

Trabajo Fin de Grado

Grado en Ingeniería Aeroespacial

Estudio comparativo de la estimación solar en un área determinada utilizando los resultados de una red neuronal artificial no supervisada

Autor: Jesús Lérída Viso

Tutora: Amparo Núñez Reyes

Dpto. Ingeniería de Sistemas y Automática
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2020



Trabajo Fin de Grado
Grado en Ingeniería Aeroespacial

Estudio comparativo de la estimación solar en un área determinada utilizando los resultados de una red neuronal artificial no supervisada

Autor:

Jesús Lériða Viso

Tutora:

Amparo Núñez Reyes

Profesora titular

Dpto. Ingeniería de Sistemas y Automática

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2020

Trabajo Fin de Grado. Estudio comparativo de la estimación solar en un área determinada utilizando los resultados de una red neuronal artificial no supervisada

Autor: Jesús Lérica Viso

Tutora: Amparo Núñez Reyes

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

El Secretario del Tribunal, Sevilla, 2020

A mi familia

Agradecimientos

En primer lugar, quiero dar las gracias a mi tutora Amparo Núñez Reyes por su entrega y su ayuda en este trabajo en una situación complicada.

A todos mis amigos y compañeros que me han acompañado en este viaje de la vida. A Elvira, Fernando, María y José María. Gracias por vuestro apoyo en los momentos más difíciles.

A mis hermanos Araceli y Paco, piezas fundamentales en mi vida.

A mi padre.

A mi ángel protector.

Jesús Lérida Viso
Alumno del Grado en Ingeniería Aeroespacial
Sevilla, 2020

Resumen

Desde el Protocolo de Kyoto en el año 1997 hasta nuestros días, se han realizado diferentes estrategias, tanto legales como obligatorias, para hacer frente al cambio climático y minimizar sus impactos en la naturaleza y en la sociedad. Las energías renovables y diferentes tecnologías de mayor eficiencia energética nacen para hacer frente a este fenómeno.

Este trabajo se enmarca en el universo de la energía solar. En grandes campos o plantas solares, es fundamental conocer los valores de la radiación solar en el área completa de estudio para así optimizar el futuro control distribuido de la planta solar.

Por ello, en el presente trabajo se han llevado a cabo investigaciones para el análisis y estimación de la radiación solar utilizando el método de interpolación IDW, de sus siglas en inglés *Inverse Distance Weighting*. Para este estudio, se ha simulado una red de sensores distribuidos en una hipotética planta solar ubicada en el entorno de las Tablas de Daimiel (Ciudad Real). Partiendo del escenario anterior y conocida la radiación solar en cada sensor a través de los datos provenientes de satélites de la Unión Europea, dentro del programa Copernicus, obtenemos nuevos escenarios con sensores distribuidos. Estos escenarios se obtienen mediante una red neuronal artificial no supervisada conocida como mapa autoorganizativo SOM, de sus siglas en inglés *Self-Organizing Map*. El procedimiento que seguirá la red es, de manera autoorganizada, agrupar o mapear rasgos comunes e incorporarlos a su estructura interna de conexiones.

La comparativa realizada de la estimación de la radiación solar entre los escenarios propuestos, permite comprobar que el uso de la red neuronal es una buena alternativa aplicable en la interpolación espacial de variables en áreas extensas.

Abstract

From the Kyoto Protocol in 1997 to the present day, different strategies, both legal and mandatory, have been carried out to face climate change and minimize its impacts on nature and society. Renewable energies and different technologies with greater energy efficiency are born to face this phenomenon.

This work is framed in the universe of solar energy. In large fields or solar plants, it is essential to know the values of solar radiation in the entire study area to optimize future distributed control of the solar plant.

Therefore, in the present work, investigations have been carried out for the analysis, optimization and estimation of solar radiation using the IDW (Inverse Distance Weighting) method. For this study, a network of sensors distributed in a hypothetical solar plant located in Tablas de Daimiel (Ciudad Real) has been simulated. Starting from the previous scenario and known the solar radiation in each sensor through the data from satellites of the European Union, within the Copernicus program, we obtain new scenarios with distributed sensors. These scenarios are the outputs of an unsupervised competitive artificial neural network known as a Self-Organizing Map (SOM).

The objective of the work is to compare the estimation of solar radiation between the proposed scenarios and to check if the use of the neuronal network is optimal.

Índice

Agradecimientos	ix
Resumen	xi
Abstract	xiii
Índice	xv
Índice de Tablas	xvii
Índice de Figuras	xix
1 Introducción	1
2 Wireless Sensor Network (WSN)	5
2.1. <i>Definición y características de WSN</i>	5
3 Inverse Distance Weighting (IDW)	9
4 Self-Organizing map (SOM)	13
4.1 <i>Fundamentos biológicos</i>	13
4.2 <i>Arquitectura de la SOM</i>	14
4.3 <i>Algoritmo de la SOM</i>	15
4.4 <i>Medidas de calidad del mapa</i>	17
5 Origen y Análisis de datos.Simulación de escenarios	19
5.1 <i>Introducción</i>	19
5.2 <i>Ubicación geográfica del área a estudiar</i>	19
5.3 <i>Escenario original</i>	21
5.4 <i>Origen de los datos</i>	22
5.5 <i>Escenarios calculados mediante las salidas de la SOM</i>	26
5.5.1 <i>Neural Network Toolbox</i>	26
5.5.2 <i>Escenario SOM1</i>	30
5.5.3 <i>Escenario SOM2</i>	32
5.5.4 <i>Escenario SOM3</i>	33
5.5.5 <i>Escenario SOM4</i>	34
6 Estimación de la radiación solar mediante IDW	37
6.1 <i>Estimación de la radiación solar mediante IDW</i>	37
6.1.1 <i>Errores de la estimación para todo el mes de junio de 2015</i>	42
6.1.2 <i>Errores de la estimación para el 12 de junio de 2015: día nublado</i>	43

6.1.3	Errores de la estimación para el 27 junio de 2015: día soleado	45
6.1.4	Errores de la estimación para un instante determinado: 10 de junio a las 7.30 a.m.	46
6.1.5	Influencia del error de cuantización de las redes SOM en la estimación IDW	49
7	CONCLUSIONES	55
8	Anexo a: Código realizado en Matlab	57
8.1	<i>Generación puntos distribuidos</i>	57
8.2	<i>Transformación de los puntos a coordenadas</i>	57
8.3	<i>Normalización datos de entrada a la SOM</i>	59
8.4	<i>Diseño de red neuronal SOM</i>	59
8.5	<i>Transformación pesos de la salida SOM en coordenadas</i>	60
8.6	<i>Error cuantización SOM</i>	61
8.7	<i>División área de estudio en puntos distribuidos regularmente</i>	61
8.8	<i>Función interpolación IDW</i>	62
8.9	<i>Función haversine</i>	65
8.10	<i>Errores MAE y RMSE</i>	66
8.11	<i>Visualizar estimación de la radiación solar</i>	67
	Referencias	69

ÍNDICE DE TABLAS

Tabla 1. Parámetros SOM1.	30
Tabla 2. Parámetros SOM2.	32
Tabla 3. Parámetros SOM3.	33
Tabla 4. Parámetros SOM4.	34
Tabla 5. Errores de la estimación de la radiación solar para todo el mes de junio.	43
Tabla 6. Errores de la estimación de la radiación solar para día 12 de junio.	45
Tabla 7. Errores de la estimación de la radiación solar para día 27 de junio.	46
Tabla 8. Errores de la estimación de la radiación solar para las 7.30 a.m en el día 10 de junio.	47
Tabla 9. Errores de cuantización de las distintas SOM en función del entrenamiento.	49
Tabla 10. Errores de la estimación con los nuevos escenarios.	52
Tabla 11. Diferencia de Errores de Cuantización.	54

ÍNDICE DE FIGURAS

Figura 1. Evolución de las emisiones de los gases CO2 en España.	1
Figura 2. Esquema de una WSN en ámbito agrario.	5
Figura 3. Subsistemas de un nodo de una WSN.	6
Figura 4. Distribución de capas de protocolo en una WSN.	7
Figura 5. Esquema del método IDW.	9
Figura 6. Método del número de vecinos fijo en IDW.	10
Figura 7. Método del radio de búsqueda fijo en IDW.	11
Figura 8. Esquema general de una neurona.	13
Figura 9. Arquitectura de una SOM.	14
Figura 10. Comparativa neuronas biológicas y artificiales.	15
Figura 11. Esquema de la fase de entrenamiento de una SOM.	16
Figura 12. Nubosidad media anual en la ubicación elegida.	20
Figura 13. Ubicación de la hipotética planta solar.	20
Figura 14. Escenario original.	21
Figura 15. Esquema de las distintas radiaciones.	23
Figura 16. Esquema del método Heliosat-4.	23
Figura 17. Interfaz de usuario del portal CAMS Service.	24
Figura 18. Ejemplo de archivo de descarga '.cv' visto en una hoja de Excel.	25
Figura 19. Ejemplo de topología hexatop, gridtop y randtop en una SOM.	27
Figura 20. Ventana de visualización de entrenamiento de una red neuronal en Matlab.	28
Figura 21. Ejemplos de visualizaciones de gráficas de una red neuronal SOM en Matlab.	29
Figura 22. Ventana de visualización de entrenamiento de SOM1.	30
Figura 23. Escenario SOM1.	31
Figura 24. Adaptación de la SOM1 a los datos de entrada.	32
Figura 25. Escenario SOM2.	33
Figura 26. Escenario SOM3.	34
Figura 27. Escenario SOM4.	35
Figura 28. Puntos a estimar en el área de estudio.	37
Figura 29. Puntos geográficos. Escenario original (ORG) y área a estimar (GRID).	38
Figura 30. Puntos geográficos. Escenarios SOM1, SOM2, SOM3 Y SOM4 y área a estimar (GRID).	39

Figura 31. Puntos de validación para la estimación de la radiación solar (VAL).	41
Figura 32. Radiación real de todo el mes de junio 2015 para el punto de validación 1.	42
Figura 33. Radiación solar real del día 12 de junio en el punto de validación 1.	43
Figura 34. Valores radiación solar 12 de junio. Valor real y valor estimado (SOM1-IDW) en punto validación 1.	44
Figura 35. Valores radiación solar 12 de junio. Valor real y valor estimado (SOM3-IDW) en punto validación 1.	44
Figura 36. Valores radiación solar 27 de junio. Valor real y valor estimado (SOM3-IDW) en punto validación 1.	45
Figura 37. Radiación solar real del día 10 de junio en el punto de validación 1. Punto rojo (TIME=31, 7.30 a.m).	46
Figura 38. Estimación radiación solar en el área de estudio con el escenario original	48
Figura 39. Estimación de la radiación solar en el área de estudio con los escenarios SOM1, SOM2, SOM3 y SOM4.	48
Figura 40. Escenario nuevo SOM1.	50
Figura 41. Escenario nuevo SOM2.	50
Figura 42. Escenario nuevo SOM3.	51
Figura 43. Escenario nuevo SOM4.	51
Figura 44. Estimaciones con escenario original, SOM1 con EC: 0.1186 y SOM1: con EC:0.1152 para instante específico.	53

1 INTRODUCCIÓN

EL progresivo desarrollo de las actividades tecnológicas e industriales y el aumento de sus efectos perjudiciales han supuesto un incremento de tensiones y conflictos con la necesidad de equilibrar los aspectos vinculados a la protección del medio ambiente y las cuestiones propias del desarrollo económico.

Por este motivo nace el Convenio Marco sobre el Cambio Climático de 1992, que concluyó en el importante Protocolo de Kyoto de 1997. Con este instrumento, los países pretenden hacer frente a los efectos perjudiciales sobre los ecosistemas producidos por las emisiones de los gases del efecto invernadero; los principales causantes del calentamiento global.

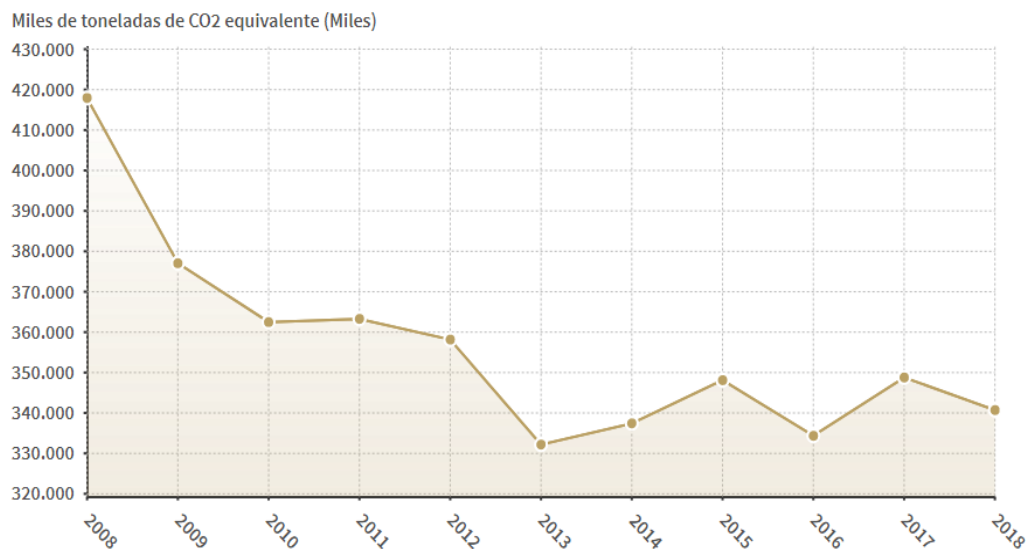


Figura 1. Evolución de las emisiones de los gases CO_2 en España [1].

De todas las energías, las renovables son aquellas que se producen de una forma continua y que son inagotables.

La energía solar es la principal, la más grande y la más importante fuente de energía renovable. Existen distintas investigaciones que utilizan energía solar térmica, fotovoltaica o termoeléctrica, donde el conocimiento de los datos de la radiación solar es importante para el desarrollo, diseño y optimización de numerosas aplicaciones, como la agrometeorología [2].

Sin embargo, las mediciones de radiación solar en un área específica no siempre se encuentran disponibles debido al alto coste y a ciertos requerimientos tecnológicos [3].

Dentro de este contexto, se encuentra este trabajo. Se va a estudiar la estimación de la radiación solar de una determinada zona a través de la interpolación IDW, obteniendo los valores estimados en las zonas donde no se ubiquen sensores.

El análisis comparativo de la estimación mediante este método de interpolación se va a realizar sobre distintos escenarios hipotéticos. El escenario original en el que se realiza el estudio es una planta solar ubicada en las proximidades del parque natural de Tablas de Daimiel, cerca de la ciudad del mismo nombre. Esta planta no requiere que sea específicamente de ningún tipo y dispone de una red de sensores distribuidos aleatoriamente por el terreno comprendido por el área de estudio. Esta planta presenta unos colectores solares que pueden ser controlados mediante un controlador avanzado de procesos. Es por ello importante conocer la radiación solar en toda el área donde se encuentran los paneles solares.

Con el uso de la interpolación IDW y conociendo los datos de la radiación real en cada uno de los nodos, es posible obtener una estimación en el conjunto del terreno. Para el trabajo, se ha simulado una red compuesta por un número de 36 sensores distribuidos irregularmente y de los que conocemos los datos, que se han obtenido de satélites disponibles de una forma libre.

Una vez se ha planteado el escenario original, el objetivo es comparar dicha estimación en otros escenarios posibles.

Para el desarrollo de nuevos escenarios se ha usado la salida de una red neuronal artificial no supervisada de tipo SOM. Estas redes, a partir de un conjunto de datos de entrada (latitud, longitud y radiación solar), calcula un número de salidas que puede ser configurable. Estas redes no simulan nada al final, sino que clasifican y obtienen patrones comunes de los datos de entrada. Debido a que los mapas autoorganizativos son herramientas visuales y que dos de los atributos de entrada son coordenadas del espacio geográfico, es posible obtener una nueva relación de coordenadas de los sensores. Con los nuevos escenarios (ubicación de sensores) se ha obtenido también la estimación de la radiación solar en la misma área.

Una vez se ha obtenido las estimaciones de los distintos escenarios, es posible la comparación entre los supuestos, viendo si el uso de la red neuronal de tipo SOM mejora el rendimiento del método IDW para estimar la radiación solar del área completa.

El principal objetivo consiste en lograr aplicar las técnicas mediante el programa informático Matlab, así como estudiar los resultados para comprobar la viabilidad que puedan tener en entornos reales.

Respecto a la organización de la memoria, se ha dividido el trabajo en 6 capítulos principales:

- **Wireless Sensor Network.** Define el concepto de las redes de sensores inalámbricos, así como sus principales usos y sus principales características.
- **Inverse Distance Weighting.** Este capítulo recoge una explicación de este método de interpolación, su funcionamiento, las ecuaciones que determinan el método y las distintas configuraciones con las que puede trabajar.
- **Self-Organizing Map.** Dentro de este capítulo se explica en detalle qué son las redes neuronales no supervisadas SOM, también llamadas mapas autoorganizativos o redes de Kohonen. Se hace una breve mención a sus fundamentos biológicos, así como se explica en detalle su arquitectura y el algoritmo con el que se llevan a cabo.
- **Origen y análisis de datos. Simulación de escenarios.** Este capítulo está compuesto por diferentes secciones relacionadas entre sí. En primer lugar, se explica como se ha elegido el área de estudio, así como la ubicación de los sensores del escenario original dentro de ella. Una vez obtenido los puntos de los sensores, se explica cómo se ha llevado a cabo la recopilación de los datos de la radiación solar. El siguiente paso consiste en alcanzar nuevos escenarios como salidas de las redes neuronales de tipo SOM. Se explica el funcionamiento de dichas redes en el entorno de Matlab, así como el método de entrenamiento para la obtención de distintos escenarios.
- **Estimación de la radiación solar mediante IDW.** Dentro de este capítulo se calcula la estimación de la radiación solar en el área de estudio con los distintos escenarios. Se hace una comparación de las estimaciones mediante la obtención de errores MAE y RMSE y se explica la influencia del error de cuantización de la salida de una SOM en el error de la estimación.

- **Conclusiones.** En este último capítulo se explican las conclusiones del trabajo y se proponen nuevas líneas de investigación para desarrollar en el futuro.

Por último, se añade un Anexo donde se plasma el código desarrollado en Matlab para obtención de los puntos clave del trabajo de investigación propuesto.

2 WIRELESS SENSOR NETWORK (WSN)

El concepto de redes de sensores inalámbricas o WSN, de sus siglas en inglés *Wireless Sensor Network*, es una tecnología actualmente en auge debido a su posible extensión a entornos naturales o aplicaciones en defensa o medicina. Su principal finalidad es la monitorización y control de fenómenos que ocurren en el mundo físico. En el futuro, se estima que su uso seguirá dando conclusiones fiables para la utilización de este mecanismo.

2.1. Definición y características de WSN

En [4] se recoge como definición de este mecanismo una red de características autoconfigurables integrada por un pequeño número de nodos sensores también llamados *moten* distribuidos espacialmente y comunicados entre sí, para los que se usan señales de radio, con la finalidad de monitorizar y entender el mundo físico para controlar diversas condiciones en distintos puntos, entre ellas, variables como pueden ser la temperatura, el sonido, presión o los contaminantes.

Las redes de sensores inalámbricas muestran una variada compilación de aplicaciones como pueden ser [5]:

- **Aplicaciones militares:** uso en monitorización de fuerzas y equipos enemigos, vigilancia en el campo de batalla, reconocimiento del terreno...
- **Aplicaciones medioambientales:** uso en agricultura de precisión, en seguimiento de animales, detección de inundaciones o sistemas sísmicos, en riego...
- **Aplicaciones médicas:** uso en control ambiental, control de inventario, administración de medicamentos, seguimiento de horarios médicos en los hospitales...



Figura 2. Esquema de una WSN en ámbito agrario [5].

Una WSN típicamente consiste en un gran número de nodos sensores inalámbricos que forman una red personalizada. Normalmente, se despliegan sobre áreas grandes para la aplicación en medición de variables, en su mayoría para la recopilación de datos.

Las características más comunes y relevantes de una WSN son [6]:

- Sensores de reducidas dimensiones y bajo coste.
- Nodos con reducido consumo de energía y memoria.
- Soportar condiciones ambientales extremas.
- Soportar fallos de nodos de la red.
- Localización de los nodos.
- Fallos de comunicación entre nodos.
- Estructura de red distribuida.
- Capacidad de autoconfiguración de los nodos.

Los nodos de los sensores pueden apreciarse como pequeños y básicos ordenadores que constan de cuatro subsistemas [6], pudiéndose observar en la figura 3:

- Una unidad central con consumo de energía y memoria limitada.
- Sensor.
- Un dispositivo de comunicador (transmisor-receptor).
- Una fuente de energía, generalmente batería.

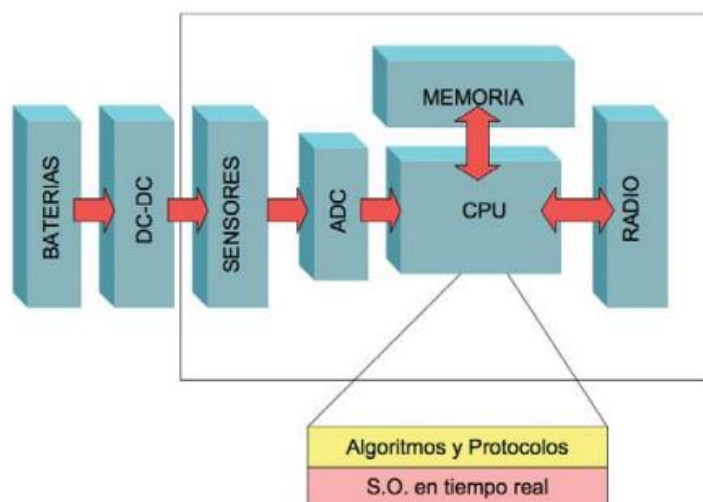


Figura 3. Subsistemas de un nodo de una WSN [6].

Dichos nodos o sensores son dispositivos muy pequeños capaces de capturar un tipo de información física. En la actualidad, se están realizando investigaciones para conseguir sensores minúsculos y de un coste de producción mínimo. Estos sensores son capaces de capturar información física sobre un espacio inmediato. La distribución de los sensores sobre un área específica crea una red inalámbrica digital. Estos nodos deben ser capaces de transformar la información sobre la variable en estudio a un cambio eléctrico que se pueda alimentar en un circuito que la utilice directamente o con mediadores, como pueden ser amplificadores. Dichos conjuntos de nodos o sensores pueden ser clasificados según [7] en dos tipos de redes: las estructuradas o distribuidas o las desestructuradas.

- **Redes estructuradas.** Se usa un mínimo número de nodos y tienen un menor costo de mantenimiento. Aún así, la localización de este tipo de redes puede ser más costosa que las desestructuradas puesto que la topología es más exacta.
- **Redes desestructuradas.** Este tipo de red distribuida contiene un número de nodos mayor que el anterior y su localización es aleatoria, por lo que su mantenimiento es más costoso. Su uso es muy utilizado en funciones medioambientales.

En la actualidad, una de las desventajas de las redes WSN es el consumo de energía debido a que el método más usual de suministro es el uso de batería, mencionado anteriormente. En este caso, la implantación de una instalación de planta fotovoltaica que use energía renovable como la solar es una manera de suavizar esta limitación.

Con respecto al consumo final de energía del sistema, las capas de protocolo tienen una gran influencia. Existen diversos protocolos [8]:

- **Protocolo de comunicación.** Se encarga de generar un código básico de envío-recepción de la información. Crea el lenguaje por el cual todos los dispositivos conectados a la red se van a comunicar.
- **Protocolo Routing.** Se encarga de generar las vías para comunicar los dispositivos que se encuentran distribuidos en la red.
- **Protocolo de transporte.** Se encarga de facilitar el tránsito de información de un nodo a otro, entre sensores, y entre éstos y la estación base.

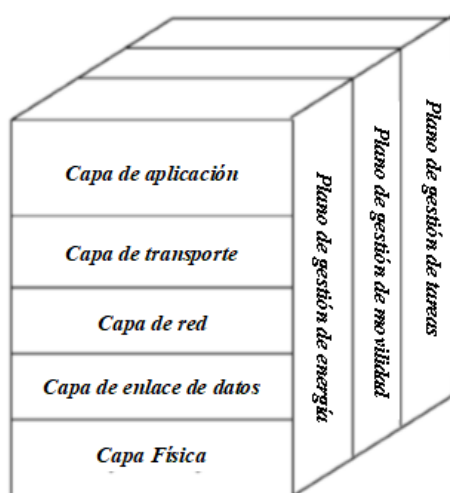


Figura 4. Distribución de capas de protocolo en una WSN. Imagen modificada de [7].

Este trabajo utiliza este mecanismo de redes para simular una red de sensores inalámbricos de tipo desestructurado en distintos escenarios sobre un área específica que se ha seleccionado. Estas redes simuladas servirán para predecir la estimación solar sobre el área mencionada a través del método de interpolación IDW. El número de nodos fijados en cada escenario ha sido establecido de antemano.

3 INVERSE DISTANCE WEIGHTING (IDW)

El método de la distancia inversa ponderada IDW, de sus siglas en inglés *Inverse Distance Weighting*, es un tipo de método determinista que se usa para la interpolación multivariante con un conjunto de datos dispersos conocidos. Los valores asignados a puntos desconocidos se calculan con un promedio moderado de los valores disponibles en los puntos conocidos. El valor de cada uno de los datos se pondera de acuerdo con el inverso de su distancia al punto o nodo de la interpolación [9].

Si se compara este método con el kriging ordinario, el método IDW es un poco más simple de programar y no requiere un proceso anticipado de selección con un modelo de variograma [10].

Este tipo de método asume que los valores más cercanos están más relacionados entre sí que los valores más lejanos. En la figura 5, se muestra el punto azul como el punto que se va a interpolar, rodeado de los puntos con valores conocidos (puntos amarillos) señalados con las distancias correspondientes entre cada punto conocido y el punto a interpolar.

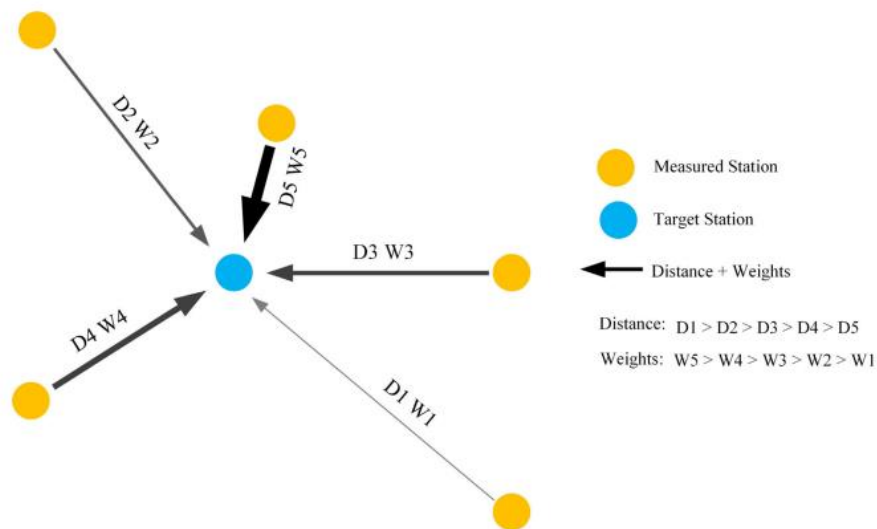


Figura 5. Esquema del método IDW [9].

En definitiva, este tipo de interpolación asume que a un punto más cercano del punto interpolado se le asigna un peso más grande (más implicación en el punto interpolado) y disminuye a medida que aumenta la distancia.

La ecuación general del método IDW [10] es la siguiente:

$$Z^*(x_0) = \frac{\sum_{i=1}^n Z(x_i) \times (W_{i0})}{\sum_{i=1}^n W_{i0}}$$

$$W_{i0} = \frac{1}{d_{i0}^k}$$

Donde:

$Z^*(x_0)$ = Valores estimados en el punto x_0 .

$Z(x_i)$ = Valores conocidos en el punto x_i .

W_{i0} = Función inversa de la distancia.

d_{i0} = Distancia entre el punto conocido x_i y el punto desconocido x_0 .

k = coeficiente de potencia.

n = Número de puntos usados en la interpolación.

El método IDW es un tipo de interpolación espacial muy flexible [11]. Puede configurar su interpolación de diferentes maneras:

- **Número de vecinos fijo**

El número de puntos que se usan para la interpolación puede fijarse de antemano. En este caso, se considera un número de vecinos fijo para interpolar a un punto desconocido. En la figura 6, se usa un número de vecinos igual a 3, siendo estos los más cercanos al punto estimado (punto morado). Véase como el punto con valor 9, el punto más lejano, no entra en consideración.

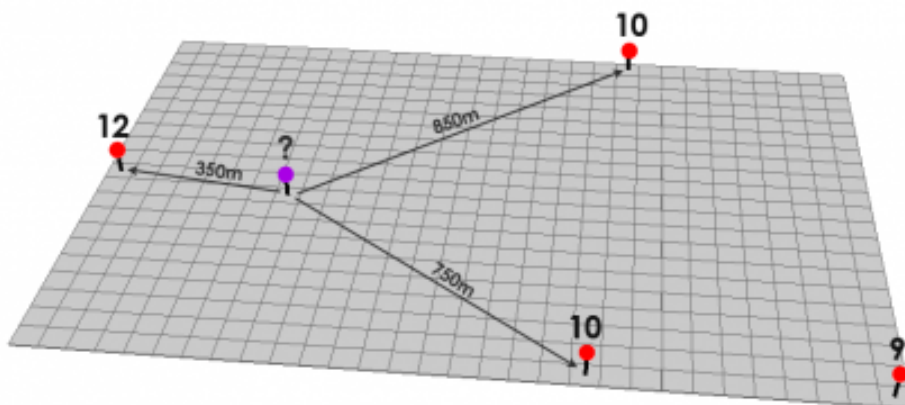


Figura 6. Método del número de vecinos fijo en IDW [11].

- **Radio de búsqueda fijo**

En esta ocasión, no se usa un número de puntos determinados para interpolar, sino que se fija un radio de búsqueda en torno al punto que se desea interpolar y en la interpolación solo participarán aquellos puntos o nodos que se ubiquen dentro del radio de búsqueda determinado. Puede verse en la figura 7, como, esta vez, solo entra en acción un punto para realizar la interpolación del punto morado.

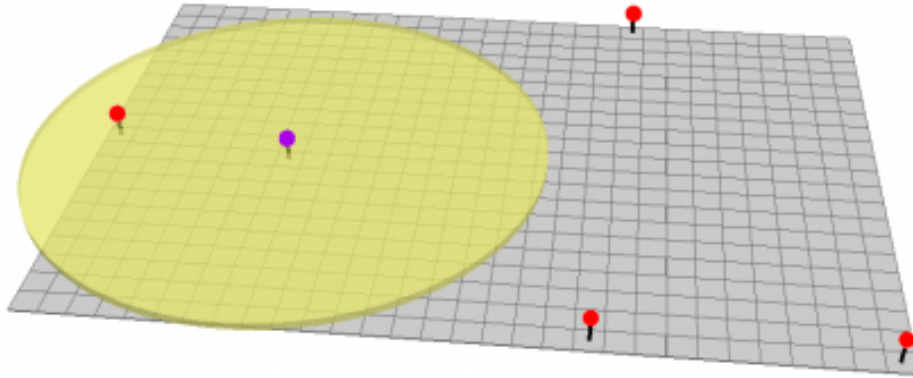


Figura 7. Método del radio de búsqueda fijo en IDW [11].

De acuerdo con [12], el valor más correcto para la potencia de la distancia inversa k es 2. Una potencia de 2 aumenta la influencia global que tiene sobre los valores conocidos, donde las diferencias significativas de los datos en diferentes puntos se estiman mejor que utilizando una potencia de 1, donde no se promedian tanto. Por ejemplo, un valor pequeño en el denominador significa más distancia, por lo que tendrá menos efecto sobre el valor interpolado.

4 SELF-ORGANIZING MAP (SOM)

En 1982, Teuvon Kohonen presentó un modelo de red denominado mapas autoorganizativos o SOM, de sus siglas en inglés *Self-Organizing Maps*. Este tipo de mapa estaba basado en ciertas evidencias descubiertas a nivel cerebral. Este tipo de red posee un sistema de aprendizaje no supervisado competitivo [13].

No existe ningún elemento externo que indique si la red neuronal está operando correctamente o no porque no existe un objetivo final hacia donde la salida deba tender.

La red autoorganizada debe descubrir rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, e introducirlos e incorporarlos en su estructura interna de conexiones. Por tanto, las neuronas deben organizarse en función de los datos procedentes del exterior, que son llamados estímulos.

En el aprendizaje competitivo, las neuronas compiten unas con otras con el fin de llevar a cabo una tarea dada. Cuando se presenta a la red un patrón o estímulo de entrada, solo una de las neuronas de salida o un grupo de vecinas se activa.

Por tanto, las neuronas compiten por activarse, quedando, finalmente, una como neurona vencedora y las demás como anuladas, que se adaptan a los valores de la neurona vencedora.

El objetivo de este aprendizaje no supervisado es categorizar los datos de entrada. Se clasifican valores similares en una misma categoría, por lo que debe activarse la misma neurona de salida.

En [14], Kohonen afirma: “Solo quería un algoritmo que mapeara efectivamente patrones similares (vectores de patrones cercanos entre sí en el espacio de la señal de entrada) en contiguas ubicaciones en el espacio de salida.”

4.1 Fundamentos biológicos

Como se ha dicho anteriormente, Kohonen fundamenta este tipo de redes en neuronas biológicas. Se observó que en el cerebro de los animales aparecen zonas donde las neuronas detectoras de rasgos se encuentran topológicamente ordenadas, de forma que las informaciones captadas del entorno a través de los órganos sensoriales se representan internamente en forma de mapas bidimensionales.

Es posible que parte de la organización neuronal esté basada en el aprendizaje, si bien en gran medida está predeterminada por la genética. Esto sugiere, por tanto, que el propio cerebro es capaz de formar o simular mapas topológicos a partir de informaciones externas.

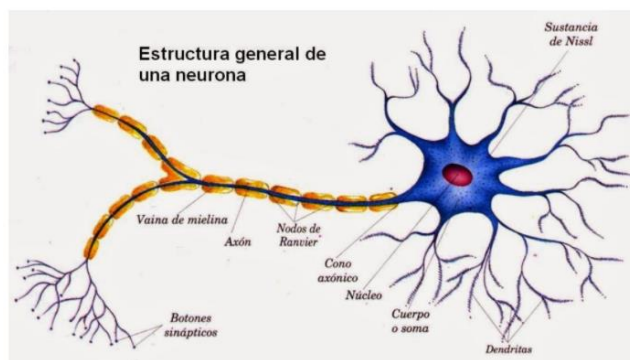


Figura 8. Esquema general de una neurona [15].

A su vez, también se ha observado que la influencia de una neurona “ i ” sobre una neurona “ j ” viene determinada por la distancia entre ellas, siendo inversamente proporcional la cantidad de influencia con respecto a la distancia que haya entre ellas. Así, una neurona “ i ” que ejerza una gran influencia sobre una neurona “ j ” significará que la distancia entre ellas es mínima.

El modelo de red autoorganizado presentado por Kohonen en 1982, pretende reproducir de forma simplificada la capacidad del cerebro de formar mapas topológicos a partir de señales recibidas.

4.2 Arquitectura de la SOM

Una red neuronal SOM está compuesta por dos capas de neuronas. La capa de entrada está formada por N neuronas, una por cada variable de entrada, y ésta se encarga de transmitir a la capa de salida la información procedente de los datos de entrada. La capa de salida está formada por M neuronas, y es la encargada de procesar la información y formar el mapa de correlaciones de los datos de entrada. En general, las neuronas de la capa de salida se organizan en forma de mapa bidimensional como se muestra en la figura 9.

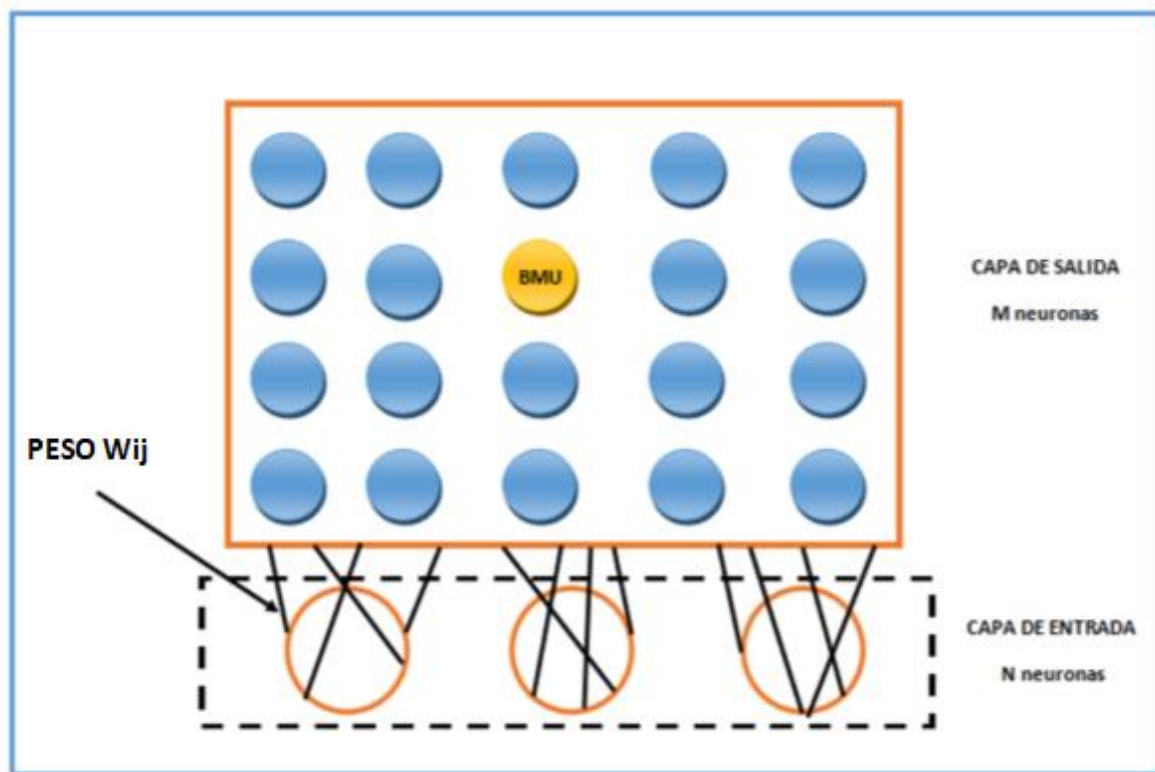


Figura 9. Arquitectura de una SOM.

Las conexiones entre las dos capas que forman la red son siempre desde la capa de entrada hacia la capa de salida. Cada neurona de entrada “ i ” está conectada con cada una de las neuronas de la salida “ j ” mediante una correlación o peso, según [13] [14] [16], w_{ij} . De esta forma, todas las neuronas de la salida tienen un vector de pesos asociados W_j llamado vector de referencia, o *codebook*, que corresponde al vector promedio de la categoría representada por la neurona de salida “ j ”. En definitiva, la SOM define una proyección desde un espacio de datos de alta dimensión a un mapa bidimensional de neuronas. En la figura 10, se puede observar la comparación entre neuronas biológicas y neuronas artificiales.

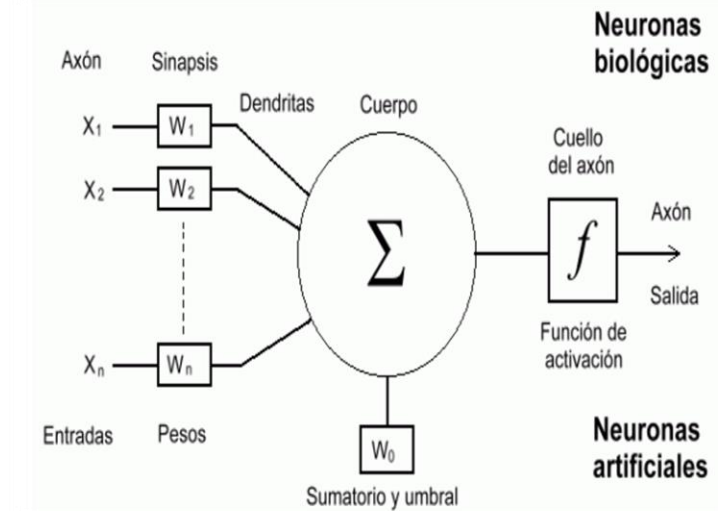


Figura 10. Comparativa neuronas biológicas y artificiales [17].

Entre todas las neuronas de salida que haya en la red neuronal existe unas conexiones de excitación, es decir, aunque no estén conectadas, todas ellas van a tener una cierta influencia sobre las demás. Esto se produce por el proceso competitivo entre ellas y por la aplicación de una función denominada vecindad, que produce el desarrollo de la tipología o estructura del mapa. Las topologías más frecuentes de la visualización de las neuronas de salida son la rectangular y la hexagonal.

La topología y el número de neuronas permanece fijo desde el principio. El número de neuronas determina la suavidad de la proyección, lo cual influye en el ajuste y la capacidad de generalización de la SOM.

Durante la fase de entrenamiento (figura 11), la SOM simula una red elástica que se pliega dentro de la nube de datos de la entrada. Los vectores de referencia del vector de pesos asociados se acercan a las áreas donde la densidad de datos es alta.

4.3 Algoritmo de la SOM

El proceso de aprendizaje de una SOM es el siguiente [15]:

- **Paso 1.** Un vector x es seleccionado al azar del conjunto de vectores de entrada y calcula la distancia, que es la similitud, a los vectores de referencia, usando, por ejemplo, la distancia euclídea.

$$\|x - W_j\| = \min \|x - W_j\|$$

Donde:

x = vector de entrada

W_j = vector de pesos de neurona "j"

- **Paso 2.** Una vez se ha encontrado el vector más próximo o BMU (véase figura 9), de sus siglas en inglés *Best Machine Unit*, el resto de los vectores del codebook se actualiza. El BMU y sus vecinos se mueven cerca del vector x en el espacio de los datos.

Mientras se va produciendo el proceso de actualización y se asignan nuevos vectores al mapa, la tasa de aprendizaje $LR(k)$, que puede ser lineal, potencial o inversa, decrece gradualmente hacia cero. Junto con ella, decrece el radio de vecindad también, $\epsilon(k)$. Hay varios tipos de funciones vecinas, y la más común es la gaussiana.

La regla de actualización para el vector de referencia dado es la siguiente:

$$w_j(k+1) = w_j(k) + LR(k)\epsilon(k)(x(k) - w_j(k))$$

Donde k representa la iteración de aprendizaje en cada instante. Al comienzo del entrenamiento, debe especificarse el número de iteraciones. Los pasos 1 y 2 deben de irse repitiendo hasta que el entrenamiento finalice. Una vez terminado el entrenamiento, el mapa ha de ordenarse en sentido topológico: "n" vectores topológicamente próximos se aplican en "n" neuronas adyacentes o incluso en la misma neurona. Así, en el mapa se asignan patrones similares a las regiones vecinas, mientras que los patrones menos similares se asignan a las regiones más lejanas.

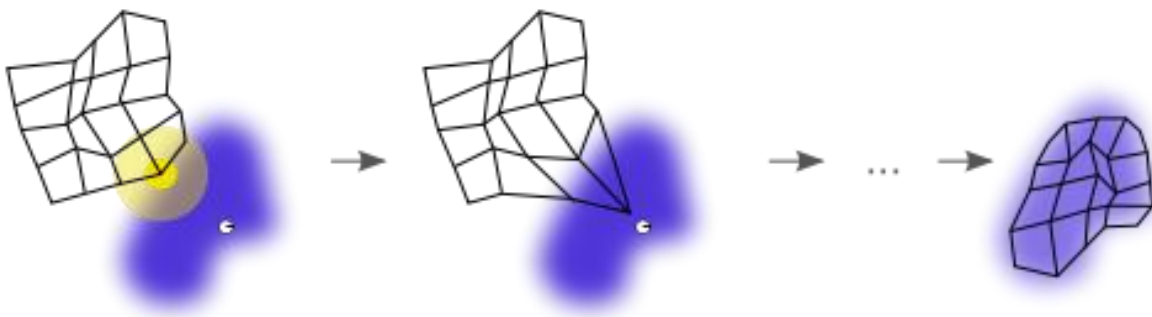


Figura 11. Esquema de la fase de entrenamiento de una SOM [18].

4.4 Medidas de calidad del mapa

Una vez que la SOM es entrenada, es importante saber si se ha adaptado adecuadamente a los datos de entrenamiento. Dos de las medidas son la precisión de la proyección y la preservación de la topología [16]:

- **Error de cuantización**

Este error mide la precisión de la proyección. Esta medida describe cómo se adaptan las neuronas de salida a los datos de entrada. En general, el número de datos de entrada es mayor que el número de neuronas de salida por lo que el error de precisión es siempre diferente de 0.

$$\varepsilon_q = \frac{1}{N} \sum_{i=1}^N \|x_i - W_j\|$$

- **Error topográfico**

Describe la manera en que la SOM preserva la topología del conjunto de datos. Esta medida considera la estructura del mapa. En un mapa donde las neuronas estén retorcidas de manera extraña, el error topográfico es grande incluso si el error de cuantización es pequeño.

$$\varepsilon_t = \frac{1}{N} \sum_{k=1}^N u(x_k)$$

Donde $u(x_k)$ es igual a 1 si el primer y segundo BMUs de x_k no están próximos el uno al otro. De otro modo, es igual a 0

En este trabajo se han usado las salidas de las redes neuronales SOM para la obtención de nuevos escenarios. Aprovechando que dos de los atributos de entrada van a ser coordenadas geográficas, se obtendrán a través de los pesos 1 y 2 las nuevas coordenadas.

5 ORIGEN Y ANÁLISIS DE DATOS.SIMULACIÓN DE ESCENARIOS

En este capítulo se explica como se ha realizado la extracción de datos de la radiación solar para la ubicación elegida así como el análisis previo de los mismos. El principal objetivo del trabajo es hacer una comparación entre distintos escenarios en una misma área del estudio de la radiación solar. Llamaremos escenarios a la ubicación de los sensores con sus medidas virtuales (datos de la radiación solar) en cada nodo o punto. Una vez se ha obtenido un escenario original con un número determinados de nodos o puntos, se simulan otros escenarios, usando el escenario original, que son las salidas de las redes neuronales SOM a través de la caja de herramientas de redes neuronales de Matlab.

5.1 Introducción

No es sencillo encontrar bases de datos de radiación solar provenientes de sensores o estaciones que tengan una resolución temporal de un minuto o inferior con el que trabajar, para detectar las variaciones de radiación solar que pueden producirse por la presencia de nubes en una ubicación exacta, por lo que se ha decidido trabajar con resoluciones mayores.

La mayoría de las bases de datos de esta variable usan datos de satélites como el Meteosat; con esto podemos comprobar la fiabilidad de estas fuentes como sustituto real para los datos recogidos mediante sensores (piranómetros), para medir la radiación solar. Si esto ocurriera, sería un sustituto mucho más económico que con respecto a los sensores, siendo menor su coste de instalación. El inconveniente principal que subyace en este tipo de datos es la inexactitud, lo que habría que comprobar si la diferencia que existiera pudiera producir grandes cambios significativos.

Antes de seleccionar los datos de la radiación solar, es necesario elegir una ubicación geográfica donde llevar a cabo el desarrollo del trabajo. Esta ubicación simulará una hipotética planta solar donde llevar a cabo nuestras predicciones, de forma que podamos extraer conclusiones a partir de los distintos escenarios que hemos simulado.

5.2 Ubicación geográfica del área a estudiar

En cuanto a la zona geográfica en la que desarrollar el trabajo, se ha decidido trabajar sobre el área comprendida entre la localidad de Tablas de Daimiel y el parque natural que lleva el mismo nombre. Esta área se encuentra en España dentro de la provincia de Ciudad Real.

Esta región se caracteriza por tener veranos muy cortos, muy calientes y secos. En cuanto a la nubosidad, el promedio de cielo cubierto con nubes varía considerablemente a lo largo del año, disminuyendo aproximadamente desde mediados del mes de junio hasta mediados del mes de septiembre.

En la figura 12, se puede observar el transcurso a lo largo de un año típico [18]. La zona de estudio es de alta radiación en España, por lo que el proyecto sería más viable. Cuanto mayor recurso, mayor producción.

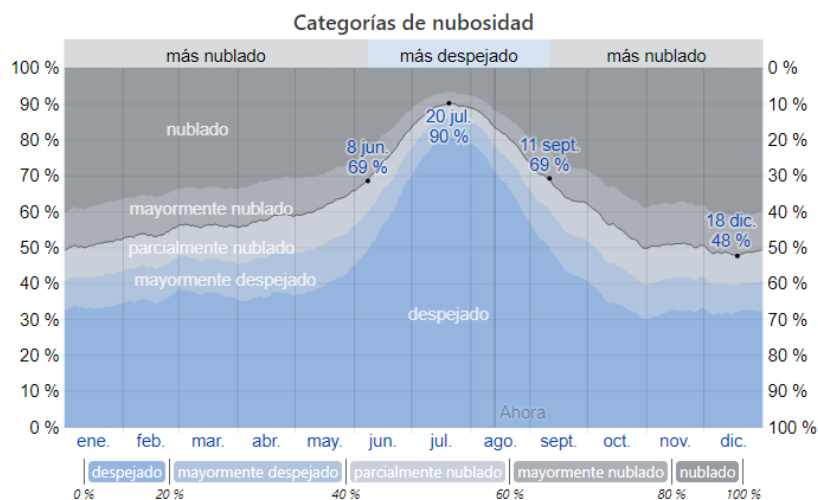


Figura 12. Nubosidad media anual en la ubicación elegida [19].

En cuanto al tamaño de área de estudio se ha tenido en cuenta otros tamaños actuales de plantas solares [20], como pueden ser Pavagada Solar Park, con 53 km², en la India, o Benban Solar Park, con 38 km², en Egipto.

Con la información anterior, se ha optado por ubicar una planta solar con un área de 36,5 km² en la zona entre la localidad de Tablas de Daimiel y el parque natural que lleva el mismo nombre. En la figura 13, se puede ver el área de estudio en la zona elegida enmarcada en el rectángulo naranja.



Figura 13. Ubicación de la hipotética planta solar [21].

5.3 Escenario original

Una vez hemos obtenido el área donde se va a ubicar la hipotética planta solar, el siguiente paso es la ubicación de sensores dentro del área mencionada. Estos sensores se van a repartir de forma aleatoria a través de un código de programación en Matlab (véase Anexo a, sección 8.1). Se ha elegido un número de sensores igual a 36. Estos sensores se reparten aleatoriamente por el área. En primer lugar, el programa ha calculado unos valores al azar en un plano 2D. Se debe hacer *run* al código varias veces para que genere un mapa de nodos que no se concentren en un mismo lado o que algunos de ellos coincidan en el mismo punto, para que a la hora de hacer la interpolación IDW no haya ningún problema.

Una vez se ha obtenido una distribución al azar que cumpla con los requisitos anteriores, las coordenadas 2D se han transformado a valores de latitud y longitud mediante una conversión lineal (véase Anexo a, sección 8.2); de esta forma podemos ver gráficamente si los puntos simulados se encuentran dentro de los límites del área y, por tanto, si cumplen con su función o no.

Los puntos de la generación escogida se han representado en la figura 14, donde se pueden ver los puntos naranjas que corresponden a los límites del área y se ve claramente, como las coordenadas de los 36 sensores (puntos azules) escogidos aleatoriamente sí se adecúa al área de la planta solar. Gracias a MyGoogleMaps [21], una herramienta de Google Maps, se pueden representar los puntos, introduciendo sus coordenadas desde un archivo Excel, que se ha guardado anteriormente desde el programa informático Matlab, y que posteriormente servirán para descargar los archivos de la radiación solar correspondiente a dichos nodos.

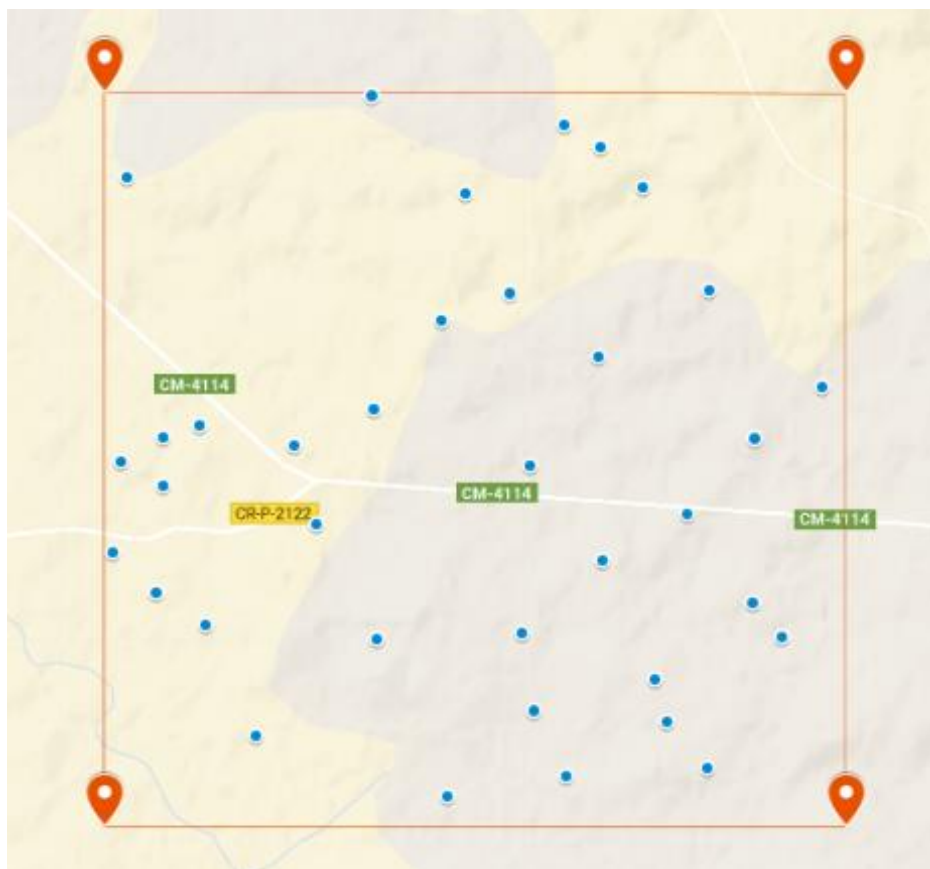


Figura 14. Escenario original [21].

5.4 Origen de los datos

Una vez se han elegido los puntos del escenario original, la siguiente parte es buscar una base de datos buena que se ajuste a la necesidad de nuestro problema, para encontrar datos de la radiación solar en esos puntos. No existen datos específicos en nuestros puntos provenientes de estaciones meteorológicas o de sensores colocados, por lo cual debemos encontrar una base de datos que nos proporcione dichos datos.

Los datos que necesitamos deben proporcionarse desde imágenes de satélites y utilizando métodos como el Heliosat-2, que se desarrolló durante los primeros años del presente siglo. Este modelo está basado en la identificación de nubes mediante el análisis de una serie temporal de imágenes [22].

Algunas bases de datos encontradas son el PVGIS [23], proyecto de investigación y desarrollo de un instrumento de políticas de apoyo a la evaluación geográfica de los recursos de la energía solar. Este proyecto forma parte de los desarrollados por SOLAREC, que forma parte de la unidad de Energías Renovables de la JCR (Joint Research Center) de la Unión Europea. PVGIS cubre el continente europeo, la cuenca mediterránea y otras zonas de África y Asia. Proporciona datos desde 1986.

Otra de la base de datos encontrada es el CM SAF [24], que ha sido desarrollado por el Servicio de Meteorología Alemán en colaboración con otros países europeos. Empezó a funcionar en su fase inicial en el año 2004. Utiliza diferentes versiones, si bien trabaja con datos de satélites de órbita polar (NOAA/METOP AVHRR) o de satélites geoestacionarios (METEOSAT/SEVIRI). Dispone de muchos productos de variables meteorológicas, incluyendo datos a tiempo real.

Cabe mencionar también, que, desde la NASA, se cuenta con una serie de productos relacionados con la radiación. Esta base de datos de la NASA, llamada SSE (Surface Meteorology and Solar Energy) [25], toma datos desde 1983 y los ofrece gratuitamente. Su punto fuerte es que tiene una cobertura de todo el mundo, aunque no mantiene una muy buena resolución. Los valores de radiación medidos en el SSE son comparados con los de la red de estaciones terrestres que tiene la NASA por todo el mundo, la BSRN (Baseline Surface Radiation Network).

El servicio de CAMS Radiation Service se enmarca en el programa Copernicus [26], cuya misión es la observación de la Tierra de la Unión Europea, que se preocupa por el medio ambiente y de nuestro planeta en general. Se utilizan enormes cantidades de datos globales procedentes de satélites y sistemas de medición terrestres. Todos los servicios de información proporcionados en este portal son de acceso gratuito y abierto para todos los usuarios, si bien, se tiene que crear un usuario de registro para poder aceptar las condiciones de uso y la licencia. Se ha decidido trabajar con este portal debido a la facilidad de descarga frente a los portales anteriormente citados.

El servicio de CAMS Radiation Service ofrece una guía de usuario [27] donde se repasan los métodos utilizados para el cálculo de la radiación solar. Este servicio ofreció el método Heliosat-2, mencionado anteriormente, con el que se construyeron unas bases de datos HelioClim-3, que comenzó en febrero de 2004 y ofrece datos hasta la actualidad.

Actualmente usan el método avanzado Heliosat-4, donde estima la irradiancia de onda corta descendiente recibida a nivel del suelo en todas las condiciones de cielo. Proporciona la irradiancia global y sus componentes directos y difusos en un plano horizontal y la irradiancia directa para un plano normal a los rayos solares.

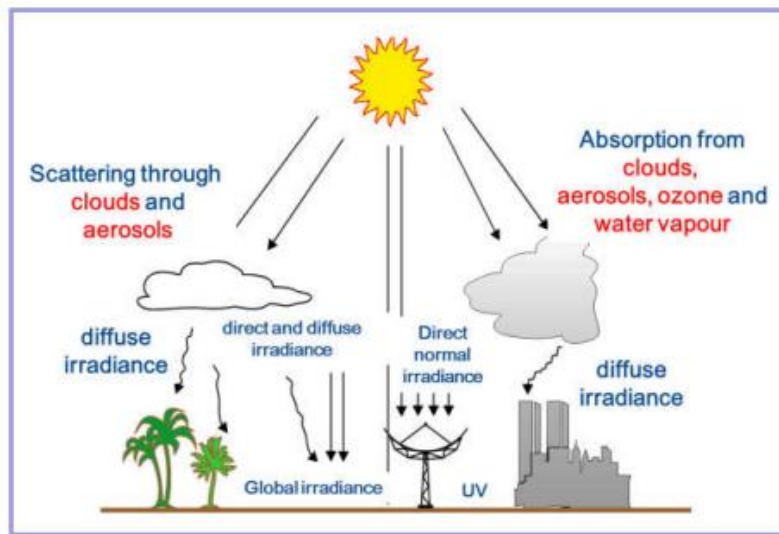


Figura 15. Esquema de las distintas radiaciones [27].

Está compuesto por dos tablas de consulta: el modelo McClear, que calcula la irradiancia en condiciones libres de nubes, y el modelo McCloud, que calcula la extinción de la irradiancia debida a las nubes. En la figura 16, se muestra un esquema de dicho método.

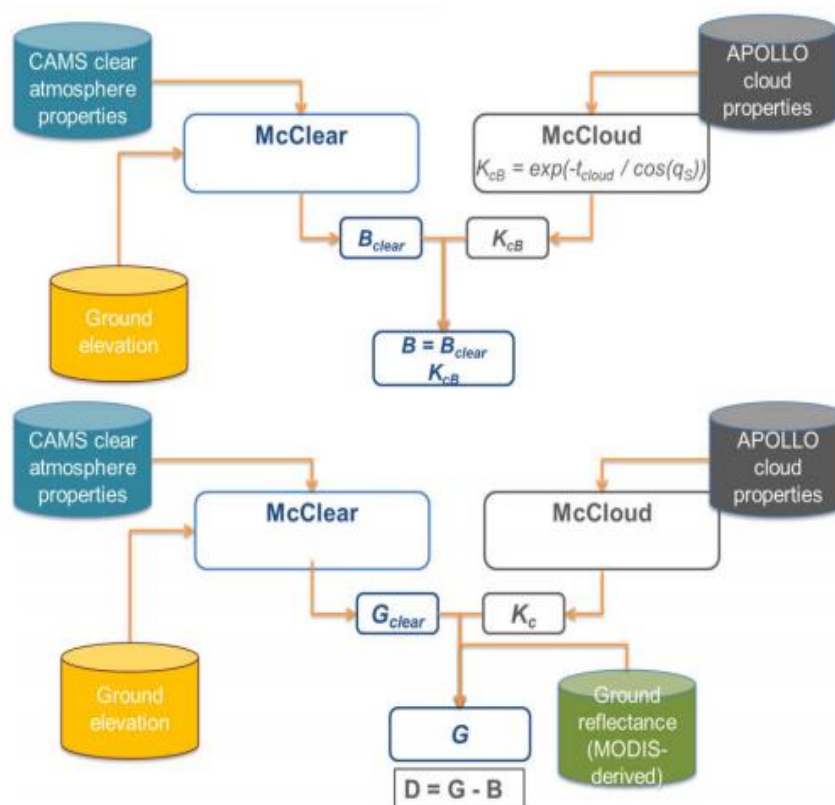


Figura 16. Esquema del método Heliosat-4 [27].

En esta guía también se destaca que los datos proporcionados simulan siempre un error tanto de sobreestimación como de subestimación, por lo que se concluye, que los datos no van a ser en su totalidad estrictamente exactos.

En la figura 17, se muestra el interfaz de usuario del portal [28]. Basta comprobar que, para descargar los datos, simplemente hay que seleccionar un punto específico del globo terráqueo, señalando latitud y longitud y se descargan los datos de radiación solar para ese punto en el intervalo que se desee, obteniendo, así como si de un sensor se tratara. Se pueden descargar los datos con un tiempo de muestreo de 1 min, 15 min, 1 hora, 1 día o 1 mes, señalando previamente la fecha en la que desea los datos.

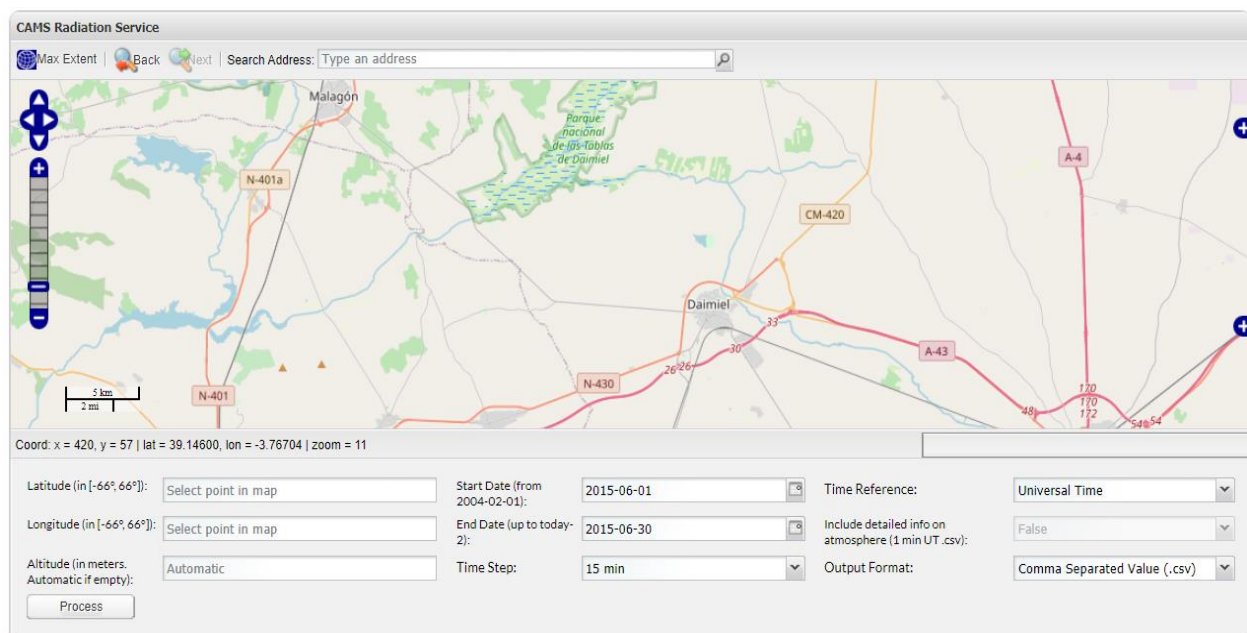


Figura 17. Interfaz de usuario del portal CAMS Service [28].

Para el trabajo, se han descargado los datos de los 36 sensores del escenario original en un período de 15 min para el mes de junio de 2015, que va a ser el mes de estudio. Los datos se pueden descargar en un archivo “.nc” o “.cv”. Se ha elegido la segunda opción por su facilidad para trabajar con el programa Excel.

Posteriormente, una vez calculadas las coordenadas de los sensores en los demás escenarios propuestos, se hará también la descarga de igual forma.

En la figura 18, se puede observar una muestra del archivo descargado para uno de los sensores del escenario original en el mes de junio de 2015.

# Observation period	TOA	Clear sky GH	Clear sky BH	Clear sky DH	Clear sky BN	GHI	BHI	DHI	BNI	Reliability	
2015-06-01T00:00:00.0/2015-06-01T00:15:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T00:15:00.0/2015-06-01T00:30:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T00:30:00.0/2015-06-01T00:45:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T00:45:00.0/2015-06-01T01:00:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T01:00:00.0/2015-06-01T01:15:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T01:15:00.0/2015-06-01T01:30:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T01:30:00.0/2015-06-01T01:45:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T01:45:00.0/2015-06-01T02:00:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T02:00:00.0/2015-06-01T02:15:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T02:15:00.0/2015-06-01T02:30:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T02:30:00.0/2015-06-01T02:45:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T02:45:00.0/2015-06-01T03:00:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T03:00:00.0/2015-06-01T03:15:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T03:15:00.0/2015-06-01T03:30:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T03:30:00.0/2015-06-01T03:45:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T03:45:00.0/2015-06-01T04:00:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T04:00:00.0/2015-06-01T04:15:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T04:15:00.0/2015-06-01T04:30:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T04:30:00.0/2015-06-01T04:45:00.0	0	0	0	0	0	0	0	0	0	0	1
2015-06-01T04:45:00.0/2015-06-01T05:00:00.0	0.5912	0.1308	0.005	0.1259	0	0.1308	0.005	0.1259	0	0.7333	
2015-06-01T05:00:00.0/2015-06-01T05:15:00.0	11.6565	3.6216	0.8521	2.7696	19.7413	3.5081	0.7758	2.7323	18.3197	0.8	
2015-06-01T05:15:00.0/2015-06-01T05:30:00.0	26.7671	10.9431	4.5001	6.443	54.1739	7.1764	0.8175	6.3588	11.697	1	
2015-06-01T05:30:00.0/2015-06-01T05:45:00.0	42.0975	20.4505	10.668	9.7825	83.1078	15.6577	1.3542	14.3035	9.6624	1	
2015-06-01T05:45:00.0/2015-06-01T06:00:00.0	57.5822	31.3536	18.6928	12.6608	107.0034	29.2601	14.9487	14.3114	84.0466	1	

Figura 18. Ejemplo de archivo de descarga ‘.cv’ visto en una hoja de Excel

Para analizar el documento descargado, debemos tener en cuenta que muestra una tabla con 11 columnas. En la primera columna, muestra para el punto deseado (longitud y latitud), los períodos de observación en 15 min para el mes de junio de 2015, donde se puede observar que esta porción de datos corresponde al primer día del mes de junio. En las siguientes columnas, se pueden observar los datos cada 15 minutos de distintos parámetros que vamos a analizar:

- TOA: Radiación en plano horizontal por encima de la atmósfera.
- Clear sky GHI: Radiación global en plano horizontal a nivel de suelo en cielo despejado.
- Clear sky BHI: Radiación directa en plano horizontal a nivel de suelo en cielo despejado.
- Clear sky DHI: Radiación difusa en plano horizontal a nivel de suelo en cielo despejado.
- Clear sky BNI: Radiación directa en plano móvil perpendicular a la incidencia solar en cielo despejado.
- GHI: Radiación global en plano horizontal a nivel de suelo.
- BHI: Radiación directa en plano horizontal a nivel de suelo.
- DHI: Radiación difusa en plano horizontal a nivel de suelo.
- BNI: Radiación directa en plano móvil perpendicular a la incidencia solar.
- Reliability: proporción fiable de datos [0-1].

Todas las variables se muestran en Wh/m². Se puede observar como existen las mismas variables en cielo despejado y en cielo con presencia de nubes (más real). Para entender los valores, podemos observar que los períodos de tiempo empiezan a las doce de la noche, por lo que hasta casi las cinco de la mañana el valor proporcional es 0. Una vez comienza la radiación solar, empiezan con valores ascendentes, si bien, si hubiera presencia de nubes, el valor decaería, viéndose la diferencia entre, por ejemplo, la columna de GHI y de Clear Sky GHI, donde ésta última seguiría aumentando sin tener en cuenta las nubes.

Para este trabajo, se trabaja con la columna de GHI, que corresponde a la suma de las radiaciones que llegan por radiación difusa y directa ya que los paneles más usados absorben el total de la radiación que llega. (véase en la figura 15).

Por tanto, se ha descargado, para el escenario original, un total de 36 documentos, correspondientes a las coordenadas donde se encuentran los sensores simulados. Cada sensor, a su vez, contiene un número de datos igual a 2880, que corresponde con la columna de GHI para todo el mes de junio 2015 en intervalos de 15 minutos, es decir, cada día de junio posee 96 valores de radiación solar en cada coordenada.

Posteriormente, para cada escenario calculado, se han descargado los datos de los sensores de igual manera, obteniendo así diferentes escenarios con distintos nodos y distintos valores de GHI en cada nodo. Para el resto del trabajo se denomina radiación solar a GHI.

5.5 Escenarios calculados mediante las salidas de la SOM

Como se ha explicado en el apartado anterior, se ha calculado el escenario original y se han descargado los datos de la radiación solar en el mes de junio de 2015 en un intervalo de 15 minutos para cada nodo de dicho escenario. Una vez obtenido el escenario original, el siguiente paso es obtener otros escenarios con nodos distintos.

Para el desarrollo de otros escenarios, se va a usar la salida de una red neuronal SOM. La red neuronal no supervisada se va a desarrollar en el entorno de Matlab con el uso de una caja de herramientas de redes neuronales *Neural Network Toolbox* de Matlab R2013a.

En primer lugar, se han definido las entradas a la SOM. Estas entradas son la longitud, la latitud y los datos de la radiación solar de los nodos obtenidos en el escenario original. Se van a simular cuatro tipos de SOM, con números de salidas igual, mayor y menor con distribución hexagonal al escenario original y una cuarta con igual número de nodos que el escenario original, pero distribución rectangular. A estos escenarios se le definirá de la siguiente manera:

- SOM1: escenario con salida de 36 nodos con distribución hexagonal.
- SOM2: escenario con salida de 64 nodos con distribución hexagonal.
- SOM3: escenario con salida de 16 nodos con distribución hexagonal.
- SOM4: escenario con salida de 36 nodos con distribución rectangular.

Por tanto, se tendrán 5 escenarios distintos, el escenario original y las cuatro salidas de las distintas SOM.

5.5.1 Neural Network Toolbox

En Matlab, existe una caja de herramientas específica para realizar distintos tipos de redes neuronales. Esta caja de herramientas posee una guía de usuario [29] donde se explica las distintas redes neuronales que puede simular como, por ejemplo, el Perceptrón, redes de base radial y la SOM. En el capítulo 7 de dicha guía se explica detalladamente como crear un mapa autoorganizativo SOM.

En primer lugar, se han de definir los vectores de entrada (atributos) de la capa de entrada.

El siguiente paso es crear una red con una distribución y una topología específica que defina la capa de salida de dicha SOM. Se ha usado la función *selforgmap* que utiliza distintos parámetros de entrada.

```
selforgmap(dimensions,coverSteps,initNeighbor,topologyFcn,distanceFcn)
```

Los parámetros de entrada son el número de nodos de la salida de la SOM (*dimensions*), que será un vector de fila que determinará el tamaño de la red. *CoverSteps* y *initNeighbor*, son el número de pasos de entrenamiento para la cobertura inicial del espacio de entrada, por defecto 100, y el tamaño del vecindario inicial, por defecto 3, respectivamente. Los dos últimos parámetros de entrada son el tipo de topología de la red y la función de distancia de los nodos con sus vecinos.

En cuanto a la topología de la red se puede elegir entre tres tipos de funciones (figura 19):

- **Hextop.** Simula las neuronas de una forma hexagonal.
- **Gridtop.** Simula las neuronas de una forma rectangular.
- **Randtop.** Simula las neuronas con un patrón aleatorio.

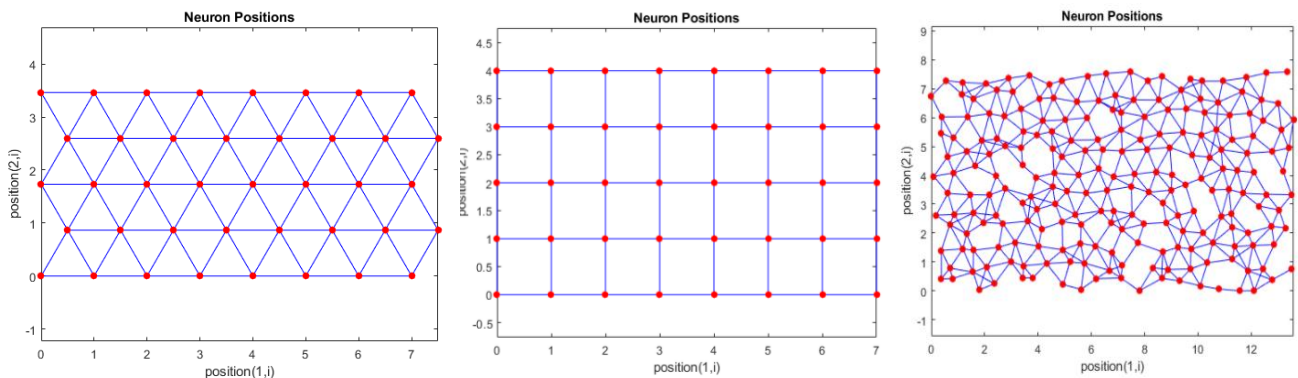


Figura 19. Ejemplo de topología hextop, gridtop y randtop en una SOM [29].

El último parámetro de entrada es el tipo de función de distancia. Las más usadas son:

- **Linkdist.** La distancia de enlace de una neurona es el número de pasos que debe tomarse para llegar a una neurona en consideración.
- **Dist.** Calcula la distancia euclideana de la capa entre neuronas.
- **Mandist.** Llamada Manhattan, calcula la distancia entre dos vectores como la suma de su diferencia en valor absoluto.

Una vez definida la red, se inicializa con los parámetros anteriores. El siguiente paso es entrenarla. Para llevar a acabo el entrenamiento de la red se usa la función *train*. Esta función usa como parámetros de entrada los datos de la capa de entrada a la SOM y la red anteriormente creada por la función *selforgmap*. Se pueden definir numerosos parámetros en la función *train* como pueden ser el número de iteraciones del entrenamiento y el método que lleva a cabo con los siguientes comandos:

```
net.trainParam=;
net.trainFcn=;
```


Hay diversos tipos de métodos de entramiento para la SOM, que pueden ser supervisadas o no supervisadas. Dentro del entrenamiento supervisado se puede elegir entre varios como puede ser *'trainr'*, entrenamiento aleatorio o *'trainc'*, entreamiento cíclico. Por su parte, dentro del entramiento no supervisado, se puede optar por un entrenamiento aleatorio *'trainru'* o entrenamiento por lote *'trainbu'*.

Una vez, que se empieza el entrenamiento se puede visualizar en Matlab la ventana de dicho proceso. En la figura 20, se puede ver un ejemplo de entrenamiento.

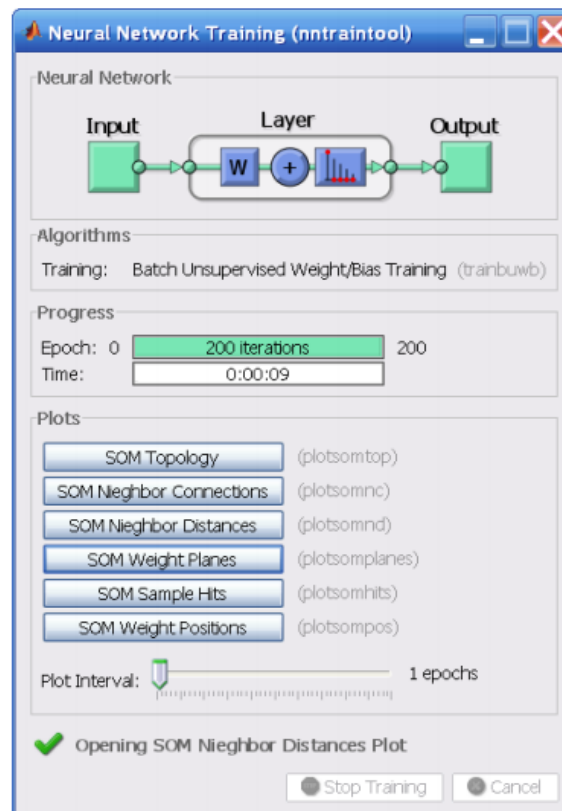


Figura 20. Ventana de visualización de entrenamiento de una red neuronal en Matlab [29].

Como puede verse en la figura 20, la ventana aporta información diversa: la función de transferencia de dicha red, el método elegido, el número de iteraciones del entrenamiento, el tiempo que tarda la red en concluir y una serie de gráficas que pueden verse concluido el entrenamiento (véase figura 21):

- **SOM Topology (*plotsomtop*)**. Se puede observar la topología de la red seleccionada en el mapa.
- **SOM Neighbor Connections (*plotsomnc*)**. Muestra las conexiones de las neuronas con sus vecinas.
- **SOM Neighbor Distances (*plotsomnd*)**. Muestra las distancias entre las neuronas vecinas. Los hexágonos azules corresponden con las neuronas. Las líneas rojas son las distancias entre ellas. Las regiones más oscuras significan unas distancias más grandes, mientras que las regiones más claras significan unas distancias más pequeñas entre ellas.

- **SOM Weight Planes (*plotsomplanes*)**. Muestra los distintos planos de pesos con respecto a los atributos de entrada. En este caso, puede observarse que existen 2 atributos de entrada. Las zonas oscuras representan unos pesos mayores.
- **SOM Sample Hits (*plotsomhits*)**. Se puede visualizar cuántos datos de entrada están asociados con cada neurona.
- **SOM Weights Positions (*plotsompos*)**. Muestra las ubicaciones de los atributos de entrada y los vectores de peso. Como puede verse en el ejemplo de la figura 21, los puntos verdes muestran las ubicaciones en el mapa de los atributos 1 y 2 de entrada, mientras que los puntos morados muestran las ubicaciones de los pesos 1 y 2 de la salida de la SOM. Es la gráfica más relevante de todas, porque muestra una visualización de cómo la red se ha adaptado a los datos de entrada.

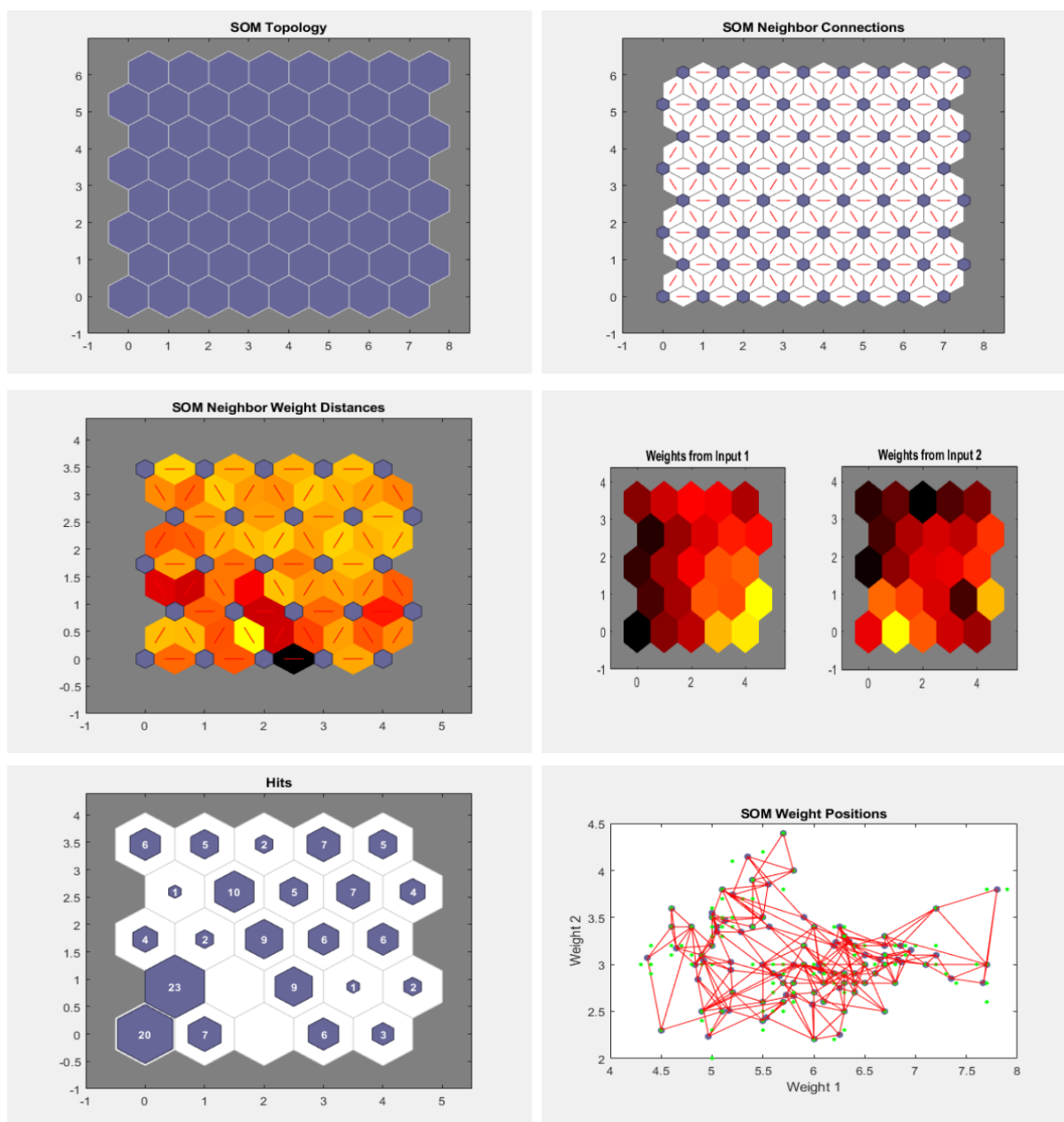


Figura 21. Ejemplos de visualizaciones de gráficas de una red neuronal SOM en Matlab [29].

Una vez se ha entrenado la SOM, si se quiere obtener los vectores de pesos de la salida de la SOM, basta con añadir la línea de comando `net.iw{I}` y te devuelve una matriz $M \times N$, donde M es el número de nodos de la salida de la SOM y N es el número de atributos de la capa de entrada.

5.5.2 Escenario SOM1

Para nuestro trabajo vamos a simular distintas SOM con distintos números de nodos y topologías para obtener nuevos escenarios. Para los datos de entrada tenemos 3 atributos que son la latitud, longitud y radiación solar del escenario original. Éste consta de 36 nodos y cada uno de ellos tiene 2880 datos de radiación solar pertenecientes al mes de estudio del trabajo, junio de 2015 en intervalos de 15 minutos. Por tanto, las entradas a la SOM son vectores de 3 dimensiones que corresponden a los 3 atributos anteriormente mencionados. La red se ha entrenado con 103680 vectores de entrada, lo que corresponde a una matriz 3×103680 y se ha estipulado un número de iteraciones igual a 1000. Para la que la SOM funcione lo mejor posible se ha normalizado cada atributo de entrada entre 0 y 1 (véase Anexo a, sección 8.3) para que los datos de entrada tengan un orden definido, ya que por ejemplo no es lo mismo la magnitud de longitud que la magnitud de radiación solar.

Los parámetros personalizados de la SOM1 se pueden observar en la siguiente tabla, siendo el resto definidos por defecto.

Dimensions	Topology	DistanceFcn	Training	Simulation
[6 6]	'hextop'	'linkdist'	'trainbu'	'negdist'

Tabla 1. Parámetros SOM1.

En el Anexo a: sección 8.4, se puede ver el script que se ha realizado para simular una red neuronal. El entrenamiento de la red se completa en unos 10 minutos aproximadamente. En la figura 22, se puede ver la función de transferencia de esta SOM.

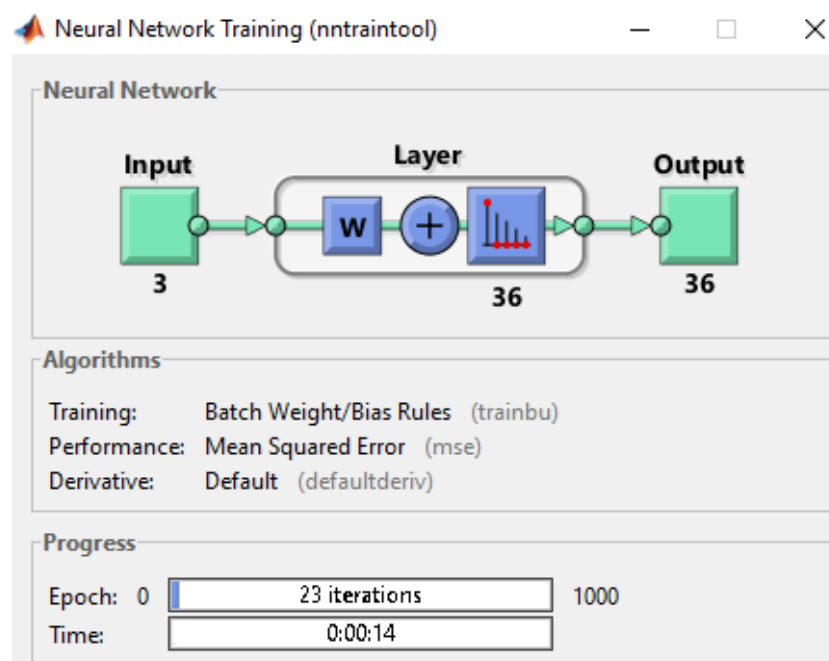


Figura 22. Ventana de visualización de entrenamiento de SOM1.

De todas las gráficas explicadas anteriormente, la mejor para visualizar el resultado de la SOM1 es *SOM Weight Positions*.

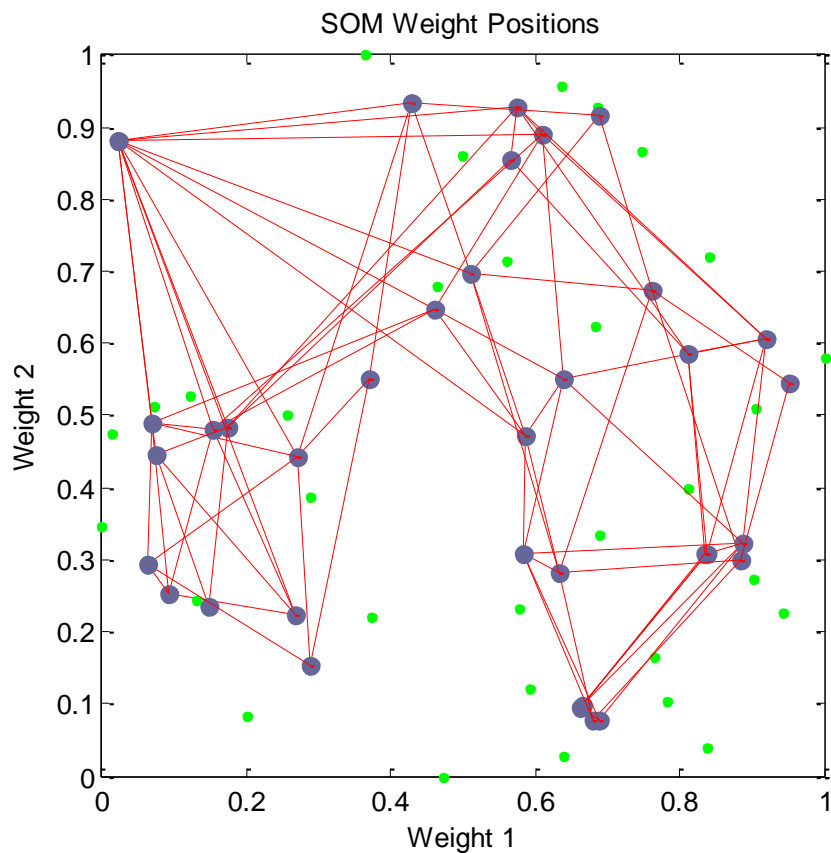


Figura 23. Escenario SOM1.

En la figura 23, se puede observar los puntos verdes que corresponden a los puntos de longitud y latitud del escenario original (véase figura 14) y los puntos morados que corresponden a los pesos 1 y 2 que son la longitud y latitud de los nodos de la salida de la SOM.

Para ver si la SOM se ha adaptado o no a los datos de la radiación solar basta con girar la imagen para ver el peso 3, que es el correspondiente a la radiación solar. En la figura 24, puede verse como las neuronas se han adaptado/amoldado a los datos de la entrada.

Para medir la calidad y/o precisión del mapa, se ha calculado el error de cuantización de la SOM1. Como se explica en el capítulo 4, sección 4.4, este error mide el valor medio de la distancia entre cada vector de entrada y la neurona ganadora.

En el Anexo a: sección 8.6, se describe un ejemplo de script donde se ha realizado dicha operación. Para ello, una vez obtenido los pesos de la salida de la SOM, a través de la función *dist* calcula la distancia euclídea de cada vector de entrada con las neuronas de salida. Como parámetros de entrada a la función se añade la matriz de pesos calculada por la red y la matriz de datos de entrada. Para este caso, la función devuelve una matriz 36x103680. Por último, se elige la menor distancia (neurona ganadora) de cada vector de entrada y se hace la media, dando un resultado de 0.1186.

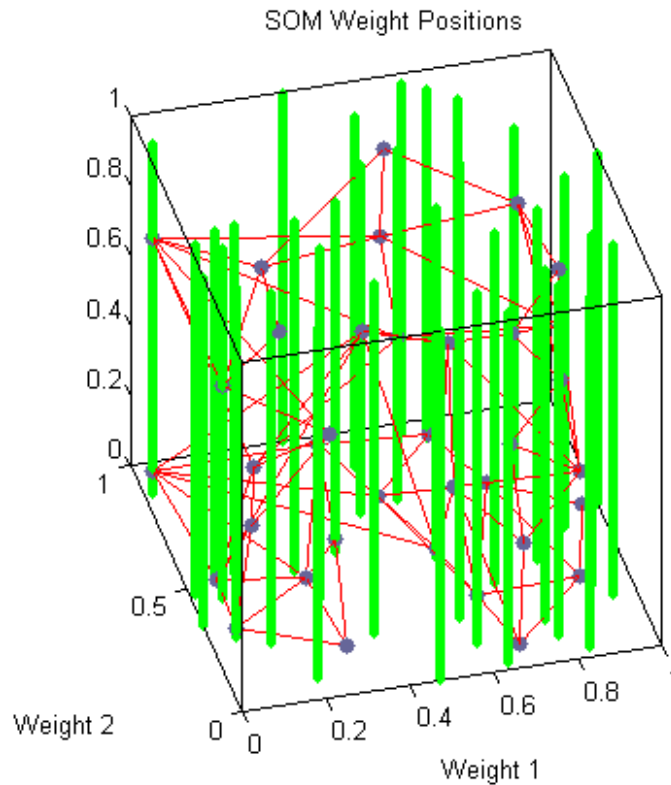


Figura 24. Adaptación de la SOM1 a los datos de entrada.

Las siguientes SOM siguen el mismo método explicado en esta sección, pero con cambios en sus parámetros de entrada.

5.5.3 Escenario SOM2

En este caso, la salida de la SOM va a ser de un número de nodos mayor que el escenario original. Todos los parámetros de entrada siguen intactos excepto el tamaño de la red. En la tabla 2, se muestran dichos parámetros. En este caso, el error de cuantización es 0.082.

Dimensions	Topology	DistanceFcn	Training	Simulation
[8 8]	'hextop'	'linkdist'	'trainbu'	'negdist'

Tabla 2. Parámetros SOM2.

Para esta SOM, el número de nodos de salida es 64. En la figura 25, se puede observar como existen neuronas de salida muy próximas entre sí.

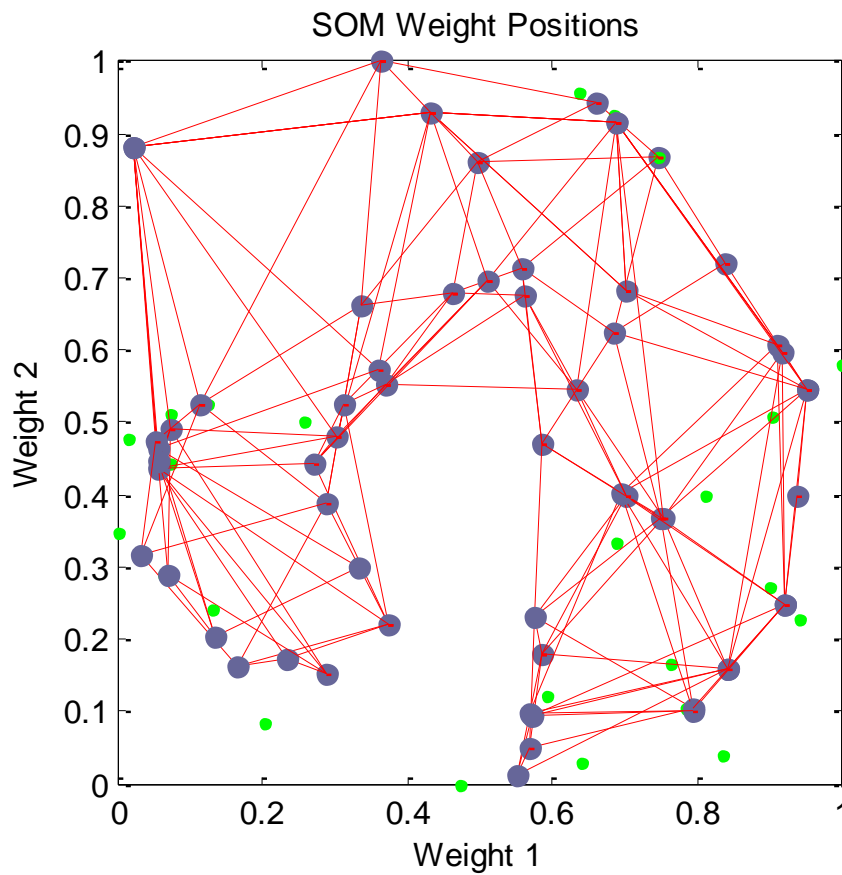


Figura 25. Escenario SOM2.

5.5.4 Escenario SOM3

En este caso, la salida de la SOM va a ser de un número de nodos menor que el escenario original. Todos los parámetros de entrada siguen intactos excepto el tamaño de la red. En la tabla 3, se muestran dichos parámetros. En este caso, el error de cuantización es 0.1735.

Dimensions	Topology	DistanceFcn	Training	Simulation
[8 8]	'hextop'	'linkdist'	'trainbu'	'negdist'

Tabla 3. Parámetros SOM3.

Para esta SOM, el número de nodos de salida es 16. En la figura 26, se han representado las posiciones de los pesos y se puede ver cómo abarca menos espacio que las salidas anteriores de las otras SOM's.

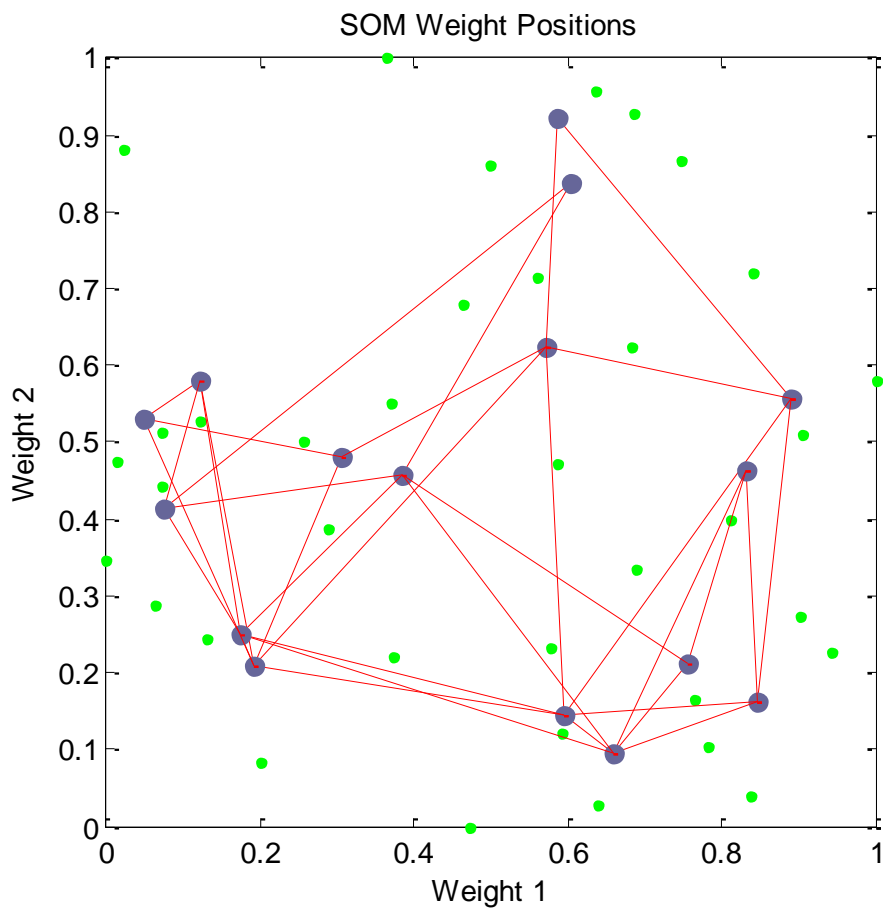


Figura 26. Escenario SOM3.

5.5.5 Escenario SOM4

En este caso, la salida de la SOM va a ser de un número de nodos igual que el escenario original y SOM1. Todos los parámetros de entrada siguen intactos excepto la topología de la red, que en este caso es rectangular (véase figura 19). En la tabla 4, se muestran dichos parámetros. En este caso, el error de cuantización es 0.1153.

Dimensions	Topology	DistanceFcn	Training	Simulation
[6 6]	'gridtop'	'linkdist'	'trainbu'	'negdist'

Tabla 4. Parámetros SOM4.

Para esta SOM, el número de nodos de salida es de 36. En la figura 27, se puede observar como se distribuyen los nodos con topología distinta a la SOM1, que pueden verse en la figura 23.

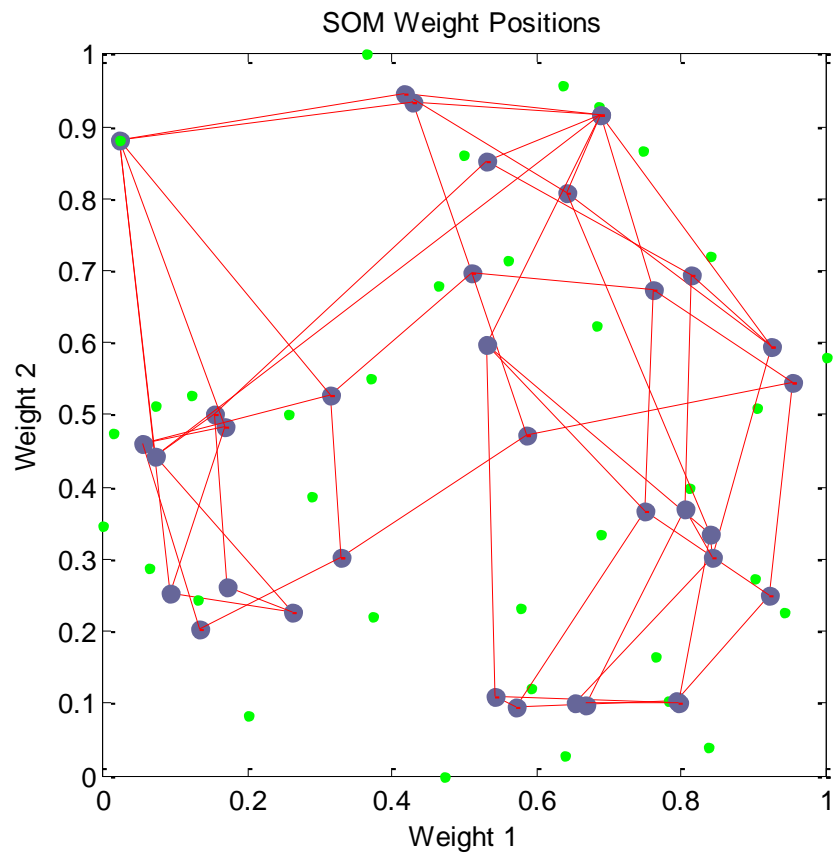


Figura 27. Escenario SOM4.

6 ESTIMACIÓN DE LA RADIACIÓN SOLAR MEDIANTE IDW

En este capítulo se explica cómo se ha realizado la estimación de la radiación solar en el área de estudio con los distintos escenarios en Matlab. Posteriormente, se comparan dichos resultados calculando índices de error y estudiando la influencia que tiene el error de cuantización de una salida SOM en los errores MAE y RMSE.

6.1 Estimación de la radiación solar mediante IDW

En el capítulo 5 se han obtenido los distintos escenarios que van a ser objeto de estudio. Una vez calculados, el objetivo es estimar los valores de radiación solar en el área de estudio (véase figura 12) con el método de interpolación de la distancia inversa ponderada para cada uno de los escenarios propuestos y compararlos.

Para calcular dicha estimación, se ha graficado en Matlab los límites del área. Una vez se tiene el área dibujada, se ha dividido en 2025 puntos espaciados regularmente por la misma. Para la realización de la obtención de puntos, se ha usado la función *meshgrid* (véase Anexo a, sección 8.7). Como resultado, se han obtenido dos matrices 45x45 correspondientes a los valores de longitud y latitud. Por último, se ha reorganizado estas matrices en una única, obteniendo una matriz 2025x2, siendo la primera columna correspondiente a la longitud y la segunda columna correspondiente a la latitud.

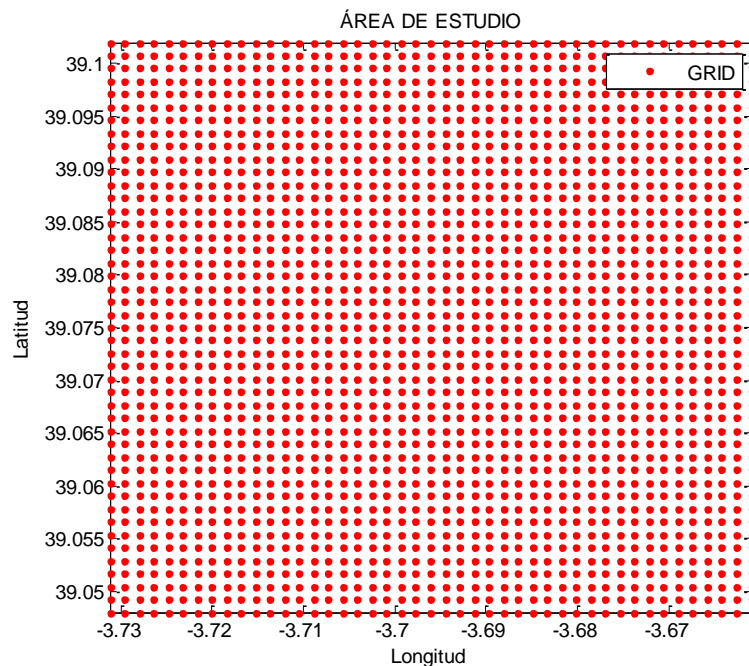


Figura 28. Puntos a estimar en el área de estudio.

En la figura 28 y 29 se muestra la división en puntos del área de estudio. Los puntos rojos son los que van a ser interpolados para así estimar la radiación solar en toda el área, mientras que los puntos azules (figura 29) corresponden a los puntos que se van a usar para realizar la interpolación (escenario original).

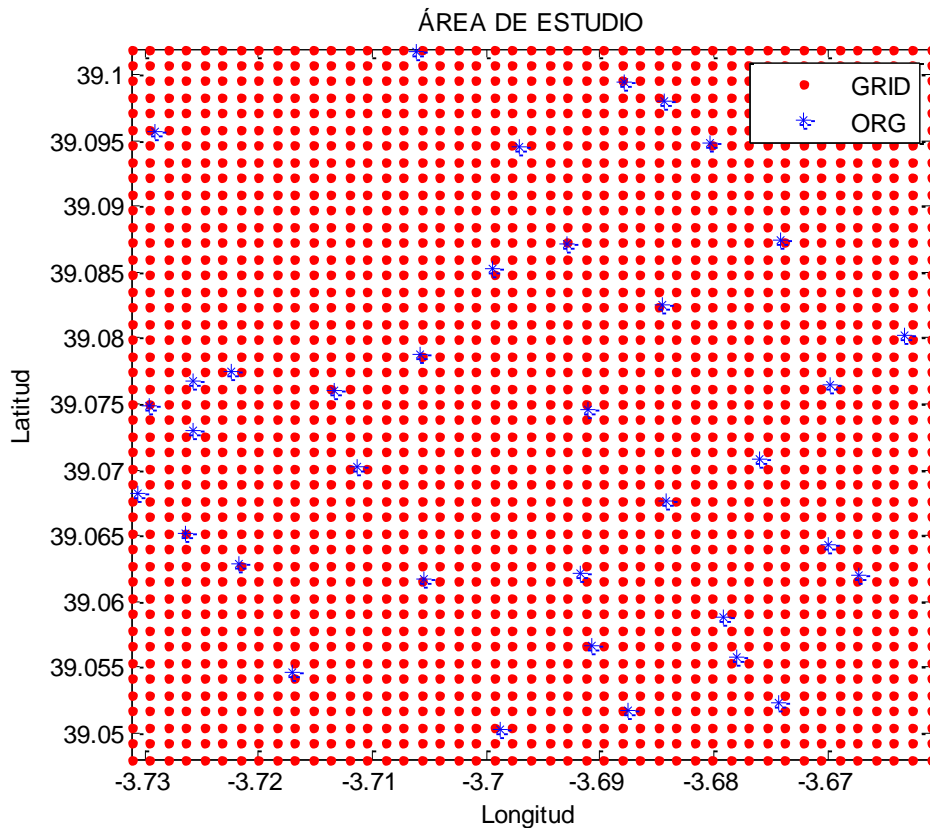


Figura 29. Puntos geográficos. Escenario original (ORG) y área a estimar (GRID).

Para poder graficar los puntos de los escenarios calculados mediante la salida de la SOM, se ha tenido que hacer una conversión lineal (véase Anexo a, sección 8.5). Los pesos 1 y 2 se han desnormalizado para así obtener coordenadas geográficas de longitud y latitud. Las coordenadas de las distintas SOM se han guardado en archivos Excel para su posterior uso.

Los distintos escenarios SOM se han dibujado en la figura 30, donde se puede ver que la conversión lineal es correcta y muestra la misma distribución que cuando los puntos estaban normalizados en las salidas de las SOM.

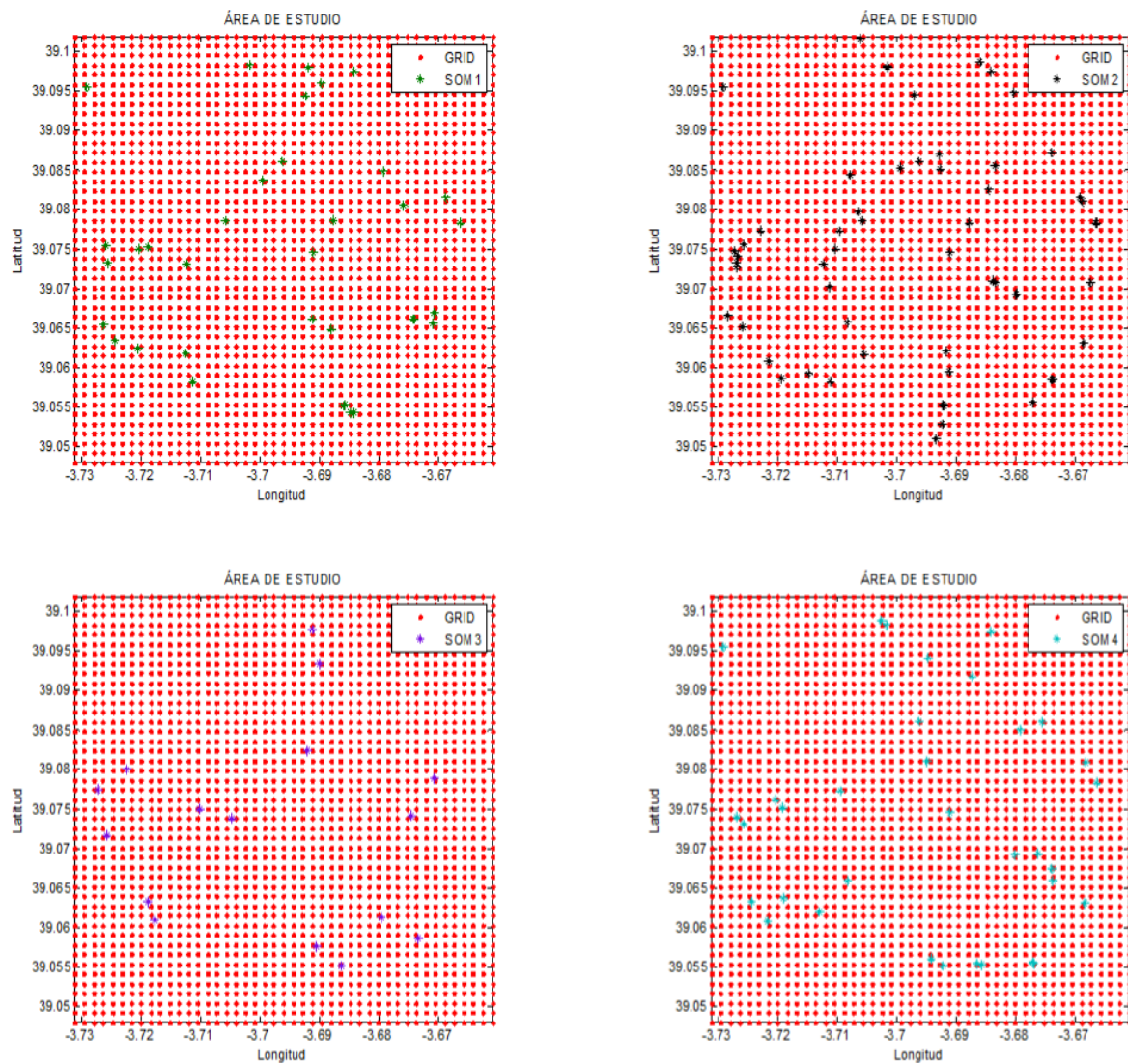


Figura 30. Puntos geográficos. Escenarios SOM1, SOM2, SOM3 Y SOM4 y área a estimar (GRID).

Una vez se han obtenido las coordenadas de los puntos en todos los escenarios, se han descargado los datos de la radiación solar para los escenarios SOM en todos los puntos de la misma forma que para los puntos del escenario original (véase capítulo 5, sección 5.4). Por lo que se tiene 5 matrices distintas con dimensiones *número_de_nodos* x 2880.

La interpolación se realiza con la función de ayuda encontrada en el entorno de archivos compartidos de Matlab [29] y modificada para este trabajo.

Esta función simula el método de interpolación IDW explicado en el capítulo 3. La función (véase Anexo a, sección 8.8), necesita unos valores de entrada previamente definidos.

$$V_i = gIDW(X_c, Y_c, V_c, X_i, Y_i, w, r_1, r_2)$$

Los valores de entrada son:

- **Xc**: coordenadas x de los puntos con valores conocidos. Longitudes de los nodos de cada escenario. Matriz [número de nodos x 1]
- **Yc**: coordenadas y de los puntos con valores conocidos. Latitudes de los nodos de cada escenario. Matriz [número de nodos x 1].
- **Vc**: valores conocidos en los puntos [Xc Yc]. Valores de radiación solar en el mes de junio 2015. Matriz [número de nodos x 2880].
- **Xi**: coordenadas x de los puntos que van a ser interpolados. Longitudes de los puntos a estimar del área de estudio. Matriz [2025 x 1].
- **Yi**: coordenadas y de los puntos que van a ser interpolados. Latitudes de los puntos a estimar del área de estudio. Matriz [2025 x 1].
- **w**: coeficiente de potencia.
- **r1**: método de interpolación. Se puede elegir con número de nodos vecinos fijados 'n' o radio de búsqueda fijo 'r'.
- **r2**: especificación del método elegido en r1.

Para este trabajo se ha definido el método de interpolar mediante IDW con número de nodos vecinos (véase figura 6, capítulo 3). El número de vecinos es igual al número de nodos que haya en cada escenario. Por ejemplo, para la interpolación en el escenario original, el número de nodos que entra en juego para calcular los valores en los puntos desconocidos es igual a 36.

En dicha función, se ha realizado una notable modificación. Para calcular la distancia real entre las coordenadas, se ha cambiado el método propuesto en la función original por la función haversine [30] (véase Anexo a, sección 8.9). Esta función calcula la distancia real en kilómetros que existe entre los nodos conocidos y los puntos a interpolar. Este cambio proporciona más verosimilitud al trabajo.

La función proporciona como salida una matriz V_i de tamaño 2025x2880, es decir, estima en cada punto del área a estudiar la radiación solar para el mes de junio de 2015 en intervalos de 15 minutos. Este proceso se ha realizado para los 5 escenarios propuestos, por lo que se tiene 5 matrices de tamaño 2025x2880.

Para comprobar la validez de la estimación en cada uno de los escenarios se ha calculado el error para 10 puntos de validación de los 2025 que hemos estimados. El error se ha ejecutado con dos medidas, el error medio absoluto (MAE), de sus siglas en inglés, *Mean Absolute Error* y el error cuadrático medio (RMSE), de sus siglas en inglés *Root Mean Square Error*.

$$MAE = \frac{1}{N} \sum_{i=1}^N |Z(x_i) - Z^*(x_i)|$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Z(x_i) - Z^*(x_i))^2}$$

Donde:

N= Número de datos.

Z= Valores observados.

Z*= Valores estimados.

A continuación, se ha elegido 10 puntos aleatorios de los puntos estimados en el área y se han descargado los datos de la radiación solar para el mes de junio de 2015 de la misma forma que se ha explicado anteriormente para los nodos de los distintos escenarios. A estos datos, los llamamos '*valores reales*'. En la figura 31, se puede observar los distintos puntos de validación que son rodeados con círculos de color negro.

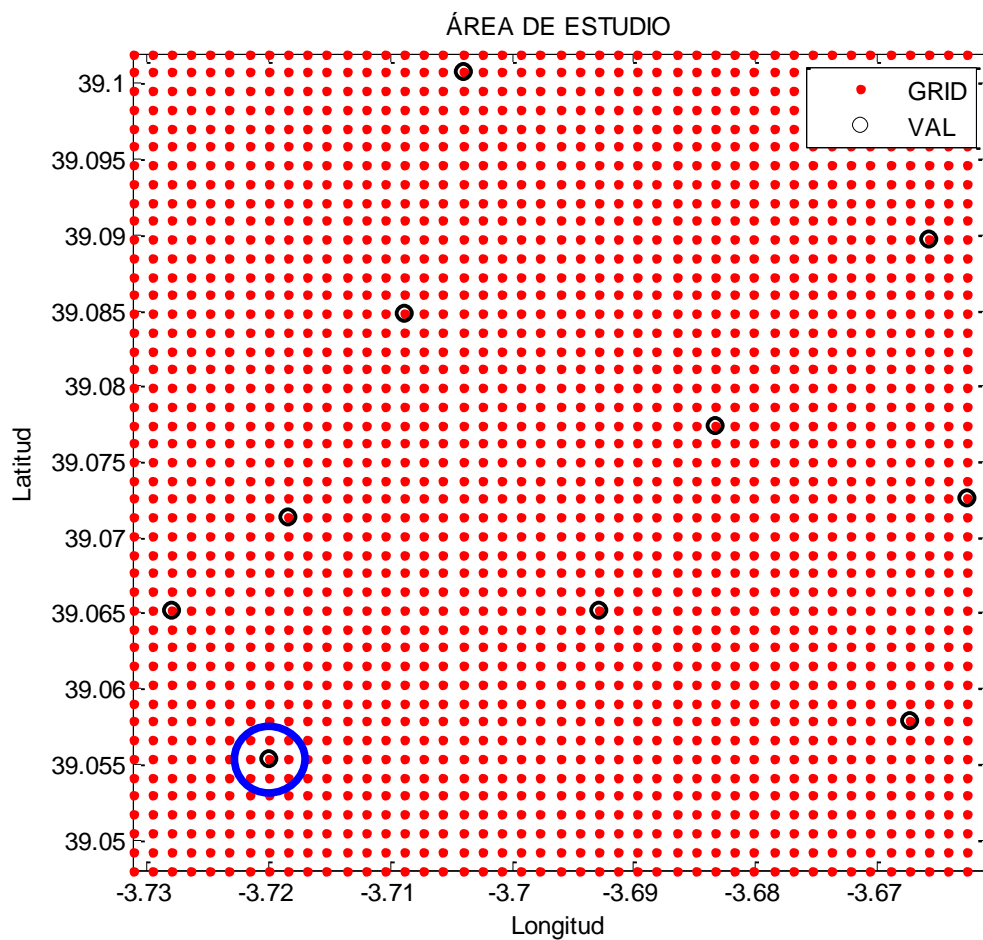


Figura 31. Puntos de validación para la estimación de la radiación solar (VAL).

Los errores se han calculado para 4 supuestos:

- Todo el mes de junio 2015.
- 12 de junio 2015: día nublado.
- 27 de junio 2015: día soleado.
- Instante determinado: 10 junio 2015, 7.30 a.m.

6.1.1 Errores de la estimación para todo el mes de junio de 2015

En este apartado se han calculado los errores MAE y RMSE de la estimación de la radiación solar para todo el mes de junio para los 10 puntos de validación calculados anteriormente y para todos los escenarios propuestos.

En la figura 32 se puede observar la radiación real para el punto de validación 1, punto rodeado en círculo azul en la figura 31.

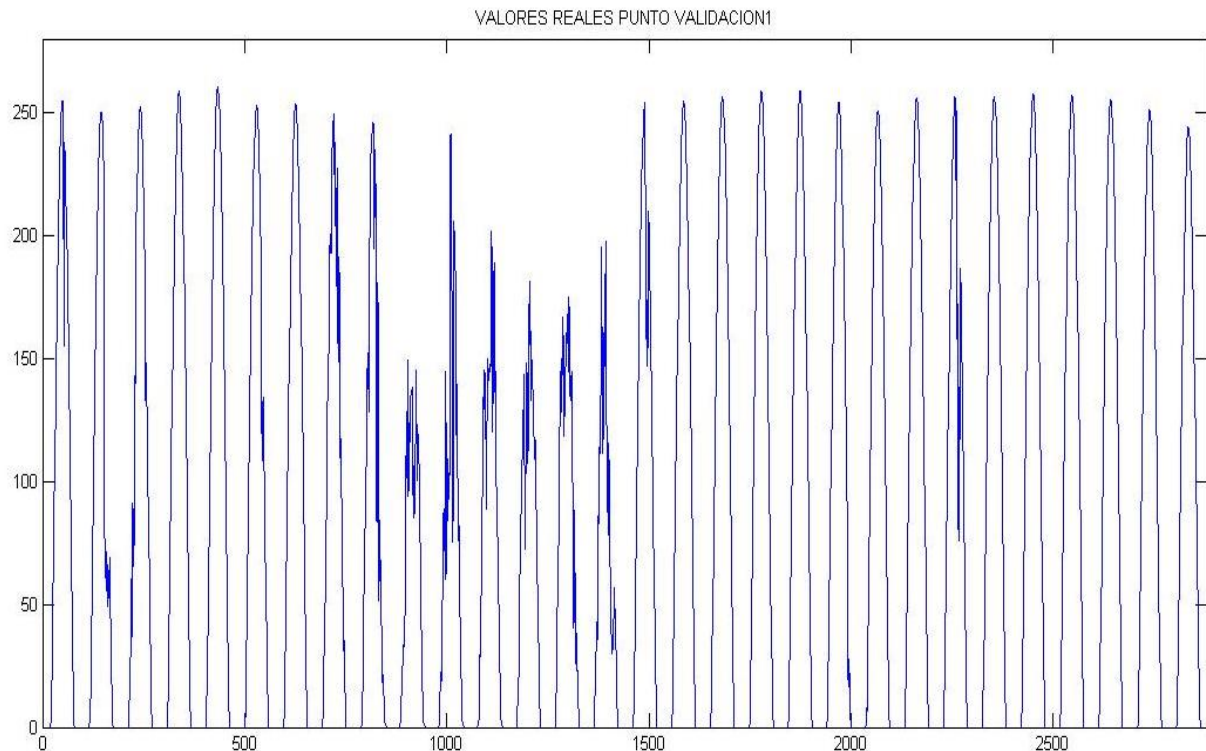


Figura 32. Radiación real de todo el mes de junio 2015 para el punto de validación 1.

Como se observa, cada curva muestra un día entero del mes de junio de 2015, correspondiente a 96 datos. De los 30 días que tiene el mes de junio se puede ver como disminuye la radiación solar entre los días 8 y 16 de junio, correspondientes a días donde la presencia de nubes es alta. Las curvas donde no se presentan cambios radicales se refieren a días totalmente soleados, donde la radiación solar dibuja una curva sin picos de cambio.

Para el mes de junio completo, obtenemos los distintos errores MAE y RMSE en los distintos escenarios para los 10 puntos de validación (véase Anexo a, sección 8.10), calculando así la diferencia entre los valores reales y los valores estimados. En este caso el número de datos que determinan los errores es 28800 para cada escenario. Los errores se pueden ver en la Tabla 5.

ESCENARIO	ORIGINAL	SOM1	SOM2	SOM3	SOM4
MAE	0.6004	0.6733	0.5410	0.8909	0.7629
RMSE	2.7745	2.8700	2.4183	3.8449	3.1615

Tabla 5. Errores de la estimación de la radiación solar para todo el mes de junio.

Los resultados de los distintos escenarios son similares entre sí, no distan mucho. Lo que demuestra que el uso de la SOM es válido para usarlo en la interpolación espacial de variables regionalizadas, en este caso la radiación solar.

Se demuestra que entre el escenario original y el uso de la SOM1 (36 sensores, topología hextop) los valores son parecidos, sin embargo, el error es un poco mayor usando la SOM1.

Si comparamos la SOM1 con la SOM4 (mismo número de nodos, pero distinta topología), se demuestra que con la topología hextop resulta un error menor que con la topología gridtop.

En cuanto a las SOM2 Y SOM3, en este caso usando un mayor número de nodos el error es menor, podemos observarlo con la SOM2 (64 nodos) que da un error menor que el escenario original y que la SOM1. Con 16 nodos, la SOM3, da el mayor error de los escenarios mostrados, justificando que cuanto menor sea el número de nodos usados en la interpolación, la estimación en el área será más imprecisa en este caso en concreto.

6.1.2 Errores de la estimación para el 12 de junio de 2015: día nublado

Una vez calculados los errores para todo el mes de junio se han realizado los mismos errores para un día nublado. En la figura 32, se han podido observar varios días donde la radiación solar tiene valores más pequeños debido a la influencia de las nubes. Se ha elegido el 12 de junio. En la figura 33, se ha representado el valor real de la radiación solar en el día señalado para el punto de validación 1.

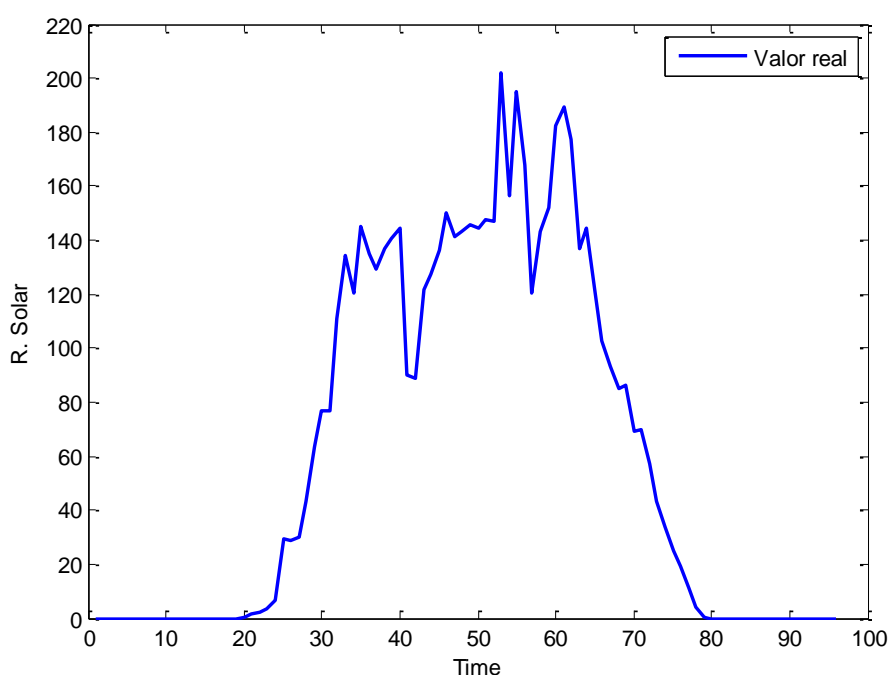


Figura 33. Radiación solar real del día 12 de junio en el punto de validación 1.

En dicha figura, se puede ver como el valor máximo de la radiación en este día es aproximadamente 200, si bien se puede comparar con la figura 32 donde se observa que los valores máximos de radiación son aproximadamente 250.

En las figuras 34 y 35, se han representado el valor real de la radiación en dicho punto de validación frente a los valores estimados mediante la interpolación IDW en dos de los escenarios propuestos. En este caso, se observa como los valores estimados con el escenario SOM1 se adapta mucho mejor a los valores reales frente a los valores estimados con el escenario SOM3, donde hay picos de diferencia considerables.

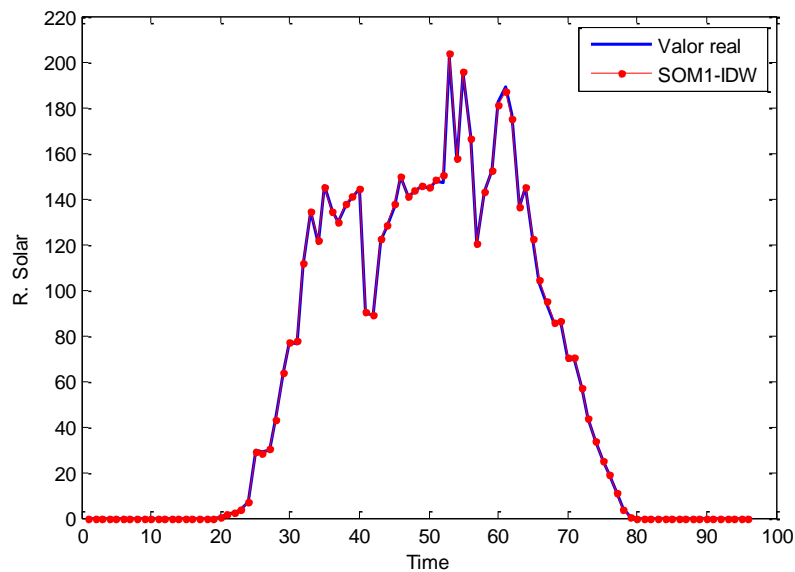


Figura 34. Valores radiación solar 12 de junio. Valor real y valor estimado (SOM1-IDW) en punto validación 1.

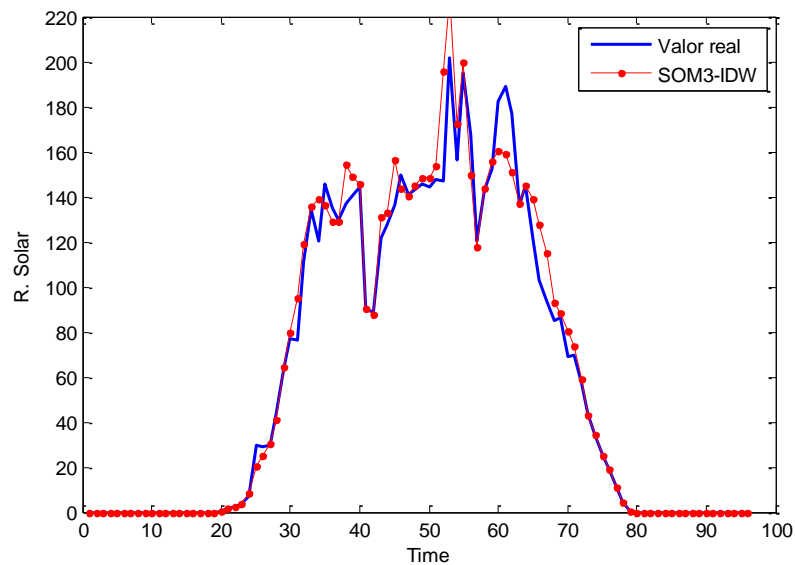


Figura 35. Valores radiación solar 12 de junio. Valor real y valor estimado (SOM3-IDW) en punto validación 1.

En la tabla 6, se han calculado los errores de la radiación solar para el día 12 de junio para los distintos escenarios en los 10 puntos de validación. En este caso el número de datos que determinan los errores es 960 para cada escenario.

ESCENARIO	ORIGINAL	SOM1	SOM2	SOM3	SOM4
MAE	1.7090	1.8912	1.4758	2.8817	2.1152
RMSE	4.1248	4.4637	3.5122	6.7602	4.8306

Tabla 6. Errores de la estimación de la radiación solar para día 12 de junio.

Se comprueba que los errores siguen la misma tendencia que para todo el mes de junio. Esto significa, que, para un día nublado, es decir, con presencia de nubes, se puede usar la interpolación IDW sin demasiados fallos. Si bien es cierto que los errores sí han aumentado con respecto a los del mes entero, debido a que estos errores son específicos para un día nublado, donde la presencia de nubes empeora la estimación de la radiación solar.

6.1.3 Errores de la estimación para el 27 junio de 2015: día soleado

En este caso, se han realizado los errores para un día de junio totalmente soleado, sin ningún tipo de nublado. A la hora de comparar estos resultados con los errores de un día nublado, se supone que los errores deberán ser más pequeños puesto que al no haber ningún tipo de nube la estimación no tiene ningún factor que la empeore.

En la figura 36, se ha representado el valor real de la radiación solar para el punto de validación 1 en el día estudiado frente al valor estimado calculado mediante IDW con el escenario SOM3. En la figura se puede observar como la radiación solar forma una curva perfecta, sin ningún pico distante pronunciado, lo que se puede visualizar perfectamente como cambia la radiación solar frente a un día nublado.

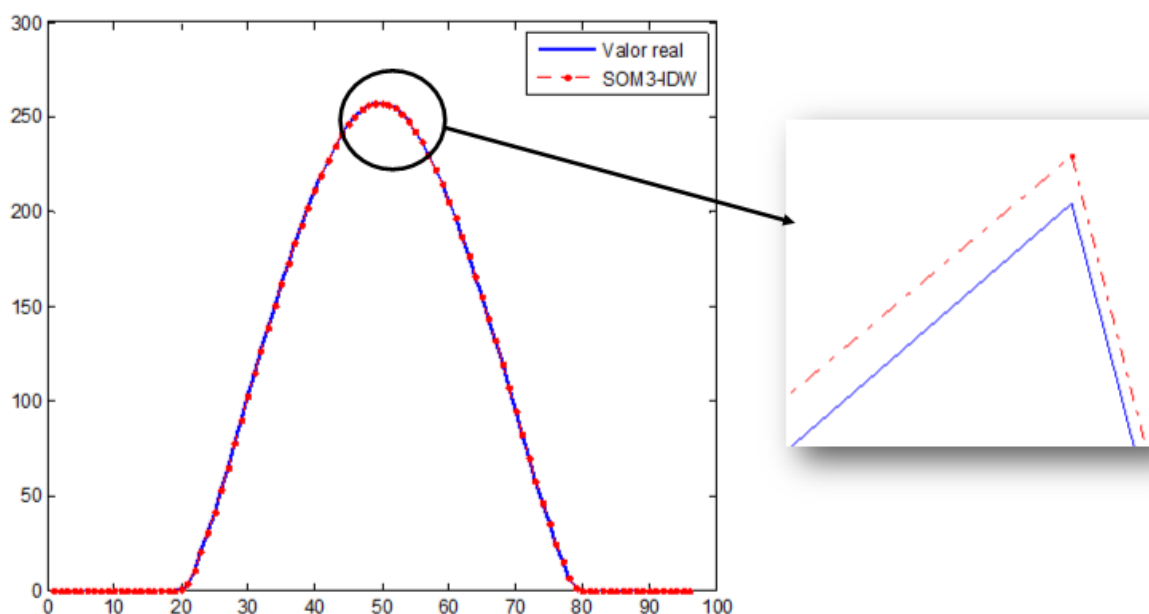


Figura 36. Valores radiación solar 27 de junio. Valor real y valor estimado (SOM3-IDW) en punto validación 1.

Se puede observar también como la estimación es muy próxima al valor real en el escenario que anteriormente había dado los mayores errores en los supuestos anteriores. Para verse con más claridad, se ha ampliado una zona concreta de la curva para poder ver la comparación.

En la tabla 7, se han calculado los errores de la radiación solar para el día 12 de junio para los distintos escenarios en los 10 puntos de validación. En este caso el número de datos que determinan los errores es 960 para cada escenario.

ESCENARIO	ORIGINAL	SOM1	SOM2	SOM3	SOM4
MAE	0.0225	0.0561	0.0359	0.0616	0.0570
RMSE	0.0760	0.1994	0.1189	0.2136	0.2001

Tabla 7. Errores de la estimación de la radiación solar para día 27 de junio.

Efectivamente, como se deducía anteriormente, los errores son más pequeños que en un día nublado. De hecho, cabe mencionar como los errores RMSE son mucho menores que en un día nublado. Se puede observar como en este caso el escenario que presenta los errores más pequeños es el original, por lo que ninguna salida de la SOM ha podido mejorar la estimación del escenario original.

6.1.4 Errores de la estimación para un instante determinado: 10 de junio a las 7.30 a.m.

En este último supuesto, se ha calculado los errores para un instante determinado a las siete y media de la mañana el día 10 de junio. Como se muestra en la figura 37, el día seleccionado es un día nublado. En el eje x, se muestra los intervalos de 15 minutos representados por puntos. El punto rojo corresponde a TIME=31, es decir, a las 7 y media de la mañana. Se han calculado los errores para ese instante determinado en los 10 puntos de validación.

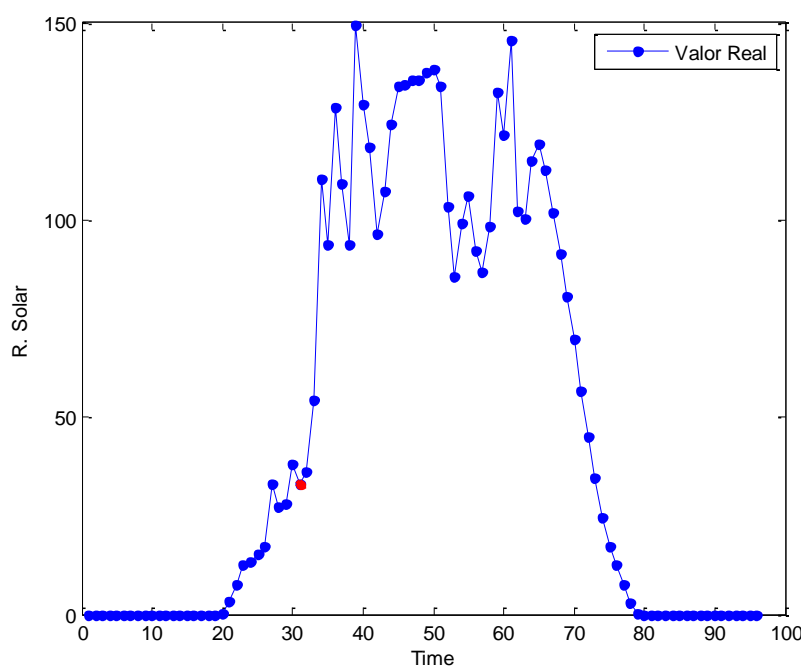


Figura 37. Radiación solar real del día 10 de junio en el punto de validación 1. Punto rojo (TIME=31, 7.30 a.m)

En la tabla 8 se muestran los errores para el TIME=31 en los distintos escenarios para los 10 puntos de validación. En este caso el número de datos que determinan los errores es 10 para cada escenario.

ESCENARIO	ORIGINAL	SOM1	SOM2	SOM3	SOM4
MAE	0.3918	0.5215	0.3814	0.7613	0.5625
RMSE	0.4767	0.6735	0.4710	0.9875	0.6588

Tabla 8. Errores de la estimación de la radiación solar para las 7.30 a.m en el día 10 de junio.

Los errores son muy similares entre sí. La SOM2 (64 nodos) es la que presenta menores errores. De nuevo, la SOM con 16 nodos (SOM3), es la que presenta unos errores más altos.

Para poder dibujar los valores de la estimación en el área con los distintos escenarios en el instante determinado, se ha usado la función *contourf* de Matlab (véase Anexo a, sección 8.11).

`contourf(X,Y,Z)`

Los parámetros de entrada son las matrices X, Y y Z. X contiene una matriz con las coordenadas x del plano 2D. Y contiene una matriz con las coordenadas y del plano 2D. Z contiene valores de altura del plano xy. La función *contourf* crea un gráfico de contorno relleno que contienen las isolíneas de la matriz Z. Matlab selecciona automáticamente las curvas de nivel que se muestran.

Para este trabajo se ha definido:

- X: matriz 45x45 con las longitudes de los puntos del área a estimar.
- Y: matriz 45x45 con las latitudes de los puntos del área a estimar.
- Z: matriz 45x45 con los datos de la radiación solar en el plano xy en el instante específico.

En la figura 38, se han representado los valores de la estimación de la radiación solar calculada con el escenario original a las 7.30 a.m del día 10 de junio. Se puede observar como los valores en el área se distribuyen desde 33 aproximadamente hasta 36.5. Las zonas azules corresponden a las zonas donde la estimación de la radiación solar presente unos valores más pequeños, quizás por la presencia de alguna nube en la zona, mientras que las zonas rojizas presentan unos valores estimados de radiación más altas.

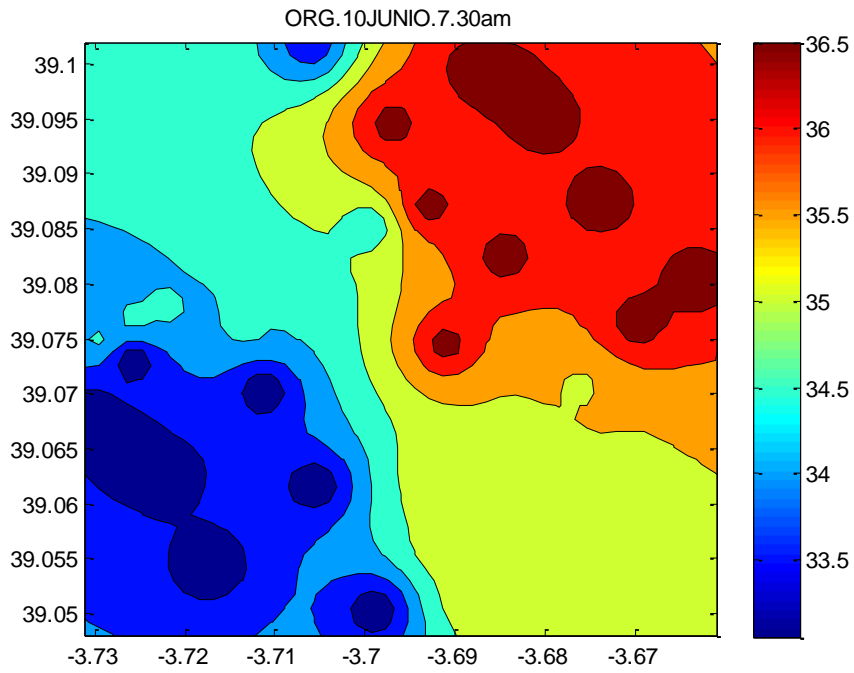


Figura 38. Estimación radiación solar en el área de estudio con el escenario original.

En la figura 39, se ha representado la estimación de la radiación solar en los demás escenarios propuestos.

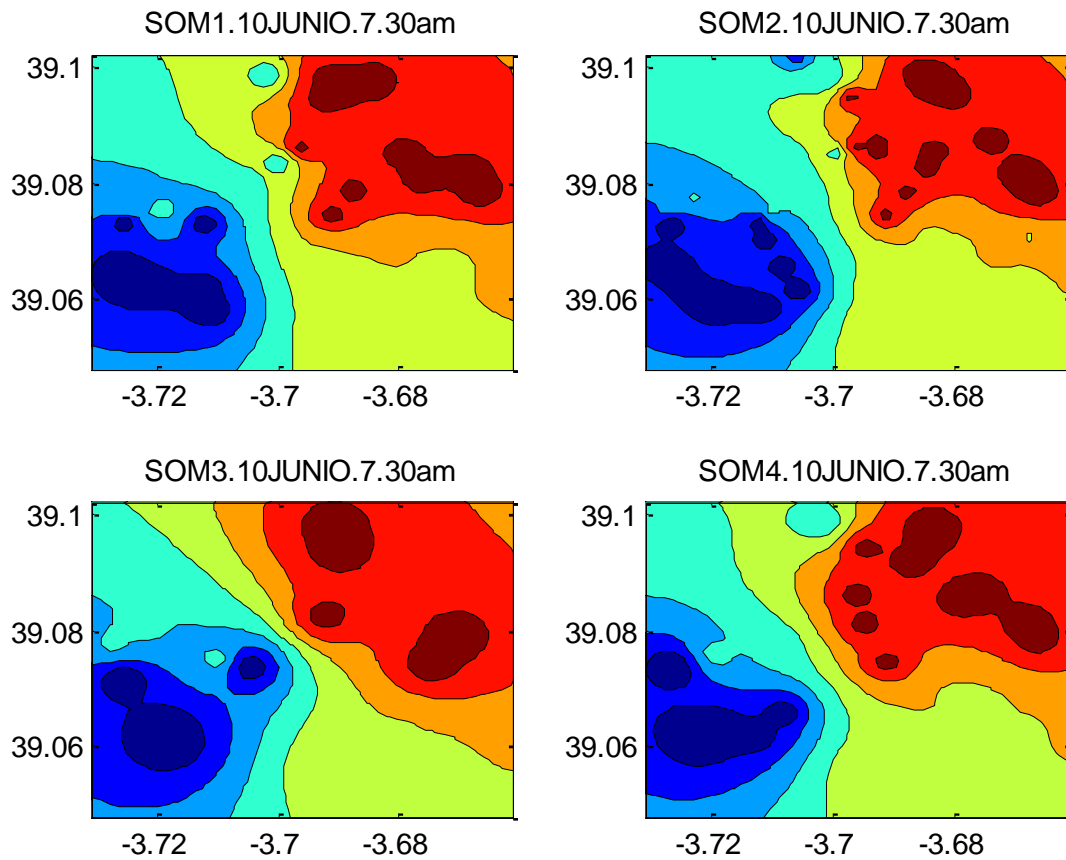


Figura 39. Estimación de la radiación solar en el área de estudio con los escenarios SOM1, SOM2, SOM3 y SOM4.

Si se comparan las figuras 38 y 39, se puede observar como las visualizaciones de los escenarios original, SOM1, SOM2 y SOM4 son los más parecidos entre sí, mientras que la estimación del escenario de SOM3 muestra unas curvas de nivel más diferentes. Si se observa la tabla 8 de los errores, las imágenes tienen sentido, ya que, al tener el error más grande con la SOM2, la visualización de la estimación es la que más diferencia establece con las demás.

6.1.5 Influencia del error de cuantización de las redes SOM en la estimación IDW

Los escenarios de las salidas de las SOM que se han visualizado anteriormente (capítulo 5, sección 5.5) son las salidas de las SOM entrenadas solo una vez. Como se sabe, cada vez que se entrena una SOM da una salida diferente, por lo que se ha decidido entrenarla cinco veces y elegir la mejor opción para cada escenario.

Para calcular qué salida es la mejor, se ha calculado la bondad de cada salida de las SOM a través del error de cuantización y se ha elegido la que presenta el menor error. Esta medida establece el ajuste de los nodos de salida a los vectores de entrada, calculando la media de las distancias entre cada vector de entrada y el nodo ganador (BMU). Anteriormente, se había calculado ya el error de cuantización en el primer entrenamiento de cada SOM.

En la tabla 9 se muestran los distintos errores de cuantización para cinco entrenamientos distintos para cada SOM. Se han señalado en color rojo los errores de cuantización menores para cada una, siendo el segundo entrenamiento el que ofrece el error de cuantización menor en las salidas de las SOM1 y SOM3, mientras que para las SOM2 y SOM4 el mejor entrenamiento es el quinto. En la tabla también se puede observar que el error de cuantización ha variado bastante en el entrenamiento de la SOM1, pasando de un error de 0.1186 a 0.1152. Posteriormente se analizará como este hecho influirá en el error de la estimación. Los errores de cuantización de las demás SOM han cambiado minimamente.

ESC. SOM	TRAINING1	TRAINING2	TRAINING3	TRAINING4	TRAINING5
EC_SOM1	0.1186	0.1152	0.1157	0.1158	0.1183
EC_SOM2	0.0820	0.0817	0.08107	0.0831	0.08106
EC_SOM3	0.173538	0.173515	0.173525	0.173543	0.173517
EC_SOM4	0.1153	0.1153	0.11522	0.1154	0.11520

Tabla 9. Errores de cuantización de las distintas SOM en función del entrenamiento.

En las figuras, 40, 41, 42 y 43, se han representado los nuevos escenarios con el mejor entrenamiento. Se puede visualizar como las posiciones de las neuronas de la SOM1 (figura 40) ha variado con respecto a la figura 23. Los nuevos escenarios de SOM2, SOM3 Y SOM4 no varían apenas la posición de las neuronas con los escenarios salientes con un solo entrenamiento (figuras 25, 26 y 27).

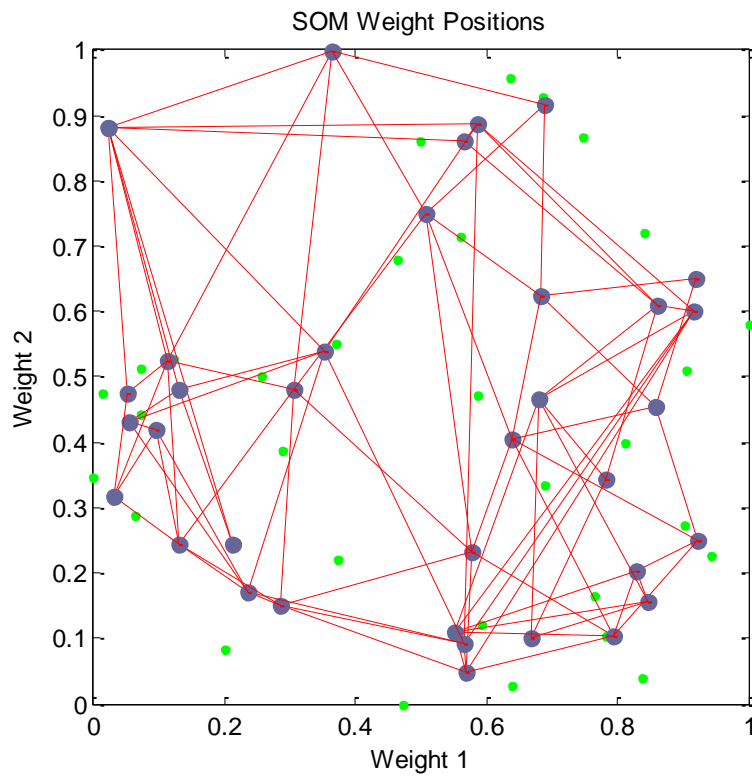


Figura 40. Escenario nuevo SOM1.

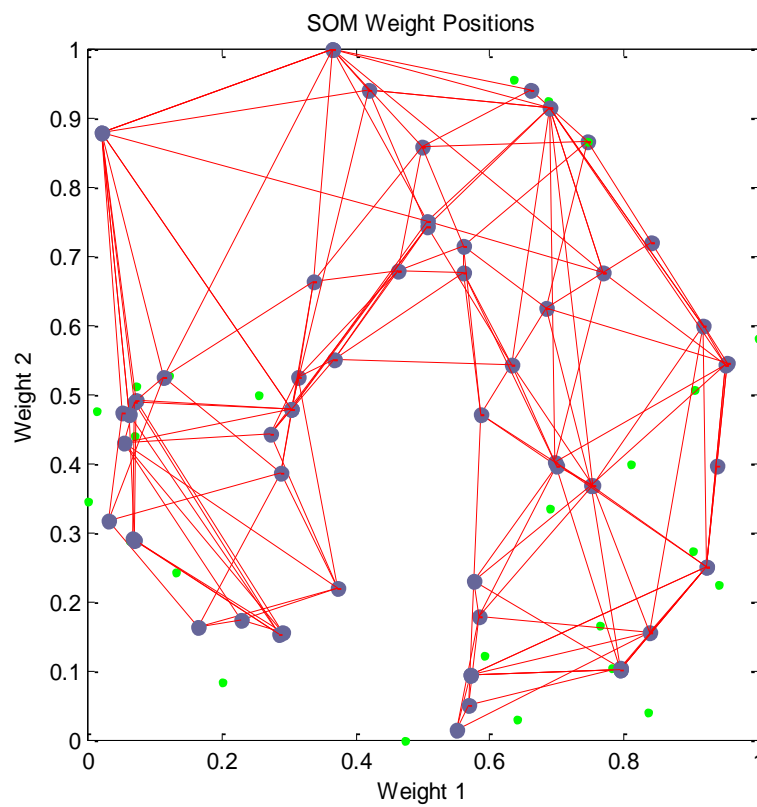


Figura 41. Escenario nuevo SOM2.

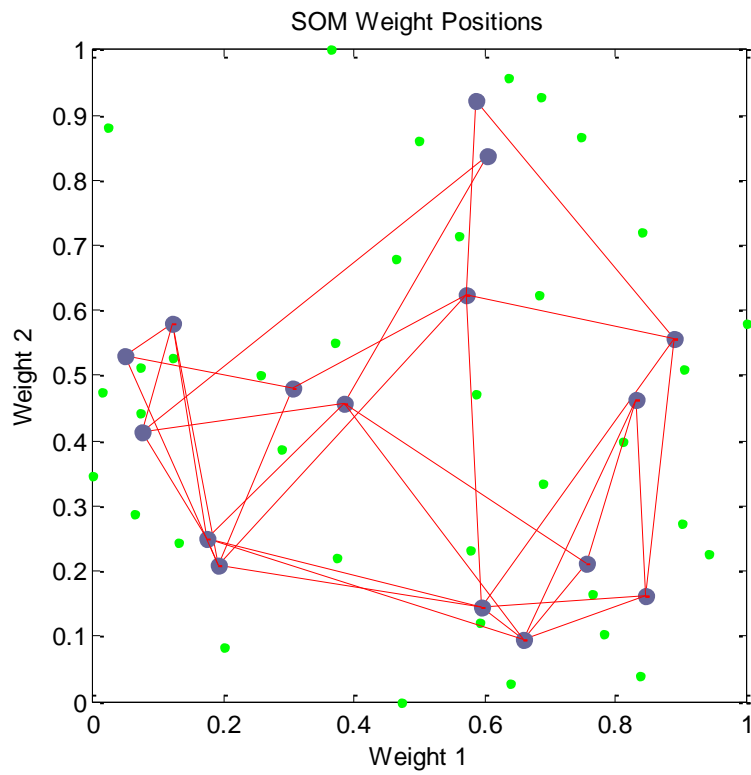


Figura 42. Escenario nuevo SOM3.

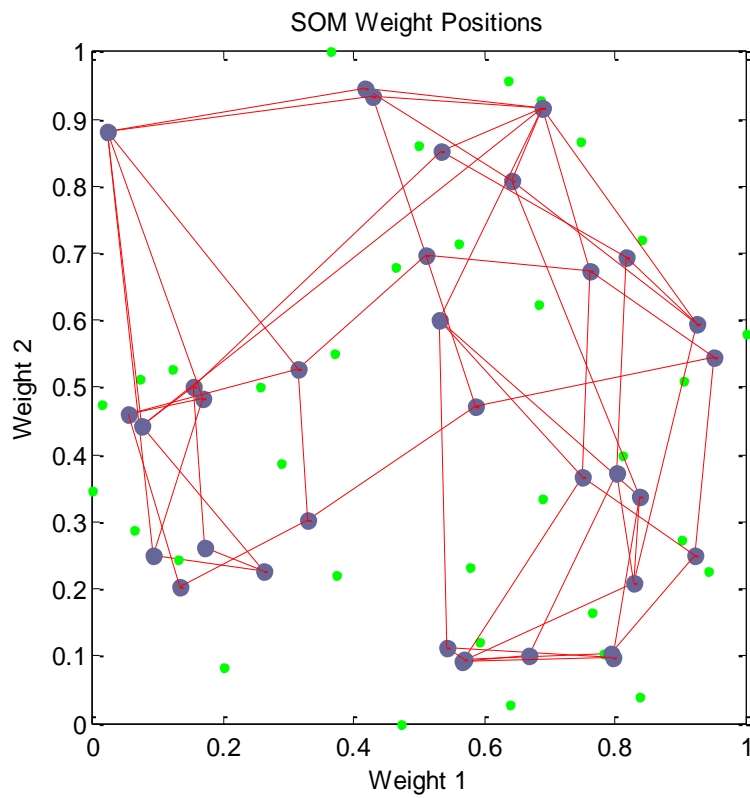


Figura 43. Escenario nuevo SOM4.

Una vez calculados los nuevos escenarios, se ha estimado la radiación de la misma forma que anteriormente y se han estimado los errores para los cuatro supuestos.

En la tabla 10, se muestran los errores para todo el mes de junio, para el día 27 de junio, para el día 12 de junio y para el instante específico, a las 7.30 a.m del día 10 de junio de los 10 puntos de validación.

TODO JUNIO 2015	MAE	RMSE
ESCENARIO ORIGINAL	0.6004	2.7745
SOM1	0.5839	2.5615
SOM2	0.5241	2.4043
SOM3	0.8906	3.8448
SOM4	0.7622	3.1552
12 JUNIO. DIA NUBLADO		
ESCENARIO ORIGINAL	1.7090	4.1248
SOM1	1.6923	3.9735
SOM2	1.4432	3.5049
SOM3	2.5714	6.0867
SOM4	2.1147	4.7880
27JUNIO. DIA SOLEADO		
ESCENARIO ORIGINAL	0.0225	0.0760
SOM1	0.0218	0.0704
SOM2	0.0201	0.0658
SOM3	0.0596	0.2069
SOM4	0.0550	0.2001
INSTANTE ESPECIFICO		
ESCENARIO ORIGINAL	0.3918	0.4767
SOM1	0.3815	0.4599
SOM2	0.3768	0.4010
SOM3	0.6622	0.8751
SOM4	0.5508	0.6522

Tabla 10. Errores de la estimación con los nuevos escenarios.

Si se comparan estos errores con los calculados anteriormente sin tener en cuenta el entrenamiento más óptimo de los cinco, se demuestra que al haber elegido la salida de cada SOM con menor error de cuantización para cada escenario los errores son menores. Es decir, existe una relación directa entre el error de cuantización y los errores de la estimación, obteniendo así una mejora en los resultados. Si bien es cierto que no se ha podido demostrar ningún patrón de comportamiento.

Por ejemplo, con respecto a la SOM1, al comparar las distintas salidas, se observa que como el error de cuantización ha pasado de 0.1186 a 0.1152, se obtiene un resultado en la estimación de la radiación solar mejor, incluso la estimación con la salida de la SOM1 es mejor que la estimación con el escenario original. La SOM1 ha conseguido obtener unos errores menores que el escenario original, que contiene el mismo número de nodos. En la figura 44, se ha dibujado la estimación de la radiación solar en el área para el instante determinado del 10 de junio a las 7.30 a.m con el escenario original, la SOM1 con error de cuantización 0.1186 y la SOM1 con error de cuantización 0.1152. Se puede observar que la SOM1 con EC: 0.1186, es la que presentaba un error un poco mayor y, por tanto, la que más se diferencia de las otras dos. Fíjase que el error MAE para el escenario original era de 0.3918, para la SOM1 con un entrenamiento de 0.5215 y el nuevo error para la SOM1 con el menor error de cuantización, 0.3815.

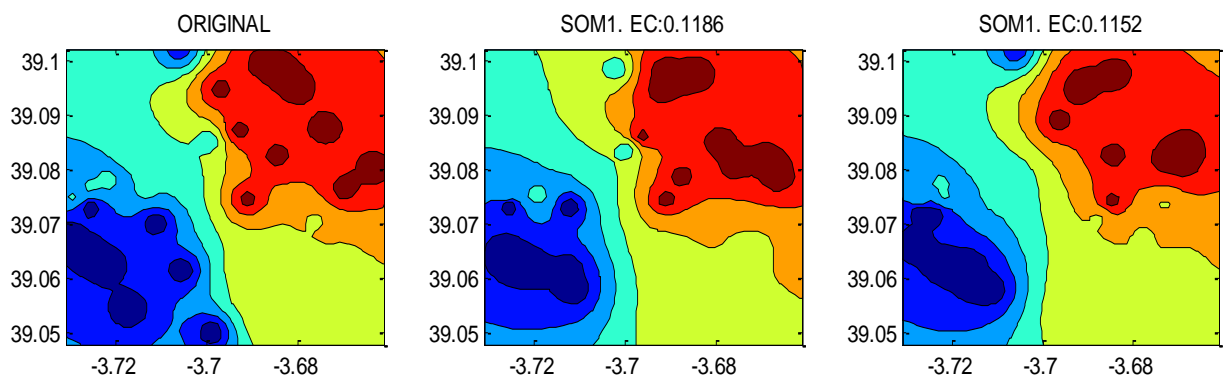


Figura 44. Estimaciones con escenario original, SOM1 con EC: 0.1186 y SOM1: con EC:0.1152 para instante específico.

Con respecto a los días soleado y nublado, los errores siguen siendo igual que los anteriores, lo que explica que la estimación en días nublados empeora por la perturbación de la presencia de nubes.

Si comparamos en función de la topología, podemos observar de nuevo como la topología hexatop con 36 nodos es mejor que la topología gridtop. Por ejemplo, si comparamos los errores del mes de junio completo observamos que con la salida de la SOM1 (hexatop) se obtiene un error MAE de 0.5839 frente a 0.7622 de la salida de la SOM4 (gridtop).

Aunque la SOM4 de 36 nodos establece unos errores más grandes que la SOM1 en ningún caso los errores superan a los de la estimación con la salida de la SOM3 de 16 nodos. Esto quiere decir, que cuanto menor sea el número de nodos de la salida de la SOM, más grandes van a hacer los errores a la hora de la interpolación puesto que cubre menos cantidad de área.

Se concluye que el escenario óptimo de los estudiados para estimar la radiación solar en nuestra área es la SOM2 de 64 nodos, es decir, la que mayor número de nodos presenta, mientras que la SOM3 de 16 nodos es la que peor estimación muestra en el trabajo.

Se ha mostrado, asimismo, la importancia que tiene el entrenamiento de una SOM para alcanzar una estimación más precisa, obteniendo así una mejora en todos los escenarios presentados.

Es cierto que al disminuir los errores de cuantización en las distintas SOM, los errores han disminuido para cada escenario, pero presentan diferentes comportamientos. Si calculamos la diferencia en valor absoluto entre los errores de cuantización de los primeros escenarios con los nuevos obtenemos la tabla 11.

ESCENARIO	EC-TRAINING1	EC-BEST TRAINING	DIFERENCIA
SOM1	0.1186	0.1152	0.0034
SOM2	0.082	0.08106	0.00094
SOM3	0.173538	0.173515	0.000023
SOM4	0.1153	0.1152	0.001

Tabla 11. Diferencia de Errores de Cuantización.

Observamos que la SOM1 es la que ha presentado una mayor mejora en el rendimiento de la red y, a su vez, la que ha presentado las mayores mejoras en los errores de la estimación. Las demás SOM han presentado menores mejoras en el rendimiento y también menores mejoras en las estimaciones. Esto se puede explicar teniendo en cuenta cómo han cambiado las posiciones de las neuronas en las salidas. Las neuronas del nuevo escenario de la SOM1 son las que presentan ubicaciones de pesos 1 y 2 más diferentes al escenario calculado con el primer entrenamiento. Las ubicaciones de las neuronas en los nuevos escenarios de las otras tres SOM son muy similares a los escenarios calculados con el primer entrenamiento, por lo que a la hora de realizar la estimación los nodos van a estar prácticamente en la misma posición y los errores de estimación van a ser prácticamente los mismos. Esto podría ser una explicación al hecho de que la relación entre el error de cuantización y errores en la estimación no siga una tendencia clara.

7 CONCLUSIONES

El objetivo del trabajo ha consistido en comprobar que el resultado de la red neuronal artificial no supervisada (SOM) proporciona resultados aplicables en la obtención de la variabilidad espacial de un área determinada, partiendo de valores de radiación solar para una planta solar hipotética ubicada en la zona de Tablas de Daimiel sin sensores colocados en el terreno. La idea ha consistido en obtener distintos escenarios donde ubicar sensores irregularmente en nuestra planta, calcular la estimación en el área con cada uno de ellos y realizar una comparación. Se han usado datos obtenidos mediante satélites de la Unión Europea, disponibles online de forma gratuita.

Una vez calculado el escenario original, se ha proporcionado una nueva alternativa para obtener la variabilidad espacial de la radiación solar basada en un mapa autorganizado. La red SOM se ha usado para obtener nuevos escenarios donde aplicar la interpolación de la radiación solar. Finalmente se han obtenido 5 escenarios distintos, el escenario original calculado aleatoriamente y 4 escenarios calculados como salida de la red SOM. Tres de los escenarios SOM se han calculado con tipología hextop y con igual, mayor y menor número de nodos que el escenario original. La cuarta SOM se ha calculado con igual número de nodos que el escenario original, pero con distinta topología, gridtop.

Una vez calculados todos los escenarios se ha estimado la radiación solar en cada uno de ellos en el área de estudio mediante el método de interpolación IDW. Al comparar las predicciones con valores reales se ha visto que los errores no son demasiado distintos entre sí. Se ha comprobado que es posible encontrar una SOM que obtenga un escenario, una nueva distribución de los sensores con los que se obtiene un error menor que con la distribución original.

Si se compara con respecto a la topología de una SOM, se ha demostrado que el uso de una topología hextop mejora los resultados que con gridtop. A su vez, se ha demostrado que cuanto mayor es el número de nodos de la red menor es la diferencia entre los valores estimados y los valores reales para este caso.

Por otra parte, se ha analizado qué influencia tiene el error de cuantización de la salida de una SOM en los errores anteriores. Se ha calculado 5 entrenamientos por cada SOM y se ha elegido la mejor opción para cada uno de ellos. Este trabajo ha demostrado que, para este caso, un error de cuantización menor en cada SOM otorga unos errores menores.

Los datos de radiación usados no se han obtenido de forma experimental mediante sensores sino mediante datos de satélite por lo que no es exactamente una aplicación a una situación real, pero puede servir de base a desarrollar para trabajos futuros con datos reales.

Con respecto al método de interpolación IDW se puede deducir que es un buen método para obtener resultados de puntos desconocidos a través de puntos conocidos.

Líneas futuras de investigación

- Como continuación de este trabajo se podría diseñar una red óptima para descubrir cual sería el mínimo número de nodos posible para cubrir toda el área y que no aumentase un error determinado. Como premisa se deduce que el error de cuantización es clave. Quizás un aumento en el número de entrenamientos pudiera ser importante.
- Otra forma de continuar el trabajo sería modificar parámetros de entrada a la SOM, aumentando los mismos. Se podría añadir variables como temperatura y ver si influye para que los errores disminuyan.

- Utilización del método de radio de búsqueda fijo de la interpolación IDW y comparar los resultados con el método de número de vecinos fijo.
- Utilización de otra medida de calidad del mapa como *Simplicity*: medida de la "expansión, envoltura" de la red, la suma de las distancias de cada nodo con todos los demás (accumulated neighbourhood distance).

8 ANEXO A: CÓDIGO REALIZADO EN MATLAB

8.1 Generación puntos distribuidos

```
% Generar los puntos del mapa
clear, close
NPUNTO=36; %Numero estaciones
AREA=36.5; %Determinar área de estudio
Z=sqrt(AREA); %AREA CUADRADA, LADO EN KM
%N números aleatorios en el intervalo (a,b) con la fórmula r=a+(b-a).*rand(N,1)
puntos=Z.*rand(NPUNTO,2);

%Dibujar puntos aleatorios

for i=1:NPUNTO
hold on
plot(puntos(i,1),puntos(i,2),'ro-')
end
axis([ 0 Z 0 Z])
title('Superficie')
xlabel('X')
ylabel('Y')
% Guardar una generacion
```

8.2 Transformación de los puntos a coordenadas

```
close
NPUNTO=36; %Numero estaciones
AREA=36.5 %Elegir area
Z=sqrt(AREA); %AREA CUADRADA, LADO EN KM
lat1=39.102;
lat2=39.048;
lon1=-3.661;
lon2=-3.7312;
diflat=lat1-lat2;
```

```
diflon=lon1-lon2;
%Puntos elegidos de forma aleatoria y guardados, se cargan.
load escenario1

%declaracion de vector de coord.

coordenadas_escenario1=ones(NPUNTO,2);

%Grafica de las estaciones en un plano
figure()
for i=1:NPUNTO
hold on
plot(estaciones1(i,1),estaciones1(i,2))
end
axis([ 0 Z 0 Z])
title('Superficie')
xlabel('X')
ylabel('Y')
%Pasamos a coordenadas con una relación lineal.
for i=1:NPUNTO
coordenadas_escenario1(i,1)=estaciones1(i,1)*diflon/Z+lon2;
coordenadas_escenario1(i,2)=estaciones1(i,2)*diflat/Z+lat2;
end

%Se guarda en Excel

filename = 'coordenadas_escenario1.xlsx';
xlswrite(filename,coordenadas_escenario1,1,'A2')

%Se dibuja
figure()
for i=1:NPUNTO
hold on
plot(coordenadas_escenario1(i,1),coordenadas_escenario1(i,2))
end
axis([ lon2 lon1 lat2 lat1])
title('Superficie')
xlabel('X')
ylabel('Y')
```

8.3 Normalización datos de entrada a la SOM

```
%Se carga los datos de longitud del escenario1 y se normaliza
load longitud.mat
min_l=min(longitud);
pos_l=longitud-min_l;
max_l=max(pos_l);
norm_l=pos_l/max_l;

%Se carga los datos de latitud del escenario1 y se normaliza
load latitud.mat
min_la=min(latitud);
pos_la=latitud-min_la;
max_la=max(pos_la);
norm_la=pos_la/max_la;

%Se carga los datos de radiación solar del escenario1 y se normaliza
load ghi.mat
min_g=min(ghi);
pos_g=ghi-min_g;
max_g=max(pos_g);
norm_g=pos_g/max_g;

%Matriz de vectores de entrada a la SOM
datosentrada=ones(103680,3)';
datosentrada(:,1)=norm_l;
datosentrada(:,2)=norm_la;
datosentrada(:,3)=norm_g;
```

8.4 Diseño de red neuronal SOM

```
CREACIÓN DE UNA RED NEURONAL SOM
%cargar los datos de entrada matriz 3x103680
load datosentrada.mat
x=datosentrada;
```

```

%creación de la SOM con los parámetros más importantes
net=selforgmap([6 6],100,3,'hextop','linkdist');

%entrenamiento no supervisado
net.trainFcn='trainbu';
%número de iterraciones
net.trainParam.epochs=1000;
%entreno la SOM
net=train(net,x);
%gráfica que relaciona los pesos 1 y 2
plotsompos(net,x)

%matriz de pesos de salida de la red
WSOM1=net.iw{1}

```

8.5 Transformación pesos de la salida SOM en coordenadas

```

%declaracion de vector de coord
%ponemos los valores minimos y maximos de latitud, longitud de los 36
%sensores distribuidos
lat1=39.1017;
lat2=39.0503;
lon1=-3.6633;
lon2=-3.7306;
diflat=lat1-lat2;
diflon=lon1-lon2;
coor_som1=ones(36,2);
%Pasamos a coordenadas con una relación lineal.
for i=1:36
coor_som1(i,1)=WSOM1(i,1)*diflon/1+lon2;
coor_som1(i,2)=WSOM1(i,2)*diflat/1+lat2;
end
%Se guarda en Excel
filename = 'coor_som1.xlsx';
xlswrite(filename,coor_som1,1,'A2')

```


8.6 Error cuantización SOM

```
%PESOS DE SALIDA SOM1
load WSOM1
%DATOS ENTRADA SOM
load datosentrada
%Calcula la distancia euclídea
Z=abs(dist(WSOM1,datosentrada));
Zmin=ones(1,103680);
for i=1:103680
    Zmin(:,i)=min(Z(:,i));
end
%error cuantización

sum(Zmin)./103680
```

8.7 División área de estudio en puntos distribuidos regularmente

```
lat1=39.102;
lat2=39.048;
lon1=-3.661;
lon2=-3.7312;
Q=45;
[Xint1,Xint2]=meshgrid(lon2:(lon1-lon2)/(Q-1):lon1,lat2:(lat1-lat2)/(Q-1):lat1);
for i=1:45
    plot(Xint1(i,:),Xint2(i:,:), 'r.')
    hold on
end
```

8.8 Función interpolación IDW

```

function Vi = gIDWbueno(Xc,Yc,Vc,Xi,Yi,w,r1,r2)
%cargo las coordenadas de los 36 sensores del escenario original
load escenario1
%cargo las coordenadas que se van a interpolar
load GRID2025
%cargo los datos de radiación solar para MES JUNIO
load datosjunio2015esc1
%coordenada x, longitudes
Xc=escenario1(:,1);
%coordenada y, latitudes
Yc=escenario1(:,2);
Vc=datosjunio2015esc1;
%coordenadas x de los puntos que se van a interpolar
Xi=GRID2025(:,1);
%coordenadas y de los puntos que se van a interpolar
Yi=GRID2025(:,2);
%peso de la distancia
w=-2;
%la estimación se va a hacer en cuanto a nodos vecinos, 'r' sería con radio
%de búsqueda
r1='n';
%elegimos el número de nodos a 36,
%sensores:36
r2=36;
%% OUTPUTS
% Vi:    (mandatory) [PxQ]      gIDW interpolated values
%
%                               --> P=1, Q=1 yields interpolation at one
%                               point
%
%                               --> P>1, Q=1 yields interpolation at a
%                               vector of points
%
%                               --> P>1, Q>1 yields interpolation at a
%                               (ir)regular grid of points
%% INPUTS
% Xc:    (mandatory) [Nx1]      x coordinates of known points
% Yc:    (mandatory) [Nx1]      y coordinates of known points

```

```

% Vc:    (mandatory) [Nx1]      known values at [Xc, Yc] locations
% Xi:    (mandatory) [PxQ]      x coordinates of points to be interpolated
% Yi:    (mandatory) [PxQ]      y coordinates of points to be interpolated
% w:     (mandatory) [scalar]    distance weight
%
%                               --> w<0, for Inverse Distance Weighted
%                               interpolation [IDW]
%
%                               --> w=0, for Simple Moving Average (only
%                               if neighborhood size is local and not
%                               global) [SMA]
%
% r1:    (optional) [string]     neighbourhood type
%
%                               --> 'n'      (default) number of neighbours
%
%                               --> 'r'      fixed radius length
%
% r2:    (optional) [scalar]     neighbourhood size
%
%                               --> number of neighbours, if r1=='n'
%                               default is length(Xc)
%
%                               --> radius length,          if r1=='r'
%                               default is largest distance between known
points
%% SYNTAX
% ===== IDW =====
% all inputs:
% Vi = gIDW(Xc,Yc,Vc,Xi,Yi,-2,'n',30);
% 6 inputs:
% Vi = gIDW(Xc,Yc,Vc,Xi,Yi,-2);
%     --> r1='n'; r2=length(Xc);
% 7 inputs:
% Vi = gIDW(Xc,Yc,Vc,Xi,Yi,-2,'n');
%     --> r2=length(Xc);
% Vi = gIDW (Xc,Yc,Vc ,Xi,Yi,-2,'r');
%     --> r2=largest distance between know points [Xi,Yi] (see D1 calculation)
% ===== SMA =====
% Vi = gIDW(Xc,Yc,Vc,Xi,Yi,0,'n',10);
% ===== Spatial Map =====
% Vi = gIDW(Xc,Yc,Vc,Xi,Yi,-2,'n',10);
%     -with Xi and Yi 2D arrays of coordinates relative to an (ir)regular
%     grid.
%INICIALIZAMOS LAS SALIDAS
Vi = zeros(size(Xi,1),2880);
D=[]; Vcc=[];

% CON r1 = 'r'

```

```

if strcmp(r1,'r')
    if (r2<=0)
        error('r2:chk','Radius must be positive!')
        return
    end
    wb = waitbar(0,mfilename);
    for i=1:length(Xi(:))
        waitbar(i/length(Xi(:)))
% if length(Xi(:))> 100, progress_bar(i, length(Xi(:)), mfilename); end
        D = sqrt((Xi(i)-Xc).^2 +(Yi(i)-Yc).^2);
        D = D(D<r2);
        Vcc = Vc(D<r2);
        if isempty(D)
            Vi(i) = NaN;
        else
            if sum(D==0)>0
                Vi(i) = Vcc(D==0);
            else
                Vi(i) = sum( Vcc.*(D.^w) ) / sum(D.^w);
            end
        end
    end
end
close(wb)

%con r2= 'n' METODO ELEGIDO
elseif strcmp(r1,'n')
    if (r2 > length(Vc) || (r2<1)
        error('r2:chk','Number of neighbours not congruent with data')
        return
    end
    wb = waitbar(0,mfilename);
    for i=1:length(Xi(:))
        for j=1:2880
            %waitbar(i/length(Xi(:)))
            if ~ mod (i, 1000); waitbar(i/length(Xi(:))); end
            %if length(Xi(:))>100, progress_bar(i, length(Xi(:)), mfilename);
            D=haversine(Yc,Xc,Yi(i),Xi(i))
            %D = sqrt((Xi(i)-Xc).^2 + (Yi(i)-Yc).^2);

            [D,I] = sort(D);
            Vcc = Vc(I,j);

```

```

    %Vcc=Vc(I);
    if D(1) == 0
        %Vi(i) = Vcc(1);
        Vi(i,j) = Vcc(1);
    else
        Vi(i,j) = sum( Vcc(1:r2).*(D(1:r2).^w) ) / sum(D(1:r2).^w);
        %Vi(i) = sum( Vcc(1:r2).*(D(1:r2).^w) ) / sum(D(1:r2).^w);
    end

end

end

close(wb)
end

return

```

8.9 Función haversine

```

function D= haversine(lat1,lon1,lat2,lon2)
%Calculo de la distancia usando haversine
R= earthRadius('km'); % 6371 (el radio es aprox)
dlat=deg2rad(lat2-lat1);
dlon=deg2rad(lon2-lon1);
lat1=deg2rad(lat1);
lat2=deg2rad(lat2);
a = (sin(dlat./2)).^2 + cos(lat1).*cos(lat2).*(sin(dlon./2)).^2;
c = 2.*asin(sqrt(a));
D=c*R;
End

```

8.10 Errores MAE y RMSE

```
%errores para mes de junio completo
%radiación solar real en los 10 puntos de validación
load valorrealpuntosval
%radiación solar estimada escenario original en los 10 puntos de validación
load estimadoescenario1
%radiación solar estimada SOM1 en los 10 puntos de validación
load estimadosom1
%radiación solar estimada SOM2 en los 10 puntos de validación
load estimadosom2
%radiación solar estimada SOM3 en los 10 puntos de validación
load estimadosom3
%radiación solar estimada SOM4 en los 10 puntos de validación
load estimadosom4

DIF0=valorrealpuntosval-estimadoescenario1;
DIF1=valorrealpuntosval-estimadosom1;
DIF2=valorrealpuntosval-estimadosom2;
DIF3=valorrealpuntosval-estimadosom3;
DIF4=valorrealpuntosval-estimadosom4;

DIF00=abs(DIF0);
DIF11=abs(DIF1);
DIF22=abs(DIF2);
DIF33=abs(DIF3);
DIF44=abs(DIF4);

MAE0=sum(DIF00(:))./28800;
MAE1=sum(DIF11(:))./28800;
MAE2=sum(DIF22(:))./28800;
MAE3=sum(DIF33(:))./28800;
MAE4=sum(DIF44(:))./28800;

DIF0CUA=DIF0(:, :).^2;
DIF1CUA=DIF1(:, :).^2;
DIF2CUA=DIF2(:, :).^2;
DIF3CUA=DIF3(:, :).^2;
```

```
DIF4CUA=DIF4(:, :).^2;

RMSE0=sqrt(sum(DIF0CUA(:))./28800);
RMSE1=sqrt(sum(DIF1CUA(:))./28800);
RMSE2=sqrt(sum(DIF2CUA(:))./28800);
RMSE3=sqrt(sum(DIF3CUA(:))./28800);
RMSE4=sqrt(sum(DIF4CUA(:))./28800);

ERRORES=ones(5,2);
ERRORES(1,1)=MAE0;
ERRORES(1,2)=RMSE0;
ERRORES(2,1)=MAE1;
ERRORES(2,2)=RMSE1;
ERRORES(3,1)=MAE2;
ERRORES(3,2)=RMSE2;
ERRORES(4,1)=MAE3;
ERRORES(4,2)=RMSE3;
ERRORES(5,1)=MAE4;
ERRORES(5,2)=RMSE4;
ERRORES
```

8.11 Visualizar estimación de la radiación solar

```
load Xint1
load Xint2
load F0
load F1
load F2
load F3
load F4
subplot(2,3,1);contourf(Xint1,Xint2,F0);title('ORG.10JUNIO.7.30am');
subplot(2,3,2);contourf(Xint1,Xint2,F1);title('SOM1.10JUNIO.7.30am');
subplot(2,3,3);contourf(Xint1,Xint2,F2);title('SOM2.10JUNIO.7.30am');
subplot(2,3,4);contourf(Xint1,Xint2,F3);title('SOM3.10JUNIO.7.30am');
subplot(2,3,5);contourf(Xint1,Xint2,F4);title('SOM4.10JUNIO.7.30am');
```


REFERENCIAS

- [1] Sitio web del INE, www.ine.es
- [2] Will, A., Bustos, J., Bocco, M., Gotay, J., & Lamelas, C. (2013). On the use of niching genetic algorithms for variable selection in solar radiation estimation. *Renewable Energy*, 50, 168–176.
- [3] Teke, A., Yıldırım, H. B., & Çelik, Ö. (2015). Evaluation and performance comparison of different models for the estimation of solar radiation. *Renewable and Sustainable Energy Reviews*, 50, 1097–1107.
- [4] Archila Córdoba, D. M., & Santamaría Buitrago, F. A. (2013). ESTADO DEL ARTE DE LAS REDES DE SENSORES INALÁMBRICOS. *Tecnología Investigación Y Academia*, 1(2). Recuperado a partir de <https://revistas.udistrital.edu.co/index.php/tia/article/view/4437>
- [5] Gascón, D. (2010). Redes de Sensores Inalámbricos, la tecnología invisible. *Tecnología y sociedad*, 53-55.
- [6] Beltrán, F., & Franco, C. (2007). Wireless sensor networks. *Bit*, 2(165), 61-64.
- [7] *Wireless sensor network survey*. Yick, Jennifer, Mukherjee, Biswanath y Ghosal, Dipak. 12, s.l. : Computer Networks, 2008, Vol. 52
- [8] Nayibe, C. C., Alexander, T. B. D., Cristina, A. Z. L., & Miguel, C. O. L. (2014). Redes de Sensores Inalámbricos. *Universidad Nacional de Colombia, Bogotá*.
- [9] Zimmerman, D., Pavlik, C., Ruggles, A., & Armstrong, M. P. (1999). An experimental comparison of ordinary and universal kriging and inverse distance weighting. *Mathematical Geology*, 31(4), 375-390.
- [10] Díaz Padilla, G., Sánchez Cohen, I., Quiroz, R., Garatuza Payán, J., Watts Thorp, C., & Cruz Medina, I. R. (2008). Interpolación espacial de la precipitación pluvial en la zona de barlovento y sotavento del Golfo de México. *Agricultura técnica en México*, 34(3), 279-287.
- [11] GisGeography. [Online] <https://gisgeography.com/inverse-distance-weighting-idw-interpolation/>
- [12] Dixon, B., & Uddameri, V. (2016). *GIS and geocomputation for water resource science and engineering*. John Wiley & Sons.
- [13] Kohonen, T. (1982). The Self-Organizing formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69, doi:10.1007/bf000337288
- [14] Kohonen, T. (1995). *The Self-Organizing Maps*. Springer-Verlag, Berlin/Heidelberg, Germany.
- [15] Azsalud. [Online] <https://azsalud.com/medicina/partes-funciones-neurona>
- [16] Kohonen, T (2001). *The Self-Organizing Maps*. Springer-Verlag.
- [17] Hernández, S. D., & Córdova, F. T. (2017). SIMULACIÓN DE UNA RED NEURONAL CON MAPAS AUTO-ORGANIZADOS (SOM) DE KOHONEN. *JÓVENES EN LA CIENCIA*, 3(2), 2650-2653.
- [18] Wikipedia. [Online] https://es.wikipedia.org/wiki/Mapa_autoorganizado
- [19] Weather Spark. [Online] <https://es.weatherspark.com/y/36810/Clima-promedio-en-Daimiel-Espa%C3%B1a-durante-todo-el-a%C3%B1o>
- [20] El periódico de la energía. [Online] (<https://elperiodicodelaenergia.com/las-10-mayores-plantas-fotovoltaicas-del-mundo/>).
- [21] Google [Online] https://www.google.com/intl/es_ES/maps/about/mymaps/
- [22] Rigollier, C., Lefevre, M. y Wald, L. (2004). The method Heliosat-2 for deriving shortwave solar radiation from satellite images. *Solar Energy* 77,159-169)

- [23] **Sitio web PVGIS**, <http://re.jrc.ec.europa.eu/pvgis/>.
- [24] **Sitio web CMSAF**, <http://www.cmsaf.eu/>
- [25] **Sitio web SSE**, <http://eosweb.larc.nasa.gov/sse/RETScreen/>
- [26] **Copernicus**, [Online] <https://www.copernicus.eu/es/sobre-copernicus>
- [27] **Marion Schroedter-Homscheidt, et al. soda-pro.** [Online] 14 07 2020. http://www.sodapro.com/documents/10157/326332/CAMS72_2015SC3_D72.1.3.1_2018_UserGuide_v1_201812.pdf/95ca8325-71f6-49ea-b5a6-8ae4557242bd.
- [28] **soda-pro.** [Online] <http://www.soda-pro.com/es/web-services/radiation/cams-radiation-service>
- [29] **Guide, M. U. S.** (2002). Neural network toolbox. *The MathWorks*.
- [30] **Giuliano Langella** (2020). Inverse Distance Weighted (IDW) or Simple Moving Average (SMA) INTERPOLATION (<https://www.mathworks.com/matlabcentral/fileexchange/27562-inverse-distance-weighted-idw-or-simple-moving-average-sma-interpolation>), MATLAB Central File Exchange. Retrieved August 1, 2020
- [31] **Fernández Rangel, M.** (2019). Estimación de la radiación solar mediante técnicas de geoestadística utilizando una red de sensores distribuidos.

