.; Cirenba, P.; Liuba, H.; Kondraguntaa, S.; Tarpleya, J.D. & Goldberga, M.D. (2008). "Remote sensing of aerosol and radiation from geostationary satellites", *Advances in Space Research*, 41(11): pp. 1882-1893.
- Linares-Rodríguez, A., Ruiz-Arias, J. A., Pozo-Vázquez, D. & Tovar-Pescador, J. (2011). "Generation of synthetic daily global solar radiation data based on ERA-Interim reanalysis and artificial neural networks", *Energy*, 36(8): 5356-5365.
- Mayo, L. (2013). "The Sun: 101!. Sun-Earth Days", NASA. Recuperado de http://sunearthday.nasa.gov/2013/articles/thesun_101.php
- Mellit, A., Benghanem, M. & Kalogirou, S.A. (2006). "An adaptive wavelet-network model for forecasting daily total solar-radiation", *Applied Energy*, 83(7): 705-722.
- Mellit, A. & Kalogirou, S.A. (2008). "Artificial intelligence techniques for photovoltaic applications: A review", *Progress in Energy and Combustion Science*, 34(5): 574-632. <http://dx.doi.org/10.1016/j.pecs.2008.01.001>.
- Mellit, A.; Kalogirou, S.A.; Hontoria, L. & Shaari, S. (2008). "Artificial intelligence techniques for sizing photovoltaic systems: A review", *Renewable and Sustainable Energy Reviews*, 13(2): 406-419. <http://dx.doi.org/10.1016/j.rser.2008.01.006>.
- Mellit A. & Pavan A.M. (2010). "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy", *Solar Energy*, 5(84): 802-821.
- MITYC (2009). Informe en relación con la cumplimentación del artículo 4.3 de la directiva 2009/28/ce, de 23 de abril de 2009, relativa al fomento del uso de energía procedente de fuentes renovables.

- Moebius, E. (2006). Solar Wind Electricity? NASA's Cosmicopia. http://helios.gsfc.nasa.gov/qa_sun.html#swelectric
- Mohandes, M.; Rehman, S. and Halawani, T. O. (1998). "Estimation of global solar radiation using artificial neural networks", *Renewable Energy*, 14(1): 179-184.
- Miorandi, D.; Sicari, S.; De Pellegrini, F. & Chlamtac, I. (2012). "Internet of things: vision, applications and research challenges", *Ad Hoc Netw*, 10(7): 1497-1516.
- Moreno, A.; Gilabert, M. & Martínez, B. (2011). "Mapping Daily Global Solar Irradiation Over Spain: A Comparative Study of Selected Approaches", *Sol. Energy*, 85(9): 2072-2084.
- Mubiru, J. (2008). "Predicting total solar irradiation values using artificial neural networks", *Renewable Energy*, 33(10): 2329-2332.
- Mubiru, J. & Banda, E.J.K.B. (2008). "Estimation of monthly average daily global solar irradiation using artificial neural networks", *Solar Energy*, 82(2): 181-187.
- Paoli, C., Voyant, C., Muselli, M. & Nivet, M.L. (2010). Forecasting of preprocessed daily solar radiation time series using neural networks, *Solar Energy*, 84(12): 2146-2160.
- Pebesma, E.J. (2004). "Multivariable Geostatistics in S: The Gstat Package", *Comput. Geosci.*, 30(7): 683-691.
- (2006). "The Role of External Variables and GIS Databases in Geostatistical Analysis", *Trans. GIS*, 10(4): 615-632.
- Perpiñan, O. (2008). "Grandes centrales fotovoltaicas: Producción, seguimiento y ciclo de vida", tesis doctoral, ETSII, UNED, Madrid, España.
- Perez, R.; Seals, R. & Zelenka, A. (1997). "Comparing Satellite Remote Sensing and Ground Network Measurements for the Production of site/time Specific Irradiance Data", *Sol. Energy* 60(3): 89-96.
- Pinker, R.T. and Laszlo, I. (1992). "Modeling Surface Solar Irradiance for Satellite Applications on a Global Scale", *J. Appl. Meteor.*, 31(2):194-211.
- Polo, J.; Zarzalejo, L.F. and Ramírez, L. (2008). "Solar radiation derived from satellite images", *Modeling Solar Radiation at the Earth's Surface*, Springer-Verlag Berlin Heidelberg, pp. 449-462.
- Pons, X. & Ninyerola, M. (2008). "Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data", *Int. J. Climatol.*, 28 (13): 1821-1834. <http://dx.doi.org/10.1002/joc.1676>.
- Posselt, R.; Mueller, R.; Stöckli, R. & Trentmann, J. (2011). "Spatial and Temporal Homogeneity of Solar Surface Irradiance across Satellite Generations", *Remote Sens.*, 3(5): 1029-1046. <http://dx.doi.org/10.3390/rs3051029>.
- Rigollier, C.; Lefèvre, M. & Wald, L. (2004). "The Method Heliosat-2 for Deriving Shortwave Solar Radiation from Satellite Images", *Solar Energy*, 77(2): 159-169.

- Reed, R. D. & Marks, R.J. (1998). *Neural smithing: supervised learning in feedforward artificial neural networks*, Mit Press.
- SoDA-is (2011). “Main Properties of the HelioClim Databases”, France. Recuperado de <http://www.soda-pro.com/help/helioclim/introduction>.
- (2013b). “HelioClim-3 Uncertainty”, France. https://web.archive.org/web/20130511070410/http://www.soda-is.com/eng/helioclim/helioclim3_uncertainty_eng.html.
- Tarpley, J.D. (1979). “Estimating Incident Solar Radiation at the Surface from Geostationary Satellite Data”, *J. Appl. Meteor.*, 18: 1172-1181.
- Toole, G.L.; Fair, M.; Berscheid, A. & Bent, R. (2010). “Electric power transmission network design for wind generation in the western United States: Algorithms, methodology, and analysis”. Transmission and distribution conference and exposition. *IEEE PES*.19-22. <http://dx.doi.org/10.1109/TDC.2010.5484549>.
- Tymvios, F.S., Michaelides, S.C. & Skouteli, C.S. (2008). “Estimation of surface solar radiation with artificial neural networks”, *Modeling Solar Radiation at the Earth's Surface*, Springer Berlin Heidelberg, pp. 221-256.
- Voyant, C.; Muselli, M.; Paoli, C. & Nivet, M.L. (2011). “Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation”, *Energy*, 36(1): 348-359.
- Voyant, C.; Randimivololona, P.; Nivet, M.L.; Paoli, C. and Muselli, M. (2014). “Twenty four hours ahead global irradiation forecasting using multi-layer perceptron”, *Meteorological Applications*, 21(3). <http://dx.doi.org/10.1002/met.1387>.
- Wang, F.; Mi, Z.; Su, S. & Zhao, H. (2012). “Short-Term Solar Irradiance Forecasting Model Based on Artificial Neural Network Using Statistical Feature Parameters”, *Energies*, 5(5): 1355-1370. <http://dx.doi.org/10.3390/en5051355>.
- Wielicki, B.A.; Barkstrom, B.R.; Harrison, E.F.; Lee, R. B.; Smith, G.L. & Cooper, J.E. (1996). “Clouds and the Earth’s Radiant Energy System (CERES): An Earth Observing System Experiment”, *Bull. Amer. Meteor. Soc.*, 77(5): 853-868.
- Yadav, A.K. and Chandel, S.S. (2012). “Artificial Neural Network based Prediction of Solar Radiation for Indian Stations”, *International Journal of Computer Applications*, 50(9): 1-4.
- Yadav, A.K. and Chandel, S.S. (2014). “Solar radiation prediction using Artificial Neural Networks Techniques: A review”, *Renewable and Sustainable Energy Reviews*, 33: 772-781.
- Zelenka, A.; Perez, R.; Seals, R. & Renné, D. (1999). “Effective Accuracy of Satellite-Derived Hourly Irradiances”, *Theoretical and Applied Climatology*, 62(3): 199-207.